FORUM FOR PUBLIC HEALTH IN SOUTH EASTERN EUROPE

Programmes for Training and Research in Public Health

METHODS AND TOOLS IN PUBLIC HEALTH

A Handbook for Teachers, Researchers and Health Professionals

Editors:

Lijana Zaletel-Kragelj and Jadranka Božikov

MetaNET and FPH-SEE Project Coordinators:

Doris Bardehle, Luka Kovačić, Ulrich Laaser and Oliver Razum

Lage, 2010

Bibliographic information published by *Die Deutsche Bibliothek*

Die Deutsche Bibliothek lists this publication in the *Deutsche Nationalbibliografie*; detailed bibliographic data is available on the Internet at http://dnb.ddb.de

This publication has been supported by the German Academic Exchange Service (*Deutsche Academische Austauschdienst – DAAD*) with funds from the Stability Pact: MetaNET Project, including Academic Programmes for Training and Research in Public Health in South Eastern Europe (FPH-SEE).

Publisher: Hans Jacobs Publishing Company
Cover design: Alma Šimunec-Jović
Printed by: C.B. print, Samobor, Croatia
Number of copies: 200

ISBN 978-3-89918-176-0

CONTENTS

PREFACE

This is the sixth out of seven books planned to be published in a series as a support to teachers and trainers in teaching public health in South Eastern Europe. Originally planned to be on the internet platform only, the Forum for Public Health in South Eastern Europe (FPH-SEE) and the MetaNET project as its continuation together with the Hans Jacobs Publishing Company decided later to publish this training material also as hard copy books. The first four books were published with the support of FPH-SEE, and the last two with the support of MetaNET. Both projects are supported by the German Academic Exchange Service (DAAD - Deutsche Academic Austauschdienst) with funds from the Stability Pact for South Eastern Europe, provided by the German Ministry of Foreign Affairs.

We are proud that this book will be published on the 10th year of the Public Health Network in South Eastern Europe.

The book **Methods and Tools in Public Health** is a collection of 47 teaching modules in 5 chapters written by 53 authors from 11 countries. The teaching modules in this book cover areas of methods of studying population health, special epidemiological methods and methods of public health interventions, methods of planning and evaluation and modules as the supportive tools and technologies. Authors had autonomy in preparation the teaching modules, they were asked to present their own teaching/training materials with the idea to be as practical and lively as possible. The role of editors was to stimulate the authors in writing modules and to collaborate with them in editing the final version of the manuscripts in order to get them as much as possible to the planned format. By preparing and publishing this teaching/training modules authors and editors expect and wish to support and improve public health education and training of public health professionals.

The editors asked and encouraged authors to incorporate in their teaching modules exercises, tests, questionnaires and other practical forms of training. We will be thankful for any comments on use of them in everyday practice.

The next and the last book will be entitled "International Public Health".

You can find all volumes on the website of the Forum of Public Health: *http://www.snz.hr/ph-see/publications.htm*, and the volumes 4-6 on the open access Literature database of the University Bielefeld: *http://biecoll.ub.uni-bielefeld.de*.

Editors and Project coordinators:

Lijana Zaletel-Kragelj, Jadranka Božikov,
Doris Bardehle, Luka Kovačić, Ulrich Laaser and Oliver Razum

October 2010

Forum for Public Health in South Eastern Europe and Hans Jacobs Publishing Company in this series published the following books:

1. Vesna Bjegović and Doncho Donev (eds). HEALTH SYSTEM AND THEIR EVIDENCE BASED DEVELOPMENT. Lage: Hans Jacobs; 2004.
2. Silvia Gabriela Scîntee and Adriana Galan (eds). PUBLIC HEALTH STRATEGIES: A TOOL FOR REGIONAL DEVELOPMENT. Lage: Hans Jacobs; 2005.
3. Lidia Georgieva and Genc Burazeri (eds). HEALTH DETERMINANTS IN THE SCOPE OF NEW PUBLIC HEALTH. Lage: Hans Jacobs; 2005.
4. Doncho Donev, Gordana Pavleković and Lijana Zaletel-Kragelj (eds). HEALTH PROMOTION AND DISEASE PREVENTION. Lage: Hans Jacobs; 2007.
5. Luka Kovačić and Lijana Zaletel-Kragelj (eds). MANAGEMENT IN HEALTH CARE PRACTICE. Lage: Hans Jacobs; 2008.
6. Lijana Zaletel-Kragelj and Jadranka Božikov (eds). METHODS AND TOOLS IN PUBLIC HEALTH. Lage: Hans Jacobs; 2010.

All books can be found at: *http://www.snz.hr/ph-see/publications.htm*, and all modules included in volumes 4-6 on the open access Literature database of the University Bielefeld: *http://biecoll.ub.uni-bielefeld.de*.

In preparation:

7. INTERNATIONAL PUBLIC HEALTH to be published in 2011.

LIST OF AUTHORS

1. **Aleksandrova Silviya, MD, MB, Assistant professor**, Faculty of Public Health, Medical University of Pleven, Bulgaria
2. **Artnik Barbara, DMD, PhD, Teaching Assistant**, Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
3. **Bardehle Doris, MD, PhD, Professor**, Faculty of Health Sciences, University of Bielefeld, Germany
4. **Bilban Marjan, MD, PhD, Professor**, Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
5. **Bjegović Vesna, MD, PhD, Professor**, Institute of Social Medicine and Centre School of Public Health, School of Medicine, University of Belgrade, Serbia
6. **Božikov Jadranka, PhD, Professor**, Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia
7. **Bulat Petar, MD, PhD, Professor**, University of Belgrade, School of Medicine, Occupational Health Department, Belgrade, Serbia
8. **Burazeri Genc, MD, PhD**, Department of International Health, Faculty of Health, Medicine and Life Sciences, Maastricht University, the Netherlands
9. **Cucu Alexandra, MD, PhD, Lecturer**, Department of Public Health and Management, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania
10. **Donev Doncho, MD, PhD, Professor**, Institute of Social Medicine, Medical Faculty, University of Skopje, Republic of Macedonia
11. **Dunn William N, PhD, Professor,** Graduate School of Public and International Affairs, University of Pittsburgh, PA, USA
12. **Dyakova Mariana, MD, PhD, Associate Professor**, Faculty of Public Health, Medical University Sofia, Bulgaria
13. **Erlih Saša, BSc, MSc**, Institute for Comprehensive Development Solutions, Slovenia
14. **Eržen Ivan, MD, PhD, Assistant Professor**, Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
15. **Fink Rok, BSc, MSc candidate, Teaching Assistant,** Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia
16. **Galan Adriana,** Public Health and Health Management consultant, Institute of Public Health, Bucharest, Romania
17. **Gjorgjev Dragan, MD, PhD, Professor**, Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius, Skopje, Republic of Macedonia
18. **Grancharova Gena, MD, PhD, Associate Professor**, Faculty of Public Health, Medical University of Pleven, Bulgaria
19. **Gudeva Nikovska Dance, MD, MPH, Teaching Assistant**, Ministry of Health, Republic of Macedonia

20. **Imamović Enida, MD, DDM**, Public Health Institute of Federation B&H, Bosnia and Herzegovina
21. **Ivartnik Matej, MSc**, Regional Institute of Public Health Ravne na Koroškem, Slovenia
22. **Janković Slavenka, MD, PhD, Professor**, School of Medicine, University of Belgrade, Serbia
23. **Jevšnik Mojca, PhD, Senior Lecturer**, Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia
24. **Jocić Neda, MD, PhD, Professor**, Faculty of medicine, Novi Sad, Serbia
25. **Karaslavova Emilia, MD, PhD, Assistant Professor**, Department of social medicine and health management, Medical University Plovdiv, Bulgaria
26. **Kasapinov Blaško, MD, Teaching Assistant**, Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia
27. **Kendrovski Vladimir, MD, PhD, Assistant Professor**, Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius, Skopje, Republic of Macedonia
28. **Kern Josipa, PhD, Professor**, Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia
29. **Klavs Irena, MD, PhD, Assistant Professor**, Institute of Public Health of the Republic of Slovenia, Ljubljana, Slovenia
30. **Kragelj Anja, Student** of Andragogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
31. **Kukec Andreja, BSc, PhD candidate**, Teaching Assistant, Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
32. **Lazarevik Vladimir, MD, MPH, Teaching Assistant**, Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia
33. **Mašić Izet, MD, PhD, Professor**, Chair for Medical Informatics, Medical faculty, University of Sarajevo, Bosnia and Herzegovina
34. **Mateev Hristo, MD**, University Hospital "Lozenetz", Sofia, Bulgaria
35. **Milevska-Kostova Neda, MSc, MCPPM**, Centre for Regional Policy Research and Cooperation "Studiorum", Skopje, Republic of Macedonia
36. **Nikšić Dragana, MD, PhD**, Medical Faculty, University of Sarajevo, Bosnia and Herzegovina
37. **Nitescu Maria, MD, PhD, Lecturer**, Department of Hygiene and Ecology, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania
38. **Pavleković Gordana, MD, PhD, Professor**, Andrija Štampar School of Public Health, Medical School, University of Zagreb, Croatia
39. **Pekmezović Tatjana, MD, PhD, Professor**, Institute of Epidemiology, School of Medicine, University of Belgrade, Serbia

40. **Plesničar Andrej, MD, PhD, Teaching Assistant**, Faculty of Health Studies, University of Ljubljana, Slovenia
41. **Roshi Enver, MD, MPH**, Faculty of Medicine, University of Tirana, Albania
42. **Rotar Pavlič Danica, MD, PhD, Assistant Professor**, Department of Family Medicine, Faculty of Medicine, University of Ljubljana, Slovenia
43. **Savova Zaharina, PhD, Associate Professor**, Faculty of Public Health, Medical University, Sofia, Bulgaria
44. **Sidjimova Dobriana, PhD, Assistant Professor**, Faculty of Public Health, Medical University, Sofia, Bulgaria
45. **Sočan Maja, MD, PhD, Assistant Professor**, National Institute of Public Health, Ljubljana, Slovenia, Ljubljana, Slovenia
46. **Sović Slavica, MD, PhD candidate**, Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia
47. **Stikova Elisaveta, MD, PhD, Professor**, National Public Health Institute and Faculty of Medicine, University "Ss. Cyril and Methodius", Skopje, Republic of Macedonia
48. **Škerget Nataša, Student** of Pedagogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
49. **Tozija Fimka, MD, PhD, Associate Professor**, Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius-Skopje, Republic of Macedonia
50. **Vitcu Anca, PhD, Assistant Professor**, Faculty of Computer Science, University "Al. I. Cuza", Iasi, Romania
51. **Vračko Pia, MD, PhD candidate, Research Assistant, Teaching Assistant**, National Institute of Public Health, Ljubljana, Slovenia, Ljubljana, Slovenia
52. **Vuletić Mavrinac Gorka, PhD**, Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia
53. **Zaletel-Kragelj Lijana, MD, PhD, Associate Professor**, Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia

Chapter 1

METHODS OF STUDYING POPULATION HEALTH

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASUREMENT OF HEALTH AND DISEASE: AN INTRODUCTION**
Module: 1.1.1	**ECTS (suggested): 0.10**
Author(s), degrees, institution(s)	**Tatjana Pekmezović**, MD, PhD, Professor Institute of Epidemiology, School of Medicine, University of Belgrade, Serbia
Address for correspondence	**Tatjana Pekmezović** Institute of Epidemiology, School of Medicine, University of Belgrade, Serbia, Višegradska 26A, Belgrade, Serbia E-mail: pekmezovic@sezampro.rs
Keywords	Incidence rate, prevalence, cumulative incidence, mortality rate, case-fatality ratio, person-time concept, adjustment of rates
Learning objectives	After completing this module students should be able to: • describe and explain basic measures of health and disease such as incidence rate, cumulative incidence, prevalence, mortality rate, case-fatality ratio; • calculate specified rates and proportions; • understand and explain persons-time concept; • describe methods for rates adjustment and understand the principles and limitations of standardization; • understand epidemiological literature that uses and refers to the concepts outlined above.
Abstract	The epidemiological research is inquiring into the frequency of occurrence of states and events of health. The first-order focus needs to be on concepts pertaining to rates of occurrence. A distinction between prevalence (of states) and incidence (of events) is made. A population at risk must be defined clearly. Any measure of occurrence is impossible to interpret without a clear statement of the period during which the population was at risk and the cases were counted. Farther on rates, one distinguishes between the overall rate and specific rates will be made. This leads to the concept of adjusted and this, in turn, to that of mutually standardized rates.
Teaching methods	The teaching method recommended: • the introduction lecture relating to basic definitions and concepts; • the distribution of the literature to small group (3-4 students); • the guided discussion within each group and added explanations; • the distribution of exercises to each group; • overall discussion.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection equipment, internet connection, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written examination with calculation of rates.

MEASUREMENT OF HEALTH AND DISEASE: AN INTRODUCTION

Tatjana Pekmezović

THEORETICAL BACKGROUND

Introduction

The fundamental epidemiological measure is the frequency with which the events of interest (usually disease, injury, or death) occur in the population to be studied. The targets in epidemiological investigations are populations.

The frequency of event can be measured in different ways, and it can be related to different denominators, depending on the purpose of the research and availability of data (1).

Ratio, proportion, rate

Measures of health and disease include are ratios. The ratio is the value obtained by dividing one quality by another; for example, sex ratio (or male to female ratio) (2). We distinguish between proportions and rates:

1. The proportion is a type of ratio in whole the numerator is included in the denominator. The ratio of a part to whole can be expressed as a (2):
    - "vulgar fraction" (1/2),
    - as a percentage (50%), and
    - as a decimal (0.5).

2. The rate is a measure of the frequency of occurrence of a disease or other health-related events. The components of a rate are:
    - the numerator (number of events),
    - the denominator (the specific period in which events occur), and usually
    - a multiplier a power of 10 (10n).

A true rate includes the sum of time units of exposure for all people at risk (person-time concept). It is useful in small populations. In large populations, a mid-period population usually can be considered a good estimate of the average number of people at risk, for the outcome during the time period. The mid-period population, as approximation, is often used as the denominator (1). It is very important to underline that the population at risk must be defined clearly. All people who are not usually resident in that area, and those who are not at risk of the event under investigation, must be excluded from denominator (3). A difference between true rate and rate in a classical epidemiological sense is presented in a separate module in this book.

The rates usually have values less than 1, and decimals are awkward to think about and discuss. Therefore, rates are usually multiplied by a constant multiplier (either 100 or else: 1,000, 10,000, 100,000, 1,000,000) in order to make the numerator larger than 1 and therefore easier to discuss (1).

Types of measures of occurrence

According to concept of incidence and prevalence

The most frequent measures of occurrence of health-related events include incidence rate (IR), prevalence (PREV), cumulative incidence (CI), mortality rate (MR), and case-fatality ratio (CFR).

1. The incidence rate.
   The incidence describes the frequency of occurrence of new cases during the time period. The incidence rate (person-time incidence rate, also called incidence density) is the number of new occurrence of disease in the study population during the time period, divided by the sum of time that each person in the population remained under observation and free of disease. In other words, the denominator of incidence represents the number of people who are at risk for developing of disease. The incidence rate is direct indicator of risk of disease in a population investigated and it is a measure of efficiency of preventive measures (4). This measure is in details presented in a separate module in this book.
2. The cumulative incidence.
   The cumulative incidence is the proportion of people who become diseased during a specified period of time. Both numerator and denominator include only those individuals who at the beginning of the period are free from the disease and therefore are at risk of getting it. The cumulative incidence depends on the incidence rate and the length of the period at risk. The cumulative incidence (risk) and the incidence rate (person-time incidence rate) can be mathematically related (Equation 1):

$$CI = 1 - e^{(-I \times t)} \qquad \textbf{Equation 1.}$$

   CI = cumulative incidence
   I = person-time incidence rate
   t = length of follow-up

   Different methods of calculation of cumulative incidence are in details presented in a separate module in this book.
   The cumulative incidence is a useful approximation of incidence rate when the rate is low or when study period is short (5).
3. The prevalence.
   The prevalence is the proportion of the population affected by a disease at a given point in time. The proportion of population that has a disease at a point in time (P) and the rate of occurrence of new disease during a period of time (I) are closely related (Equation 2):

$$P = I \times t \qquad \textbf{Equation 2.}$$

   P = point prevalence
   I = incidence
   t = length of duration of disease

Prevalence doesn't show a risk. This measure is helpful in assessing the need for health care and the planning of health services (4). In the medical and public health literature, the word prevalence is often used in two ways:

- point prevalence: it is prevalence of the disease at a point in time;
- period prevalence: number of people have had the disease at any time during a certain period of time.

4. The mortality rate.
   The mortality rate is a number of deaths in a specified period of time in a specified population (a mid-period population). Mortality is a measure of risk of death in population and efficiency of preventive measures (5). The same principles mentioned in the discussion of incidence apply to mortality: for a rate to make sense, anyone in the group represented by the denominator must have the potential to enter the group represented by the numerator.
5. The case-fatality ratio.
   The case-fatality ratio is a number of deaths from a disease in a specified period of time, divided by number of diagnosed cases in the same period. The case-fatality ratio is a measure of the severity of disease and efficiency of treatment procedures (6). In other words, the case-fatality ratio is a percentage of people diagnosed as having a certain disease who die within a certain time after diagnosis.

According to different type of adjustment

There are three broad categories of measures according to different type of adjustment: crude measures, specific measures, and standardized measures.

1. Crude measures.
   The measures that apply to an entire population, without reference to any characteristics of the individuals in it are crude measures (for example, annual mortality rate from all causes of death in country).
2. Specific measures.
   Specific measures may be specific according to age, sex, cause or some other characteristic (for example, annual mortality rate from breast cancer in females).
3. Standardized measures.
   Standardized measures are very useful in case when we compare two populations with different age structure. In this way, effect of age as a confounding variable may be controlled. The essential of standardization is comparing the investigational populations with standard population with known age structure. The standard population is a hypothetical population, and choice of it depends on purpose of the analysis. For international comparisons, European or World standard populations are favoured (7).

   There exist two methods of standardization of epidemiological measures, direct and indirect:

   - in the direct method of standardization, the age-specific measures of two (or more) populations to be compared are applied to a reference

population known as the standard. This is done be multiplying each age-specific measures of a population to be compared by the number of persons in the corresponding age group of the standard population. This way, one derives the expected numbers of deaths that would have occurred in populations being compared. Dividing each of the total expected numbers by the standard population leads to the adjusted or standardized measures (8). The procedure is in details presented in a separate module in this book.

- indirectly standardized measures compare the actual number of events in an area with the expected number of events based on mortality measures of a standard population. This method is often used to look at differences in mortality rates, and is often referred to as standardized mortality ratio (SMR). The standardized mortality ratio is ratio of observed to expected number of deaths, expressed as a percentage. A SMR greater than 100 indicates that the observed number of deaths exceeds the expected number, and a SMR less than 100 indicates that the observed number of deaths is less than the expected number. It can also be used to look at other events such as, for example, hospital activity. The observed figures comes from the local area, and the expected from applying in the death rate in the standard population to the local population. The following steps were used to calculate the SMR:
  - find the age-specific death rates in the standard population;
  - find the age-specific populations in observed area;
  - calculate the expected deaths in each of the age groups by multiplying the population in area A by death rate in the reference population;
  - add up the number of deaths in each age group to get the total number of expected deaths.

  Indirect standardisation is more robust with small numbers and avoids the distortions caused by direct standardisation based on unstable age-specific rates (3,7).

The decision to use crude, standardized, or specific measures depends on the information that an investigator is trying to obtain or impart:

- crude measures represent the actual experience of the population and provide data for the allocation of health resources and public health planning. Although they are easy to calculate and widely used for international comparisons, the fact that the values may be confounded by differences between underlying population structures make any observed differences in crude measures difficult to interpret.
- specific measures are un-confounded by that factor and provide the most detailed information about the pattern of the disease in a population.
- standardized measures provide a summary value that removes the effect of the differences in population structure to allow for valid comparison between groups or over time. The actual value of the standardized measures is

meaningless, however, since it has been statistically constructed based on the choice of a standard.

Finally, depending on the nature of the information required, one or a combination of different measures can be chosen (9).

CASE STUDY: MORTALITY RATES FROM BREAST CANCER IN WOMEN IN TWO UNITS IN BELGRADE

Introduction

In two urban units in Belgrade, mortality rates from breast cancer (BC) in women were as presented in Tables 1 and 2.

Table 1. Mortality rates (Mt) from breast cancer (BC) in women from urban unit A in Belgrade.

Age group	No. of women	No. of deaths from BC	Mt/100,000
0-19	25,138	0	0
20-29	14,961	1	6.68
30-39	18,249	3	16.64
40-49	17,251	8	46.37
50-59	16,849	23	136.51
60-69	13,187	13	98.58
70+	9980	9	90.18
All ages	115,615	57	49.30

Table 2. Mortality rates (Mt) from breast cancer (BC) in women from urban unit B in Belgrade.

Age group	No. of women	No. of deaths from BC	Mt/100,000
0-19	6722	0	0
20-29	3545	0	0
30-39	5832	1	17.15
40-49	5173	3	57.99
50-59	4770	5	104.82
60-69	6485	7	107.94
70+	5554	9	162.04
All ages	38,081	25	65.65

Comparison of overall and age-specific mortality rates

First, we will show how overall and age-specific mortality rates from BC could be computed and compared.

Following questions could be posed:

1. Are mortality rates higher in units A or unit B?
2. Are there reasons for this situation?

3. How can the difference between age-specific and crude mortality rates be explained?
4. How could the problem of comparability be overcome?
5. What is essential in standardization?
6. How can the standard population be chosen?

In example mentioned above, we chose World population as a standard. Calculation of standardized mortality rates for BC in units A and B are summarized in Table 3.

Table 3. Calculation of standardized mortality rates for breast cancer in urban units A and B in Belgrade.

		Unit A		Unit B	
(1)	**(2)**	**(3)**	**(2 × 3)**	**(5)**	**(2 × 5)**
Age group	**Standard population**	**Mt/100,000**	**No. of expected deaths**	**Mt/100,000**	**No. of expected deaths**
0-19	40,000	0	0	0	0
20-29	16,000	6.68	1.07	0	0
30-39	12,000	16.64	1.97	17.15	2.06
40-49	12,000	46.37	5.56	57.99	6.96
50-59	9,000	136.51	12.29	104.82	9.43
60-69	7,000	98.58	6.90	107.94	7.56
70+	4,000	90.18	3.61	162.04	6.48
No. of all expected deaths			31.40		32.49

Computation of standardized mortality ratios (SMR)

In continuation, we can pose a question, what are the standardized mortality rates from BC in units A and B?

For answering to this question, we will use data from the following table (Table 4) (Adapted from Hennekens&Buring, 1987) (9):

Table 4. Computation of standardized mortality rates. Adapted from Hennekens &Buring, 1987 (9).

Age group	**Population**	**Mt/100,000**	**No. of expected deaths**	**No. of observed deaths**
(1)	**(2)**	**(3)**	**(2 × 3)**	**(4)**
10-19	74,598	12.26	9.14	10
20-29	85,077	16.12	13.71	20
30-39	80,845	21.54	17.41	22
40-49	148,870	33.96	50.55	98
50-59	102,649	56.82	58.32	174
60-69	42,494	75.23	31.96	112
Total	534,533		181.09	436

Final calculation is presented in Equation 3.

$$SMR_{10-69\,years} = \frac{436}{181.09} \times 100 = 241$$ **Equation 3.**

EXERCISE

Teaching methods for this topic, among others, would be included the distribution of different exercises in small groups of students and calculation and explanation of different measures of health-related events. All tasks are adapted from Gordis (6).

Task 1

In 1997, there were 39 cases of myocardial infarction in town A among people aged 50-54 years. The number of person-time was 515,212 in that age group. Calculate the incidence rate of myocardial infarction.

Task 2

A sample including 2368 women at the age group 70-74 years was selected from the population of town B. After examination, 80 were assigned the diagnosis of rheumatoid arthritis. Calculate the prevalence of this disease.

Task 3

Of 229,400 children born in a given region, 411 had one congenital malformation at birth. Which measure of occurrence of congenital malformation can be calculated? Calculate this.

Task 4

Assume that in a population of 100,000 persons, 20 have disease X. In one year, 18 people die from that disease. Calculate the mortality rate and case-fatality ratio. Explain why the same disease has low mortality rate and case fatality ratio?

Task 5

In a study in the country A, the frequency of stroke was measured in 228,525 women who were 30-45 years of age and free from coronary heart disease, stroke and cancer in 1997. A total of 546 stroke cases were identified in the 10 years of follow-up. Calculate cumulative incidence.

ASSESMENT OF STUDENTS (type of questions)

1. The incidence rate of disease is 5 times greater in women than in men, but the prevalence shows no sex difference. The best explanation is:
   The crude all-cause mortality rate is greater in women.
   A. The case-fatality ratio for this disease is greater for women.
   B. The case-fatality ratio for this disease is lower for women.
   C. Risk factors for developing the disease are more common in women.

2. Which of the following is a good measure of the severity of a acute disease:
   A. cause-specific death rate
   B. survival rate
   C. case-fatality ratio
   D. standardized mortality rate
   E. non of the above.

3. Age-adjusted death rates are used to:
   A. correct death rates for errors in the statement of age.
   B. determine the actual number of deaths that have occurred in specified age groups in a population.
   C. correct death rates for missing age information
   D. compare deaths in person of the same age group
   E. eliminate the effects of difference in the age distributions of populations in comparing death rates.

REFERENCES

1. Jekel JF, Elmore JG, Katz DL. Epidemiology, biostatistics and preventive medicine. Philadelphia: W.B. Saunders Company; 1996.
2. Last JM, editor. A dictionary of epidemiology. New York, Oxford, Toronto: Oxford University Press; 2001.
3. Dos Santos Silva I. Cancer epidemiology: principles and methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Mausner JS, Kramer S. Epidemiology - an introductory text. Philadelphia: W.B. Sounders Company; 1985.
5. Coggon D, Rose G. Barker DJP. Epidemiology for the uninitiated. London: BMJ Publishing Group; 1997.
6. Gordis L. Epidemiology. Philadelphia: W.B. Sounders Company; 2004.
7. Jensen OM, Parkin DM, Mc Lennan R, Muir CS, Skeet RG. Cancer registration. Principles and methods. Lyon: IARC; 1991.
8. Radovanovic Z, Jevremovic I, editors. Epidemiology. Practical sessions. Belgrade: School of Medicine, University of Belgrade; 2000.
9. Hennekens CH, Buring JE. Epidemiology in medicine. Boston/Toronto: Little, Brown and Company; 1987.

RECOMMENDED READINGS

1. Beaglehole R, Bonita R, Kjellstrom T. Basic epidemiology. Geneva: World Health Organization; 2007.
2. Dos Santos Silva I. Cancer epidemiology: principles and methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
3. Jekel JF, Elmore JG, Katz DL. Epidemiology, biostatistics and preventive medicine. Philadelphia: W.B. Saunders Company; 1996.
4. Last JM, editor. A dictionary of epidemiology. New York, Oxford, Toronto: Oxford University Press; 2001.
5. Rothman KJ, Greenland S. Modern epidemiology. Philadelphia: Lippincott Raven Publishers; 1998.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**PROBABILITY – BASIC CONCEPTS**
Module: 1.1.2	**ECTS (suggested): 0.10**
Author(s), degrees, institution(s)	**Jadranka Božikov**, PhD, Professor Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia
Address for correspondence	**Jadranka Božikov** Andrija Štampar School of Public Health, Rockefellerova 4 10000 Zagreb, Croatia E-mail: jbozikov@snz.hr
Keywords	Probabiliy, set theory, probability theory, Bayes' theorem
Learning objectives	After completing this module students should: • repeat or understand definition of the following terms: random event, probability of an event, elementary event, sample space. • use Venn's diagrams to show union, complement, and intersection of events. • use rules for summarizing and multiplication of probabilities i.e. computation of complex events' probabilities • understanding of the concept of conditional probabilities, proof of Bayesian theorem.
Abstract	Basic concepts of probabilities, theoretical background of sets theory, use of Venn's diagrams for probability presentation. Elementary and complex events, complementary probability, proof of Bayesian theorem.
Teaching methods	Teaching methods include introductory lecture, exercises, and interactive methods such as small group discussions. Students after introductory lectures first carefully read the recommended sources in age standardization. Afterwards they discuss standardization as method of controlling confounding with other students. In continuation, they in practice in groups of 2-3 students perform the procedure of direct standardization using the programme tool (e.g. MS Excel) on given data. At the end they compare and discuss their results.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: face to face teaching in lecture room; • equipment: computers with LCD projection; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaires (MCQ).

PROBABILITY – BASIC CONCEPTS

Jadranka Božikov

THEORETICAL BACKGROUND

Random event is any event that does not have to occur unconditionally in a given moment but can occur with some **probability** (1). Intuitively, probability is the measure of how likely an event is. It can be undersood also as frequency of occurrence of a phenomenon (or outcome) in a large number of attempts or during prolonged time. Thus, for example, we find that the probability of "head" as an outcome of a coin toss is equal to 1/2 (just like a "tail"). What we individually consider to be random (i.e. the outcome of "head" in coin toss), in large mass (after a large number of experiments) lose character of coincidence and behaves according a certain rules. With this interpretation we can define probability as a frequency (relative frequency) of the occurrence of an outcome or a set of outcomes. If we want to be more precise we will define it as the limit value (*limes*) of relative frequency (when it comes to very large, borderly endless number of experiments).

If the subject of our study is final population (final set) of n events or outcomes, and the event X is expected to happen m times we assign to it P(X) (P after Probability) (Equation 1).

$$P(X) = \frac{m}{n} \qquad \textbf{Equation 1.}$$

We say that m is the number of favourable events (i.e. outcomes of an experiment), n is the number of possible events (the probability is therefore the relative frequency of occurrence of some event or outcome). It is obviously true that :

$$0 \leq (X) \leq 1$$

Often we express the probability as percentage (relative frequency multiplied by 100). If $P(X) = 0$ we say that X is impossible event. X is assured event (must-happend) when $P(X) = 1.$

Random experiment is each process that results in one of several possible outcomes (1). This may be a coin toss (two possible and equally probable outcomes), but also gender of the next comming patient who may be a woman or a man with a certain probability, or outcome of treatment of some malignant diseases (possible outcomes are "died" or "cured"). Individual outcomes of such experiments are called **elementary events (**or **atomic events)**. All elementary events together make a **full (comprehensive) set of elementary events** also called **sample space**. This set is obviously a universal set for a particular experiment. Elementary events are mutually excluded (are disjoint or *disiunct*) and all together exchauste the entire sample space.

Terms elementary event and event should be distinguished. **Event** is defined as any subset of sample space. According to the definition each event is union of elementary events and

every elementary event is an event but not vice versa. Events, elementary and non-elementary, are sets and theory of sets is thus correspondingly applicable to the probability theory and Venn diagrams are used for the presentation of both (Examples 1 and 2).

Example 1.

Sample space (or set of all elementary events) in a coin toss has only two elements, "head" and "tail". If the coin has been tossed twice, then the sample space has four elements, i.e. U=(HH, HT, TH, TT), where H means "head" a T "tail."

Examples of the events are: "at least one head" equals to {HH, HT, TH}, "two same outcomes" {HH, TT}, etc.

If the experiment is a single 6-sided die roll, sample space is a set {1, 2, 3, 4, 5, 6}. Examples of events are: "even number" ({2, 4, 6}⊆{1 ,2, 3, 4, 5, 6}), "odd number" ({1, 3, 5}⊆{1, 2, 3, 4, 5, 6}), "number is less or equal to 3" ({1, 2, 3}⊆{1, 2, 3, 4, 5, 6}), "3 or 5" ({3, 5}⊆{1, 2, 3, 4, 5, 6}) etc.

Example 2.

In the case of coupling of laboratory animals with genotypes Bb and Bb, sample space is (BB, Bb, bb, bb). The event can be "descendant heterozygote", "descendant homozygote" or "descendant, dominant gene carrier."

Analoguous to sets operations we are defining union and intersection of events as well as the complement of an event. The union of two sets A and B is a set obtained by combining all the members of the sets A and B i.e. the union of two sets is the set of elements which are in either set (1). The union of two events is an event consisted of all elementary events belonging to either event. The intersection of two sets is the set of elements which are in both sets and analoguous the intersection of two events consists of all elementary events contained in both sets. Complement of an event A consists of all elementary events that do not belong to the event A. Symbolically this is shown in Figure 1.

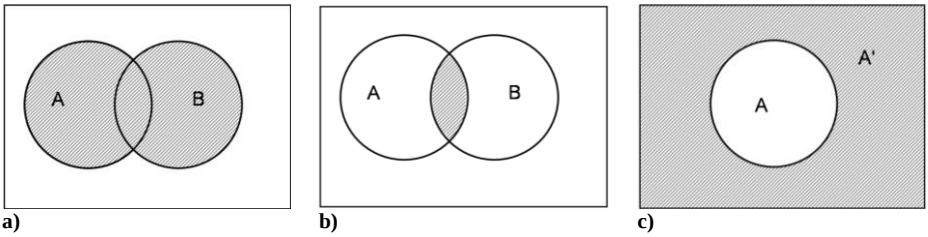


Figure 1. Venn's diagrams for ossible relations of events: a) union of events A and B, b) intersection of events A and B, c) complement of an event A'.

From the definition of complementary event arises: if the probability of an event, A is equal to $P(A)$, then the probability that the event would not occur (also called the complementary event A') is equal to $1-P(A)$ (Equation 2).

$$P(A') = 1 - P(A)$$ **Equation 2.**

This is easy to prove because event A' consists of all elementary events that do not belong to A and their number must be n-m (where m indicates the number of elementary events in A, and n total number of elementary events in a sample space), then (Equation 3):

$$P(A') = \frac{n-m}{n} = 1 - P(A)$$ **Equation 3.**

Following are Axioms and theorems of the probability theory. E_i, i=1,...,n denote elementary events.

Probability theory is based on three axioms (some of them we already mentioned defining probability) (Equations 4-6):

$$0 \le P(E_i) \le 1$$ **Equation 4.**

$$\sum_{i=1}^{n} P(E_i) = 1$$ **Equation 5.**

$$P(E_i \cup E_j) = P(E_i) + P(E_j)$$ **Equation 6.**

It reads:

1. Probability of each elementary event is greater or equal to 0 and less or equal to 1.
2. Sum of all probabilities of elementary events is 1. In other words this means that one of the elementary events must occur.
3. For every two elementary events valid is that probability of occurance of either one of them is equal to the sum of their probabilities.

From these axioms are derived the rules related to events in general (not necessarily elementary events):

Additivity

If A and B are **disjoint events** (there is no elementary event to realize and A and B in the same time), A∪B denotes new event that happend when either A or B occur and its probability is given with (Equation 7):

$$P(A \cup B) = P(A) + P(B)$$ **Equation 7.**

Consequently, for series of k mutually disjoint events A_i is (Equation 8):

$$P(A_1 \cup A_2 \cup A_3 \cup ... \cup A_k) = P(A_1) + P(A_2) + P(A_3) + ... + P(A_k)$$ **Equation 8.**

It is important to notice: it is valid only for mutually disjoint events. It is not valid generally for any events. Generally it is (Equation 9):

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$ **Equation 9.**

Multiplication

Suppose that two events A and B do not exclude each other but they may come at the same time. Occurrence of events A and B at the same time is a new event, realised by elementary events that are in the intersection of A and B. The probability of this new event is (Equation 10):

$$P(A \cap B) = P(A) \times P(B)$$ **Equation 10.**

Conditional probability

Often, we are interested in the probability of occurence of an event if another event previously occurred. If A and B are events then the probability of event A under the condition that event B occurred previously is called conditional probability and is denoted as P(A|B) what reads "probability of A given B", where by definition of probability (the ratio of the number of favourable outcomes and possible outcomes) gives (Equation 11):

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$ **Equation 11.**

In presentation of events by Venn diagrams we can imagine that area of an event (set) is appropriate to its probability (where sample space or complete set of elementary events is denoted by U and its probability is 1, i.e. P(U)=1) the above expression means that the probability of event A given B is equal to the ratio of area A∩B and the area of set B. In other words, it is the area of intersection A with B (A∩B) taken

relatively to area B, and it is equal to non-conditional probability P(A∩B) in case B is sample space (complete set of elementary events).

In order to better understand the concept of conditional probabilities consider the following example (Example 3):

Example 3.

Random sample of 1500 examinee of both genders were tested on colour blindness and results presented in Table 1 were obtained

Table 1. Results of testing on colour blindness between sexes.

		Sex		
		Males	**Females**	**Total**
Colour blindnes	Yes	65	10	75
	No	735	690	1425
	Total	800	700	1500

Example 3. cont.

Random sample of 1500 examinee of both genders were tested on colour blindness and results presented in Table 1 were obtained.
Events are: M (to be male) and S (to be colour blind).
We can calculate:

1. *The probability that someone is colour blind, if he is male:*
   P(S|M)=P(S∩M)/P(M)=65/800=0.08125(8.125%)
2. *The probability that someone is colour blind, if she is female.*
   P(S|M')=P(S∩M')/P(M')=10/700=0.01429(1.429%)

We can determine:

1. *What are complementary events to the above events:*
   P(S'|M)=1-P(S|M)=735/800=0.91875
   P(S'|M')=1-(S|M')=690/700=0.98571
2. *P(S) in the general population (probability of colour blindness in general population). However, we should notice that it depends on the population composition i.e. proportion of men/women in the population!*

It is clear that measures of test validity described in Module 2.4.1 in this book are conditional probabilities as follows (Equations 12 and 13):

$$sensitivity = P(O \mid B) = \frac{P(O \cap B)}{P(B)}$$ **Equation 12.**

$$specificity = P(O' \mid B') = \frac{P(O' \cap B')}{P(B')}$$ **Equation 13.**

Analogous the complementary probabilities to the sensitivity and specificity, false negative rate (FNR) and false positive rate (FPR) are (Equations 14 and 15):

$$FNR = P(O'|\, B) \qquad \textbf{Equation 14.}$$

$$FPR = P(O \,|\, B') \qquad \textbf{Equation 15.}$$

Bayes' theorem

Let's suppose that B_1, $B_{2,...,}B_k$ is a series of disjoint events. It is necessary to clarify that the particular sequence means that union of sets B_i, i=1,...,k exhausted or fill in the entire sample size i.e. $B_1 \cup B_2 \cup \ldots \cup B_k = U$, and since the B_i are mutually disjoint the sum of their probabilities is equal to 1, i.e. $P(B_1)+P(B_2)+...+P(B_k)=1$.

If O is an event $P(O|B_i)$ is called likelihood (usually these conditional probabilities are known $P(O|B_i)$, i=1,...,k) and then (Equation 16):

$$P(B_i \,|\, O) = \frac{P(O \,|\, B_i) \cdot P(B_i)}{P(O)} \qquad \textbf{Equation 16.}$$

where (Equation 17):

$$P(O) = P(O \,|\, B_1) \cdot P(B_1) + P(O \,|\, B_2) \cdot P(B_2) + ... + P(O \,|\, B_k) \cdot P(B_k) \qquad \textbf{Equation 17.}$$

This theorem was formulated by Thomas Bayes in 1764. This is only one form of Bayes' theorem (for one event O and there is a more complex form, too). Bayes' theorem shows the relation between two conditional probabilities which are the reverse of each other.

Confirmation is simple: according to definition of conditional probability (Equations 18 and 19):

$$P(B_i \,|\, O) = \frac{P(B_i \bigcap O)}{P(O)} \qquad \textbf{Equation 18.}$$

as well as

$$P(O \,|\, B_i) = \frac{P(O \bigcap B_i)}{P(B_i)} \qquad \textbf{Equation 19.}$$

From the above equations hence is follows that (Equations 20 and 21):

$$P(B_i \bigcap O) = P(B_i \mid O) \times P(O) \qquad \textbf{Equation 20.}$$

and

$$P(O \bigcap B_i) = P(O \mid B_i) \times P(B_i) \qquad \textbf{Equation 21.}$$

Since left sides of Equations 20 and 21 are equal (intersection operation is commutative i.e. $P(B_i \cap O) = P(O \cap B_i)$), it follows (Equation 22):

$$P(B_i \mid O) \cdot P(O) = P(O \mid B_i) \cdot P(B_i)$$
$$\Rightarrow$$
$$P(B_i \mid O) = \frac{P(O \mid B_i) \cdot P(B_i)}{P(O)} \qquad \textbf{Equation 22.}$$

This is statement that should be proven. It is evident that it is also valid (Equation 23):

$$P(O) = P(O \mid B_1) \cdot P(B_1) + P(O \mid B_2) \cdot P(B_2) + \ldots + P(O \mid B_k) \cdot P(B_k) \qquad \textbf{Equation 23.}$$

since B_i, i=1,…,k, are mutually disjoint events and exhausted out entire sample space.

That is why all $O \cap B_i$, i=1,…,k are mutually disjoint and exhausted out the entire O (since $O=O \cap (B1 \cup B2 \cup \ldots \cup Bk)=(O \cap B1) \cup (O \cap B2) \cup \ldots \cup (O \cap B_k)$ and P(O) is equal to the sum of probabilities $P(O \cap Bi)$, i=1,…,k) (Equation 24).

$$P(O) = P(O \bigcap B_1) + P(O \bigcap B_2) + \ldots + P(O \bigcap B_k) \qquad \textbf{Equation 24.}$$

If we put $P(O \cap B_i)=P(O|B_i)*P(B_i)$ we will get above form for P(O) as follows (Equation 25):

$$P(O) = P(O \mid B_1) \cdot P(B_1) + P(O \mid B_2) \cdot P(B_2) + \ldots + P(O \mid B_k) \cdot P(B_k) \qquad \textbf{Equation 25.}$$

With the use of Venn's diagram we can present this as shown in Figure 2 for k=4.

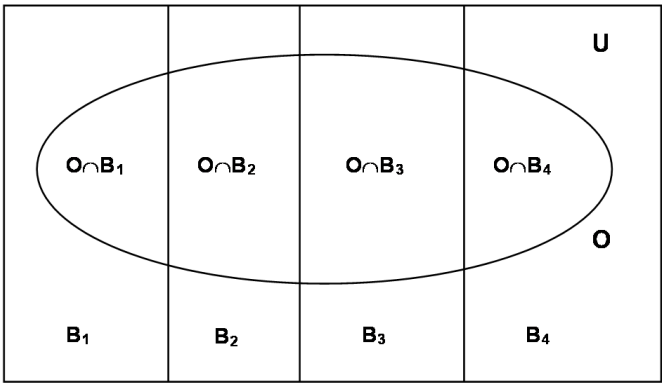


Figure 2. Venn diagram for Bayes' theorem.

For case k=2 there are two sets, B1 i B2 that are complementary, so we can write B1=B and B2=B', so Bayes' theorem in that simplest form looks like (Equation 26):

$$P(B \mid O) = \frac{P(O \mid B) \cdot P(B)}{P(O \mid B) \cdot P(B) + P(O \mid B') \cdot P(B')} \qquad \textbf{Equation 26.}$$

That is the form used as measure of test validity if marks: B mean sick, B' healthy, O positive test result, O' negative test result (see Module 2.4.1).

EXERCISE

Task 1

Try the following:

- check that additivity of probability for disjoint events is only special case of general additivity.
- calculate the probability of complex events, "get a number less or equal to 3" or "get even number" in the above-mentioned example with die roll.
- calculate the probability of event "birth of carrier of the dominant gene" for different combinations of genotypes of parents in the example mentioned above.

Task 2

If the probability of blue hair is 0.30 and the probability of black eyes 0.20. calculate the probability of occurrence of blue hair and black eyes.

Task 3

Please interpret the probabilities: P(O|B), P(O|B'), P(B) i P(B').

REFERENCE

1. Grinstead CM, Snell JL. Introduction to probability. Introduction to Probability, 2nd edition. Providence, RI: American Mathematical Society; 2003. Available from URL: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/pdf.html and www.astrohandbook.com/ch17/intro_probability.pdf. Accessed: May 20, 2010.

RECOMMENDED READINGS

1. Bayes' Theorem. Standford Encyclopedia of Phylosophy. Available from URL: http://plato.stanford.edu/entries/bayes-theorem/. Accessed: June 24, 2010.
2. Bayes' Theorem. Wikipaedia– the free encyclopaedia. Available from URL: http://en.wikipedia.org/wiki/Bayes'_theorem. Accessed: June 24, 2010.
3. Grinstead CM, Snell JL. Introduction to probability. Introduction to Probability, 2nd edition. Providence, RI: American Mathematical Society; 2003. Available from URL: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/pdf.html and www.astrohandbook.com/ch17/intro_probability.pdf. Accessed: May 20, 2010.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**ORGANIZING AND DESCRIBING DATA**
Module: 1.2.1	**ECTS (suggested): 0.40**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associated Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Data organization, data description, statistical distribution, typical value, ratio
Learning objectives	After completing this module students should: • know how to organize data for statistical and epidemiologic description; • be familiar with basic statistical description of data (frequency distribution, typical values of distribution); • be familiar with basic epidemiologic description of data, and • be aware of existence of different ratios, used in epidemiology.
Abstract	As in every other profession also in public health (PH) the research process (in this profession the research issue are different kinds of health problems of a population and their determinants) takes a very important role. Organizing and describing data is the very beginning of this process. The module is describing basic principles of statistical and epidemiologic description of data
Teaching methods	An introductory lecture gives the students first insight in characteristics of organization and description of data, statistical and epidemiologic. The theoretical knowledge is illustrated by three case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards they on provided data set in pairs perform two extensive tasks. They use computer programme to complete their exercise. They are stimulated to compare results with results of other pair and discuss the differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and statistical programmes (recommended SPSS); • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on analysis of a given data set.

ORGANIZING AND DESCRIBING DATA

Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Introduction

Why to organize and describe the data in public health

As in every other profession also in public health (PH) the research process takes a very important role. In PH the research issue is different kinds of health states/problems of a population (i.e. diseases, disabilities, injuries, deaths), and their determinants (1-6) Organizing and describing data is the very beginning of this process.

Research process in public health

The phases of this process are similar to other research processes in medicine (i.e. in clinical medicine, laboratory medicine etc.) and in fact represent very important part of total process of solving health problems. The phases are as follows (Figure 1):

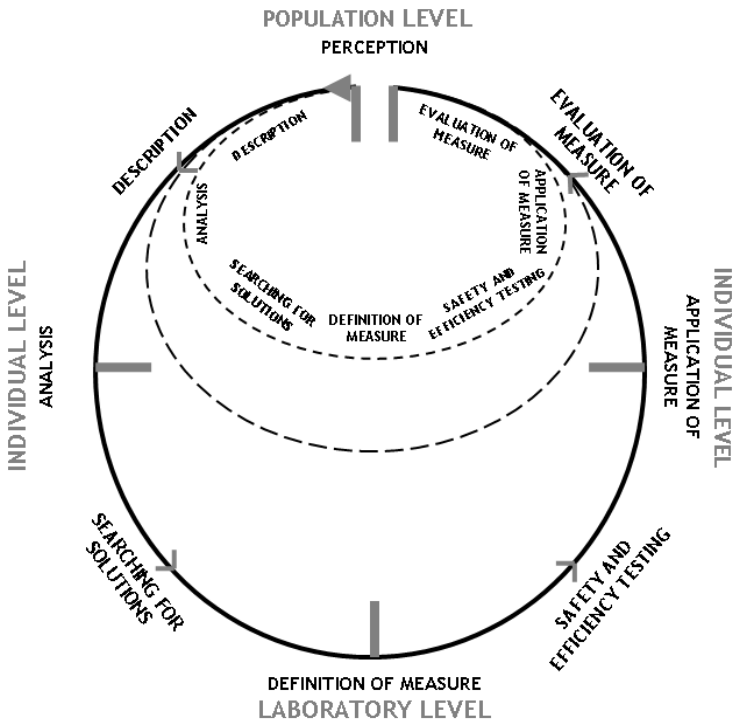


Figure 1. The levels of research process in public health.

1. perception of a health problem,
2. description of the problem,
3. analysis of the problem,
4. searching for possible measure for solving the problem,
5. the final definition of the measure,
6. testing safety and efficiency of the measure on human beings,
7. mass application of the measure on individuals,
8. observation of a long-term safety and efficiency of the measure.

In the first part of the process we are describing and analyzing the problem from the population level through level of an individual to the laboratory level (Figure 1) aiming at discovering the most appropriate measure for solving it. In the second part, first testing of safety and efficiency of the measure at the level of an individual before mass application takes its role, and afterwards the evaluation of efficiency at the population level. In solving some health problems the individual level could be skipped (in phenomena which could not be measured at an individual level like different kinds of environmental or community phenomena).

In process of organizing and describing of the data in PH research statistical methods take very important role. The relationship between PH, epidemiology and statistics is as follows:

1. PH is defined as one of the efforts organized by society to protect, promote, and restore the people's health. It is the combination of sciences, skills, and beliefs that is directed to the maintenance and improvement of the health of all the people through collective or social actions (1,4,7). One of these sciences also being one of important branches of medicine science itself is epidemiology.
2. Epidemiology in its broadest sense is defined as the study of the distribution of health states (different kinds of diseases or other phenomena related to the health of the people) and their determinants in specified populations, and its application to the control of health problems (1,4,7). Statistical methods represent one of the most powerful tools in epidemiology.
3. Statistics is defined as the science and art of collecting, summarizing and analyzing data that are subject to random variation (1). It is represented by a huge set of different methods adequate for different situations. Statistical methods take their role in (1,4):
    - description of health phenomena - descriptive statistics - and are used in descriptive epidemiology (activities to study occurrence of disease or other health-related characteristics in human population; it is concerned in where, when and how frequent such phenomena are), and
    - analyzing of health phenomena - the methods of analytical statistics - and are used in analytic epidemiology (usually concerned with identifying or measuring the effects of risk factors, or with the health effects of specific exposure).

Organizing data

Data matrix

For quality research it is of basic importance to have data well organized and prepared for both description and analysis. As the methods for both kinds of activities are statistical methods it is very important to follow the rules of preparing the data in an adequate structure for statistical analysis.

The appropriate structure is data matrix (1). This is the structure in which data of all observational units and all observed attributes of units are organized in a table (Figure 2). The basic element of this table is a cell. The cells are organized in a matrix with rows and columns. The meaning of elements of this table is as follows:

1. Cell – the record of a piece of information (lat. datum) on single attribute (variable) of a single unit of observation (statistical unit),
2. Row - the record of values of all variables for a single unit,
3. Column - the record of values of all units for a single variable.

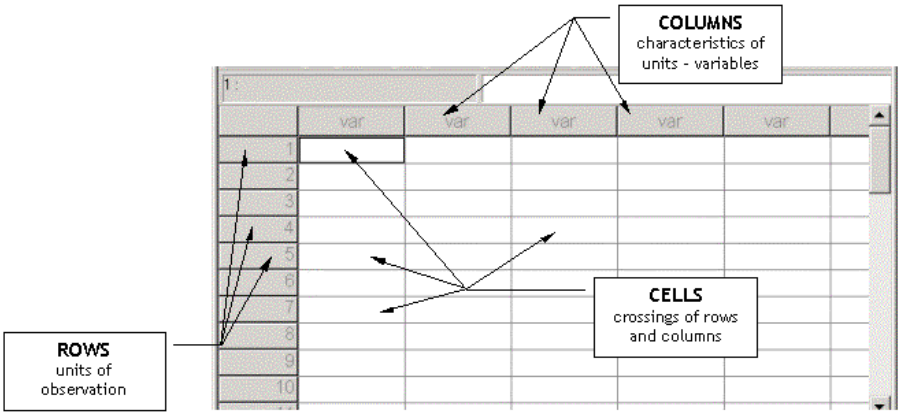


Figure 2. Organization of data for statistical description and analysis.

Statistical description of data

Overview of foundations of statistics

Basic statistical concepts

There exist four basic concepts in statistics (1,4,8,9):

1. Statistical population– the whole collection of units of phenomenon under study subjected to statistical methods,
2. Statistical unit – every single element (member) of statistical population,
3. Statistical variable – every single characteristic (attribute, phenomenon) of statistical unit under study,

4. Statistical sample – a selected subset of a statistical population; it may be random or non-random; may be representative or non-representative.

All these concepts are closely related to each other:

- when we are performing statistical observation of a certain phenomenon, the subject of interest is a whole mass of members, called statistical population,
- one single member of this mass is called unit of observation or statistical unit,
- the units have the attributes of their own. As these attributes can have different values (they vary), we call them variables,
- usually we cannot observe the whole population under study, so we draw a sample from the population. In that case we describe first the statistical features of the sample and then we generalize them to the population.

The relationship among basic concepts is also shown in Figure 3.

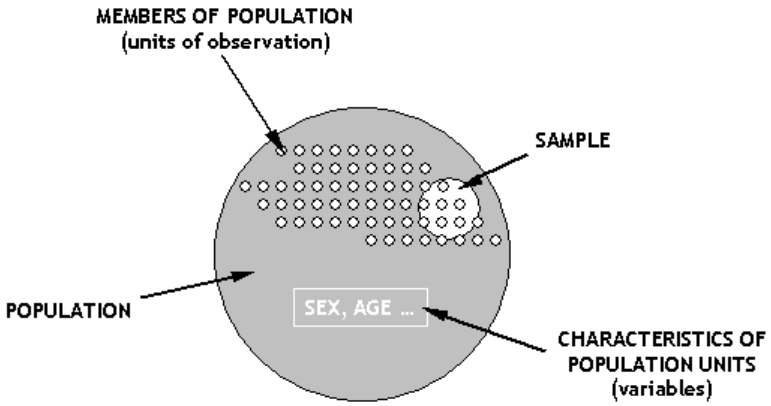


Figure 3. Basic concepts in statistics and the relationship among them.

The key concept in statistics is the concept of statistical variable or more precisely the concept of random variable or variate (1,10,11). According to Last et al., a variate is a variable that may assume any of a set of values, each with a pre-assigned probability (1).

Concepts, related to statistical activities

In statistics we can perform the following kinds of activities (8,9):

1. Statistical description – the process of summarizing the characteristics of data under study (at the sample or population level); we call this process descriptive statistics,
2. Statistical relationship analysis - the process of analysis of relationship between dependent (effect) and one or several independent (causes) variables (phenomena),

3. Statistical inference – the process of generalization from sample data to population, when the observation is not performed in a total population, but only in (representative) sample, usually with calculated degrees of uncertainity; we call this process inferential statistics.

If we observe the total population we perform only the methods of descriptive statistics. When only the sample is available we usually need to perform description and inference while relationship analysis could be performed in both situations.

Which methods are to be used depends on statistical features of variables under research.

Concepts, related to statistical variables

Statistical description and inference are closely related to the concept of statistical variable. Here we shall introduce some other concepts, also closely related to it.

Values of variables and their distribution. The first two important concepts are:

1. variable values – every single variable can take two or more different values,
2. distribution of variable values – the complete summary of the frequencies of the values of a single variable (some of the values are more frequent than the others); it can tell the number or the proportion of the whole group of observations to be of each value out of all observations.

Classifying variables. The variables, or more precisely their values, could have various statistical features. Regarding these features they could be classified in several ways (1,4,12):

1. Regarding the expression of their values to:
    - numerical variables – variables, values of which are expressed by numbers (e.g. weight, number of patients per day),
    - categorical (qualitative, attributable) variables or attributes – variables, values of which are expressed only by description (e.g. sex),
2. Regarding the possibility of infinitive number of their values to:
    - continuous variables – variables with potentially infinite number of possible values along a continuum (e.g. weight, height),
    - discrete variables – variables values of which could be arranged into naturally or arbitrarily selected groups of values (e.g. number of patients per day),
3. Regarding the ordinality of values to:
    - ordinal variables – variables values of which are classified into ordered categories (e.g. social class),
    - nominal variables - variables values of which are classified into unordered categories only by equality or inequality (e.g. race, religion, country),
4. Regarding the number of distinct values to:
    - dichotomous or binary variables – variables with only two possible values, often contain information of having the characteristic of interest or not,
    - polytomous variables – variables with more than two possible values,
5. Regarding the interrelationship between two or more variables to:

- dependent variables – variables values of which are depending on the effect of other variables (independent variables) in the relationship under study,
- independent variables – variables that are hypothesized to influence the values of other variables (dependent variables) under study.

All these classifications could be related to each other. When we put the classification on numerical and categorical variables in the central position and link it to all other classifications, then we get (Figure 4):

- numerical variables are continuous or discrete, only ordinal and polytomous and they could be dependent or independent,
- categorical variables are only discrete, dichotomous or polytomous, ordinal or nominal and they could be dependent or independent.

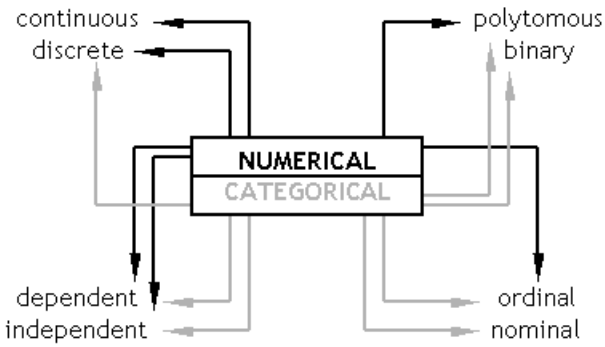


Figure 4. Various classifications of variables and the linkage of classification into numerical and attributable variables with all other classifications.

This linkage leads to formation of types of variables.

Types of variables. Usually we are classifying variables into four main types of variables (5,13):

1. Numerical continuous variables,
2. Numerical discrete variables,
3. Categorical ordinal variables, and
4. Categorical nominal variables.

This sequence of types of variables represents also their hierarchy regarding the amount of information encompassed in each of them. In the direction from numerical continuous to categorical nominal variables, the amount of information is decreasing. In this direction also the transformation from one type to another is possible but it is not possible in the opposite way (Figure 5).

What does the amount of information means and what means reducing it, would be more understandable in an example presented in Case study 2.

In the phase of planning and designing the study it is very important to be aware that the information could be reduced anytime but could never become more precise unless we acquire it once again.

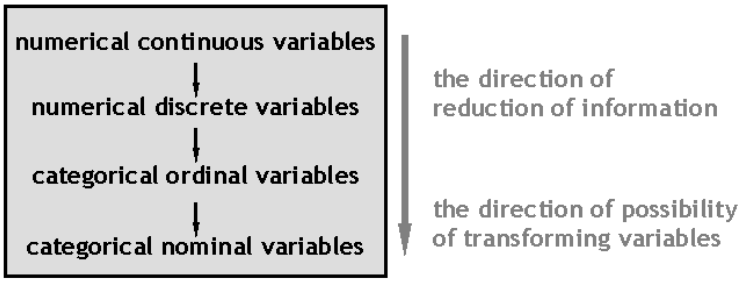


Figure 5. Hierarchy of types of variables regarding the amount of information encompassed in each of them.

Classifying the variables into right types is very important for deciding which method is to be used in describing variables, and examination of the relationship between them.

Concepts related to probability distributions

Some of variable values are more frequent than the others – they are more probable. The way how frequent the values of the particular variable are is called probability law of the variable. The distributions of values of variables are therefore called probability distributions. According to Last et al., probability distribution for a discrete random variable is the function that gives the probabilities that the variable equals each of a sequence of possible values, while for a continuous random variable is often used synonymously with the probability density function – the frequency distribution of a continuous random variable (1). We could roughly classify probability distributions into two groups:

1. Empirical probability distributions – distributions observed in real situation,
2. Theoretical (mathematical) probability distributions – mathematical idealization of distributions observed in real situations.

   By far the most important theoretical probability distribution is known as Normal or Gaussian distribution (1,4,8,9,11,14).

   Other also important theoretical distributions (all of them are families of similar distributions, varying with regard to the number of observations) are (1,4,8,9):

   - Student's t distribution,
   - binomial distribution,
   - chi-square distribution,
   - Poisson distribution,
   - Fisher's F distribution.

Theoretical probability distributions are very important as many statistical methods are based on the assumption that the observed data are a sample drawn from a population with known distribution. If such assumption is reasonable (it could never be checked whether it is true) the use of statistical methods becomes simple. But here we have to warn that if the assumption of distribution under study is not reasonable and we proceed with the activities, we could make the misleading conclusions.

Process of statistical description of data

Statistical description of data is a set of consecutive procedures used for describing the empirical distributions in an agreed way. The result of these procedures is:

1. Description of a shape of distribution, and
2. Determination of measures which summarize the features of the shape.

When statistically describing data we can choose between methods that make assumptions on theoretical probability distributions, called parametric methods (the origin of this term will be discussed later) and those which make no such assumptions, called non-parametric or distribution-free methods (8,9,14).

Presenting Data

Presentation of data could be numerical or graphical.

1. Numerical data presentation - ordered series and frequency distribution.
   The very first step in describing data statistically is to put data in order by making first an array and then a frequency distribution table:
   - ordered series or an array (10) - arrangement of values of a variable in order, usually from the lowest to the highest value,
   - frequency distribution - we summarize the frequency of every single value of a variable in ordered series in a table in which we usually insert two kind of frequencies:
     - absolute frequency called usually simple frequency - the number of units with particular value of a variable,
     - relative frequency - ratio between the number of units with particular value of a variable and all units under study; it could be expressed as a proportion (decimal fraction) or as a percentage (different ratios are discussed later in this chapter),
     - in some statistical programs also the third kind of frequency could be found, called cumulative frequency. This is the number of units with values less than or equal to each value. It could be expressed also as a relative measure - relative cumulative frequency.

   The example of a frequency distribution is given in Case study 2.
2. Graphical data presentation.
   The frequency distribution table could be useful for determination of some of data distribution features like the lowest and the highest value and thus also of the range of values, but not for all of them. The graphical data presentation is thus obligatory.

Basic graphical presentation of data in statistics is a chart of bars organized in such a way that values (categories) of a variable are listed along the x axis of the chart and their frequencies (absolute or relative) along the y axis. The area of every single bar is proportional to the frequency of the value it represents. It could be divided in two main forms (1,4,9,11):

- ordinary bar chart – the bars are lying separately; it is mostly used for presentation of attributable data,
- histogram – the bars are connected one to another; it is used for presentation of frequency distributions of numerical data; if there are many different values (continuous data) it is desirable to group observations before constructing a histogram in order to get a better visual impression of the observed distribution. An example of histogram is given in Case study 2.
- if we connect the centres of bars of a histogram at their upper part, we get the polygon called frequency polygon. When the bars are very numerous and very narrow (continuous data arranged in very small intervals) we can smooth the polygon. So we get the curve called probability density curve.

For understanding of the principles of statistical methods, it is the most important: when the relative frequency (a proportion or percentage) is used in graphical presentation of a distribution the sum of areas of all bars equals to 1 or 100%, regardless the type of bar chart is used – the ordinary one or histogram. Also the entire area under every probability density curve equals to 1 or 100%.

Data presentation by graph shows us clearly the shape of the distribution under study. This step of data presentation is very important for deciding which statistical methods are to be used for statistical description or/and inference in numerical variables.

Describing a distribution

When the graphical presentation of the shape of a distribution is done it should be described. The shape itself depends on a number and features of the place of highest density (peak). We say that distributions have diverse statistical features. Regarding these features they could be classified in several ways (4,8,9,11,14):

1. Regarding the number of peaks to:
    - unimodal – distributions with a single peak,
    - bimodal – distributions with two peaks,
    - polymodal – distributions with more than two peaks,
2. Regarding the shape of the peak to:
    - bell shaped – distributions in which extreme values tend to be less likely than values in the middle of the ordered series,
    - uniform – distributions in which all values have the same frequency,
3. Regarding the symmetry to:
    - symmetrical,
    - asymmetrical,

4. Regarding the inclination of the peak or skewness (when the distribution is not symmetrical) to distributions with:
    - positive skewness – distributions with an extended right hand tail (lower values more likely),
    - negative skewness – distributions with an extended left hand tail (higher values more likely),
5. Regarding the flatness or peakedness in symmetrical distributions to:
    - platykurtic – distributions with more flat peak than in normal distribution,
    - mesokurtic – distributions with the similar flatness of peak as in normal distribution,
    - leptokurtic – distributions with higher and more slim peak than in normal distribution.

Usually we are the most interested in first four features.

By representing the distribution graphically we would like to get the impression if the empirical distribution under study is similar the normal distribution which is unimodal, bell shaped and symmetric. If it is case, then in determination of the measures which summarize the features of the shape parametric methods for statistical description will be used, otherwise the non-parametric ones will be used.

Summarizing the distribution features

When describing the distribution of a numerical variable, continuous or discrete, we summarize its features also by special summary measures called typical values or measures of location of a distribution or shortly measures of location (4,8,9,11,14).

1. Types of typical values.
   The most well known typical values are the following ones (4,5,8,9,11,14):
    - measures of central tendency – the term includes several characteristics of the distribution of sets of values at or near the middle of the set; the principal measures of central tendency are:
        - mean (average) – the sum of values of a variable for each observation, divided by the number of observations,
        - median – a point in the ordered series which divides it into two parts of equal number of units, half of them falling below and half above this point,
        - mode – the most frequent value in the set of observations,
    - measures of dispersion or variation or spread of units around the centre of the distribution:
        - minimum and maximum - the lowest and the highest value of a distribution,
        - range – the difference between the minimum and the maximum,
        - variance – sum of the squares of the deviations from the mean, in population divided by the number of observations,
        - standard deviation – positive square root of the variance.
        - subgroups, based on an array, with equal number of units; in any case the number of quantiles is one less then the number of

corresponding equal parts; centiles are dividing the ordered series to hundredths and there are ninety-nine of them, deciles are dividing it to tenths and there are nine of them, quartiles are dividing it to quarters and there are three of them (median is also a quantile, dividing an ordered series to halfs); for describing the spread we usually use quartiles (1^{st} and 3^{rd}) or certain centiles (25^{th} and 75^{th}),

We could classify typical values also in parametric and non-parametric ones:

- parametric typical values – (measures that are basing on normal distribution) mean as a measure of central tendency and variance and standard deviation as measures of dispersion are called parametric measures,
- non-parametric typical values – not basing on theoretical distributions.

Which set of typical values is the most appropriate for certain distribution is to be decided after observing the shape of the distribution shown by the histogram. The decision should be made not only on the shape of the distribution but also on the number of observations and whether the inferential methods would be performed. The summary about possible decision in some typical situations is shown in Table 1. An example of presentation of typical values is presented in Case study 2.

Table 1. Which typical values could be chosen in some typical examples of distributions.

Shape of distribution	**Other important characteristics**	**Typical values**	
		Measure of central tendency	**Measure of dispersion**
Symmetrical or almost symmetrical Bell shaped		Mean	Standard deviation Minimum and maximum
Slightly asymmetrical Bell shaped	Large number of units	Mean	Standard deviation Minimum and maximum
	Small number of units	Median	Quartiles Minimum and maximum
Strongly asymmetrical	Only description	Mode	Minimum and maximum
	Inference planned	Median	Quartiles Minimum and maximum

2. Typical values in populations and samples.
We can perform statistical description in populations as well as in samples. So we determine the typical values as the summary measures at both levels. Here we have to emphasize that typical values at the sample level are not the same as the typical values at the population level. In fact, in a process of inferential statistics we infer from sample characteristics to population characteristics from which the sample was drawn, that means that we infer from the values of statistics to the values of parameters. To distinguish these measures between both levels we have different

names for them, and also the labelling is different. Some selected representatives of typical values are (Figure 6):

- statistics - typical values in samples:
  - mean (average) of a sample, usually labelled as " $\bar{X}$ ",
  - standard deviation of a sample, usually labelled as "s",
  - proportion of a sample, usually labelled as "p",
- parameters - typical values in populations:
  - mean (average) of a population, usually labelled as "μ",
  - standard deviation of a population, usually labelled as "σ",
  - proportion of a population, usually labelled as "Π".

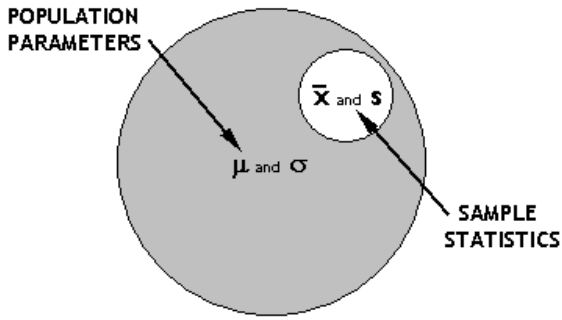


Figure 6. Labeling of some typical values in populations and samples.

Strict distinguishing between statistics and parameters is a basis for understanding the methods of statistical inference.

Epidemiologic description of data

Mathematical foundations of epidemiologic measurement

Basic tool for any kind of epidemiologic observation or research is quantification of frequency of health phenomena. In principle it is very similar to statistical process, but in epidemiologic measurement the emphasis is on discrete type of data, usually binary (e.g. disease is present or not, people are exposed to the certain risk factor or not) (3-5,7,15).

The frequency of a binary event could be expressed as:

- an absolute frequency or
- a relative frequency.

In public health, both, absolute and relative frequency measures convey important information, although relative measures seem to be frequently used. The probable reason is that relative frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations etc.), while absolute

measures are important in health care planning (e.g. number of hospital beds needed for treatment of certain group of health states).

Relative frequency measures are defined as a ratio between two data. In its broadest sense the ratio is a result of dividing one quantity by another (1,15,16). One quantity is representing a numerator and the other a denominator in this relationship. The term »ratio« is a general term of which rate, proportion, percentage, etc., are subsets (1,15). The numerator and denominator need not be related (5). One of the most important features is if the numerator is included in the denominator in calculation of the ratio. Regarding the relationship of the numerator and the denominator there exist different types of ratios which could be grouped in two main groups (Figure 7) (1,5,15,16):

1. Ratios in which the numerator is included in the denominator:
    - proportion,
    - proportion with multiplier (e.g. percentage in which multiplier is 100), and
    - rate in epidemiologic sense or epidemiologic rate.
2. Ratios in which the numerator is not included in the denominator:
    - rates as »true rates«,
    - ratios in a narrow sense.

The difference between different ratios will be presented using the same set of data in Case study 3.

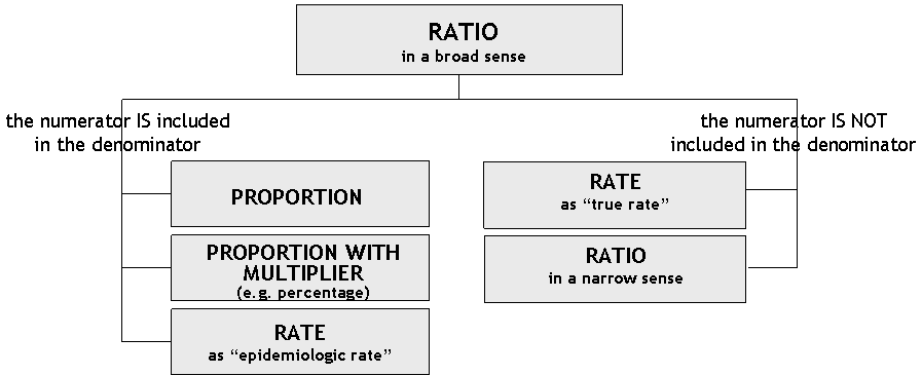


Figure 7. Types of relative frequency measures in epidemiology.

Types of relative frequency measures in epidemiology

Ratios in which the numerator is included in the denominator

1. Proportion.

   Proportion is the most simple relative frequency measure (17). It is the ratio of a part to the whole (1,5,17). According to Last et al., the important difference between a proportion and a ratio is that the numerator of a proportion is included in the denominator, whereas this is not necessarily so for a ratio (1). Proportion is calculated by using following equation (Equation 1):

$$Proportion = \frac{N_{events}}{N_{total}}$$ **Equation 1.**

N_{events} = number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)

A proportion could be expressed as a »vulgar fraction« (e.g. ½) or as a »decimal fraction« (e.g. 0.5) (1).

By definition, a proportion (p), if decimal, must be in the range $0 \leq p \geq 1$ (1).

A proportion is dimensionless since numerator and denominator have the same dimension, obtained through algebraic cancellation (1). If numerator and denominator are based upon counts (e.g. in our dataset), the originals are also dimensionless.

Calculating of this kind of relative frequency measure is presented in Case study 3.

2. Proportion with a multiplier.

A proportion could be multiplied by a factor K (1,5,17). A multiplier is usually a power of 10 (100, 1.000, 10.000…). Its role is mainly to convert the decimal fraction to a whole number. It is calculated by using following equation (Equation 2):

$$Proportion_{with\ a\ multiplier} = \frac{N_{events}}{N_{total}} \times K$$ **Equation 2.**

N_{events} = number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)
K = multiplier (100, 1000, 10.000, 100.000…)

Which multiplier is to be used depends on a given situation (e.g. in World Health Organization Health for All Database multiplier 100.000 is mostly used).

Typical representative of this kind of relative frequency measures is a percentage, in which multiplier is 100 (17) (Equation 3):

$$Percentage = \frac{N_{events}}{N_{total}} \times 100$$ **Equation 3.**

N_{events} = number of events of observed phenomenon (part of a whole)
N_{total} = number of all possible events of observed phenomenon (a whole)

Calculation of this kind of relative frequency measure is presented in Case study 3.

3. Rate in a classic epidemiologic sense.
   Before trying to explain what means the term »rate in a classic epidemiologic sense« we need to discuss the term »rate« itself. To a non-epidemiologist, rate means how fast something is happening or going, for example, the speedometer of a car indicates the car's speed or rate of travel in miles or kilometres per hour (5). This rate is always reported per some unit of time. Consecutively, some epidemiologists restrict use of the term »rate« to similar measures that are expressed per unit of time. For these epidemiologists, a term »rate« describes how quickly disease occurs in a population. These measures convey a sense of the speed with which disease occurs in a population (5). But this kind of ratio is a ratio in which the numerator is not included in the denominator and it will be discussed later.

   Other epidemiologists use the term »rate« more loosely, referring to proportions with case counts in the numerator and size of population in the denominator as rates (5). For this kind of ratios we are using the term »rate in a classic epidemiologic sense«. If these rates are referring to a specified period of time, they are calculated as a proportion with multiplier and specified period of time as a compulsory element by using following equation (Equation 4):

$$Rate_{epidemiologic} = \frac{N_{events(in\ a\ specified\ time\ period)}}{N_{total}} \times K \qquad \textbf{Equation 4.}$$

N_{events}= number of events of observed phenomenon (part of a whole) in a specified time period
N_{total} = number of all possible events of observed phenomenon (a whole) – at risk for occurrence of the event at the beginning of a specified time period
K = multiplier (100, 1000, 10.000, 100.000…)

An example of this kind of rate is an incidence rate in a classic epidemiologic sense. This measure will be in details discussed in a separate module of this book.

In this loose usage the time component is not always referring to a period of time in which the outbreak of new cases of health phenomenon under observation is followed-up. Sometimes is referring to a number of cases in a specific point in time (Equation 5):

$$Rate_{epidemiologic} = \frac{N_{cases(in\ a\ specified\ point\ of\ time)}}{N_{total}} \times K \qquad \textbf{Equation 5.}$$

N_{cases}= number of events of observed phenomenon (part of a whole) in a specified point of time
N_{total} = number of all possible events of observed phenomenon (a whole) in a specified point of time
K = multiplier (100, 1000, 10.000, 100.000…)

An example of this kind of rate is a prevalence rate. This measure will also be discussed in details in a separate module of this book.

Calculation of this kind of relative frequency measure is presented in Case study 3.

In epidemiology, and especially in vital statistics, this kind of measures are essential for comparing health phenomena between different populations (1,12).

Ratios in which the numerator is not included in the denominator

In ratios in which the numerator is not included in the denominator could be of two kinds. First are those in which the numerator and denominator are completely different variables (5). For this kind of ratios we use the term »true rate«. In others, the numerator and denominator are different categories of the same variable. We use the term »ratio in a narrow sense« for this kind of ratios.

1. True rate.
   This kind of rates refers to ratios representing changes in two quantities, where the two are separate and distinct quantities. In its precise usage a rate is the ratio of a change in one quantity to a change in another quantity, with the denominator quantity often being time (18,19). A classic example of a rate is velocity, which is a change in location divided by a change in time. Dimensionality of this kind of ratio is obtained through combination of dimensions of the numerator and the denominator (e.g. km/h). In epidemiology a representative of this kind of ratio is for example so called incidence rate as a true rate. The detailed description if this measure is out of the scope of this module. It will be presented in a separate module of this book.
2. Ratio in a narrow sense.
   There exist also ratios that could not be classified in none of the previously presented ratios. For example, in epidemiology ratio in which the numerator and denominator are different categories of the same variable is rather frequent kind of measure. It could be simply the ratio between males and females, or persons 20-29 years and 30-39 years of age (5). The other example is a ratio in which we are relating events of an observed health phenomenon to non-events (ratio between the number of people with observed phenomenon and the number of people without it) (Equation 6):

$$Ratio = \frac{N_{events}}{N_{non\text{-}events}} \qquad \textbf{Equation 6.}$$

N_{events}= number of events of observed phenomenon (part of a whole)
$N_{non\text{-}events}$= number of non-events of observed phenomenon (part of a whole)

Calculation of this kind of relative frequency measure is presented in Case study 3.

Important considerations in epidemiologic research

When we are observing the frequency of specified health phenomenon, we have first precisely to define:

1. If the study is cross-sectional or longitudinal.
   Cross-sectional study examines the phenomena in a point of time or very short period of time (e.g. a couple of weeks) while longitudinal examines it over a long period. In cross-sectional studies we are usually studying the frequency of all cases of observed phenomenon, while in longitudinal the frequency of new cases (1,4,5,7,15,17).
2. Which quantity represents the numerator and which the denominator in the equation and if the numerator is included in denominator.
3. What is the unit under observation. In epidemiology it is not necessary that the unit of observation is a person, it could be for example an episode of a health state. One should be aware in interpretation. Frequently health indicators are measuring health care services load (that is dependent also on health care services availability and accessibility, and health care services use demands of the population) and not a burden of disease in the population

Some important epidemiologic concepts

Outcome and exposure

First two important concepts are the concept of "outcome", and the concept of "exposure".

1. Outcome.
   Outcome is any possible disease or other health phenomenon or event related to health. It is a result of influence of an exposure to another phenomenon (1,4,5,7).
2. Exposure.
   Exposure is a process by which an agent (risk factor) comes into contact with a person, and provoke the relevant outcome, such as a disease (1,4,5,7).

Case, control and cohort

Other important concepts are the concepts of "case", the concept of "control", and the concept of "cohort".

1. Case (case-patient).
   In epidemiology a case is mostly defined a person identified as having the health condition under observation (1,4,5,7).
2. Control (control person).
   Controls are a group of persons with whom comparison is made in certain types of epidemiologic studies (e.g. in »case-control« studies and »randomized clinical trials«) (1).
3. Cohort.
   In it broad sense, the tem »cohort« describes any designated group of persons followed over a period of time, as in a »cohort study« (1). We distinguish two types of cohorts in this sense (1):

- fixed cohort – in which no additional membership is allowed after beginning of the study, and
- dynamic cohort – which gains and/or loses its members during the observation time.

In other meaning it is a part of population, born during a particular period and identified by date of birth.

Probability, risk and odds

At the end, we need to present also the concepts of probability, risk, and odds.

1. Probability.
   In a statistical sense probability is quantification of likelihood of an event or a quantitative description of the likely occurrence of a particular event (9,12,20,21). It is conventionally expressed on a scale from 0 to 1 (a rare event has a probability close to 0, while a very common event has a probability close to 1). The probability of an event has been defined as its long-run relative frequency, defined as a ratio between number of events and total number of all possible events (Equation 7) (1,9,20):

$$p_{event} = \frac{N_{events}}{N_{total}} \qquad \textbf{Equation 7.}$$

   P_{event} = *probability for occurrence of observed phenomenon*
   N_{events} = *number of events of observed phenomenon (part of a whole)*
   N_{total} = *number of all possible events of observed phenomenon (a whole)*

   In fact it is a ratio of a type »proportion« (numerator is included in denominator) and as such could be expressed as a vulgar fraction, a decimal fraction, or as a percentage. Relative frequency expressed as a proportion of a sample is an estimate of the probability of observed phenomenon in a population.

   Calculation of probability is presented in Case study 3.
2. Risk in a statistical sense.
   In a statistical sense risk is probability that the expected event does not occur. It could be expressed as (Equation 8) (12):

$$R_{statistical} = p_{non-event} = 1 - p_{event} \qquad \textbf{Equation 8.}$$

   $R_{statistical}$ = *risk in a statistical sense*
   $P_{non\text{-}event}$ = *probability for non-occurrence of observed phenomenon*
   P_{event} = *probability for occurrence of observed phenomenon*

   This measure could be expressed as a vulgar fraction, a decimal fraction, or a as percentage as well.

The sum of probability for expected event to occur and probability that it does not occur (risk) is 1 or 100%.

3. Risk in a classic epidemiologic sense.
   In an epidemiologic sense the definition of »risk« is different. It is defined as a probability for an unfavourable health outcome (e.g. disease), or some other unfavourable phenomenon related to health (e.g. smoking or other unhealthy behaviour), to occur (Equation 9) (1,9).

   If we are more precise, in epidemiology, the term »risk« is generally used to mean the probability that an unfavourable event (e.g., that a person will be affected by, or die from, an illness, injury, or other health condition) will occur in a given time interval (5,18). In its epidemiologic usage, risk is a conditional probability, because it is the probability of experiencing an event or becoming a case conditional on remaining »at risk« (eligible to become a case) and »in view« (available for the event to be detected) (5,18). In its narrowest sense is related to the incidence concept.

$$R_{unfavourable\ health\ outcome} = p_{unfavourable\ health\ outcome} \qquad \textbf{Equation 9.}$$

$R_{unfavourable\ health\ outcome}$ = *risk for an unfavourable health outcome*
$p_{unfavourable\ health\ outcome}$ = *probability for an unfavourable health outcome*

This measure is presented in details in a separate module in this book.

4. Odds.
   Odds are defined as the ratio of the probability of occurrence of an event to that of non-occurrence (or the ratio of the probability that something is so to the probability that it is not so) (1) (Equation 10).

$$O_{event} = \frac{p_{event}}{1 - p_{event}} \qquad \textbf{Equation 10.}$$

O_{event} = *odds for occurrence of observed phenomenon*
P_{event} = *probability for occurrence of observed phenomenon*

In epidemiology, if we define the probability of occurrence of an unfavourable event as a risk, it is also defined as the ratio of the risk of occurrence of a disease to that of non-occurrence (Equation 11):

$$O_{unfavourable\ health\ outcome} = \frac{R_{unfavourable\ health\ outcome}}{1 - R_{unfavourable\ health\ outcome}} \qquad \textbf{Equation 11.}$$

$O_{unfavourable\ health\ outcome}$ = *odds for an unfavourable health outcome*
$R_{unfavourable\ health\ outcome}$ = *risk for an unfavourable health outcome*

As in the presented equation the quantity »total number of possible events« is included in both, numerator and denominator, it could be reduced through algebraic cancellation. In this case we get a new equation (Equation 12):

$$O_{event} = \frac{N_{events}}{N_{non\text{-}events}}$$ **Equation 12.**

O_{event} = odds for occurrence of observed phenomenon
N_{events} = number of events of observed phenomenon (part of a whole)
$N_{non\text{-}events}$ = number of non-events of observed phenomenon (part of a whole)

From this equation we can clearly see, that odds are the ratio of a type »ratio in a narrow sense« - numerator is not included in denominator, and the numerator and denominator are different categories of the same variable.

Calculation of odds is presented in Case study 3.

Odds are very powerful analytical tool in epidemiology (8,22). Technically we distinguish between odds in different kind of situations (in different types of study design):

- when we are observing the presence of exposure in case-control studies we calculate »exposure-odds«,
- when we are observing in a cross-sectional study the frequency of all cases of a disease versus all non-cases we are talking about odds for having a disease in a specified point of time or »prevalence-odds«,
- when we are observing in a longitudinal study the occurrence of new cases of disease versus non-occurrence we are talking about odds for getting a disease in a specified period of time or »disease-odds« which are the estimate of risk-odds in the sense of incidence-odds (the concepts of incidence and prevalence are out of the scope of this module, and are discussed in a special module).

At the end we need to stress that the mathematical properties of odds make them advantageous for various uses. Whereas probabilities are restricted to the 0-1 interval, odds can be any nonnegative number. The logarithm of the odds can therefore be any real number. The natural logarithm of the odds (called the »logit«) is relatively widely used in biostatistics and epidemiology (8,18,22).

Application of frequency measures in epidemiology

Both, absolute and relative frequency measures represent the basic tool in epidemiology. They could be classified in three big groups of epidemiologic measures (1,4-7,16,18,19,23-29) (Figure 8):

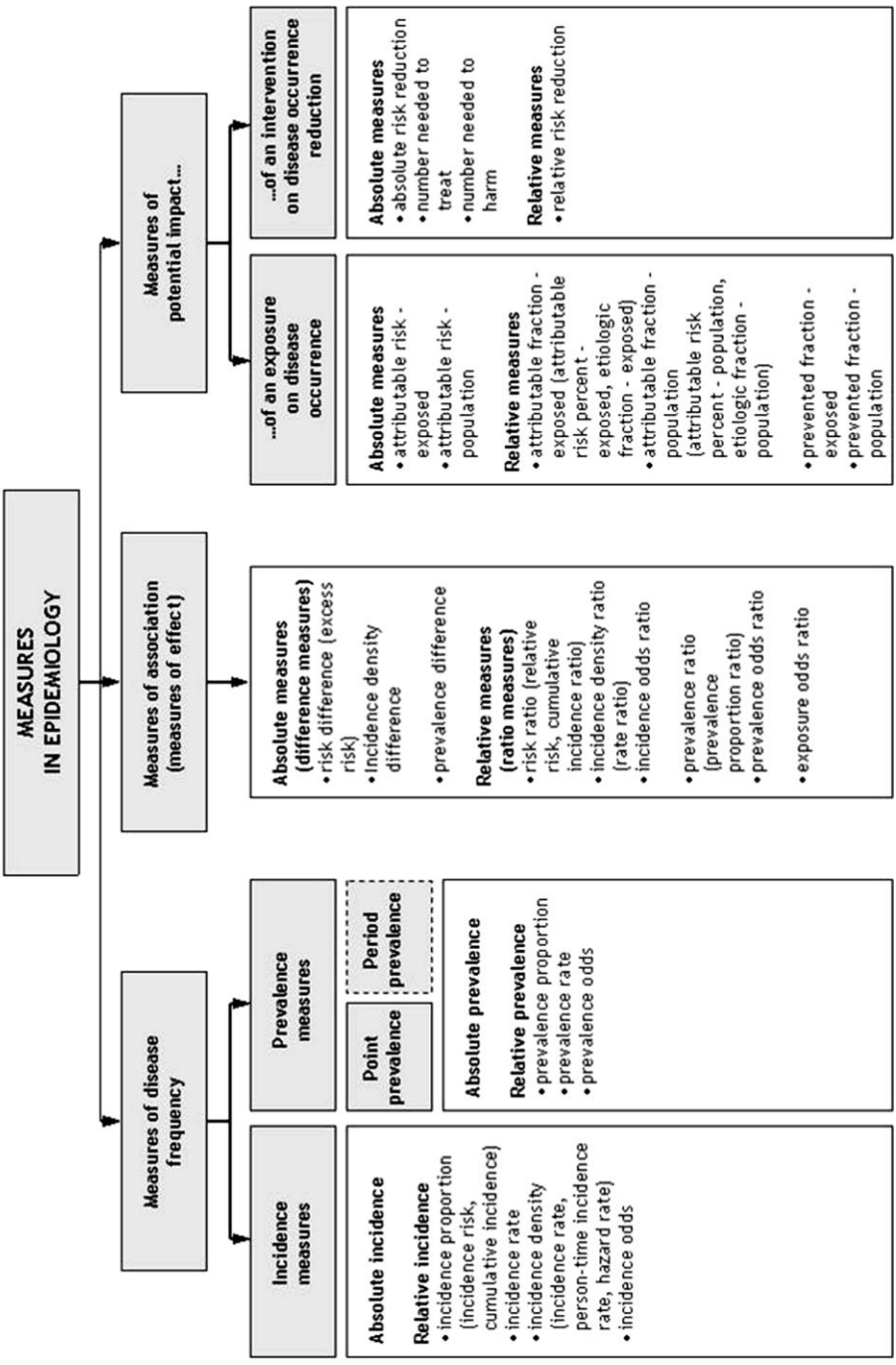


Figure 8. An overview of frequently used measures in epidemiology.

1. Frequency measures.
   These measures (Figure 7) are also called measures of disease occurrence, measures of occurrence of disease and other health related events, or measures of extent. They are trying to answer to the question how often do happen the observed phenomena (diseases, death etc.) in the population.
2. Measures of association.
   These measures (Figure 7) are also called measures of effect. They are trying to answer to the question why do happen the observed phenomena more often is some population groups than in others. To be able to answer to this question we relate different phenomena to each other. Thus these measures express the extent of association between two (or more) phenomena one of them usually being negative health phenomenon and the other risk factor (putative cause) for the first one.
3. Measures of potential impact.
   In this group there are two groups of measures (Figure 7). Measures of the first group express potential impact of risk factor on occurrence of observed health phenomenon among exposed persons or in population. These measures are common in public health. Measures of the second group express potential impact of an intervention on disease occurrence reduction. They are much more common in clinical epidemiology than in public health.

Some of these measures are discussed in details in three separate modules of this book.

In this place we need to give a warning about terminology in the field of epidemiologic measures. A common problem in epidemiology is existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (18).

CASE STUDIES

Case study 1: Organizing data

Introduction to data set

For this case study, real data are used.

In Slovenia for already several years for the purpose of teaching epidemiologic methods in public health, comprising also statistical methods, data collection which enables learning such methods in much comprehensive way has been created. These are the data collected in the frame of the Perinatal Informational System of Slovenia (PISS) (29), which is considered to be one of the permanent of health data-bases of the highest quality with many years' tradition in the country. It was started in 1987, when collection of perinatal data started according to a common protocol in all fourteen Slovene maternity hospitals.

The basic data material for all epidemiologic and statistical activities is composed of 6,356 statistical units, representing the model of a population. For

teaching different epidemiologic and statistical methods, samples of various sizes are then randomly selected from the population database. The data set used in this example is composed of 100 observational units.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal data have been removed, and moreover, the data are selected from the whole collection which shall be used only for the teaching purpose.

Variables in a data set

In a maternity hospital, data on 100 deliveries were collected. Which characteristics of mother and her child (unit of observation) were observed it is shown in Table 2. Data were organized for description and analysis in a data matrix (Figure 9).

Table 2. Description of variables, their values and codes in demonstrational data set.

Column in a data set	Short name of a variable	Information the variable is giving about	Variable values and their codes in data set
1	IDN	unit identification number	
2	BWEIGHT	birth weight (child)	(weight in grams)
3	SEX	sex (child)	1 = boy 2 = girl
4	GESTAGE	gestational age (child)	(age in weeks of pregnancy)
5	MOTHAGE	age at delivery (mother)	(age in years)
6	SMOKING	smoking habits during pregnancy (mother)	0 = no 1 = up to 10 cigarettes/day 2 = 10 cigarettes or more/day
7	HYPIRUT	hyperactivity of uterus (hyper irritable uterus) during pregnancy (mother)	0 = no 1 = yes
8	EBP	elevated blood pressure during pregnancy (mother)	0 = no 1 = yes
9	MEDVIT	consumption of vitamins preparations during pregnancy (mother)	0 = no 1 = yes
10	MEDFE	consumption of iron preparations during pregnancy (mother)	0 = no 1 = yes
11	MEDAB	consumption of antibiotics during pregnancy (mother)	0 = no 1 = yes

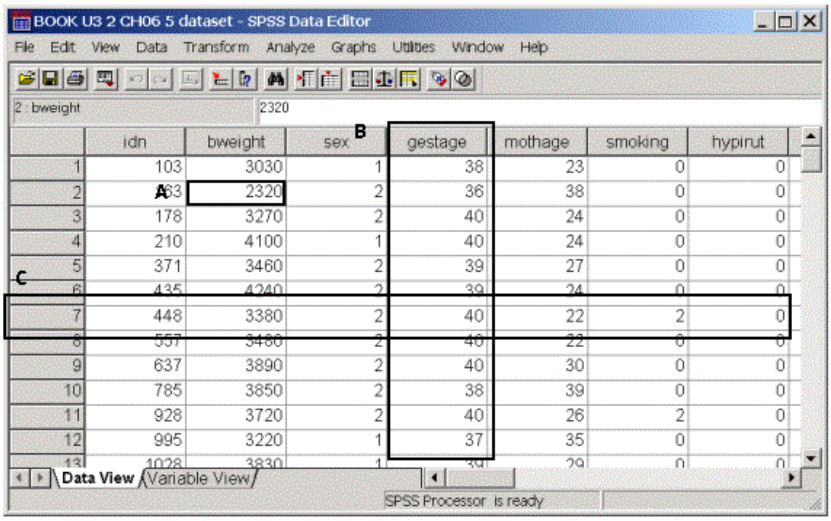


Figure 9. Organization of data in data set for Example 1. A – value of varable BWEIGHT for unit 2; B – values for variable GESTAGE for the first twelve units; C - values for the first seven variables (IDN - HYPIRUT) for unit number 7.

Case study 2: Statistical description of data

This case study is basing on the same set of data as Case study 1.

Defining variables type

From the Table 2 we can see that variables BWEIGHT and MOTHAGE are numerical variables, while all other variables are categorical. We will now analyze the variable MOTHAGE - age at delivery (mother).

From the Figure 8 could be seen the record for first 12 units and first 7 variables. A record for all seven variables for unit No.7 (ROW 7), and record for variable GESTAGE for first twelve units (COLUMN GESTAGE) is accentuated. The crossing of ROW7 and COLUMN GESTAGE has value 40 (value of variable GESTAGE for unit 7).

Analysis of variable "MOTHAGE"

In its origin; "AGE" is a numerical continuous variable. Theoretically, the smallest interval between two values of this variable depends on the precision of the device for measuring it, but in practice we are never interested in such precise information so we always reduce it:

- the information is usually limited (reduced) to intervals of 1-year; in observation of different health states the intervals of 1-day, 1-week, 1-month (neonatology, paediatrics) or 5-years (public health) are also used,
- in public health, the information is often reduced even more when we classify (group) the values according to age groups (periods of life) (e.g. babies, preschool

children, primary school children, adolescents, adults, aged people); the difference between this and the previous level of reduction of information is that intervals of present level are no more equal while on the previous are,

- the information about age is the most reduced when we divide the whole scale in two parts (e.g. adults of 25 years or more: yes or no).

Variable MOTHAGE is in our case interval numerical variable, with 1-year interval width.

Frequency distribution and histogram

Again, we will use variable MOTHAGE - age at delivery (mother). After arranging of values in ordered series, the next step is to summarize this ordered series in a frequency distribution table. The frequency distribution table of this variable is presented in Figure 9.

AGE AT DELIVERY (MOTHER)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18	4	4,0	4,0	4,0
	19	3	3,0	3,0	7,0
	20	4	4,0	4,0	11,0
	21	5	5,0	5,0	16,0
	22	9	9,0	9,0	25,0
	23	9	9,0	9,0	34,0
	24	9	9,0	9,0	43,0
	25	9	9,0	9,0	52,0
	26	8	8,0	8,0	60,0
	27	6	6,0	6,0	66,0
	28	7	7,0	7,0	73,0
	29	4	4,0	4,0	77,0
	30	5	5,0	5,0	82,0
	31	4	4,0	4,0	86,0
	32	4	4,0	4,0	90,0
	34	3	3,0	3,0	93,0
	35	1	1,0	1,0	94,0
	38	1	1,0	1,0	95,0
	39	2	2,0	2,0	97,0
	40	2	2,0	2,0	99,0
	42	1	1,0	1,0	100,0
	Total	100	100,0	100,0	

Figure 10. An example of a frequency distribution table of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30).

In Figure 11, the graphic presentation of frequency distribution for variable MOTHAGE, the histogram, is presented.

Next step is to calculate or to determine the typical values of this distribution.

Figure 11. An example of a histogram of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30).

Typical values

The description of the distribution of values of variable MOTHAGE - age at delivery (mother) of example dataset are as follows

- the distribution is bell shaped (Figures 9 and 10),
- it is slightly asymmetrical (Figure 9 and 10),

As the number of units is rather large, it could be summarized by mean and standard deviation or median and quartiles/percentiles. The mean and the median value are, because slight asymmetry of the distribution, similar but not the same. The decision which set of typical values to choose is up to investigator. In Figure 12, typical values for the distribution shown in Figure 11 are presented.

Statistics

AGE AT DELIVERY (MOTHER)

N	Valid	100
	Missing	0
Mean		26,19
Median		25,00
Std. Deviation		5,256
Minimum		18
Maximum		42
Percentiles	25	22,25
	50	25,00
	75	29,00

Figure 12. An example of set of typical values for the distribution of variable MOTHAGE - age at delivery (mother) of example dataset (the SPSS statistical programme printout) (30).

Case study 3: Epidemiologic description of data

Introduction to data set

This case study as well, is basing on real data.

In Slovenia in 2001 the survey aiming at assessing the prevalence of health behaviours (smoking, nutrition, alcohol consumption, physical activity and traffic safety) was performed for the first time. This survey is conceptually a part of a wider international project in the frame of the Countrywide Integrated Non-communicable Diseases Intervention (CINDI) program, supported by the World Health Organization, CINDI Health Monitor.

The stratified random sample was drawn from the Central Population Registry of the Republic of Slovenia. The sample size was 15,379 with the age range 25-64 years. The sampling was performed by the Statistical Office of the Republic of Slovenia.

The data were collected in late spring 2001 by using a self-administered postal questionnaire, based on the CHM Core Questionnaire (31).

Out of 15,379 inhabitants included in the sample 15,153 were contacted (226 were excluded because of changed domicile, severe illness or death). The response rate was 63.8% (9,666 responses). The respondents did not differ statistically from non-respondents in age distribution or distribution of size of settlements of permanent residence, but the response to the survey was slightly

lower among men (47.0%) than among women (53.0%) at a ratio 1:1.1 (according to population data in 2001 the ratio was 1:1). The questionnaires of 9,034 respondents were eligible for analysis (eligibility criteria: sex and age provided by Statistical Office of the Republic of Slovenia).

For the purpose of this module, we have chosen observation of smoking behaviour.

Variables in a data set

In CINDI Health Monitor survey in Slovenia in 2001 (CHMS 2001) 8,904 respondents reported their current smoking status. The answers grouped regarding the sex of respondents are shown in Table 3.

Table 3. Smoking status in CINDI Health Monitor survey in Slovenia in 2001 in total sample and by sex.

	SEX		Total
SMOKER	**Male**	**Female**	
No	2,931	3,859	6,790
Yes	1,143	971	2,114
Total	4,074	4,830	8,904

Ratios

Proportion and percentage

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey (Table 3). The proportion of smokers could be calculated as a vulgar fraction (Equation 13):

$$Proportion = \frac{2{,}114}{8{,}904} \qquad \textbf{Equation 13.}$$

or as a decimal fraction (Equation 14):

$$Proportion = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 14.}$$

or as percentage (Equation 15):

$$Proportion = \frac{2{,}114}{8{,}904} \times 100 = 23.7\% \qquad \textbf{Equation 15.}$$

Rate in classic epidemiologic sense

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the moment of the survey (Table 3). As the survey is a representative of cross-sectional studies the time component is a point in time (a moment) (Equation 16):

$$epidemiologic\ Rate_{(at\ the\ moment\ of\ a\ survey)} = \frac{2{,}114}{8{,}904} \times 1{,}000 = 237 \qquad \textbf{Equation 16.}$$

The epidemiologic rate (prevalence rate) has value 237 per 1,000 population.

Ratio in a narrow sense

We could express several ratios in a narrow sense using this example:

1. In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey and the other 6,790 that they were not (Table 3). The ratio between non-smokers and smokers is (Equation 17):

$$Ratio = \frac{6{,}790}{2{,}114} = 3.21 \qquad \textbf{Equation 17.}$$

   The ratio is 3.21 to 1, what means that in Slovenia in 2001 there were 3.21-times more non-smokers than smokers (or there were 3.21 non-smokers to one smoker).
2. If we now turn the ratio and observe the ratio of smokers to non-smokers we get (Equation 18):

$$Ratio = \frac{2{,}114}{6{,}790} = 0.31 \qquad \textbf{Equation 18.}$$

   The rate is 0.31 to 1, what means that in Slovenia in 2001 there was 0.31 of a smoker to one non-smoker.
3. In CHMS 2001, 2,114 respondents reported that they were smokers at the time of the survey. Among them there were 1,143 men in 971 women (Table 3). The ratio between men and women among smokers was (Equation 19):

$$Ratio = \frac{1{,}143}{971} = 1.18 \qquad \textbf{Equation 19.}$$

   The ratio was 1.18 to 1, what means that in Slovenia in 2001 there were 1.18-times more male smokers than female smokers (or to one female smoker there was 1.18 male smoker).

Probability and odds

Probability

In CHMS 2001, 2,114 out of 8,904 respondents stated that they were smokers at the time of the survey (Table 3). The probability for being a smoker at the time of survey could be calculated as:

- a vulgar fraction (Equation 20):

$$p = \frac{2{,}114}{8{,}904} \qquad \textbf{Equation 20.}$$

- as a decimal fraction (Equation 21):

$$p = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 21.}$$

- or as a percentage (Equation 22):

$$p = \frac{2{,}114}{8{,}904} \times 100 = 23.7\% \qquad \textbf{Equation 22.}$$

The probability for being a smoker at the moment of the survey CHMS 2001 was, expressed as percentage, 23.7%.

Odds

The odds for being a smoker for data presented in Table 3 could be calculated in two different ways.

1. Through calculating first the probability for being a smoker at the moment of the CHMS 2001 survey (also interpreted as risk in an epidemiologic sense - a probability for unfavourable health behaviour in this case) and the probability for being a non-smoker. The probability for being a smoker was (Equation 23):

$$p = \frac{2{,}114}{8{,}904} = 0.237 \qquad \textbf{Equation 23.}$$

while the probability for being a non-smoker was (Equation 28):

$$1 - p = 1 - 0.237 = 0.763 \qquad \textbf{Equation 24.}$$

The odds for being a smoker can be calculated now as (Equation 25):

$$O = \frac{0.237}{1-0.237} = \frac{0.237}{0.763}0 = 0.31 \qquad \textbf{Equation 25.}$$

2. Through algebraic cancellation of total number of possible events from calculation of odds:

$$O = \frac{\frac{2{,}114}{8{,}904}}{1-\frac{2{,}114}{8{,}904}} = \frac{\frac{2{,}114}{8{,}904}}{\frac{8{,}904}{8{,}904}-\frac{2{,}114}{8{,}904}} = \frac{\frac{2{,}114}{8{,}904}}{\frac{6{,}790}{8{,}904}} = \frac{2{,}114}{6{,}790} = 0.31 \qquad \textbf{Equation 26.}$$

The odds for being a smoker versus non-smoker are 0.31. This means that in Slovenia in 2001 there was 0.31 of a smoker to one non-smoker, or the ratio non-smokers to smoker is 1 to 0.31. This is exactly the same result as in Equation 25.

EXERCISE

Task 1: Statistical description of data

1. From the table with description of the basic data set (Table 2) find out how many variables are in the data set, their names and which information they contain.
2. Find out which of the variables could play the role of "the effect" and which ones the role of "the cause".
3. Carefully read the following statements and determine how many variables are included and which could be their values:
   - after the fractures children recover faster than adults,
   - men with inflammation of joints differ in response to therapy from women,
   - men more often get lung complications after the heart operation then women.

4. In the table with raw data set (APPENDIX, Table A1) find out how many units are there in this sample.
5. Find out what is the unit under study.
6. Enter the data for the first 20 units in data matrix.
7. For the variable MOTHAGE make the ordered series.
8. Make the frequency distribution table (with absolute and relative frequencies in percentages) for this variable, too.

9. Find out for the same variable if any value exists with absolute frequency equal to 0.
10. Make the frequency distribution table for the same variable (MOTHAGE) also with your statistical program and compare it with the table you made in Exercise 8.
11. Make frequency distribution tables with your statistical program also for the following variables: BWEIGHT, SEX, GESTAGE, SMOKING and HYPIRUT.
12. For variables BWEIGHT, SEX, GESTAGE, MOTHAGE, SMOKING and HYPIRUT find out:
    - ordinality of their values
    - how many values can you find for each one in a frequency distribution table,
    - find out if could you classify these values as continuous or discrete,
    - classify each variable according to type,
    - for numerical variables (continuous and discrete) find out from the frequency distribution table the lowest, the highest and the most frequent value/s (if there are more than one, list all of them),

13. Find out from the frequency distribution of variables MOTHAGE and GESTAGE for each one:
    - where the density of distribution is the highest,
    - is the distribution symmetrical or not and if it is not, to which direction it is skewed,
    - where would you locate the centre of the distribution.

    Compare both frequency distributions.
14. For variables MOTHAGE and GESTAGE draw a histogram using the absolute frequency from the frequency distribution table; consider also all values with the frequency equal to 0.
15. Draw a histogram for variable BWEIGHT, too; as there are many different values, group observations in intervals of 250gr of width in order to get a better visual impression of the observed distribution.
16. Find out in histogram for each variable:
    - is the distribution bell shaped or not,
    - is the distribution symmetrical or not and if it is not, to which direction it is skewed,
    - where would you locate the centre of the distribution.

    Compare all three histograms.
17. Compare these conclusions with your conclusions in Exercise 12.
18. Make the histograms also with your statistical program and compare them to histograms you made by yourself.
19. For attributable variables SEX, SMOKING and HYPIRUT draw the ordinary bar charts by yourself.
20. Make the bar charts for these variables also with your statistical program and compare them to bar charts you made by yourself in Exercise 18.

21. For numerical variables BWEIGHT, GESTAGE and MOTHAGE make a decision which typical values would be the most appropriate for summarizing the features of distribution of their values and prove it.
22. For these variables determine the typical values by the means of your statistical program
23. What can you say about the distribution of a variable for which the mean and the median differ significantly
24. Find out if it is sensible to determine typical values for any type of attributable variables
25. Carefully read the following statement and decide if it is true or false:
    In attributable variable with only two values, 0 and 1, the proportion of units with value 1 equals to hidden arithmetic mean of this type of variable
26. Check your decision with statistical program: calculate the proportion of units with value 1 and the arithmetic mean of the variable HYPIRUT.

Task 2: Epidemiologic description of data

This task is basing again on PISS data (28).

Table 4. Description of newly designed variables, created from original ones, their values and codes in demonstrational data set.

Short name of a variable	Original variable (cut-off point)	Information the variable is giving about	Variable values and their codes in data set
LBW	BWEIGHT (2500 g)	Low birth weight (2500 g or less)	0 = no 1 = yes
SMOKING1	SMOKING	Smoking during pregnancy	0 = no 1 = yes

Frequency of low birth weight (LBW) was observed in two groups of newborns according to smoking of mother during pregnancy. The results are presented in Table 5. All following exercises are referring to this table.

Table 5. Frequency of low birth weight (LBW) in newborns in two groups divided according to smoking of mother during pregnancy based on PISS data (29).

Low birth weight	Smoking of mother during pregnancy		Total
	No	Yes	
No	558	191	749
Yes	35	16	51
Total	593	207	800

1. Calculate following ratios in which numerator is included in denominator:
   - a proportion of LBW newborns as a decimal fraction in the total sample of newborns,
   - a percentage of LBW newborns in the total sample of newborns,

2. Calculate the ratio in a narrow sense of LBW in the total sample of newborns.
3. Calculate probability of LBW in the total sample of newborns, in the group of smoking mothers, and in the group of non-smoking mothers during pregnancy.
4. Calculate odds of LBW in the total sample of newborns, in the group of smoking mothers, and in the group of non-smoking mothers during pregnancy.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
3. Timmreck TC. An introduction to epidemiology. 3rd ed. Sudbury, MA: Jones&Bartlett Publishers; 2002.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Co.; 2001.
5. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25]. Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 24, 2009.
6. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
7. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
8. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
9. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
10. ISI - International Statistical Institute. ISI Multilingual Glossary of Statistical Terms [homepage on the Internet]. Available from URL: http://isi.cbs.nl/glossary.htm. Accessed: June 24, 2009.
11. StatSoft, Inc. Electronic statistics textbook. Tulsa, OK: StatSoft; 2007. Available from URL: http://www.statsoft.com/textbook/stathome.html. Accessed: June 24, 2009.
12. Košmelj B, Arh F, Doberšek-Urbanc A, Ferligoj A, Omladič M. Statistical dictionary (in Slovene). Ljubljana: Statistično društvo Slovenije: Študentska založba; 2002.
13. Swinscow TDV. Statistics at square one. Ninth Edition. London: BMJ Publishing Group; 1997. Available from URL: http://www.bmj.com/collections/statsbk/. Accessed: June 24, 2009.
14. Hicks C. Research for physiotherapists: project design and analysis. Edinburgh: Livingstone; 1995.
15. Friis RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
16. Kue Young T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
17. Babuš V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
18. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.

19. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
20. Easton VJ, McColl JH. Statistics glossary. V 1.1. Probability [homepage on the Internet]. Available from URL: http://www.stats.gla.ac.uk/steps/glossary/alphabet.html. Accessed: June 25, 2009.
21. Organization for Economic Co-operation and Development - OECD. Glossary of statistical terms. Probability [homepage on the Internet]. Available from URL: http://stats.oecd.org/glossary/detail.asp?ID=3794. Accessed: June 25, 2009.
22. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 2000.
23. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
24. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
25. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
26. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
27. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1247374&blobtype=pdf. Accessed: June 24, 2009.
28. Jenicek M. Foundations of evidence-based medicine. London: Parthenon publishing; 2003.
29. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia [homepage on the Internet]. Availbale from URL: http://www.ivz.si/index.php?akcija=podkategorija&p=44 (Accessed July 24, 2008).
30. SPSS Base 7.5 for Windows. User's Guide. SPSS Inc. Chicago; 1997.
31. Zaletel-Kragelj L, Fras Z, Maučec Zakotnik J, editors. Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001. II. Risky behaviours (in Slovene). Ljubljana: CINDI Slovenija; 2004. p.107-564.

RECOMMENDED READINGS

1. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1993.
2. Babuš V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
3. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25]. Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 24, 2009.
4. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
5. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
6. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
7. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.

APPENDIX

Table A1. Raw perinatal data. Perinatal Informational System of Slovenia (PISS) (29).

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
1.	103	3030	1	38	23	0	0	0	0	1	0
2.	163	2320	2	36	38	0	0	0	0	1	0
3.	178	3270	2	40	24	0	0	0	0	0	0
4.	210	4100	1	40	24	0	0	0	0	0	1
5.	371	3460	2	39	27	0	0	0	1	0	0
6.	435	4240	2	39	24	0	0	0	0	0	0
7.	448	3380	2	40	22	2	0	0	0	0	0
8.	557	3480	2	40	22	0	0	0	1	1	0
9.	637	3890	2	40	30	0	0	0	0	0	0
10.	785	3850	2	38	39	0	0	0	0	0	1
11.	928	3720	2	40	26	2	0	0	0	1	0
12.	995	3220	1	37	35	0	0	0	1	1	1
13.	1028	3830	1	39	29	0	0	0	1	1	0
14.	1034	3680	2	40	21	0	0	0	0	0	0
15.	1048	3160	2	37	23	0	0	0	0	0	0
16.	1142	3250	1	39	29	0	0	0	0	0	0
17.	1143	4130	1	40	26	0	0	0	0	1	1
18.	1171	2980	1	39	24	0	0	0	1	0	0
19.	1209	4900	1	41	26	1	0	0	0	0	0
20.	1258	1880	2	37	30	0	0	0	0	0	0
21.	1365	2100	2	37	26	1	0	0	0	0	0
22.	1397	5000	2	42	28	0	0	0	1	1	0
23.	1424	3430	2	39	25	0	1	0	0	1	0
24.	1426	3590	1	38	42	0	0	0	0	0	0
25.	1472	3680	2	39	23	0	0	0	0	0	0
26.	1473	3320	2	41	23	0	0	0	0	0	0
27.	1576	3560	1	39	19	0	0	0	0	0	0
28.	1588	3430	1	41	31	0	0	0	0	1	0
29.	1604	1840	2	36	27	0	0	1	0	0	1
30.	1620	3170	1	40	19	0	0	0	0	1	0
31.	1642	3740	1	41	32	0	0	0	0	0	1
32.	1706	3130	2	41	28	0	0	0	1	1	0
33.	1808	3460	2	39	22	0	0	0	1	1	1

Table A1. Cont.

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
34.	2021	3710	1	40	28	1	0	0	0	1	0
35.	2031	3120	1	41	27	0	0	0	0	0	0
36.	2096	2200	2	34	22	2	0	0	0	0	0
37.	2166	1530	1	32	25	0	0	0	1	1	1
38.	2264	2820	1	40	23	1	0	0	0	1	0
39.	2269	3880	2	38	27	1	0	0	0	0	0
40.	2346	3870	1	39	24	0	0	0	0	0	1
41.	2499	3450	2	40	29	1	0	0	0	0	0
42.	2632	2770	1	35	18	0	0	0	0	1	0
43.	2668	3480	1	41	25	0	0	0	1	1	0
44.	2747	3300	1	38	19	0	0	0	0	0	0
45.	2786	3920	1	39	25	0	0	0	0	1	0
46.	2799	4140	2	40	28	0	0	0	0	1	0
47.	2871	3470	2	41	23	1	0	0	1	1	0
48.	2965	3700	2	40	31	0	0	0	0	1	0
49.	3092	2420	2	37	29	0	0	0	1	0	0
50.	3127	3400	1	38	30	0	0	0	0	0	0
51.	3156	2770	2	38	23	1	0	0	0	0	0
52.	3170	3590	2	40	21	0	0	0	1	1	0
53.	3220	3800	1	40	22	0	1	0	1	1	1
54.	3286	3370	1	40	32	2	0	0	0	1	1
55.	3314	3200	2	39	26	1	0	0	0	0	0
56.	3333	3480	2	38	26	0	0	0	0	1	0
57.	3379	2920	1	38	34	1	1	1	0	0	0
58.	3417	3850	1	40	25	0	0	0	0	1	0
59.	3430	3160	2	41	20	2	0	0	0	0	0
60.	3469	3500	1	39	30	0	0	0	1	0	0
61.	3471	2970	2	39	27	0	0	0	1	0	0
62.	3498	3640	2	39	34	0	0	0	0	0	0
63.	3501	2440	2	36	18	0	0	0	0	0	0
64.	3567	2660	1	39	28	0	0	0	0	0	0
65.	3604	3210	2	40	24	1	0	0	0	0	0
66.	3621	3260	1	39	20	0	0	0	1	1	0
67.	3732	3730	1	41	23	0	0	0	1	0	0
68.	3851	3040	2	39	22	0	0	0	1	1	0

Table A1. Cont.

	IDN	BWEIGHT	SEX	GESTAGE	MOTHAGE	SMOKING	HYPIRUT	HBP	MEDVIT	MEDFE	MEDAB
69.	3918	2940	1	38	25	0	0	0	0	1	0
70.	3923	2920	2	40	31	0	0	0	1	0	1
71.	4019	3460	2	39	40	0	0	0	1	0	0
72.	4034	3360	2	40	24	0	0	0	0	1	1
73.	4145	3140	1	41	25	0	1	0	0	0	0
74.	4193	3800	2	37	40	0	1	0	0	1	0
75.	4206	3350	1	39	25	1	0	0	1	0	0
76.	4209	3350	2	40	24	1	0	0	1	1	0
77.	4386	3810	2	41	32	0	0	0	0	1	0
78.	4421	2420	1	35	30	0	0	0	0	0	0
79.	4522	3050	2	39	22	1	0	0	0	0	0
80.	4598	2840	1	39	39	0	1	0	1	0	0
81.	4672	4170	1	41	25	0	0	0	1	0	0
82.	4944	2410	2	38	26	0	0	0	0	1	0
83.	4957	3780	1	40	22	0	0	0	0	0	0
84.	5002	4130	1	42	21	1	0	0	0	0	0
85.	5122	2830	1	39	18	0	0	0	0	0	0
86.	5249	3960	2	40	18	0	0	0	0	0	0
87.	5409	3280	2	39	21	0	0	0	0	1	0
88.	5433	3020	2	37	28	0	1	0	0	1	0
89.	5445	1270	1	30	22	0	0	0	0	0	0
90.	5564	3100	1	40	21	0	1	0	0	1	1
91.	5656	3870	1	38	23	0	0	0	0	1	0
92.	5678	3470	1	40	24	2	0	0	0	0	0
93.	5761	4000	1	40	26	0	0	0	0	0	0
94.	5820	2960	2	40	20	0	0	0	1	0	0
95.	5862	3670	1	40	27	0	0	0	0	1	0
96.	5871	3400	2	38	32	0	0	0	1	1	0
97.	5887	2840	1	38	20	0	0	0	1	0	0
98.	5907	3060	1	38	34	0	0	0	1	1	0
99.	6013	3230	2	39	31	0	1	0	0	1	0
100.	6131	3230	2	40	28	0	0	0	0	0	1

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**FREQUENCY MEASURES: PREVALENCE AND INCIDENCE**
Module: 1.2.2	**ECTS (suggested): 0.40**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloska 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Frequency measures, prevalence, incidence
Learning objectives	After completing this module students should: • know and understand the difference between prevalence and incidence measures of frequency, • be familiar with different types of prevalence measures, and know how to calculate them; • be familiar with different types of incidence measures, and know how to calculate them.
Abstract	Frequency measures are quantities that express frequency of health phenomena. Prevalence and incidence are two main groups of frequency measures. The most distinctive difference between these two groups is that by prevalence measures we are observing the transversal section through the situation of phenomenon under observation at designated time (e.g. in a point of time) while by incidence measures we are observing its dynamics (by performing regular observation of breaking out of new cases of phenomenon under observation durig every of specified equal time periods) in a specified population.
Teaching methods	An introductory lecture gives the students first insight in features and types of frequency measures. The theoretical knowledge is illustrated by case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards they on provided data set perform extensive tasks on calculation of different types of measures. They are stimulated to compare results with results of each other and discuss the differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which detailed description of process of calculation is described.

FREQUENCY MEASURES: PREVALENCE AND INCIDENCE

Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Introduction to frequency measures

Prevalence and incidence are two main groups of frequency measures in epidemiology. We should be conscious that there are in fact two families of measures under each term although we are frequently talking about only two measures. This is the reason of great deal of misunderstanding and misinterpretation of frequency measures. The most distinctive difference between these two groups of measures is the fact that by prevalence measures we are observing the transversal section through the situation of phenomenon under observation most usually at a designated time (e.g. in a point of time) while by incidence measures we are observing its dynamics (by performing regular observation of breaking out of new cases of phenomenon under observation durig every of specified equal time periods) in a specified population (1-22). General equations for these two families of measures are as follows (Equation 1 and Equation 2) (1-3):

$$P = \frac{N_{d+all\ cases(dt)}}{N_{all\ persons(dt)}}$$ **Equation 1.**

P = prevalence
$N_{d+\ all\ cases\ (dt)}$ *= number of all persons with the disease under observation (cases) at designated time*
$N_{all\ persons\ (dt)}$ *= number of all persons under observation at designated time*

$$I = \frac{N_{d+newcases(gp)}}{N_{all\ persons\ at\ risk\ (bgp)}}$$ **Equation 2.**

I = incidence
$N_{d+\ new\ cases\ (gp)}$ *= number of new cases of the disease under observation during a given period*
$N_{all\ persons\ at\ risk\ (bgp)}$ *= number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*

Both families could be theoretically classified regarding various features what will be discussed later.

The process of explanation of differences between both families of measures and the differences between measures inside both families will be illustrated in the case studies.

Prior starting with explanation of concepts of incidence and prevalence, both extremely important epidemiologic concepts, it could be worthy to emphasize a common problem in epidemiology is existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (21)

At the beginning it could be useful for students to emphasize that there exist different frequency measures on one hand and different study designs on the other. Although we could on a theoretical level show that we have several analogous frequency measures, not all of them are of "common sense" and are not meaningful in all situations. One should be aware that it is important to use the measure that is most appropriate for the current task (22). Sice we think that one can choose the most appropriate measure if he/she fully understands the differences between them.

Most of textbooks on epidemiologic methods first concentrate on incidence measures, and only afterwards on prevalence measures. We will do the opposite way since we think that the concept of prevalence might be easier to understand.

Prevalence

Definition

Prevalence is a common term for a group of measures which are quantifying the state (situation) of a given health phenomenon (e.g. a disease, a disorder, an unhealthy health behaviour etc.) at a designated time (at a specified moment, or at any time during a specified period) irrespective of weather the cases of observed phenomenon are old or new (1-6,12-17,21,22). Prevalence measures are measuring thus the burden of disease or any other health condition in a population or its power (praevaleo lat. to be very strong) at a designated time. A special problem could be the fact that the term »prevalence« could denote various prevalence measures, mostly as a synonym for »prevalence rate« (in a classic epidemiologic sense).

Prevalence measures could be classified according to various characteristics. Three different classifications are summarized in Figure 1.

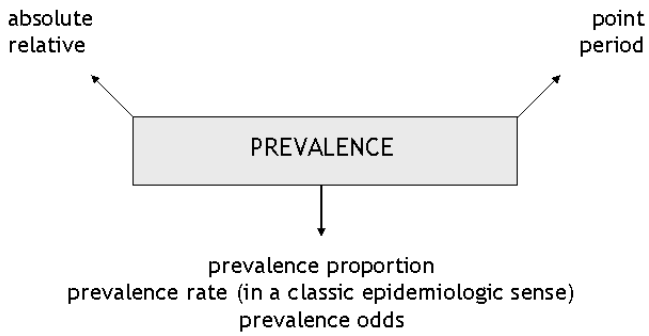


Figure1. Classifications of prevalence measures according to various characteristics.

Classification of prevalence measures according to type of frequency measure upon which are based

According to type of frequency measure upon which are based we distinguish between absolute and relative prevalence.

1. Absolute prevalence, prevalent number or prevalence.
   This measure is simply a number of all cases of observed phenomenon at a designated time (at a specified moment, or at any time during a specified period).

   Absolute prevalence frequency measures are important in health care planning (e.g. number of hospital beds needed for treatment of certain group of health states).

   Calculation of this measure in practice is presented in Case study 1.
2. Relative prevalence.
   Relative prevalence are several measures based upon relative frequency (based on different measures of relative frequency) - prevalence as a proportion, prevalence as rate (in a classic epidemiologic sense), and prevalence as odds. All three measures will be in details discussed later, only the first one, the prevalence as a proportion, is presented in this place to be contrasted to absolute prevalence (Equation 3):

$$P_{rel\,(proportion)} = \frac{N_{d+\,all\ cases(dt)}}{N_{all\ persons(dt)}} \qquad \textbf{Equation 3.}$$

$P_{rel\ (proportion)}$ = *relative prevalence as a proportion*
$N_{d+\ all\ cases\ (dt)}$ = *number of all persons with the disease under observation (cases) at designated time*
$N_{all\ persons\ (dt)}$ = *number of all persons under observation at designated time*

Prevalence proportion is a probability of having a disease at a time t, or a probability that an individual will be a case at time t (7).

Relative prevalence frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations etc.)

Calculation of this measure in practice is presented in Case study 1.

Classification of prevalence measures according to the type of a designated time of observation

According to the type of a designated time of observation, which could be a specified moment, or a specified period, we distinguish between point and period prevalence. When used without any qualification, the term prevalence refers usually to point prevalence.

1. Point prevalence (1-3).
   It is the measure which expresses the burden of the disease under observation at a specified point in time. It could be absolute (absolute point prevalence) or relative (relative point prevalence). A specified point in time could be:

- a specified point in calendar time (e.g. calendar day, calendar week, calendar month), or
- a specified point in the natural course of the disease (e.g. the point of onset of the symptoms), or a specified event that may be associated with or produce changes in human health (e.g. specified event in a life course, which could be different by calendar time for any of individuals under observation, like onset of puberty, menopause, the beginning of retirement, third post-operative day etc.).

Relative point prevalence could be expressed as a proportion, rate (in a classic epidemiologic sense), or odds. Point prevalence expressed as prevalence proportion is a probability for having a disease under observation at a specified point in time (7,22), and it could be calculated (Equation 4):

$$_{point}P_{rel\,(proportion)} = \frac{N_{d+\,all\ cases(point\ in\ time)}}{N_{all\ \ persons(point\ in\ time)}} \qquad \textbf{Equation 4.}$$

$_{point}P_{rel\ (proportion)}$ = *point prevalence as proportion (probability) for having a disease at a specified point in time*
$N_{d+\ all\ cases\ (point\ in\ time)}$ = *number of all persons with the disease under observation (cases) at a specified point in time*
$N_{all\ persons\ (point\ in\ time)}$ = *number of all persons under observation at a specified point in time*

Calculation of this measure in practice is presented in Case study 1.

2. Period prevalence (1-3,7,10).
   Another prevalence measure is period prevalence which is much less frequently used as point prevalence. It is the measure which expresses the probability that an individual in a population will be a case any time during a period of time (7,10), and it could be calculated as a ratio (not a proportion) as follows (Equation 5):

$$_{period}P_{rel} = \frac{N_{d+\,all\ cases(period\ of\ time)}}{N_{all\ \ persons(period\ of\ time)}} = \frac{N_{d+_0} + N_{d+\,new cases(period\ of\ time)}}{N_{all\ \ persons(period\ of\ time)}} \qquad \textbf{Equation 5.}$$

$_{period}P_{relative}$ = *period prevalence as a probability for having a disease at any time during a specified period*
$N_{d+\ all\ cases\ (period\ of\ time)}$ = *number of all persons with the disease under observation (cases) at any time during a specified period*
$N_{all\ persons\ (period\ of\ time)}$ = *number of all persons in the population for this same period*
$N_{d+\ 0}$ = *number of persons with the disease under observation (cases) at the beginning of the specified period*
$N_{d+\ new\ cases\ (period\ of\ time)}$ = *number of new cases of the disease under observation during a specified period*

This measure requires the assumption of a stable dynamic population for estimation (7,10). If the study population is unstable, this measure has little practical value.

Period prevalence is most often used in situations when the exact time of the onset of a phenomenon under observation for individual cases is not known (7).

Calculation of this measure in practice when the assumption of a fixed cohort is met is presented in Case study 1.

In continuation we will discuss in details only a point prevalence.

Classification of relative point prevalence measures according to type of relative frequency measure upon which are based

According to type of relative frequency measure upon which is based point prevalence measure, we distinguish between relative point prevalence as a proportion, as a rate, and as odds.

1. Measures in which the numerator is included in the denominator.

   Prevalence proportion (7,22)**.** This measure expresses the probability of having a disease at a designated time under observation. We have already presented this measure (Equation 4), but since here we introduce the notation usually used in epidemiologic textbooks, we repeat it as a new equation (Equation 6):

$$P = \frac{N_{d+\,all\ cases(point\ in\ time)}}{N_{all\ \ persons(point\ in\ time)}} \qquad \textbf{Equation 6.}$$

P = prevalence proportion
$N_{d+\ all\ cases\ (point\ in\ time)}$ *= number of all persons with the disease under observation (cases) at designated time*
$N_{all\ persons\ (point\ in\ time)}$ *= number of all persons under observation at designated time*

Calculation of this measure in practice is presented in Case study 1.

Prevalence rate (in a classic epidemiologic sense) (1,2,12,23)**.** This measure is very similar to the first one. The only difference is that it has additional components – the multiplier and a time component. It is a rate in a classic epidemiologic sense and when it is a point prevalence it is calculated as follows (Equation 7):

$$PR = \frac{N_{d+\,all\ cases(point\ in\ time)}}{N_{all\ \ persons(point\ in\ time)}} \times K \qquad \textbf{Equation 7.}$$

PR = prevalence rate
$N_{d+\ all\ cases\ (dt)}$ *= number of all persons with the disease under observation (cases) at designated point in time*
$N_{all\ \ persons\ \ (dt)}$ *= number of all persons under observation at designated point in time*
K = multiplier (100, 1000, 10.000, 100.000…)

Calculation of this measure in practice is presented in Case study 1.

2. Measure in which the numerator is not included in the denominator.
   Prevalence odds (14,22). Prevalence odds are the ratio of the probability of having a disease to that of not having it at a point in time, or when through algebraic cancellation of total number of possible events the reduction is performed, the ratio of cases to non-cases of the disease under observation at a point in time. They could be calculated as follows (Equation 8):

$$PO = \frac{N_{d+\ all\ cases(dt)}}{N_{d-\ all\ cases(dt)}} \qquad \textbf{Equation 8.}$$

PO = prevalence oddse
$N_{d+\ all\ cases\ (dt)}$ = *number of all persons with the disease under observation (cases) at designated time*
$N_{d-\ all\ cases\ (dt)}$ = *number of all persons without the disease under observation (non-cases) at designated time*

Calculation of this measure in practice is presented in Case study 1.

Incidence

Incidence is a common term for a group of measures which are quantifying a break out of new cases of a health phenomenon (e.g. a disease) under observation (incido in morbum lat. to fall ill) during a specified period in a specified group of persons (e.g. population) (1-6,12-17,21,22). A special problem is that the term »incidence« is used to denote various incidence measures.

By performing regular observation of breaking out of new cases of phenomenon under observation durig every of specified equal time periods we are observing its dynamics in a specified population.

Incidence measures could be classified according to various characteristics. Four different classifications are summarized in Figure 2.

Classification of incidence measures according to type of frequency measure upon which are based

According to type of frequency measure upon which are based we distinguish between absolute and relative incidence. The later is frequently used than the first. When used without any qualification, the term incidence refers usually to absolute incidence, though sometimes is used to mean relative incidence, mostly incidence rate.

1. Absolute incidence, incident number or incidence (1).
   This measure is simply a number of new cases of observed phenomenon during a specified period in a specified group of persons (e.g. population).

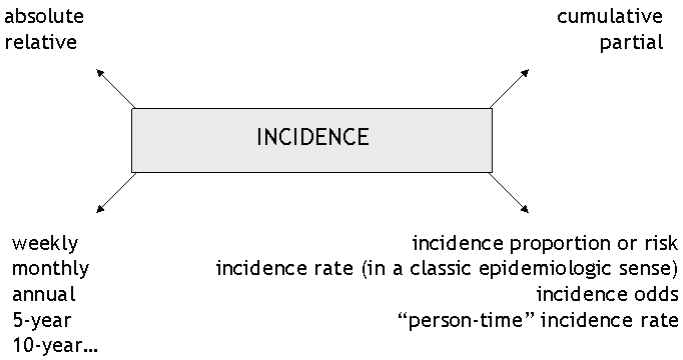


Figure 2. Classifications of incidence measures according to various characteristics.

Absolute incidence frequency measures, similarly as absolute prevalence frequency measures, are important in health care planning.

Calculation of this measure in practice is presented in Case study 2.

2. Relative incidence.

Relative incidence are several measures based upon relative frequency (based on different measures of relative frequency) - incidence as a proportion, as a rate, as odds, and as an incidence density. All four measures will be in details discussed later, only the first one, the incidence proportion, is presented in this place to be contrasted to absolute incidence (Equation 9):

$$I_{rel\,(proportion)} = \frac{N_{d+\,new\,cases(gp)}}{N_{all\ \ persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 9.}$$

$I_{rel\,(proportion)}$ = *relative incidence as proportion (proportion of candidate individuals who developed the disease during the given period)*

$N_{d+\,new\,cases\,(gp)}$ = *number of new cases of the disease under observation during a given period*

$N_{all\ persons\,at\,risk\,(bgp)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*

Relative incidence frequency measures are important in comparisons (e.g. between two or more population groups, between two or more populations etc.)

Calculation of this measure in practice is presented in Case study 2.

Classification of incidence measures according to that if the measure expresses the incidence in total observation time or in several parts of it

Although in epidemiology the term »cumulative incidence« is commonly referring to incidence proportion, its intrinsic meaning is referring to cumulation of something (similarly as in statistics). According to this characteristic we distinguish between cumulative and partial, usually annual incidence. The later is the most frequently used measure among possible partial measures.

1. Cumulative incidence.
   This measure is the number or proportion of individuals under observation in which the onset of observed disease was registered during the entire specified period of observation. Usually it is expressed as a proportion, and it is calculated as follows (equation 10):

$$ {}_{cum}I_{rel\,(proportion)} = \frac{N_{d+\,new\,cases(gp)}}{N_{all\ persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 10.} $$

$ {}_{cum}I_{rel\,(proportion)} $ = *cumulative incidence as proportion (proportion of candidate individuals who developed the disease during the entire given period)*

$ N_{d+\,new\,cases\,(gp)} $ = *number of new cases of the disease under observation during a given period*

$ N_{all\ persons\,at\,risk\,(bgp)} $ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*

   The period of observation (the beginning and the end of the period) has to be exactly stated. The beginning could be based upon callendar time or upon some event in a life-course of individuals under observation (the time of the diagnosis of the disease under observation, or exposure to an agent). This interval is generally the same for all members of the group of individuals under observation what is true only for fixed cohorts. When withdrawals are present, calculation of this measure is more complicated (7,14,22). Usually in cohort studies, there are several losses of individuals under observation from follow-up. This is the situation in which the occurrence of the event of interest is uncertain because of different reasons. A situation in which the time-to-event is unknown is called censoring (24). Detailed discussion on this issue is beyond the scope of this module, and is being worked out in a separate module in this book.

   Calculation of this measure in practice is presented in Case study 2.

2. Partial incidence.
   Total period of observation could be split to several parts in order to get more correct estimate of incidence, especially the when frequency is varying over time. Annual incidence is usually a representative of a partial incidence (one should note that annual incidence could be also a cumulative incidence if a course of a phenomenon under observation is relatively rapid). This measure is

the number or proportion of individuals under observation in which the onset of observed disease was registered during the 1-year period. If it is expressed as a proportion, it could be calculated as follows (Equation 11):

$$ann\,I_{rel\,(proportion)} = \frac{N_{d+new\,cases(1\text{-}year\,period)}}{N_{all\ persons\,at\,risk\,(beginning\,of\ 1\text{-}year\,period)}} \qquad \textbf{Equation 11.}$$

$_{ann}I_{rel\ (proportion)}$ = *annual incidence as a proportion (proportion of candidate individuals who developed the disease during the given period)*

$N_{d+\ new\ cases\ (1\text{-}year\ period)}$ = *number of new cases of the disease under observation during1-year period*

$N_{all\ persons\ at\ risk\ (beginning\ of\ 1\text{-}year\ period)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given 1-year period*

Calculation of this measure in practice is presented in Case study 2.

Classification of relative incidence measures according to type of relative frequency measure upon which are based

According to type of relative frequency measure upon which are based relative incidence measures we distinguish between relative incidence as a proportion, as a rate, as odds, and as an incidence density.

1. Measures in which the numerator is included in the denominator.

   Incidence proportion (21)**.** Incidence proportion is a proportion of individuals under observation who developed the disease under observation during a period of observation out of all individuals under observation who were free of disease at the beginning of the specified period of observation (but at risk for getting the disease). Here we need to introduce two very important terms, being »risk« and »cumulative incidence« (1,2,7,9,10,21). Frequently it seems that risk and cumulative incidence are the same measure, although this is true under very restricted conditions (7). Both terms are closely related to incidence proportion.

   Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period (7). Thus, risk is a conditional probability, with values varying between zero and one. It is dimensionless (7). It usually refers to the first occurrence of the disease for each initially disease-free individual, although it is possible to consider the risk of developing the disease under observation within a specified period more than once (7).

   In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and most widely used (7). For a cohort of subjects followed for a given period of time, risk is often estimated by

calculating the proportion of candidate subjects who develop the disease during the observation period. This measure is usually referred as the cumulative incidence (CI) (7). One should be aware that in this case the term »cumulative incidence« is in a role of a technical term, used more in a meaning of incidence proportion than in a meaning of incidence, cumulated over time. The observation period has to be clearly stated since the value of the measure is increasing with the prolongation of period of observation. This period could be based upon a callendar time or not (e.g. first year after the exposure, first year after surgery etc.). Generally cumulative incidence is estimated only for first occurrence of the disease. If the durations of the individual follow-up periods for all non-cases are equal, the cumulative incidence is equivalent to the average risk for members of the cohort. This means thait under the condition of a fixed cohort cumulative incidence is good estimate of risk. This is the reason that cumulative incidence and risk are frequently equalized. But once again, because risk is, by its definition, a conditional probability, it cannot be accurately estimated by calculating cumulative incidence unless all subjects in the observed candidate population are followed for the entire follow-up period or are known to develop the disease (or other observed phenomenon) during the period (7).

We have already presented the equation for calculation of incidence proportion (Equation 9), but since usually this measure is frequently denoted as risk (R), we repeat it as a new equation (Equation 12). In this module, this notation will be used from now on.

$$R = \frac{N_{d+\,new\,cases(gp)}}{N_{all\ persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 12.}$$

R = *risk*
$N_{d+\ new\ cases\ (gp)}$ = *number of new cases of the disease under observation during a given period*
$N_{all\ persons\ at\ risk\ (bgp)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*

Calculation of this measure in practice is presented in Case study 2.

Risk could be estimated using different methods (simple cumulative, actuarial, density, or Kaplan Meier method) (7,14,21), which will be discussed in a separate module in this book.

Incidence rate (in a classic epidemiologic sense) (1-3,23)**.** Under the term »incidence rate« many types of ratios are frequently referred, including proportions (21). One of them is incidence rate in a classic epidemiologic sense. This measure is a ratio between new cases of the diasease under observation in a given period and total number of the population at risk for getting a disease at the beginning of the given period, with suitable multiplier. Mostly is calculated by the equation (Equation 13):

$$IR = \frac{N_{d+\,new\,cases(gp)}}{N_{all\ persons\,at\,risk\,(bgp)}} \times K$$ **Equation 13.**

IR = incidence rate
$N_{d+\ new\ cases\ (gp)}$ *= number of new cases of the disease under observation during a given period*
$N_{all\ persons\ at\ risk\ (bgp)}$ *= number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*
K = multiplier (100, 1000, 10.000, 100.000…)

Calculation of this measure in practice is presented in Case study 2.

2. Measures in which the numerator is not included in the denominator.

Incidence odds (1,2,22,25)**.** Incidence odds (also known as disease, or risk odds) is the measure of odds of getting ill during the period of observation. It is a ratio of conditional probability of developing the disease (risk) to conditional probability of not developing it (1-risk) (7). If we abridge the elements of this ratio, odds is a ratio of new cases of the disease under observation to persons who remained non-cases during the period of observation (Equation 14):

$$IO = \frac{R}{1-R} = \frac{N_{d+\,new\,cases(gp)}}{N_{d-\,(gp)}}$$ **Equation 14.**

IO = incidence odds
$N_{d+\ new\ cases\ (gp)}$ *= number of new cases of the disease under observation during a given period*
$N_{\ d-\ (gp)}$ *= number of all persons without the disease under observation (non-cases) during a given period*

Calculation of this measure in practice is presented in Case study 2.

This measure has no practical value, since in an incidence (cohort) study we can calculate an incidence proportion, or person-time incidence rate.

Incidence density (1,2,7,14,16,21,22)**.** Although, as it was emphasized previously, many types of ratios (including proportions) are frequently referred to as »rates«, in its precise usage a »rate« is the ratio of a change in one quantity to a change in another quantity, with the denominator quantity often being time (21).

In measurement of incidence, there exist a measure that measures how rapidly new cases of a phenomenon under observation are developing (when a death is a phenomenon under observation, how rapidly persons with a disease under observation are dying), or that measures the change in frequency of a health phenomenon to a change per unit of time. Some epidemiologists use the term »incidence rate« to denote this instantaneous measure (7,10), while others have referred to this concept as an instantaneous risk (7), or hazard rate (especially when death is the event under observation) (7,26,27). This measure is measuring the

instantaneous potential for change in disease status (from being disease-free to being diseased) per unit of time, relative to the size of the candidate population (disease-free population) at a given moment in time (1,7,27). If this measure is contrasted to incidence risk (incidence proportion), it is an instantaneous measure, which refers to a point in time and not to a period. Also, the incidence risk is dimensionless while person-time incidence rate is expressed in units of 1/time or $time^{-1}$ (e.g. $years^{-1}$) (7). In fact is the probability of the event under observation occurring within the time unit (e.g. day, month, year) under observation, given that it did not occur to that time unit (e.g. to that day).

The problem of this measure is that we usually cannot express the size of the population at risk under observation (population free of disease at the beginning of the observation period) as a mathematical function of time, and thus we cannot express instantaneous incidence rates. Instead we estimate an average incidence rate for a given period. This is analogous to the use of speed as an estimate of average velocity) (7,10). The speedometer in a car is measuring how fast we are travelling at the moment of time we are looking at the speedometer. This does not mean that we are travelling with the same velocity all the time. The velocity is an example of an instantaneous rate. If we would read the velocity every few seconds for an hour, we could obtain an average velocity per hour. But there exist another measure, called speed that estimates the average velocity. The speed is change in location divided by a change in time (we look at the kilometers counter at the beginning of the one-hour trip and at the end). The speed is an example of an average rate. Comming back to epidemiologic data, it is much easier to determine an average rate than an instantaneous rate. Incidence density is an average rate for estimating average of instantaneous incidence rates (7,26). For this measure other terms are used as well, being incidence rate, person-time incidence rate, average incidence rate, and force of morbidity (26)[1].

Technically, incidence density is the rate between the number of new cases which occur during the period under observation, and the quantity known under the term person-time (PT). This measure is expressed in units of 1/time or $time^{-1}$ as well. It could be calculated as (Equation 15):

$$ID = \frac{N_{d+newcases(gp)}}{PT} \qquad \textbf{Equation 15.}$$

[1] In this place we need to give a warning about terminology used for this group of measures. A common problem in epidemiology of existence of multiple terms for the same concept is very explicit here. Also, a single term is applied to different concepts. The usage of terms differs from one author to another. On one hand, for example, Kleinbaum et al (7) are using the term »incidence rate« to denote the instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation, and denote the average rate for estimating average of instantaneous rates as incidence density. The term »hazard rate« is used as a synonym for »incidence rate«. On the other hand, for example, Benichou and Palta (26) are using the term »hazard rate« to denote the instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation, and denote the average rate for estimating average of instantaneous rates as »incidence rate«. The term »incidence density« is used as a synonym for »incidence rate«.

ID = incidence density
$N_{d+ \text{ new cases (gp)}}$ *= number of new cases of the disease under observation during a given period*
PT = person-time

The measure is interpreted as average incidence rate for a cohort during the period under observation. Incidence density is the measure among family of incidence measures, which could play a role of an autonomous measure, or a role of intermediate measure in the process of estimating of incidence risk, what will be discussed in a separate module in this book.

As an element for calculation of incidence density, the quantity person-time or, person-time at risk, is introduced. It is the quantity which encomprises the information on number of individuals under observation at risk for getting the disease under observation (free of disease at the beginning of the observation period), and the exact time interval of this risk (the time between the beginning of the observation and the moment of break-out of the disease) (1). Mathematically it is the sum of the periods of time at risk for each of the individuals under observation. This method enables to take into account how much of time exactly contributes each individual under observation to the population at risk, and thus to measure incidence rate over extended and variable time periods in a dynamic cohort in which there are several censored observations (deaths of other causes, change of domicile etc.). Usually the time period is one year, and the measure is person-year (PY). In this concept each individual under observation contributes to the population at risk that many years as much as he/she was under observation before the disease under obsevation broke out (an individual under observation that is observed 1 year contributes 1 person-year, an individual under observation that is observed 9 months 0.75 person-year etc.). PY could be calculated as (Equation 16):

$$PY = t(y)_{ob1} + t(y)_{ob2} + ... + t(y)_{obn} \qquad \textbf{Equation 16.}$$

PY = person-years
$t(y)_{ob1}$ *= No. of years at risk for individual under observation No.1*
$t(y)_{ob2}$ *= No. of years at risk for individual under observation No.2*
$t(y)_{obn}$ *= No. of years at risk for individual under observation No.n*

Calculation of PY and incidence density in practice is presented in Case study 2.

Special incidence measures

Mortality. Mortality is one of the most important epidemiologic and demographic measures which could be classified in the family of incidence measures (8,19). It is a ratio between number of deaths during a given period (usually 1 callendar year)

and number of all persons at risk of dying during given period at the beginning of this period (usually number of the population, usuallly estimated at the middle of the year of the observation) (1,2,8). In fact, mortality is the incidence of death.Technically it is usually expressed as rate (in a classic epidemiologic sense) (mortality rate, or death rate), and could be calculated as (Equation 17):

$$Mo = \frac{N_{deaths(gp)}}{N_{all\ persons\ at\ risk\ (bgp)}} \times K \qquad \textbf{Equation 17.}$$

Mo = mortality rate
$N_{deaths\ (gp)}$ *= number of deaths during a given period*
$N_{all\ persons\ at\ risk\ (bgp)}$ *= number of all persons at risk of dying during given period at the beginning of this period*
K = multiplier

Hazard rate. We have already introduced the term »hazard rate« when we were introducing the concept of measuring instantaneous potential per unit of time for event under observation to occur, given that the individual has survived up to the time (moment) of observation. In the case when the observed event is a death from a disease, this measure is usually known as »hazard rate« (1,7,26,27). Similar to idea of velocity, a hazard function h(t) gives the instantaneous potential at time t for getting an event. Estimate of average of these instantaneous potentials could be calculated as follows (27) (Equation 18)

$$h = \frac{N_{deaths(gp)}}{PT_d} \qquad \textbf{Equation 18.}$$

h = hazard rate
$N_{deaths\ (gp)}$ *= number of deaths during a given period*
PT_d *= person-time (sum of periods at risk for death for each individual)*

When the unit of time under observation is a day, this measure could be interpreted as the risk of dying for the person on that particlar day, given that he/she has survived to that day (27).

Relationship between prevalence and incidence

Prevalence and incidence are very closely related (2,3). This relationship is ilustrated in Figure 3.

Input to the prevalence pool represent incident cases (new cases of a disease under observation), while output represent recoveries and deaths. Indirectly the prevalence depends on duration of the disease. If the recovery rate is low, and the mortality is low as well, the chronicity of the disease is high, and consecutively the prevalence is high. In such a case even low incidence brings to the high prevalence.

Prevalence thus depends on incidence and duration of the disease. When both quantities are stable, and the prevalence of the disease is low (e.g. in cancer), this relationship could be expressed as follows (Equation 19) (2):

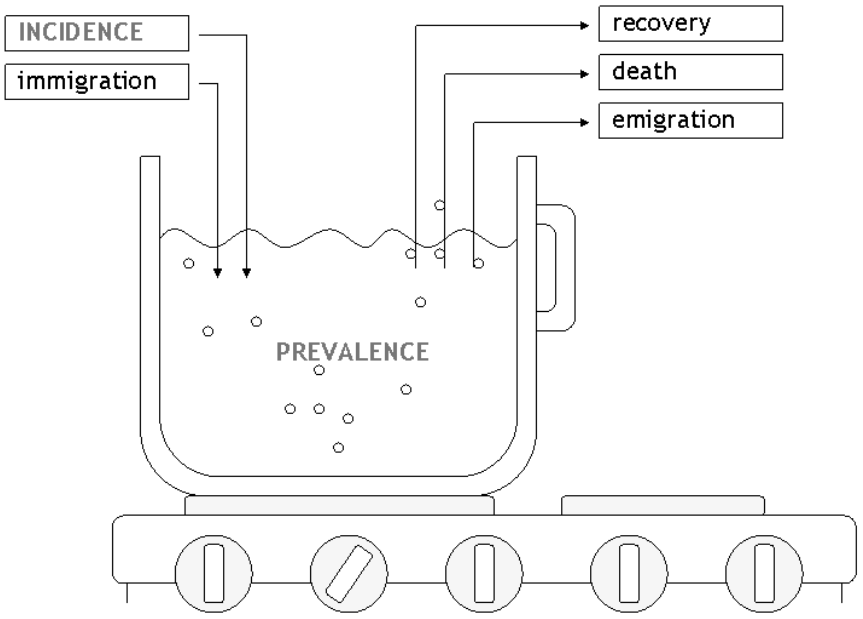


Figure 3. The ralationship between prevalence and incidence of a health phenomenon.

$$P = ID \times average\ duration\ of\ the\ disease$$ **Equation 19.**

P = prevalence
ID = incidence density

But incidence, recovery rate and mortality rate of the disease under observation are not the only factors which influence the prevalence. The smaller part of the prevalence pool input contribute also the immigrated cases, while the smaller part of the output contribute also emigrated cases. Also there exist the influence of competitive factors like mortality of extraneous factors (deaths because other causes than disease under observation e.g. traffic accidents). Because of a lot of possible influences prevalence always has to be interpreted cautiously.

CASE STUDIES

Data set

The illustration of differences between families of prevalence and incidence measures, and the differences between measures inside both families is be based upon an ideal set of example data (Figure 4). The example data-set could be described as follows:

1. We have 20 individuals under observation which are all followed up for exactly 5 years. The course of events during 5-year observation time is shown in Figure 4. The most important example data characteristics are as follows:
    - at the beginning of the study all individuals under observation are without disease under observation, and
    - all of them are exposed to the effect of the same noxious agent,
    - some of them get ill and some not, and
    - all cases of disease are diagnosed.
2. Other important characteristics (for easier understanding of measures) are also:
    - all members enter the study at the same time (at the beginning of the study), and
    - nobody gets out the study (because of recovery, death, or change of domicile) – our cohort is a fixed cohort, all are followed-up exactly the same time,
    - the disease under observation is supposed to be chronic (there is no recovery after becoming diseased).

Case study 1: Prevalence measures

Absolute and relative prevalence

For calculating the absolute prevalence let us choose the point 2 years after beginning of the study (Figure 5). Results of counting of existing cases of observed disease exactly 2 years after beginning of the study (Figure 5) show that there exist 9 persons with the disease (e.g. cases of the disease). Thus absolute prevalence or prevalent number, or simply prevalence, of the observed disease 2 years after beginning of the study is 9.

For calculating the relative prevalence let us choose again the point 2 years after beginning of the study (Figure 5). Results of counting of existing cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there exist 9 individuals under observation with the disease (e.g. cases of the disease) among a whole group of 20 individuals under observation. Relative prevalence as prevalence proportion or probability for having a disease under observation at point 2 years after beginning of the study is, when calculated according to Equation 3 as a decimal fraction, (Equation 20):

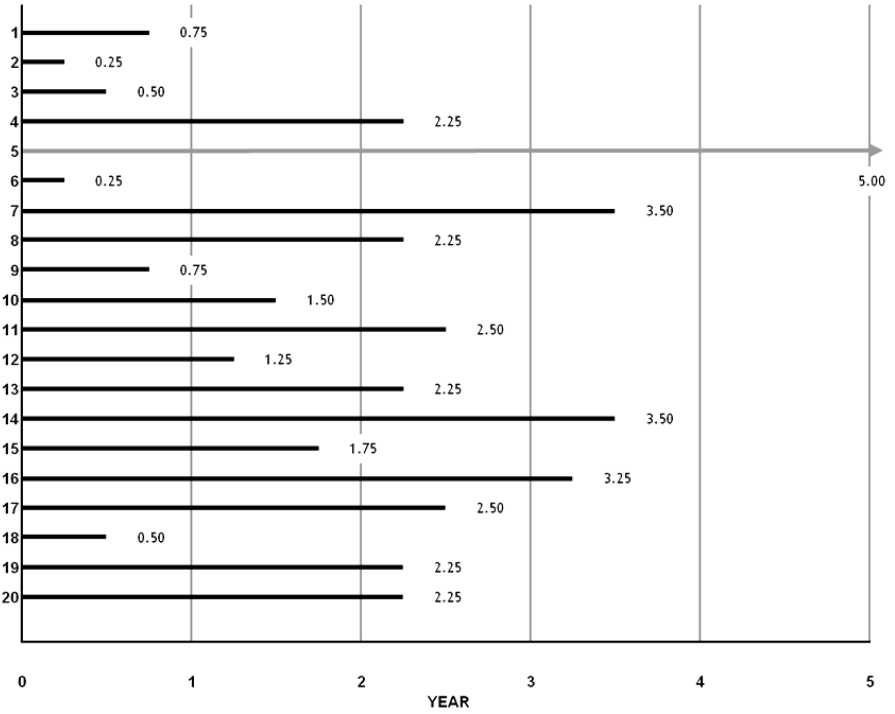


Figure 4. The course of events during 5-year observation time of 20 individuals under observation, exposed to the effect of the same noxious agent, in the example data set. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

$$P_{rel\,(proportion)} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 20.}$$

The relative prevalence expressed as prevalence proportion (probability of having a disease under observation) at point 2 years after beginning of the study is 0.450 or 45.0%.

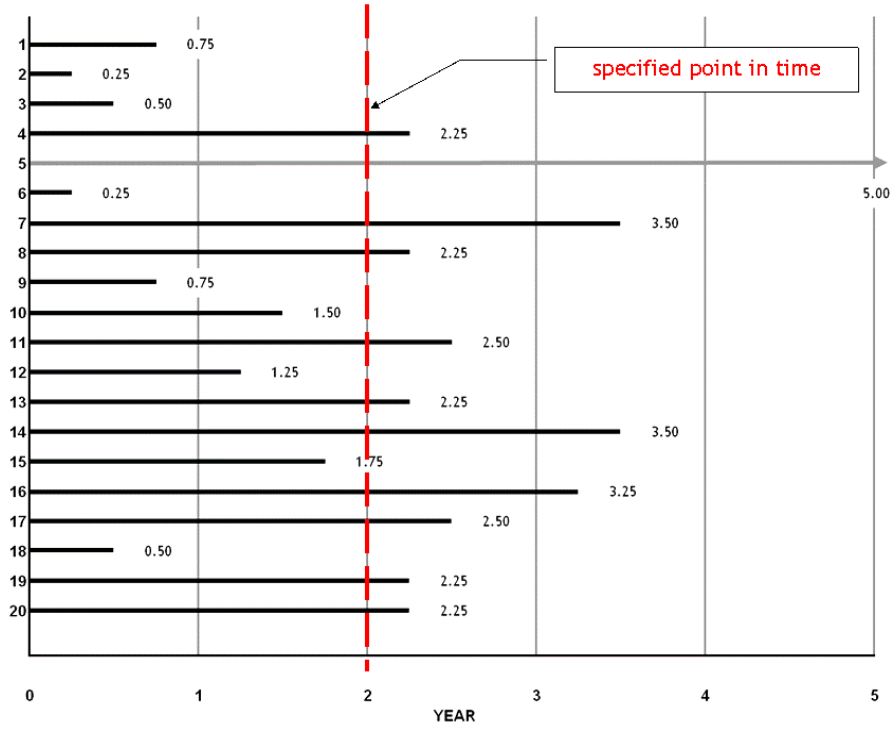


Figure 5. Graphic presentation of point prevalence exactly two years after beginning of the observation period on example data. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Point and period prevalence

For calculating the point prevalence let us choose again the point 2 years after beginning of the study (Figure 5). The point prevalence in this point expressed as absolute point prevalence is 9, while expressed as a relative point prevalence (as a proportion) is according to Equation 4 (Equation 21):

$$_{point}P_{rel\,(proportion)} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 21.}$$

For calculating the period prevalence under assumption of fixed cohort, and assumption of chronicity of the disease under observation (once an individual gets the disease he/she cannot recover) let us choose the period of the second year of the study (Figure 6, dashed frame). The period prevalence in this period expressed as an absolute period prevalence is 9 - six individuals (No. 1, 2, 3, 6. 9 and 18) already had a disease at the beginning of the second year of the observation, while three of them got the disease during the second year period (No. 10, 12 and 15). The period prevalence, expressed as a relative period prevalence (as a proportion) is according to Equation 5 (Equation 22):

$$_{period}P_{rel\,(proportion)} = \frac{6+3}{20} = \frac{9}{20} = 0.450 \qquad \textbf{Equation 22.}$$

If the observed period is only the first half of the second year of the study (Figure 6 gray filled part of the dashed frame), the period prevalence in this period expressed as an absolute period prevalence is 7 - six individuals already had a disease at the beginning of the second year of the observation, while one of them got the disease during the first half of the second year period. The period prevalence, expressed as a relative period prevalence (as a prevalence proportion) is (Equation 23):

$$_{period}P_{rel\,(proportion)} = \frac{6+1}{20} = \frac{7}{20} = 0.350 \qquad \textbf{Equation 23.}$$

Relative point prevalence measures

Prevalence proportion. Results of counting of existing cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there exist 9 persons with the disease (e.g. cases of the disease) among 20. We have already calculated the prevalence proportion (Equations 19 and 20). We repeat this equation with a new notation according to Equation 6 (Equation 24):

$$P = \frac{9}{20} = 0.450 \qquad \textbf{Equation 24.}$$

Prevalence rate. Relative prevalence could be also expressed as prevalence rate (in a classic epidemiologic sense). In this case, it is calculated when the multiplier is 1,000 according to Equation 7 as follows (Equation 25):

$$PR = \frac{9}{20} \times 1{,}000 = 0.450 \times 1{,}000 = 450 \qquad \textbf{Equation 25.}$$

The relative prevalence of the disease under observation expressed as prevalence rate at point 2 years after beginning of the study is 450 per 1,000.

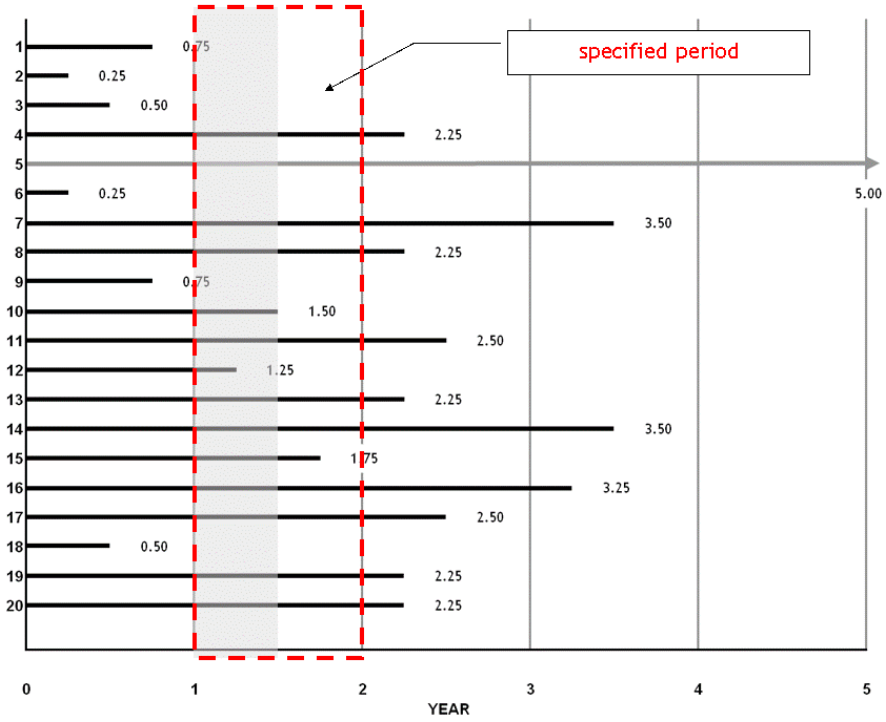


Figure 6. Graphic presentation of period prevalence in the whole second year (dashed frame) or in the first half of the second year of the study (gray filled part of the dashed frame) on example data. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Prevalence odds. Results of counting of cases and non-cases of observed disease exactly 2 years after beginning of the observation period (Figure 5) show that there existed 9 persons with the disease (e.g. cases of the disease) and 11 without it (non-cases). Relative prevalence as prevalence odds for having a disease at this point in time is calculated according to Equation 8 as follows (Equation 26):

$$PO = \frac{9}{11} = 0.818 \qquad \textbf{Equation 26.}$$

The relative prevalence of the disease under observation expressed as prevalence odds at point 2 years after beginning of the study is 0.818. This means that 2 years after exposure there is 0.818 of a person with disease to 1 person without it (or if we calculate reverse odds – 1.222 of a person without a disease to 1 with it).

Case study 2: Incidence measures

Absolute and relative incidence

For calculating the absolute incidence let us choose the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease which broke out during the 5-year period (Figure 7) show that there were 19 cases, thus absolute incidence of the observed disease in a 5-year period of the study is 19.

For contrasting the relative incidence let us choose again the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease in which onset of this disease was registered during the 5-year time of observation (Figure 7) show that there were 19 cases among 20 individuals under observation. Relative incidence as incidence proportion during the 5-year period of observation is when calculated according to Equation 9 as a decimal fraction (Equation 27):

$$I_{rel\,(proportion)} = \frac{19}{20} = 0.950 \qquad \textbf{Equation 27.}$$

The relative incidence expressed as incidence proportion during the 5-year period of observation is 0.950 (or when expressed as a percentage, 95.0%).

Cumulative and partial incidence proportion

For contrasting the cumulative and partial incidence let us choose again the entire 5-year period of observation (Figure 7). Results of counting of cases of observed disease which broke out during the entire 5-year time of observation (Figure 7) show that absolute cumulative incidence is 19, and relative cumulative incidence expressed as risk according to Equation 10 is (Equation 28):

$${}_{cum}I_{rel\,(proportion)} = \frac{19}{20} = 0.950 \qquad \textbf{Equation 28.}$$

Thus the cumulative incidence proportion for the 5-year period is 0.950 (or when expressed as a percentage, 95.0%).

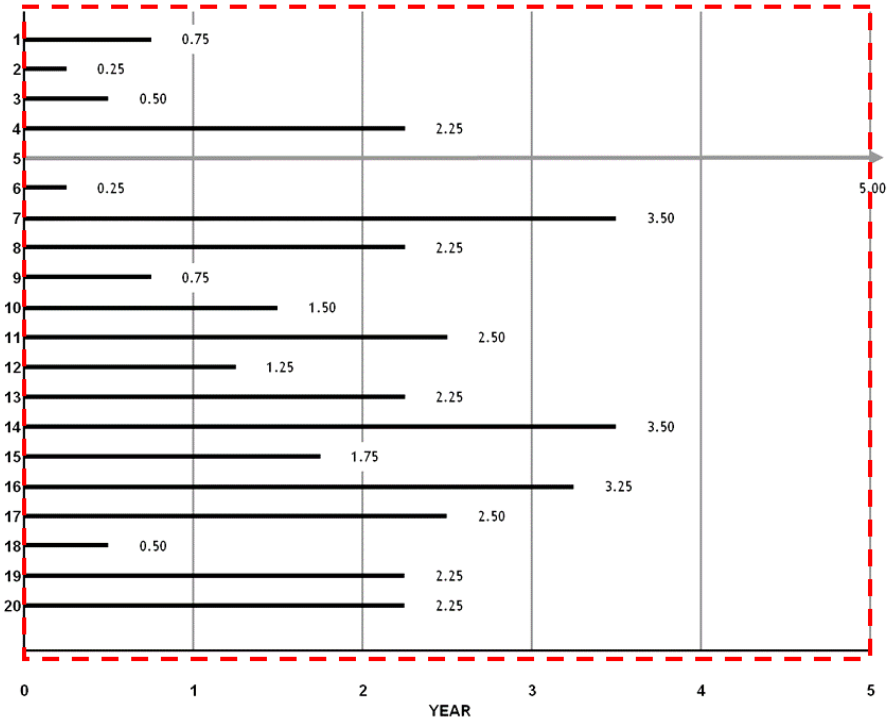


Figure 7. Graphic presentation of incdence in the whole period of study on example data. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Results of counting of cases of observed disease which broke out during the first year of observation (Figure 8, dashed frame) show that there 6 cases of the disease appeared within this period thus absolute annual incidence in the first year of observation is 6. Relative annual incidence expressed as a proportion according to Equation 11 is (Equation 29):

$$ann\,I_{rel(proportion)(year1)} = \frac{6}{20} = 0.300 \qquad \textbf{Equation 29.}$$

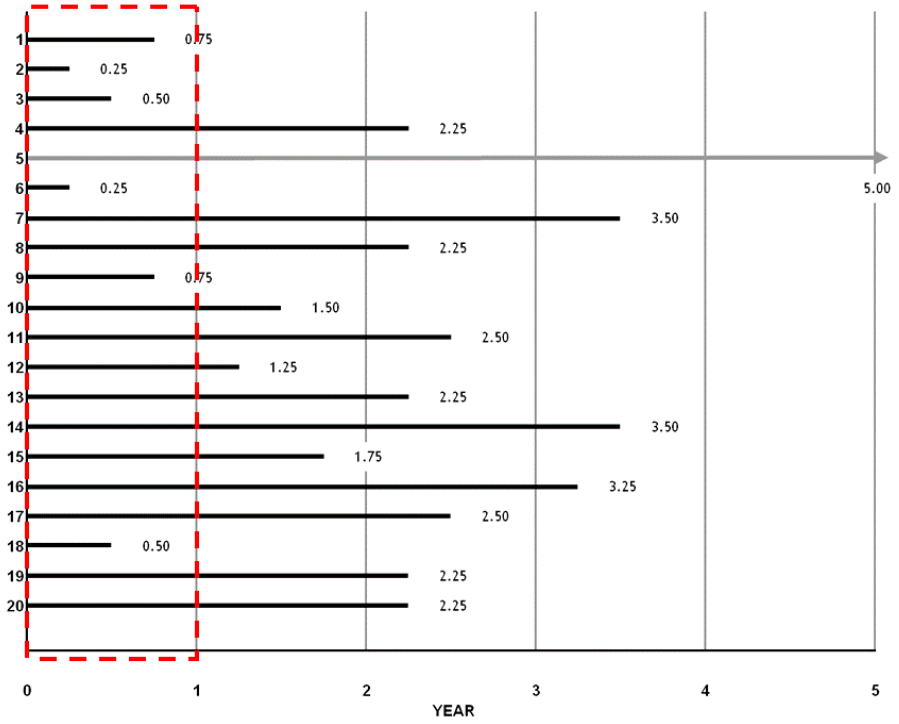


Figure 8. Graphic presentation of annual incidencein the first year of the study (red frame) on example data. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Thus the annual incidence proportion for the the first 1-year period is 0.300 (or when expressed as a percentage, 30.0%).

Relative annual incidences expressed as a proportion for the following four years are (Equations 30-33):

$$_{ann}I_{rel\,(proportion)(year2)} = \frac{3}{14} = 0.214 \qquad \textbf{Equation 30.}$$

$$_{ann}I_{rel\,(proportion)(year3)} = \frac{7}{11} = 0.636 \qquad \textbf{Equation 31.}$$

$$_{ann}I_{rel(proportion)(year4)} = \frac{3}{4} = 0.750$$ **Equation 32.**

$$_{ann}I_{rel(proportion)(year5)} = \frac{0}{1} = 0.000$$ **Equation 33.**

Relative incidence measures

Incidence risk (estimate). Results of counting of cases of observed disease in which onset of this disease was registered during the 5-year time of observation (Figure 7) show that there were 19 cases among 20 individuals under observation. According to Equation 12, estimate of incidence risk is calculated as follows (Equation 34):

$$R = \frac{19}{20} = 0.950$$ **Equation 34.**

Thus, the incidence risk for the 5-year period of observation estimated by calculating cumulative 5-year incidence proportion is 0.950 (or when expressed as a percentage, 95.0%).

Incidence rate (in a classic epidemiologic sense). Relative incidence could be also expressed as incidence rate (in a classic epidemiologic sense). In this case, it is calculated when the multiplier is 1,000 according to Equation 13 as follows (Equation 35):

$$IR = \frac{19}{20} \times 1{,}000 = 0.950 \times 1{,}000 = 950$$ **Equation 35.**

The relative incidence of the disease under observation expressed as incidence rate during the 5-year period of observation is 950 per 1,000.

Incidence odds. Results of counting of cases and non-cases of observed disease at the end of the 5-year period of observation show, that there exist 19 persons with the disease (e.g. cases of the disease) and 1 without it (non-case). Relative incidence as incidence odds of getting a disease during 5-year period according to Equation 14 is (Equation 36):

$$IO = \frac{19}{1} = 19.000$$ **Equation 36.**

The relative incidence expressed as incidence odds at the end of the 5-year period of observation is 19.000. This means that in 5-year interval there will be 19

persons with disease to 1 person without it (or if we calculate reverse odds – 0.053 of a person without a disease to 1 with it).

Incidence density. The last relative incidence measure to be presented is incidence density. To calculate this measure we need first to calculate a quantity called person-years (Equation 16). Table 1 (based on Figure 4) is presenting the exact duration of time in whict each individual under observation was under observation before the disease under observation broke out (time of being at risk). In total all 20 individuals under observation were exposed (at risk) in 5-year period 39 person-years.

Table 1. Data for calculation of person-years.

Id. number	Onset of the disease (0=no, 1=yes)	Time of being at risk* (Years)
1	1	0.75
2	1	0.25
3	1	0.50
4	1	2.25
5	0	5.00
6	1	0.25
7	1	3.50
8	1	2.25
9	1	0.75
10	1	1.50
11	1	2.50
12	1	1.25
13	1	2.25
14	1	3.50
15	1	1.75
16	1	3.25
17	1	2.50
18	1	0.50
19	1	2.25
20	1	2.25
Total	Diseased: 19	Person-years: 39.00

* time in which an individual under observation is exposed to effect of noxious agent (is at risk of an event: getting ill)

In continuation, the incidence density for data presented in Figure 4 and Table 1 could be calculated for 5-year period according to Equation 15 as (Equation 37):

$$ID = \frac{19}{39} = 0.4872 \qquad \textbf{Equation 37.}$$

If we than multiply the ID with 1,000 we get the value 487, which could be interpreted as: on average in 5-year interval 487 individuals under observation got ill per 1,000 population with the disease under observation if they are exposed to the effect of the noxious agent.

EXERCISE

Data set 1

A cohort of 20 individuals initially without a disease under observation, were followed up for 5 years. Times of events are presented in Figure 9.

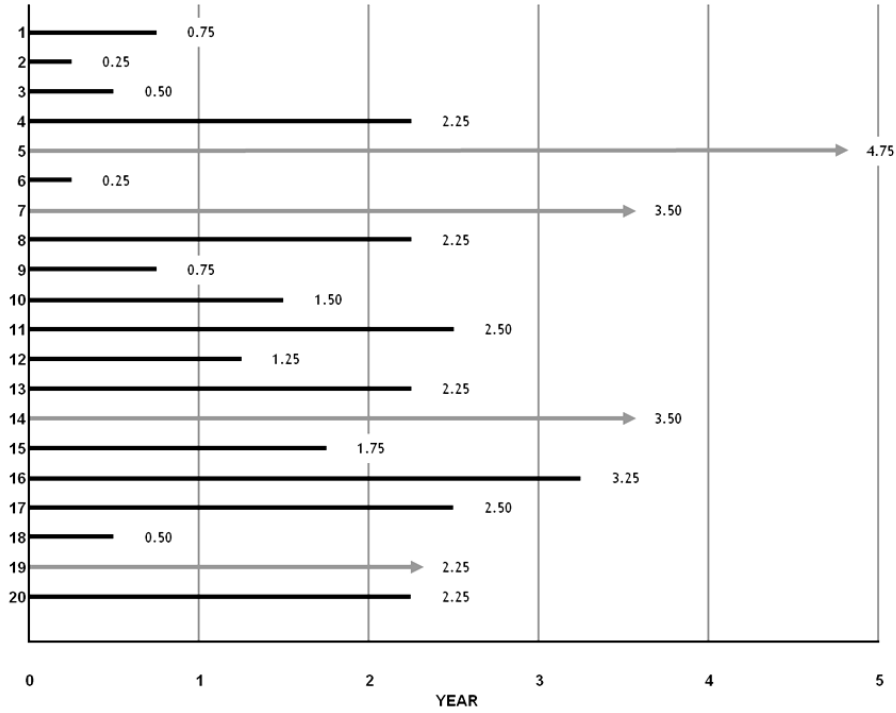


Figure 9. Graphic presentation of events in a cohort of 20 people. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Task 1

For the example set of data presented in Figure 9, please, calculate:

- absolute prevalence and relative prevalence as prevalence proportion at the point one years after beginning of the study.
- relative prevalence as prevalence proportion and prevalence odds at the point two years after beginning of the study.

Task 2

For the example set of data presented in Figure 9, please, calculate:

- cumulative incidence as a proportion for 5-year interval; discuss whether this measure is reliable or not,
- annual incidences as incidence proportion for each year of observation,
- incidence density for 5-year interval; discuss whether this measure is reliable or not.

Data set 2

In Figure 10, another imaginary data-set is presented. Again, a cohort of 20 individuals initially without a disease under observation, were followed up for 5 years.

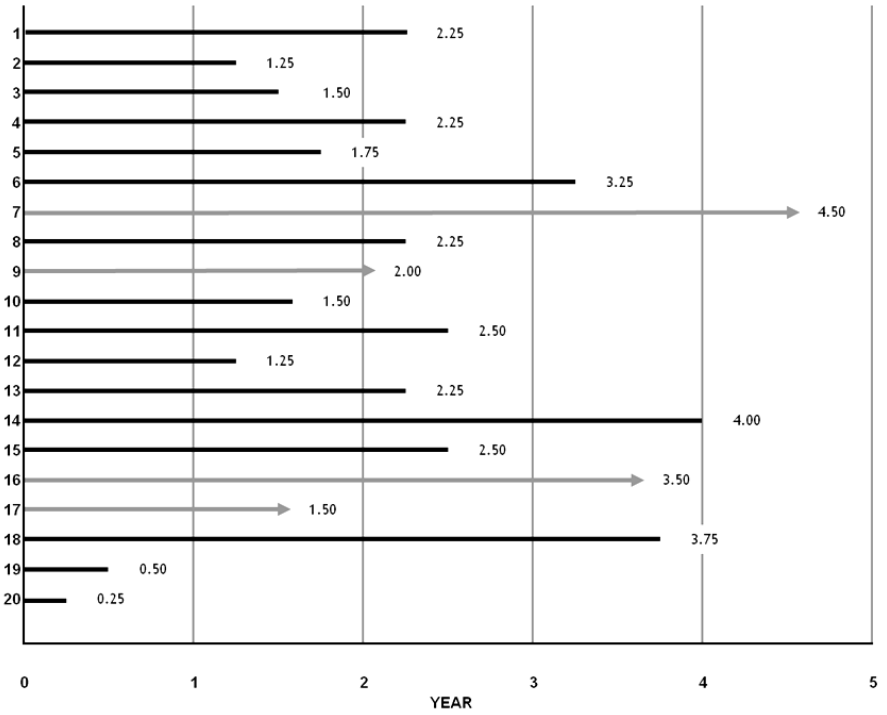


Figure 10. Graphic presentation of events in a cohort of 20 people. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease

under observation before an event occurred) in individuals that developed the disease under observation; — the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Task 3

For the example set of data presented in Figure 10, please, calculate:

- cumulative incidence as a proportion for 5-year interval; discuss whether this measure is reliable or not,
- annual incidences as incidence proportion for each year of observation,
- incidence density for 5-year interval; discuss whether this measure is reliable or not.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
4. Anonymous. Prevalence and incidence. WHO Bull, 1966;35:783-784.
5. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
6. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
7. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
8. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing Group; 1997. Available from URL: http://www.bmj.com/epidem/epid.html. Accessed: August 24, 2008.
9. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
10. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
11. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
12. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
13. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
14. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
15. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
16. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25].

Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 25, 2009.
17. Babuš V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
18. Friis RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
19. Kue Youtng T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
20. Timmreck TC. An introduction to epidemiology. 3rd ed. Sudbury, MA: Jones&Bartlett Publishers; 2002.
21. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.
22. Pearce N. A short introduction to epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003. Available from URL: www.massey.ac.nz/~wwwcphr/publications/introepi.pdf. Accessed: June 25, 2009.
23. Bayona M, Olsen C. Measures in epidemiology. Washington, D.C.: Young Epidemiology Scholars Program YES; 2004. Available from URL: http://www.clepprepcenter.net/prod_downloads/yes/4297_MODULE_09.pdf. Accessed: June 25, 2009.
24. Kleinbaum DG, Klein M. Survival analysis. A self-learning text. Second edition. New York: Springer; 2005.
25. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1247374&blobtype=pdf. Accessed: June 24, 2009.
26. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
27. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.

RECOMMENDED READINGS

1. Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005.
2. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**AGE STANDARDIZATION PROCEDURE: DIRECT METHOD**
Module: 1.2.3	**ECTS (suggested): 0.15**
Author(s), degrees, institution(s)	**Jadranka Božikov**, PhD, Professor Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Doris Bardehle,** MD, PhD, Professor Faculty of Health Sciences, University of Bielefeld, Germany
Address for correspondence	**Jadranka Božikov** Andrija Štampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: jbozikov@snz.hr
Keywords	Confounding, standardization, age standardization, direct standardization, indirect standardization, cumulative rate
Learning objectives	After completing this module students should: • understand a role confounding in epidemiologic studies; • increase knowledge about methods of controlling of confounding in epidemiologic studies; • understand principles of direct standardization; • be capable to calculate age-standardized death rates using direct method.
Abstract	Basic theoretical background of standardization as one of methods for controlling the effect of confounding in epidemiology is presented. Direct method of standardization as most common standardization method is presented in details, using a case study. Step by step the procedure is described using simple spreadsheet computer tool for facilitating it.
Teaching methods	Teaching methods include introductory lecture, exercises, and interactive methods such as small group discussions. Students after introductory lectures first carefully read the recommended sources in age standardization. Afterwards they discuss standardization as method of controlling confounding with other students. In continuation, they in practice in groups of 2-3 students perform the procedure of direct standardization using the programme tool (e.g. MS Excel) on given data. At the end they compare and discuss their results.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1 computer per 2-3 students), LCD projection, access to the Internet; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire (MCQ) and case-study.

AGE STANDARDIZATION PROCEDURE: DIRECT METHOD

Jadranka Božikov, Lijana Zaletel-Kragelj, Doris Bardehle

THEORETICAL BACKGROUND

Population diversity and confounding

When examining the health of populations one of the fundamentals of this process is the comparison of health indicators among across and/or across different population subgroups within the countries.

Whenever we want to compare epidemiologic measures, irrespective of what they represent: morbidity (e.g. incidence or prevalence measures), mortality or other measure, across different populations or population groups we should take into account their diversity (1). Namely, populations/population groups are heterogeneous in regard to various health related characteristics (e.g. age, gender, education, religion, genetic and geographic factors, etc.) (2).

When the epidemiological measures are calculated without taking into account this diversity, such kind of epidemiological measures are called crude measures. The potential influence of the diversity could be imagined if the procedure of calculation of crude values is taken into consideration - the value of crude population measure is in fact an average of the values for the individual subgroups within a population (e.g. subgroups according to age), weighted by their relative sizes (1). This means, the larger the subgroup (e.g. age subgroup), the more influence it will have on the crude measure. The comparison of crude measures across populations (or population groups) can be thus misleading because they can be greatly affected by the influence of such characteristics (e.g. different age distributions in the populations/population groups being compared).

In statistical terms, these characteristics are so called confounders. Confounding (from the Latin "confundere" that means to mix together) is according to Last et al. defined as an effect which appears when the measurement of the effect of an exposure on a risk is distorted by the relation between the exposure and other "extraneous" factor (or multiple factors) that also influence the outcome under study (3). In this context extraneous factors are considered as factors other than the relationships between two phenomena under study. But not every characteristic meets the criteria for being confounder. A confounding factor (or confounder) must meet three criteria:

- to be a known risk factor for the result of interest (4),
- to be a factor associated with exposure but not a result of exposure (4), and
- to be a factor that is not an intermediate variable between them.

Thus, when crude rates are interpreted, this interpretation would have been confounded by differences in the populations being compared (e.g. differences in age distribution). We therefore need to control for the effects of confounders in order to remove the confounding effect.

Controlling for the effects of confounding

There exist several procedures for controlling for the effects of confounding. Some of them could be performed in the designing and plannig phase of a study, and the others in the phase of data analysis (5-7). The first group of procedures (e.g. randomization, restriction, matching) is usually performed in experimental studies while the second group (stratification, standardization, statistical modelling) in observational studies (5-7). This concept of control of confounding in epidemiology derives from the limited opportunities for experimental control in non-experimental design of studies.

In practice, age is the factor that is most frequently controlled or adjusted for confounding. In an older population higher rates of certain diseases that more frequently appear in older age-groups (e.g.cancers) could be observed not because of the presence of risk factors, but because of the higher age itself (8). Traditionally in controlling for age confounding, standardization is applied (8).

Standardization

Definition and description

Standardization of health indicators is a classic epidemiological method defined as:

- a set of techniques used to remove as much as possible the effects of differences in age or other confounding variables when comparing two or more populations (3),
- a method that removes the confounding effect of variables that we know (or think) that they could influence the comparison between two or more populations (5,6),
- a statistical method for deriving measures that are comparable across populations that differ in age and other demographic characteristics (9).

Standardization provides an easy-to-calculate and easy-to-use summary measures e.g. standardized mortality (abbreviated sometimes as SMR[2]) or standardized death rate (abbreviated as SDR[3]) when the outcome is death, or a standardized morbidity measure when the outcome is disease occurrence (e.g. standardized incidence rate in the case the morbidity measure is incidence - abbreviated sometimes as SIR[4]). These measures can be useful for information users, such as decision-makers.

Types of standardization

Two approaches to standardization could be used, direct and indirect (1,3,5-9). They are used in different situations what will be described in continuation.

---

[2] We should be aware that this abbreviation, SMR, is also used in the case of standardized mortality ratio as an outcome measure in indirect standardization procedure which is not the subject of this module.

[3] The term standardized death rate (abbreviated as SDR) is commonly used in Health for All Data Base of WHO, European Region (10).

[4] The same as under 1.

Direct standardization

Direct standardization is a procedure that forms a weighted average of age specific rates or risks, using as weights the distribution of a scpecified standard population (1,3,8,9).

The method is called "direct" because it uses the actual morbidity or mortality rates of the populations being compared (9).

In the direct standardization method, according to Last et al. (3), the directly standardized rate represents what the crude rate would have been in the observed population if that population had the same structure as the standard population with the respect to the variable (or more variables) for which the standardization was performed.

Thus, these rates are hypothetical and by themselves they are not meaningful because they are not real. These rates are useful only if they are used in comparisons of populations in the case that standardized rates in all compared populations are derived by the same procedure using the same standard population.

Direct standardization could be used to compare observed populations for which the specific crude rates are known and statistically stable. It is commonly used in reports of vital statistics (e.g., mortality) or major disease incidence trends (e.g., cancer incidence).

Indirect standardization

Indirect standardization is used to compare observed populations for which the specific crude rates are unknown or statistically unstable (3). This is frequently in small populations or when the observed phenomenon is rare.

It is different from direct standardization in both, method and interpretation. Instead of using the structure of the standard population, we utilize its specific rates and apply them to the populations under comparison, previously stratified by the variable to be controlled. The total of expected cases is obtained this way. The SDR is then calculated by dividing the total of observed cases by the total of expected cases. This ratio allows comparison of each population under study to the standard population. A conclusion can be reached by simply calculating and looking at the SDR. A SDR higher than one (or 100% if expressed in percentage) indicates that the risk of dying in the observed population is higher than what would be expected if it had the same experience or risk than the standard population. On the other hand, a SDR lower than one (or 100%) indicates that the risk of dying is lower in the observed population than expected if its distribution were the same as the reference population.

Indirect standardization plays a major role in studies of occupational disease.

Age standardization

Although age standardization is not a special type of standardization we think it is worthy to emphasize it. As already mentioned, age is the factor that is most frequently standardized for, since the age is one of the most important confounders. Compared populations could have very different age structure that can influence the interpretation of differences in crude rates of observed phenomenon.

Age-standardized rates calculated using the direct method represent what the crude rate would have been if the population had the same age distribution as the standard population.

Age-standardization is particularly used in comparative mortality studies, since the age structure has an important impact on a population's overall mortality.

Limitations of standardization

It is important to know that standardization as method for controlling confounding has some limitations. Any summary measure can hide patterns that might have important public health implications. For example, with age standardization, one might fail to detect age-specific differences in risk across time or place. This might arise if a disease is displaying an increasing incidence due to a birth cohort effect (people at younger ages might have a higher risk in recent years compared to previous years, while older people could have the opposite pattern). An age-standardized rate could hide these trends. Despite this risk, standardized rates have proved to be very useful summary measures.

The procedure of direct age standardization

Entry data for the procedure

For accomplishing the procedure of direct age standardization we need three sets of data:

1. Number of cases of a health phenomenon (death, disease) to be standardized.
   We need absolute frequency (number of cases) of a health phenomenon to be standardized across the age groups.

   These data are usually derived from registration of health phenomena (mortality, morbidity data) - health statistics of a country. Usually are administered by national public health institutes. Mortality data are usually available, while morbidity data (e.g. cancer incidence) are more difficult to obtain. In Slovenia for example cancer incidence for several sites could be obtained from a high quality Cancer Registry of the Republic of Slovenia. The Registry's annual reports, Cancer Incidence in Slovenia, are one of the regular ways of disseminating information of this registry. They are publicly available from their homepage as PDF files (11).
2. Observed population data.
   Next set of data that is needed for direct standardization is distribution of population according to age.

   These data are usually derived from on-going registration of population and/or population censuses. They are usually provided by every country's statistical office. For example, for Slovenia these data are provided by the Statistical Office of the Republic of Slovenia. They are publicly available in Office's annual reports, Statistical Yearbook, from their homepage as PDF files (12).

   For most of countries of the world these data could also be obtained from the U.S. Census Bureau International Data Base Entry (13).
3. Standard population data.

An important step in direct standardization is the selection of a standard population (4), since value of the adjusted rate depends on the standard population used.

The standard population may come from the populations under study - average for example. In this case however, it is important to ensure that the populations do not differ considerably in their size, since a larger population may influence the adjusted rates (14). The standard population may also be a population without any relation to the data under study, but in general, its distribution with regard to the adjustment factor should not be radically different from the populations we wish to compare.

In European region of World Health Organization, for comparison across countries within this region, age-standardized death rates are calculated using the European standard population while in other regions other standard populations. The detailed description of the European standard population could be obtained from the European Health for All Database manual (15). The age distribution of four different hypothetical standard populations is presented in Table 1 (15,16).

Table 1. Some standard populations. Adapted from Health for All database Manual (15) and Surveillance, Epidemiology and End Results (SEER) Program homepage (16).

Age group	European standard population (100,000)	European standard population (million)	World standard population (million)	1996 Canadian standard population (million)	2000 US standard population (million)
0	1,600	16,000	24,000	12,342	13,818
1-4	6,400	64,000	96,000	53,893	55,317
5-9	7,000	70,000	100,000	67,985	72,533
10-14	7,000	70,000	90,000	67,716	73,032
15-19	7,000	70,000	90,000	67,841	72,169
20-24	7,000	70,000	80,000	67,761	66,478
25-29	7,000	70,000	80,000	72,914	64,529
30-34	7,000	70,000	60,000	87,030	71,044
35-39	7,000	70,000	60,000	88,510	80,762
40-44	7,000	70,000	60,000	80,055	81,851
45-49	7,000	70,000	60,000	71,847	72,118
50-54	7,000	70,000	50,000	55,812	62,716
55-59	6,000	60,000	40,000	44,869	48,454
60-64	5,000	50,000	40,000	40,705	38,793
65-69	4,000	40,000	30,000	37,858	34,264
70-74	3,000	30,000	20,000	32,589	31,773
75-79	2,000	20,000	10,000	23,232	26,999
80-84	1,000	10,000	5,000	15,424	17,842
85+	1,000	10,000	5,000	11,617	15,508
Total	100,000	1,000,000	1,000,000	1,000,000	1,000,000

The European standard population which will be used in our case study is also presented in Figure 1.

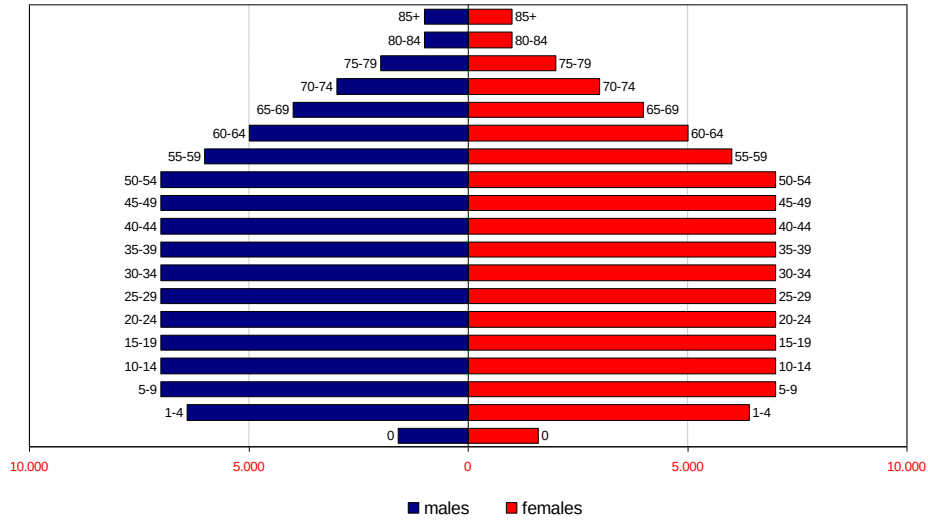


Figure 1. European standard population (100,000). Adapted from Health for All database Manual (15).

For all three sets of enty data the same age distribution is needed.

The procedure

Directly standardized rate is, in general, calculated by dividing the number of deaths by the actual local population in a particular age group multiplied by the standard population for that particular age group and summing across the relevant age groups. The rate is usually expressed per 100,000. The exact procedure for calculating standardized death rates in 4 steps is as follows:

1. Step 1 - calculation of the specific crude death rate for every (specific) age group. The crude specific death rate for every age group is obtained by dividing the number of deaths in every specific age group by the observed (actual local) population in this age group multiplied by a multiplier (usually 100,000) (Equation 1):

$$ {}_{crude}DR_{(spec.group)} = \frac{N_{deaths(\,spec.group)}}{N_{pop(\,spec.group)}} \times 100{,}000 \qquad \textbf{Equation 1.} $$

${}_{crude}DR_{(spec.\ group)}$ = *crude death rate in a specific population group*
$N_{deaths(spec.\ group)}$ = *number of deaths in a specific population group*
$N_{pop(spec.\ group)}$ = *number of population in a specific population group*

2. Step 2 - calculation of the crude rate for total population.
   The crude rate for total population is calculated using the similar formula as in calculating specific death rate for every age group (Equation 1), except that in this calculation totals of number of cases and population are used (Equation 2).

$$_{crude}DR_{(total\ pop)} = \frac{N_{deaths(total\ pop)}}{N_{pop(total\ pop)}} \times 100{,}000 \qquad \textbf{Equation 2.}$$

$_{crude}DR_{(total\ pop)}$ = *crude death rate in a total population*
$N_{deaths\ (total\ pop)}$ = *number of deaths in a total population*
$N_{pop\ (total\ pop)}$ = *number of population in a total population*

   These totals need to be calculated prior calculation of the crude rate for total population.
3. Step 3 - calculation of the expected number of deaths in the standard population for every specific age group.
   The expected number of deaths in a specific age group is calculated by multiplying the result obtained in step 1 by the number of population in standard population in this specific age group and dividing it by the multiplier used in step 1 (usually 100,000) (Equation 3):

$$N_{exp.deaths(spec.group)} = \frac{_{crude}DR_{(spec.group)} \times N_{stand.pop(spec.group)}}{100{,}000} \qquad \textbf{Equation 3.}$$

$N_{exp.deaths\ (spec.\ group)}$ = *number of expected deaths in the standard population in a specific population group*
$_{crude}DR_{(spec.\ group)}$ = *crude death rate in a specific population group*
$N_{stand.\ pop(spec.\ group)}$ = *number of population in a specific population group of a standard population*

   The result of this step, the expected number of deaths in every specific age group, is in fact the standardized death rate in this particular age group.
4. Step 4 - calculation of the standardized death rate in a total population.
   Finally, the standardized death rate is obtained by summation of expected number of deaths in a specific age group across all age groups (Equation 4).

$$_{stand}DR_{(total\ pop)} = \sum N_{exp.deaths(spec.group)} \qquad \textbf{Equation 4.}$$

$_{stand}DR_{(total\ pop)}$ = *standardized death rate in a total population*
$N_{exp.\ deaths\ (spec.\ group)}$ = *number of expected deaths in the standard population in a specific population group*

CASE STUDY: THE PROCEDURE OF DIRECT AGE STANDARDIZATION OF DISEASE D MORTALITY IN CROATIA

Entry data

For accomplishing the procedure of direct age standardization of disease D mortality in Croatia we need following sets of data:

1. Number of deaths of a disease D to be standardized.
   In table 2 the number of cases of the disese D in every age group for the male population for year 2000 is presented. The data were obtained from National Health Institute of Croatia (17).

Table 2. Number of death cases (absolute incidence) of the disese D in Croatia for male population for every age group for year 2000. Source: National Health Institute of Croatia (17).

Age group	Number of cases	Age group	Number of cases
0	0	45-49	80
1-4	0	50-54	143
5-9	0	55-59	237
10-14	2	60-64	258
15-19	3	65-69	249
20-24	4	70-74	200
25-29	15	75-79	253
30-34	14	80-84	159
35-39	36	85+	68
40-44	52	Total	1773

2. Observed population data.
   In Table 3 the number of population in every age group of the Croatian population is presented. The 1991 census data are used (Table 3, Figure 2). The data were obtained from Central Bureau of Statistics of Republic of Croatia (18).

Table 3. Croatian population (census 1991) in figures. Source: Central Bureau of Statistics of Republic of Croatia (18).

Age group	Males	Females	Age group	Males	Females
0	26361	26361	45-49	147304	148308
1-4	130000	148272	50-54	155474	161793
5-9	168031	159688	55-59	146177	162304
10-14	166573	159218	60-64	105909	146527
15-19	162383	155564	65-69	69655	113449
20-24	169107	164779	70-74	43815	71653
25-29	179330	175245	75-79	44536	75999
30-34	192397	184039	80-84	23986	44564
35-39	184654	174497	85+	12844	27651
40-44	142937	139918	Total	2271473	2439829

Tha same data are presented also in Figure 2.

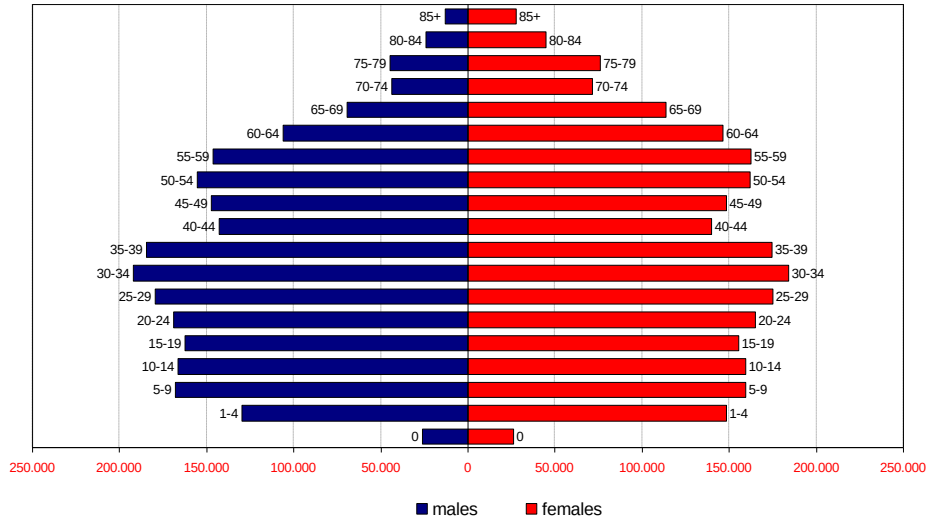


Figure 2. Number of population by sex in nineteen age groups of the Croatian population, the 1991 census data. Source: Central Bureau of Statistics of Republic of Croatia (18).

3. Standard population data.
   Given the fact that we want to compare the mortality rate of disease D in Croatia to other countries of World Health Organization (WHO) European region mortality rates, the European standard population is the best choice for standard population. This standard population has been already presented (Table 1, Figure 1).

Setting the frame table for the standardization procedure

The procedure for standardization of rates could be authomatised by using an appropriate computer programmes. Spreadsheet programme like Microsoft Excel could be used. The frame table for the procedure should contain following columns:

- Age group,
- Number of cases of deaths,
- Population (observed),
- Rate per 100,000,
- European standard population, and
- Expected cases of deaths (in European standard population).

In Figure 3 this frame is presented while in Figure 4 in this frame entry data are already filled in.

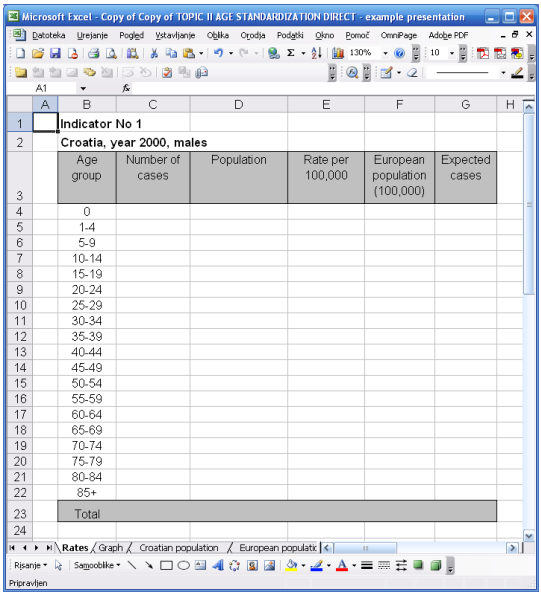


Figure 3. The frame table for the standardization procedure in Microsoft Excel computer programme.

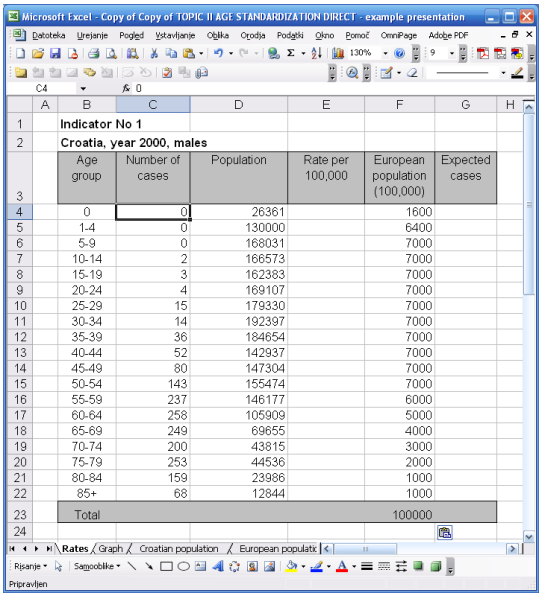


Figure 4. The frame table for the standardization procedure in Microsoft Excel computer programme filled in with entry data.

The procedure

The four steps are as follows:

1. Step 1 - calculation of the specific crude death rate for every (specific) age group.
   The crude specific death rate for every age group is calculated by using the Equation 1. In Figure 5 the equation for calculating the crude specific death rate for the age group 0 using corresponding cells for number of deaths and the observed (actual local) population in this age group in a spreadsheet is presented.

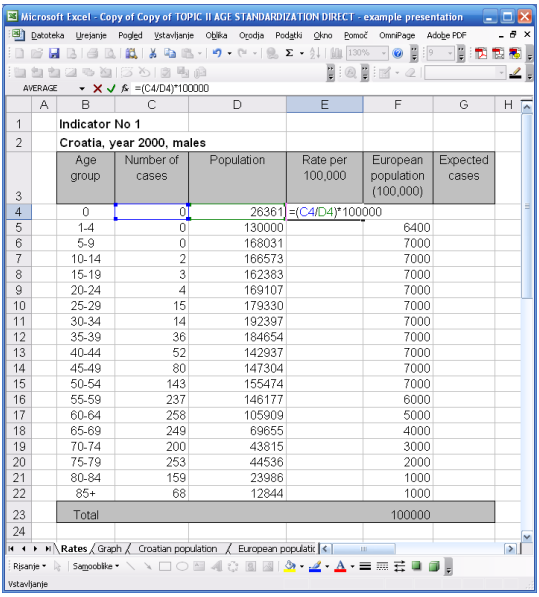


Figure 5. Calculation of the specific rate for age group 0 using corresponding cells for number of deaths, and the observed (actual local) population in this age group in a spreadsheet.

In Figure 6 the results of this step in the procedure is presented.

In Equation 5 the procedure for calculating the crude specific death rate for the age group 40-44 is presented, as well as the result.

$$_{crude}DR_{age\,40\text{-}44} = \frac{52}{142{,}937} \times 100{,}000 = 36.38 \qquad \textbf{Equation 5.}$$

In continuation, the formula from the cell containing the function for calculating the crude specific death rate for the age group 0 is copied to other cells in the same column by dragging the right lower corner of the cell and extending it to the last age group cell (Figure 7).

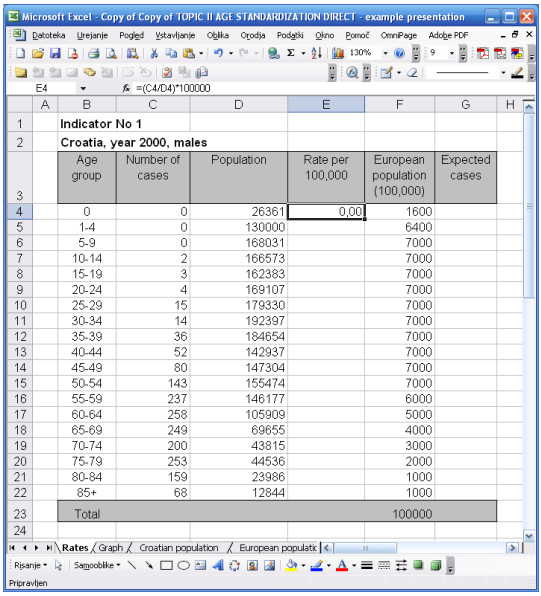


Figure 6. The result of calculation of the specific rate for age group 0.

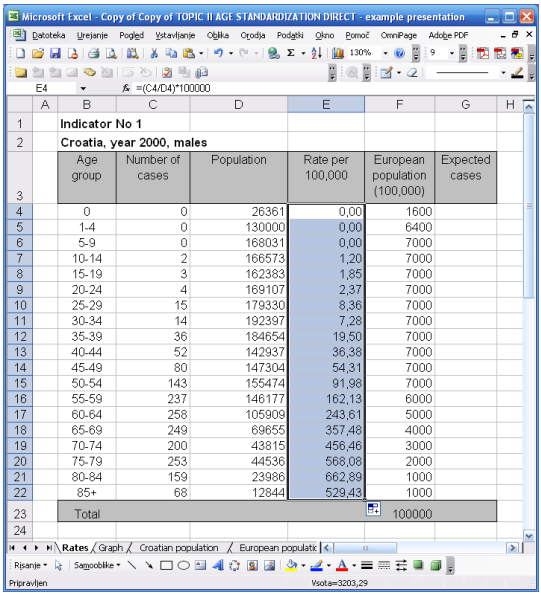


Figure 7. Copying of the function used for calculation of the specific rate for age group 0 to all age groups by dragging the right lower corner of the cell and extending it to the last age group cell.

In Figure 7 we can verify if the result of calculation of the crude specific death rate for the age group 40-44 is correct.

2. Step 2 - calculation of the crude rate for total population.
   Prior calculating the crude rate for total population, the totals of number of cases and population need to be calculated. Figure 8 presents the procedure for calculating the totals by using the SUM function.

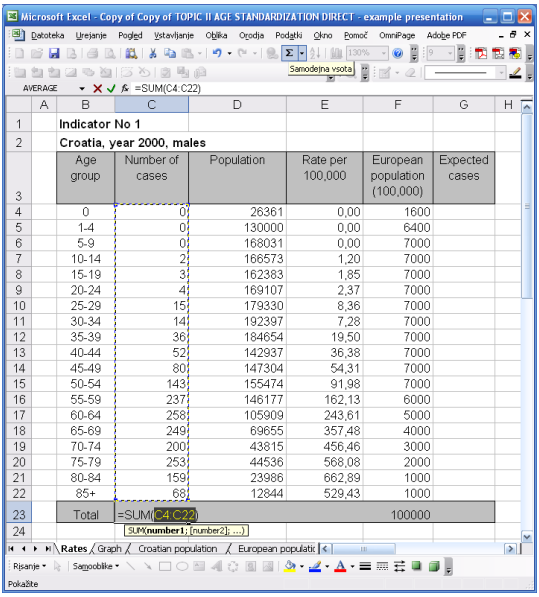


Figure 8. The procedure of calculation of totals for number of deaths and for Croatian population.

In Figure 9 the results of this procedure are presented. By comparing the totals in Tables 2 and 3 we can verify if they are correct.

In continuation, the crude rate for total population is calculated by using the Equation 2 (Equation 6).

$$_{crude}DR_{(total\ pop)} = \frac{1,773}{227,1473} \times 100,000 = 78.06 \qquad \textbf{Equation 6.}$$

The procedeure of calculation of the crude rate for total population using corresponding cells for number of deaths and the observed (actual local) total population in the spreadsheet is presented in Figure 10, while the result in Figure 11.

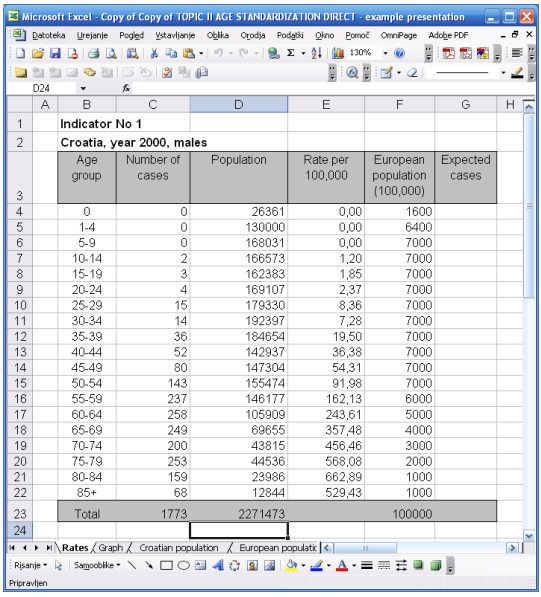


Figure 9. The result of the procedure of calculation of totals for Croatian population, number of cases and European standard population.

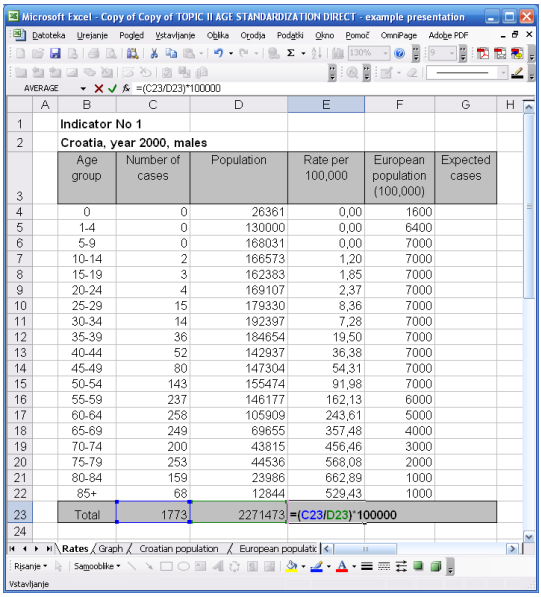


Figure 10. The procedeure of calculation of the crude rate for total population using corresponding cells for number of deaths and the observed (actual local) total population.

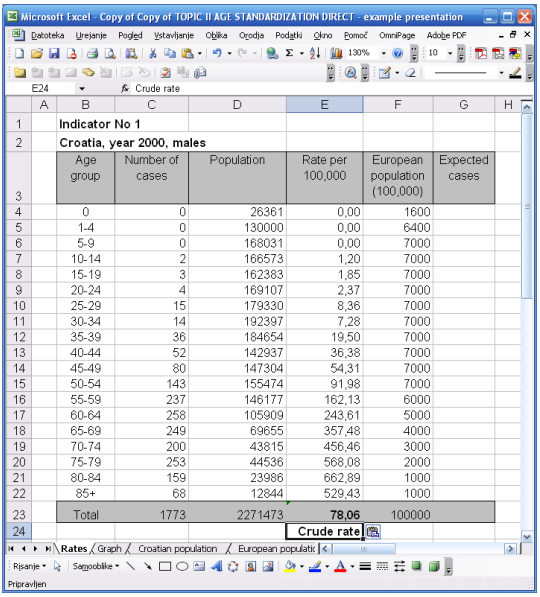


Figure 11. The result of calculation of the crude rate for total population Calculation of the specific rate for age group 0.

3. Step 3 - calculation of the expected number of deaths in the standard population for every specific age group.
   In the next step, the expected number of deaths in a specific age group is calculated by using Equation 3. In Figure 12 the equation for calculating the expected number of deaths in the standard population for the age group 0 using corresponding cells for crude death rate and the standard population in this age group in a spreadsheet is presented. In Figure 13 the results of this step in the procedure is presented.

   In Equation 7 the procedure for calculating the expected number of deaths in the standard population for the age group 40-44 is presented, as well as the result.

$$N_{expected_{age\,40\text{-}44}} = \frac{36.38 \times 7{,}000}{100{,}000} = 2.55 \qquad \textbf{Equation 7.}$$

In continuation, the equation from the cell containing the function for calculating the crude specific death rate for the age group 0 is copied to other cells in the same column by dragging the right lower corner of the cell and extending it to the last age group cell (Figure 14).

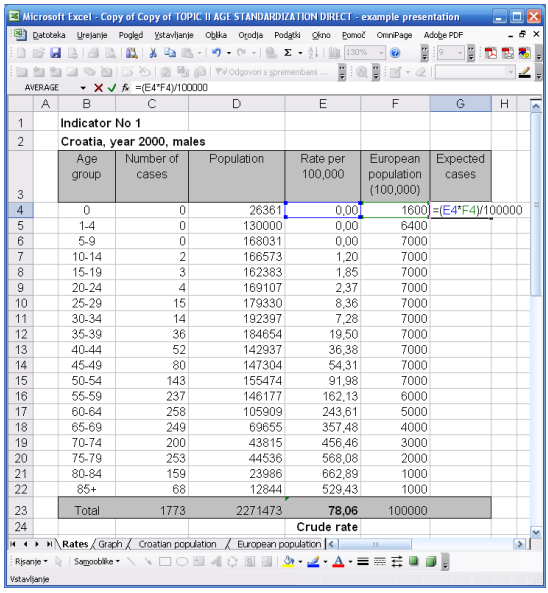


Figure 12. Calculation of the number of expected number of deaths in the standard population for age group 0 by using corresponding cells for crude death rate and the standard population in this age group in a spreadsheet.

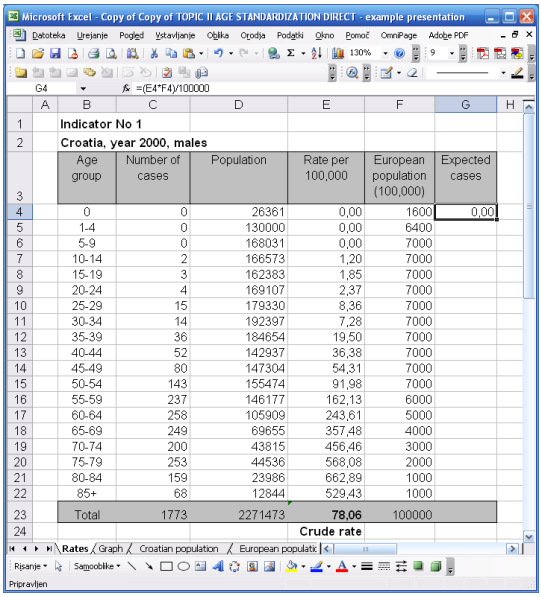


Figure 13. The result of calculation of expected number of deaths in the standard population for age group 0.

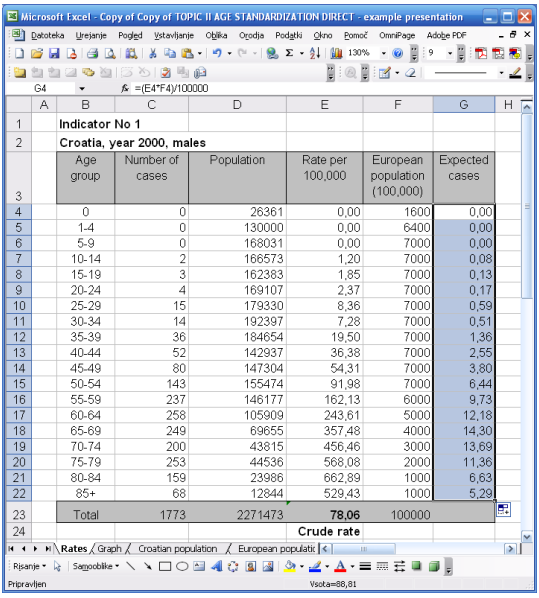


Figure 14. Copying of the function used for calculation of the number of expected number of deaths in the standard population for age group 0 to all age groups by dragging the right lower corner of the cell and extending it to the last age group cell.

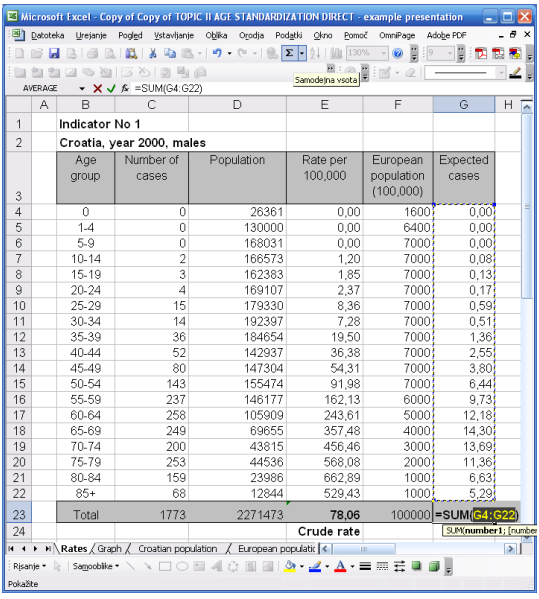
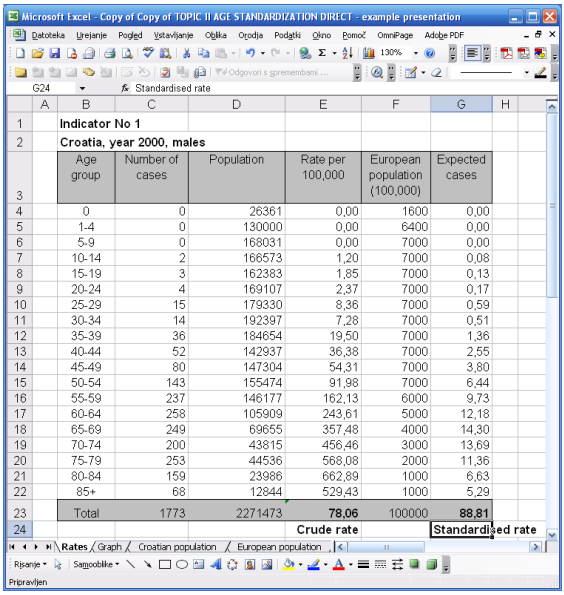


Figure 15. The procedeure of calculation of the crude rate for total population using the SUM function in MS Excel programme.

4. Step 4 - calculation of the standardized death rate in a total population.
   Finally, the standardized death rate is obtained by summation of number of expected cases in standard population across all age groups. Figure 15 presents the procedure for calculating the totals by using the SUM function in MS Excel programme, while Figure 16 presents the final result of the procedure.



	A	B	C	D	E	F	G	H
1		Indicator No 1						
2		Croatia, year 2000, males						
3		Age group	Number of cases	Population	Rate per 100,000	European population (100,000)	Expected cases	
4		0	0	26361	0,00	1600	0,00	
5		1-4	0	130000	0,00	6400	0,00	
6		5-9	0	168031	0,00	7000	0,00	
7		10-14	2	166573	1,20	7000	0,08	
8		15-19	3	162383	1,85	7000	0,13	
9		20-24	4	169107	2,37	7000	0,17	
10		25-29	15	179330	8,36	7000	0,59	
11		30-34	14	192397	7,28	7000	0,51	
12		35-39	36	184654	19,50	7000	1,36	
13		40-44	52	142937	36,38	7000	2,55	
14		45-49	80	147304	54,31	7000	3,80	
15		50-54	143	155474	91,98	7000	6,44	
16		55-59	237	146177	162,13	6000	9,73	
17		60-64	258	105909	243,61	5000	12,18	
18		65-69	249	69655	357,48	4000	14,30	
19		70-74	200	43815	456,46	3000	13,69	
20		75-79	253	44536	568,08	2000	11,36	
21		80-84	159	23986	662,89	1000	6,63	
22		85+	68	12844	529,43	1000	5,29	
23		Total	1773	2271473	78,06	100000	88,81	
24					Crude rate		Standardised rate	

Figure 16. Final result of the procedure of calculating standardized death rate.

EXERCISE

Task 1

Carefully read the theoretical background of this module and discuss the confounding phenomenon with other students.

Task 2

Compare and interpret the crude rate and the SDR for disease D for male population of Croatia (Figure 16) given the disease is cancer (all sites). What do you think such a result mean?

Task 3

Perform the standardization procedure for female population for the disese D mortality in Croatia. Number of death cases (absolute incidence) is presented in Table 4.

Table 4. Number of death cases (absolute incidence) of the disese D in Croatia for female population for every age group for year 2000 (8). Source: National Health Institute of Croatia

Age group	Number of cases	Age group	Number of cases	Age group	Number of cases
0	0	30-34	10	65-69	270
1-4	0	35-39	29	70-74	240
5-9	0	40-44	24	75-79	330
10-14	1	45-49	67	80-84	238
15-19	0	50-54	136	85+	102
20-24	2	55-59	158	Total	1825
25-29	9	60-64	209		

Follow the procedure presented in this paper from Step 1 to Step 4 (Figures 3 thru 16)[5].

Task 4

Compare:

- your results obtained in the Task 3 to results of other students,
- the results of female part of the population (Task 3) to the male part of the population (Case study),
- try critically to discuss the differences.

Task 5

Critically discuss strengths and limitations of standardization procedure in controlling the confounding phenomena.

REFERENCES

1. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.
2. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25]. Available from URL:

[5] The answer:
Crude death rate: 74.80
Standardized death rate: 62.53

http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 24, 2009.
3. Last JM, editor. A Dictionnary of epidemiology. Fourth Edition. New York, New York: Oxford University Press; 2001.
4. Gordis L. Epidemiology. Second Edition. Philadelphia: W.B. Saunders Company; 2000.
5. Dos Santos Silva I. Dealing with confounding in the analysis. In: Dos Santos Silva I. Confounding Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.305-332. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
6. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
7. Perace N, Greenland S. Confounding and interaction. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.371-97.
8. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
9. Milic N. Standardization. In: Kirch W, editor. Encyclopedia of public health. Leipzig: Springer Verlag; 2008. p.1335-8.
10. World Health Organization, Regional Office for Europe. European health for all database. Copenhagen: World Health Organization, Refional Office for Europe, 2002. Available from URL: ftp://ftp.euro.who.int/hfadb. Accessed: September 5, 2009.
11. Cancer Registry of the Republic of Slovenia [homepage on the Internet]. Available from URL: http://www.onko-i.si/en/sectors_of_the_institute_of_oncology/epidemiology_and_cancer_registries/cancer_registries/cancer_registry_of_rs/. Accessed: September 5, 2009.
12. Statistical Office of the Republic of Slovenia. Statistical Yearbooks of the Republic of Slovenia [homepage on the Internet]. Available from URL: http://www.stat.si/eng/pub_letopis_prva.asp. Accessed: September 5, 2009.
13. U.S. Census Bureau. International Data Base Entry [homepage on the Internet]. Available from URL: http://www.census.gov/ipc/www/idb/informationGateway.php. Accessed: September 5, 2009.
14. Kramer S. Clinical epidemiology and biostatistics. A primer for clinical investigators and decision-makers. Berlin: Springer-Verlag; 1988.
15. World Health Organization, Regional Office for Europe. European health for all database. Manual. Copenhagen: World Health Organization, Refional Office for Europe, 2002. Available from URL: ftp://ftp.euro.who.int/hfa/hfa-db.pdf. Accessed: September 5, 2009.
16. U.S. National Institutes of Health, National Cancer Institute, Surveillance, Epidemiology and End Results (SEER) Program. Standard Populations (Millions) for Age-Adjustment [homepage on the Internet]. Available from URL: http://seer.cancer.gov/stdpopulations/index.html. Accessed: June 24, 2009.
17. Croatian National Institute of Public Health. Croatian health service yearbook 2000. Zagreb: Croatian National Institute of Public Health; 2001.
18. Central Bureau of Statistics of Republic of Croatia. Population by sex and age, by settlements. Census of population, households and dwellings, 2001. Zagreb: Central Bureau of Statistics of Republic of Croatia, 2001. Available from URL: http://www.dzs.hr/default_e.htm. Accessed: April 4, 2008.

RECOMMENDED READINGS

1. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25]. Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 24, 2009.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon, IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
3. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Bown and Company; 1987.
4. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASURES OF LOCATION: MEASURES OF CENTRAL TENDENCY AND DISPERSION**
Module: 1.2.4	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Gena Grancharova**, MD, PhD, Associate Professor Faculty of Public Health, Medical University of Pleven, Bulgaria **Silviya Aleksandrova**, MD, MB, Assistant professor Faculty of Public Health, Medical University of Pleven, Bulgaria
Address for correspondence	**Gena Grancharova** Medical University of Pleven 1, Kliment Ochridski Str. Pleven 5800, Bulgaria e-mails: dean-ph@mu-pleven.bg; gegran@optisprint.net
Keywords	Mean, median, mode, standard deviation, variance, inter-quartile range
Learning objectives	After completing this module students should be able to: • explain why summary measures are needed in medicine and public health; • define the meaning and compute the mean, median and mode of a given set of data (grouped and ungrouped); • define the meaning and compute the range, inter-quartile range, variance, standard deviation and coefficient of variation; • describe the limitations, advantages and disadvantages of different measures of central tendency and dispersion; • select appropriate measures of cental tendency and dispersion for a given data situation; • discuss the concept of normality of health data in terms of measures of central tendency and dispersion.
Abstract	Central tendency and dispersion measures (CTDM) are essential for summarizing any data set of individual scores. This process is based on two main characteristics of quantitave data - its variability and tendency to some typical level. This section is devoted to the numerical approach of data summarising with the objective to underline the meaning of different measures, to present some simple basic methods of converting the raw data into meaningful summary statistics. The advantages and disadvantages of different CTDM are underlined in relation to different scales of data presentation, and to different form of frequency distributions. The module is also describing the main uses of different CTDM and the concept of normality of health data.
Teaching methods	Two-hour lecture introduce the students to the main concepts of CTDM. After the lecture students read and discuss in groups all the material presented in this section and individually answer the multiple choice questionnaire.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1-2 student), LCD projector, access to the Internet and statistical package software; • training materials: recommended readings or other related readings; • target audience: bachelor and master students in public health.
Assessment of students	Multiple choice questionnaire (MCQ) - minimum 70% success.

MEASURES OF LOCATION: MEASURES OF CENTRAL TENDENCY AND DISPERSION

Gena Grancharova, Silviya Aleksandrova

THEORETICAL BACKGROUND

Introduction

In the chapter on organizing and describing data it was demonstrated how raw data could be organized and presented in a meaninful way. A frequency distribution gives a good general picture of the pattern of the observations but sets of measurements cannot be adequately described only by the values of all individual measurements. For many purposes, the overall summary of a group's characteristics is of utmost importance (1,2). The researcher usually asks questions such as, "What is the average fruit consumption in the group under study?" or "What is the average age of women delivering low birth weight infants?" Such questions seek a single number that best represents the whole distribution of the corresponding data values. It is obvious that further summarization is necessary, particularly before inferences or generalizations are drawn from the data under observation (3-6).

The process of summarization is based on two main characteristics of quantitave data.

1. firstly, the principal problem encountered in working with medical and public health data is its variability. Even when we follow all the strict requirements for random sampling we would almost certainly obtain different values of the variables studied in particular populations and samples, and
2. secondly, despite the individual fluctuations, many of the variables used in the behavioural or life sciences are distributed in such a way that most scores fall in the middle, with fewer scores falling on either side, in the "tails" of the distribution (7). In other words, the values of the most quantitative variables tend to some typical "middle" level (central point or the most characteristic value) around which all the values are distributed. Because an index of typicalness is more likely to be representative if it comes from the center of a distribution than if it comes from either extreme, such measures are referred to as measures of central tendency (2,6,8).

The central tendency is due to determining factors and causes inherent in all cases of a given sample or population while the variability or dispersion is due to specific factors which may occur in some cases but may be absent in others.

There are two basic methods of summarization: numerical and graphical. The objective of the numerical approach is to convert masses of numbers (raw data) into meaningful summary statistics (indices), reduced to a single number, that convey information about the average (typical) degree of a given variable and the degree to which observations differ (the degree of dispersion or spread).

Measures of central tendency

Before presenting the specific measures of central tendency, it is important to be familiar with some basic terms. An array of a set of numbers is simply those numbers in (algebraically) ordered sequence from the lowest to the highest (5). Each array has the following basic components:

- x - each individual raw score in a sample or in a population;
- n - the number of cases in a sample;
- N - the number of cases in a population;
- f - frequency (the number of observations with the same value);
- range - the difference between the largest and the smallest value in an array;
- Σx - the sum of all values in a sample or in a population

Before presenting the specific measures of central tendency, it is important also to know the shape of the distribution (there are distributions that do not assume a "normal" distribution) and the dispersion of the scores in order to interpret the data correctly.

Measures of central tendency are measures of the location of the middle or the center of a distribution. The definition of "middle" or "center" is purposely left somewhat vague so that the term "central tendency" can refer to a wide variety of measures. The three most commonly reported measures of central tendency are (9,10):

- the arithmetic mean,
- the median, and
- the mode.

Mean

Arithmetic mean

The mean is the most commonly used measure of central tendency. When the word "mean" is used without a modifier, it can be assumed that it refers to the arithmetic mean.

In a sample arithmetic mean is denoted by $\bar{x}$ while in a population it is denoted by μ.

Mathematically the mean is the sum of all the scores divided by the number of scores. In Equation 1 calculation of mean in a sample is presented :

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum x}{n} \qquad \textbf{Equation 1.}$$

$\bar{x}$ = *mean of a sample*
n = *number of units in a sample studied*
x_n = *value of a single unit*

If we are studying a population, then the calculation is exactly the same only denotation is different (Equation 2):

As the public health investigations are mostly based on samples from which the conslusions and generalizations for the populations are made, further on in the text we will use $\bar{x}$ as a symbol of arithmetic mean.

$$\mu = \frac{x_1 + x_2 + ... + x_n}{N} = \frac{\sum x}{N}$$ **Equation 2.**

μ = *mean of a population*
N = *number of units in a population studied*
x_n = *value of a single unit*

The approaches to compute the arithmetic mean depend on the way on which the initial data are presented (raw or groupped data), and, if we have no statistical computer programme or hand calculator available, on the number of cases (statistical units) (3).

Calcualtion of the arithmetic mean in practice when raw data are available is presented in Case study 1, Case 1. The procedure of calculation has already been presented in Equations 1 and 2.

When only groupped data are available the mean can also be calculated. The procedure is different when group intervals are equal, or when they are not:

- groups of unequal intervals (Equation 3):

$$\bar{x} = \frac{\sum xf}{\sum f}$$ **Equation 3.**

$\bar{x}$ = *arithmetic mean*
$\sum xf$ = *the sum of products of values of variable X* ($x_1, x_2, x_3, ... x_n$) *and the absolute frequency f for each value of variable X*
$\sum f$ = *the sum of absolute frequency f for each value of variable X that equals to the number of cases*

Calcualtion of the arithmetic mean in practice when only groupped data are available, and intervals of groups are not equal is presented in Case study 1, Case 2.

This approach could be quite boring when the sample is big and the variable has many values. In such situations the transformation of the grouped data into equal width intervals (classes) is preferred;

- groups of equal intervals (classes) (Equation 4):

$$\bar{x} = \frac{\sum cf}{\sum f}$$ **Equation 4.**

$\bar{x}$ = *arithmetic mean*
c = *the mid-point of the class interval*
$\sum f$ = *the number of cases*
$\sum cf$ = *is the sum of the products of c and* Σf

Calcualtion of the arithmetic mean in practice when only groupped data are available, and intervals of groups (classes) are equal is presented in Case study 1, Case 3.

Weighted arithmetic mean

When the research design includes observation of a sample consisting of two or more groups with different number of cases in each group the weighted arithmetic mean will be the most appropriate measure of central tendency for the whole sample (6,11,12).

This approach is similar to the computing of the mean for grouped data but instead of the interval mid-points the true group means are used in calculation. Thus, the weigted mean takes into account the different number of cases (that is, unequal weight) and the real mean for each group. The calculation of the weighted mean follows 2 steps:

- firstly, we have to multiply the mean for each group by the corresponding number of cases in each group, and add up the totals;
- secondly, the sum of totals (obtained in the first step) is divided by the total number of cases in the sample

Data values with larger weights contribute more to the weighted mean and data values with smaller weights contribute less to the weighted mean. The formula for calculation is (Equation 5):

$$\bar{x}_w = \frac{\sum wx}{\sum w}$$ **Equation 5.**

$\bar{x}_w$ = *weighted arithmetic mean*

w = *weight assigned to each data value*

Calcualtion of the weighted arithmetic mean in practice is presented in Case study 1, Case 4.

Characteristics of the mean

At the end we think it is worthwhile to summarize and present basic characteristics of the aritmetic mean:

1. The arithmetic mean is unquestionably the most widely used measure of central tendency. It has the advantage to substitute by one single number all the individual values of a given variable and describe the typical level of a variable in a data set.
2. For normal or roughly symmetric distribution, the mean is the most efficient and therefore the least subject to sample fluctuations of all measures of central tendency. But can be misleading in skewed distributions since it can be greatly influenced by scores in the tail.

Therefore, other statistics such as the median may be more informative for skewed distributions. Also, the geometric mean is less affected by extreme values than is the arithmetic mean and is useful as a measure of central tendency for some positively skewed distributions.

3. Substantial disadvantage of the mean is that it can be drastically affected by the presence of a small number of outliers, e.g. observed values that are strikingly different from the rest (either unusually large or small). Such extreme values can distort the mean, particularly if there is a small number of subjects. Nowever, there are methods that allow to eliminate some extreme values and compute a new mean, which will be more typical for a particular empirical distribution. Such method is the criterion U, calculated as the ratio of the difference between the outlier (x_i) and the mean $\bar{x}$ and the standard deviation s. The computed criterion U is then compared with the table critical values of u_t and if $u \geq u_t$, the extreme value x_i is disgarded as unusual.
4. The sum of the deviations of the scores in the distribution from the mean always is equal to zero. This is true because, by definition, the mean is the mathematical centre of the data. Thus, half of the distribution is above and half below the mean.
5. The sum of the squares of the deviations around the mean is smaller than the sum of squares around any other value. This characteristics of the mean underlies the calculation of the "least squares", which is used in applying some other statistical methods, such as regression analysis.
6. If to each value of the frequency distribution the same number is added or subtracted, then the mean is increasing or decreasing by the same number.
7. The mean is not generally a "real" value and this makes the acceptance and interpretation of the data sometimes more difficult – for example, a mean number of children in a sample might be 2.4, or an average number of limbs 3.997.

Median

The median (Me) is the measure of central tendency which is not calculated but identified or determined. It is a member of a family of measures of location called quantiles that are in details describen in a separate module.

It is the measure that tells what is the value of the middle observation when data have been arranged in ordered series from the lowest to the highest value. Half of units of a sample (or population) lie above the median and half below the median. The procedure to identify the median is to:

- rearrange all observations (units) in order of magnitude (from the smallest to the largest) in an ordered series (one must be sure to list all data values even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the median is simply the value of the middle observation (unit) in the ordered series;

- when the number of observations is even, the median is just a halfway of values of the two middle observations.

Identification of the median in practice is presented in Case study 1, Case 5.

Because the median is less sensitive than the mean to the biasing influence of extreme scores on a data set, it is best used when a distribution is known to be asymmetric or when its shape is otherwise unknown.

The median is particalarly suitable for scales of measurement having ordinal characteristics and when the validity of assumotions about in the size of the intervals between data points is questionable (9,13).

Characteristics of the median

Characteristics of the median are as follows:

1. The median is usually a realistic value, or at worst measured in half-units (when the number of observations is even).
2. A better advantage over the mean is that the median is more robust towards outliers (extreme scores). The persense of a few extreme observations in the sample (from whatever cause) do not affect the middle values (8). This makes the median a better measure than the mean for highly skewed distributions.
3. The only disadvantage of the median is that it does not include all the individual values of a variable. It reflects only one value in odd number of cases or two values in even number of cases.
4. The median is preferred measure of central tendency when (11):
    - the lowest and highest values of a quantitative variable are far off of the rest values;
    - there is uncertainty in some values;
    - it is not possible to determine the exact shape of the distribution or when the distribution is highly asymmetric;
    - the number of cases is small.

Mode

The mode (Mo) is the observation in an array with the highest frequency of occurrence. The advantage of the mode as a measure of central tendency is that its meaning is obvious and it is the simplest to determine of the three measures of central tendency. Actually, the mode is not computed but rather is determined through inspection of a frequency distribution.

Although it is common for most distributions to contain exactly one mode (as in a normal distribution and large homogenious samples), it is possible for more than one mode to exist. A distribution having one mode is called unimodal. A distribution having two modes is called bimodal.

Identification of the mode in practice is presented in Case study 1, Case 6.

Characteristics of the mode

Characteristics of the mode are as follows:

1. The mode is a quick and easy method of determining the most popular score at a glance.
2. It is the only measure of central tendency that can be used with nominal data.
3. The mode is the weakest measure of central tendency. It often provides a crude and limited representation of the characteristics of a distribution as compared to the mean and median. This is true because, in some cases, the mode may be the lowest or the highest value in the distribution.
4. It is greatly subject to sample fluctuations and is rather unstable, e.g. the modes tend to fluctuate widely from one sample drawn from a population to another sample drawn from the same population. Therefore, it is not recommended to be used as the only measure of central tendency. The mode is seldom used in research reports except in association with other measures of central tendency.
5. As the weakest measure of central tendency, the use of the mode is restristed to nominal scales of measurement and is seldom reported except in association with other measures of central tendency.
6. A further disadvantage of the mode is that many distributions have more than one mode. These distributions are called multimodal.
7. Nevertheless, the mode has a true meaning and this is very important in medicine and public health. For example, it is more important to determine which group has higher risk for some disease, e.g. to determine the mode in the age distribution instead of calculating the mean age of persons with the disease.

Other measures of central tendency

Trimmed mean

A trimmed mean is a systematic method for avoiding outliers when calculating means by discarding or "trimming off" a certain percentage of the lowest and the highest scores and then computing the mean of the remaining scores (14,15). After removing the specified observations, the trimmed mean is found using an arithmetic averaging formula. For example, a mean trimmed 50% is computed by discarding the lower and higher 25% of the scores and taking the mean of the remaining scores.

A trimmed mean is obviously less susceptible to the effects of extreme scores than is the arithmetic mean. It is therefore less susceptible to sampling fluctuation than the mean for extremely skewed distributions. It is less efficient than the mean for normal distributions. This method is best suited for data with large, erratic deviations or extremely skewed distributions.

A trimmed mean is stated as a mean trimmed by X%, where X is the sum of the percentage of observations removed from both the upper and lower bounds. For example, to trim the mean by 40% it means that we remove the lowest 20% and the highest 20% of values.

In contrast to the arithmetic mean, the trimmed mean is a robust measure of central tendency. For example, a small fraction of anomalous measurements with abnormally large deviation from the center may change the mean value

substantially. At the same time, the trimmed mean is stable in respect to presence of such abnormal extreme values, which get "trimmed" away.

Trimmed means are often used in Olympic scoring to minimize the effects of extreme ratings possibly caused by biased judges, where the extreme scores are often discarded before computing the score for a particular performance.

Calculation of trimmed mean in practice is presented in Case study 1, Example 7.

Trimean

The trimean is a measure of central tendency computed by using quantiles (percentiles or quartiles and median) that are described in details in a separate module in this book. If we use percentiles for calculation, the procedure is as follows (Equation 6):

- adding the 25^{th} percentile,
- plus twice the 50^{th} percentile,
- plus the 75^{th} percentile, and
- dividing by four.

$$TM = \frac{P_{25} + (2 \times P_{50}) + P_{75}}{4}$$ **Equation 6.**

TM = trimean
P_{25} = 25^{th} percentile
P_{50} = 50^{th} percentile
P_{75} = 75^{th} percentile

Exactly the same results we obtain if we use the quartiles, or quartiles and median) for calculation (Equation 7):

$$TM = \frac{Q_1 + (2 \times Q_2) + Q_3}{4} = \frac{Q_1 + (2 \times Me) + Q_3}{4}$$ **Equation 7.**

TM = trimean
Q_1 = 1^{st} quartile
Q_2 = 2^{nd} quartile
Q_3 = 3^{rd} quartile
Me = median

The trimean is a good measure of central tendency but it is not used as much as it should be (16). It is almost as resistant to extreme scores as the median and is less subject to sampling fluctuations than the arithmetic mean in extremely skewed distributions. It is less efficient than the mean for normal distributions.

Calculation of trimean in practice is presented in Case study 1, Example 8.

Geometric mean

The geometric mean is a type of mean or average, which indicates the central tendency or typical value of a set of numbers. It is similar to the arithmetic mean, except that instead of adding the set of numbers and then dividing the sum by the count of numbers in the set, *n*, the numbers are multiplied and then the n^{th} root of the resulting product is taken. It is rarely used in biomedical research. However, it should be obligatory applied when the values of the variable increase in geometric progression or the distribution of frequencies by the logaritms of observed values of the variable are approximately symmetric (17).

The geometric mean only applies to positive numbers. It is also often used for a set of numbers whose values are meant to be multiplied together or are exponential in nature, such as data on the growth of the human population.

Calculation of the geometric mean is easy – it is just the n^{th} root of the product of the scores (Equation 8) (16):

$$GM = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots \times x_n} = \left(x_1 \times x_2 \times x_3 \times \ldots \times x_n\right)^{\frac{1}{n}} = \left(\prod x\right)^{\frac{1}{n}} \qquad \textbf{Equation 8.}$$

GM = geometric mean
n = number of units in a sample studied
x_n = value of a single unit
Πx = product of all scores

The geometric mean can also be computed by taking the logarithm of each number, or computing the arithmetic mean of the logarithms

Calculation of geometric mean in practice is presented in Case study 1, Example 9.

The geometric mean must be used when working with percentages (which are derived from values), whereas the standard arithmetic mean will work with the values themselves (16).

Comparison of the measures of central tendency

Of the five measures of central tendency, the mean is the most stable. If repeated samples were drawn from a given population, the means would vary ot fluctuate less than the modes or medians. Because of its stability, the mean is the most reliable estimate of the central tendency of the population.

The arithmetic mean is the most appropriate measure in situations in which the concern is for totals of combined performance of a group. When the primary concern is to learn what a typical value is, then the median would be preferred (2).

Of the five measures of central tendency discussed, the mean is by far the most widely used. It takes every score into account, is the most efficient measure of central tendency for normal distributions and is mathematically tractable making it possible for statisticians to develop statistical procedures for drawing inferences about means. On the other hand, the mean is not appropriate for highly skewed distributions and is less efficient than other measures of central tendency

when extreme scores are possible. The geometric mean is a viable alternative if all the scores are positive and the distribution has a positive skew.

The median is useful because its meaning is clear and it is more efficient than the mean in highly-skewed distributions. However, it ignores many scores and is generally less efficient than the mean, the trimean, and trimmed means.

The mode can be informative but should almost never be used as the only measure of central tendency since it is highly susceptible to sampling fluctuations.

The trimean and trimmed means are both examples of statistics developed to resist sampling fluctuations. It is highly recommended that at least one of these two be computed in addition to the mean (14,15,18).

The level of measument plays an important role in determining the appropriate index of central tendency that can be used to describe a variable. In general, the mode is appropriate for nominal measures. The median is appropriate for ordinal measures. The mean is appropriate for interval and ratio measures (7).

However, the higher the level of measurement, the greater the flexibility we have in choosing a descriptive statistic. Variables measured on an interval or ration scale can use any of the three measures of central tendency, although it is usually preferable to use the mean (2).

In skewed distributions, the values of the mode, median, and mean differ. The mean is always pulled in the direction of the long tail (17). This means that the mean is typically higher than the median and mode in positively skewed distributions (Mo<Me<$\bar{x}$) (Figure 1). The mean is lower than the median and mode in negatively skewed distributions (Mo>Me>$\bar{x}$) (Figure 2) (3,17). When a distribution is symmetric and unimodal, the three measures of central tendency – the mean, the median and the mode – coincide (3,17).

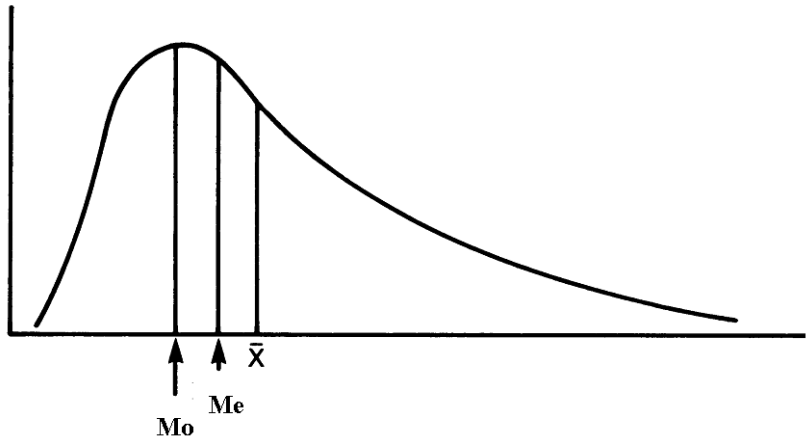


Figure 1. Arithmetic mean ($\bar{x}$), mode (Mo) and median (Me) in a positively skewed distribution.

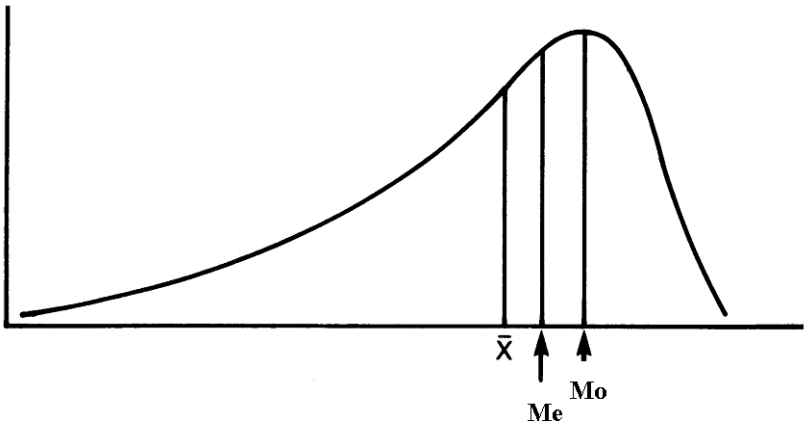


Figure 2. Arithmetic mean ($\bar{x}$), mode (Mo) and median (Me) in a negatively skewed distribution.

To illustrate the difference between arithmetic and geometric mean, the best is to think in terms what question answer both of them. While arithmetic mean answers the question, "If all the units in a sample (population) had the same value, what would that value have to be in order to achieve the same total?", the geometric mean answers the question, "If all the units in a sample (population) had the same value, what would that value have to be in order to achieve the same product?"

Measures of dispersion (variability)

Why do we need measures of dispersion?

We need measures of dispersion because measures of central tendency do not give a total picture of a distribution.

Two sets of data with identical means could be different from one another. Two distributions with the same means could be very different in shape: for example, they could be skewed in opposite directions (2). Secondly, even when two sets of data have equal means, medians, modes, and the same form of distribution, they could be different from one another. Knowledge of a single summary figure for describing the location of a center of a sample or population is not enough without a measure of the extent of variability or spread of the measurements around this summary index. Illustration is given in Case study 2, Case 1.

Health workers often have to decide whether to classify an individual as healthy ot sick, suffering from a particular disease or not, needing treatment or not, etc. For this task, the co-called "normal" values of certain measurements provide the necessary yardstick. But the word "normal" value is a statistical

concept and depends, to a great extent, on the distribution of the classifying attribute in the population. Measuresures of spread or dispersion or variability are, therefore, a complete description of a given health data set. No description of any health data by summary measures is complete without the measures of variability (6). Several such measures have been developed, the most common of which are:

- the range,
- standard deviation,
- variance,
- inter- and semiquartile range, and
- coefficient of variation.

Range

The range is simply the difference between the extreme values (the highest and the lowest) of the variable in a given empiric distribution (12,19). It is usually denoted by *d* (difference) and is express as (Equation 9):

$$d = x_{max} - x_{min} \qquad \textbf{Equation 9.}$$

d = the range
x_{max} *= the lowest value*
x_{max} *= the highest value*

The chief virtue of the range is the ease with which it can be computed.

As an index of variability, the shortcomings of the range outweigh this modest advantage. The range being based on only two scores (and the two most unusual at that), is a highly unstable index. The drawback to this simple measure is also the fact that a single outlier may have a large impact on the range (8) (Case study 2, Case 2).

Another difficulty with the range is that it ignores completely variations in scores between the two extremes. Surely, a value that seeks to measure variation within a group of individuals should reasonably be expected to be based on information gathered from all the individuals under study, not merely on two selected and unrepresentative ones.

For these reasons, the range is used only as a gross descriptive index and is typically reported in conjunction with, not instead of, other measures of variability (6,10,13,20).

Standard deviation and variance

The standard deviation (denoted by SD or s for a sample and σ for a population) is the most commonly reported measure of variability, especially with interval- or ratio-level data. It is a statistic that describes the degree of variarion among the individual observartions in the sample (21). Like the mean, the standard deviation

considers every score in a distribution. For this reason, means and standard deviations are generally reported together, whether in the text or in tables (7).

The standard deviation represents the average deviation of scores in a given distribution around the central score, e.g. the mean, which serves as a reference point or baseline for the entire data set. The calculation of SD includes the following steps (3,10):

- firstly, we need to calculate how much each individual varies from the mean by simply subtracting the mean from each individual value - $\mathbf{x}-\bar{x}$ **;**
- measuring the overall variation present in the study group by adding the individual variations together - this sum is denoted by $\Sigma(x-\bar{x})$. In order to calculate the average deviation the sum $\Sigma(x-\bar{x})$ should be divided by the number of scores n**.** We would expect this to be large if the variation is large, and small if most individuals are very similar to the mean value and hence show relatively little variation. Unfortunately, totaling the differences of deviations in this way tells us nothing at all. The reason is that those individuals who have values larger than the mean (e.g. $\mathbf{x}-\bar{x}$ is a positive value) will simply cancel out those who have values below the mean (e.g. $\mathbf{x}-\bar{x}$ is a negative value). So, any data set, highly consistent or highly variable, will result in a zero value when all the deviations are added together - $\Sigma(x-\bar{x})$ will always be equal to zero;
- to get around the problem of positive and negative variation, we can square each difference so that both positive and negative values end up as positive and no longer cancel out one another. The new measure of the variation in a group of individuals is therefore $\Sigma(x-\bar{x})^2$ and it is referred to as the sum of squares or the sum of squared deviations from the mean;
- the sum of squares as a measure of variation has some limitations related to the number of results under study. To allow fair comparisons between studies of different sizes, and to provide a truly universal measure of variation, it is reasonable to take the study size into account by calculating an "average" variation. This measure is called the variance - s^2 and is calculated as (Equation 10):

$$s^2 = \frac{\sum (x-\bar{x})^2}{n-1} \qquad \textbf{Equation 10.}$$

s^2 = variance of a sample
x = value of a single unit
$\bar{x}$ = arithmetic mean of a sample
n = number of units in a sample studied
n-1 = degrees of freedom (df)

- finaly, the variance s^2 has one major drawback as a truly useful measure of variation. It is based on squared differences, and hence it measures variation in squared units which is not convenient (for example, it is nonsense to say that the variation in blood pressure study is in squared millimeters of mercury). To

solve this problem we need to take the square root of the variance and we will come to the most meaningful and most widely used measure of variability known as the standard deviation - s or SD (Equation 11):

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$ **Equation 11.**

s = standard deviation of a sample
x = value of a single unit
$\bar{x}$ *= arithmetic mean of a sample*
n = number of units in a sample studied
n-1 = degrees of freedom (df)

Calcualtion of variance and standard deviation in practice is presented in Case study 2, Case 3.

The standard deviation can also be used in interpreting individual scores from within a distribution. Using the basic principle of normal distribution (Figure 3), we can determine the limits of different groups of normality to assess and interpret the individual results.

When the distribution of scores is normal or nearly normal, it is possible to say even more about the standard deviation. There are about 3 SDs above and 3 SDs below the mean with normally distributed data. Further more, in a normal distribution, as shown in Figure 3, a fixed percentage of cases falls within certain distances from the mean (2,3):

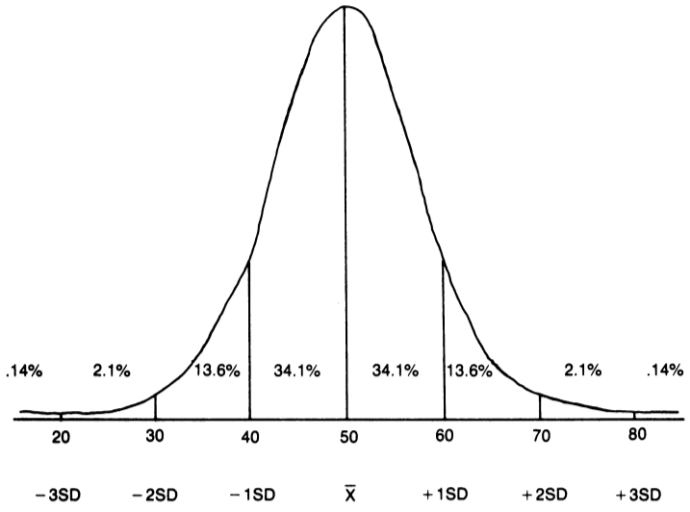


Figure 3. Standard deviations in a normal distribution.

- 68% percent of all cases fall within 1 SD of the mean (34% above and 34% below the mean);

- 95% of the scores fall within 2 SDs from the mean;
- only a handful of cases – about 2% at each extreme – lay more than 2 SDs from the mean.

Using this principle, we can easily create "normal limits" and interpret individual scores for many clinical and laboratory tests.

In case of skewed distributions, the "normal limits" can be established based on percentiles, as described in a separate module in this book.

In summary, the standard deviation is a useful index of variability that can be used to describe an important characteristic of a distribution and also can be used to interpret the score of the performance of an individual in relation to others in the sample. Like the mean. The standard deviation is stable estimate of a population parameter and also is used in more advanced statistical procedures. The standard deviation is the preferred measure of distribution's variability but it is appropriate only for variables measures on the interval or ratio scale (2,3,22,23).

Interquartile and semiquartile ranges

The interquartile range

The interquartile range (IQR) is the difference between the third (Q_3) and the first (Q_1) quartiles in a dataset (where quartiles are the values that divide the data into four equal sized parts). Quartiles are described in details in a separate module. Characteristics of IQR are:

- the advantage of the IQR over the range is that it is quite robust to outliers;
- the IQR is commonly quoted in conjunction with the sample median.

Calcualtion of IQR in practice is presented in Case study 2, Case 4.

The semiquartile range

The semiquartile range (SQR), used as a term in many statistical texts instead of IQR, is half of the distance between Q_1 and Q_3.

Because these two measures of variability are based on middle cases rather than extreme scores, they are considerably more stable than the range (2,8).

Coefficient of variation

The standard deviation *s* and the variance s^2 have the same measurement units as the mean $\bar{x}$ and because of this they are not appropriate for comparing the relative variability of different distributions where the variables are measured in different units (height, weight, blood pressure, cholesterol level, etc.) This problem can be overcome by calculating another measure of variation called the coefficient of variation (denoted by CV), also known as relative variability. It expresses the sample standard deviation in a proportion or percentage of the mean value (1). When expressed as percentage it is calculated by the formula (Equation 12):

$$CV = \frac{s}{\bar{x}} \times 100$$ **Equation 12.**

CV = coefficient of variation
$\bar{x}$ = arithmetic mean of a sample
s = standard deviation of a sample

The main advantage of the coefficient of variation is its independence of any unit of measurement, and thus useful for comparison of variability in two or more distributions having variables expressed in different units (1,6). For example, if we measure height and weight in a sample, it is not possible to say which variable varies greatly because these two variables have different measurement units. Using the coefficient of variation we can transform the standard variations in comparable units, expressed in percents. The interpretation is as follows (3):

- when the value of the coefficient of variation is less than 10% we could say that the degree of variation is low and the sample is quite homogeneous;
- in a situation when 10%<V<30% - the variation is moderate;
- with V>30% the variation is considerable, and this is a clear evidence of heterogeneity of the sample or population under study

Calcualtion of coefficient of variation in practice is presented in Case study 2, Case 4.

Summary

In this module two essential measures of location for describing and representing frequency distributions were discussed: measures of central tendency and variability.

The measures of central tendency outlined were the mode and the median for discrete data, and the mean for continuous data. The weighted mean, trimean and trimmed mean were shortly presented as well.

Measures of variability were shown to be the range, the variance, the standard deviation, the inter- and semiquartile range, and the coefficient of variation.

All these statistics are appropriate for processing the data to the point that a distribution of raw data can be meaningfully represented by only two statistics. We have seen that the mean and the standard deviation are the most appropriate for interval or ratio data when the distribution is similar enough to normal distribution. The median and the inter- or semiquartile range are used when the data was measured on an ordinal scale, or when interval or ratio data is found to have a highly skewed distribution.

The content of the module focused on the meaning and the use of these concepts, rather than stressing on calculations that nowadays are made usually by statistical package software.

CASE STUDY 1: MEASURES OF CENTRAL TENDENCY

Case 1 – calculation of arithmetic mean when raw data are available

Suppose the age at first birth for a sample of 10 mothers is (24):

18 21 23 23 25 27 27 28 30 33

Then the mean age is given according to Equation 1 by (Equation 13):

$$\bar{x} = \frac{18 + 21 + 23 + ... + 33}{10} = \frac{255}{10} = 25.5 \qquad \textbf{Equation 13.}$$

Case 2 – calculation of arithmetic mean when only groupped data are available and intervals of groups are of unequal size

Suppose the values of height for 100 female urban liveborns are those presented in Table 1.

Table 1. Elements of calculation of arithmetic mean when only groupped data are available and intervals of groups are of unequal size.

Height in cm *(x)*	**Frequency** *(f)*	**Product** *(xf)*
46	2	92
47	6	282
48	7	336
49	20	980
50	30	1,500
51	20	1,020
52	8	416
53	5	265
54	2	108
	$\Sigma f = n = 100$	$\Sigma xf = 4{,}999$

The arithmetic mean can be computed according to Equation 3 through the following algorithm (Equation 14):

- each value of the variable is multiplied by its frequency and the product *xf* is recorded in the appropriate row and column;
- the products in column 3 are summed and recorded as Σxf;
- the sum Σxf is divided by the number of cases $(\Sigma f = n)$ to come to the mean $\bar{x}$.

$$\bar{x} = \frac{4{,}999}{100} = 49.99 = 50 \qquad \textbf{Equation 14.}$$

Case 3 – calculation of arithmetic mean when only groupped data are available and intervals of groups are of equal size

Let us take the same example of height for 100 female urban liveborns and regroup the data into 3 equal class intervals (Table 2).

Table 2. Elements of calculation of arithmetic mean when only groupped data are available and intervals of groups are of equal size.

Height in cm	**Frequency**	**Mid-point of the class interval**	**Product**
(x)	***(f)***	***(c)***	***(cf)***
46 - 48	15	47	705
49 - 51	70	50	3,500
52 - 54	15	53	795
	$\Sigma f = n = 100$		$\Sigma cf = 5{,}000$

The algorithm for computing the mean is according to Equation 4 the following (Equation 15):

- defining the interval width;
- regrouping the data into equal width intervals and summing the frequences for each interval;
- defining the mid-point of each interval;
- multiplying the frequency of each interval by its mid-point;
- recording the products *cf* in the corresponding rows;
- summing the products *cf* and recording the sum Σcf in the bottom line;
- dividing the sum Σcf by the number of cases Σf to come to $\bar{x}$.

$$\bar{x} = \frac{5{,}000}{100} = 50.00 = 50 \qquad \textbf{Equation 15.}$$

It is obvious that the arithmetic mean is the same as in case 1 but with intervals the computing is easier.

Table 3. Elements of calculation of arithmetic mean when only groupped data are available and intervals of groups are of equal size.

Systolic blood pressure in mmHg (class interval)	**Frequency**	**Mid-value of the class**	**Product**
(x)	***(f)***	***(c)***	***(cf)***
90-99,9	4	95	380
100-109,9	16	105	1,680
110-119,9	18	115	2,070
120-129,9	40	125	5,000
130-139,9	66	135	8,910
140-149,9	56	145	8,120
150-159,9	34	155	5,270
160- 169,9	6	165	990
Total	$\Sigma f = n = 240$	-	$\Sigma cf = 32{,}420$

Let us now consider another set of data presented in Table 3 - Systolic blood pressure for 240 men (data modified from Lwanga et al. (6). The arithme3tic mean is then calculated according to Equation 4 as follows (Equation 16):

$$\bar{x} = \frac{32{,}420}{240} = 135.1 \qquad \textbf{Equation 16.}$$

Case 4 - calcualtion of the weighted arithmetic mean

Suppose mean ages of preschool children in 5 different villiages (data modifeid from Lwanga et al. (6)) (Table 4). On the basis of available data we can, according to Equation 5, calculate weighted arithmetic mean (Equation 17).

Table 4. Elements of calculation of weighted arithmetic mean.

Village No.	**No. of children** (w)	**Mean age (in months)** (x)	**Product** ***(wx)***
1	35	61.5	2,152.5
2	40	62.0	2,480.0
3	45	62.5	2,812.5
4	50	63.5	3,175.0
	$\Sigma w = n$ = 170		Σwx = 10,620

$$\bar{x}_w = \frac{10{,}620}{170} = 62.47 \qquad \textbf{Equation 17.}$$

Thus, the weighted mean is 62.47 months.

Case 5 - identification of the median

Suppose the age at first birth for a sample of 10 mothers is (8):

18 21 23 23 25 27 27 28 30 33

In this set of data where the number of units in the sample is even (ten), the median age at first birth is 26 – the mean of the two middle numbers 25 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation we add to the observed set of data one case more:

18 21 23 23 25 27 27 28 30 33 33

In this set of data where the number of units in the sample is now odd (eleven), the median age at first birth is 27 - simply the value of the middle observation in the ordered series of data:

18 21 23 23 25 27 27 28 30 33 33

Case 6 - identification of the mode

Suppose the age at first birth for a sample of 10 mothers is (8):

18 21 23 23 25 27 27 28 30 33

In this set of data, the variable "age at first birth" has two modes - 23 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation we change the set of data: the case number 3 has value 22 instead of 23:

18 21 22 23 25 27 27 28 30 33

In this set of data, the variable "age at first birth" has one mode - 27.

18 21 22 23 25 27 27 28 30 33

Case 7 - calculation of the trimmed mean

As a first example we have calculation of mean score of referees in a figure skating competitition. A figure skating competition produces the following scores:

6.0 8.1 8.3 9.1 9.9

A mean trimmed 40% by using Equation 1 and trimming 40% of values would equal (Equation 18):

$$\bar{x}_{trimmed} = \frac{8.1 + 8.3 + 9.1}{3} = 8.5$$ **Equation 18.**

while ordinary mean would be (Equation 19):

$$\bar{x} = \frac{6.0 + 8.1 + 8.3 + 9.1 + 9.9}{3} = 8.28$$ **Equation 19.**

We could see that trimmed mean is larger than the ordinary arithmetic mean. To trim the mean by 40%, we remove the lowest 20% and highest 20% of values, eliminating the scores of 6.0 and 9.1. As shown by this example, trimming the mean can reduce the effects of outlier bias in a sample (18).

Let us take now a set of final exam scores for a 40-question test of 20 students (modified from Weiss (14,15)). The results of the test are as follows:

2 15 16 16 19 21 21 25 26 27 4 15 16 17 20 21 24 25 27 28

The following steps in calculating the trimmed mean include:

1. Exploration of the data in order to find the outliers. For accomplishing this step we need first to arrange the data into an ordered series:

2 4 15 15 16 16 16 17 19 20 21 21 21 24 25 25 26 27 27 28

In this example there are two outliers - the low values of 2 and 4.

2. Computation of the usual mean for the data (Equation 20):

$$\bar{x} = \frac{2 + 4 + 15 + ... + 27 + 27 + 28}{20} = 19.3$$ **Equation 20.**

3. Computation of the 10% trimmed mean for the data (Equation 21):

$$\bar{x}_{trimmed} = \frac{15 + 15 + 16 + ... + 27 + 27 + 28}{20} = 20.2$$ **Equation 21.**

4. If we compare the two means obtained we can conclude that the trimmed mean provides better measure of central tendency for this set of data.

Case 8 - identification of the trimean

Let's take the values for the following 50 measurements (23):

50 50 50 50 50 50 50 50 50 50
51 51 51 51 51 51 51 51 52 53
53 53 55 55 55, 55 56 56 56 58
58 59 60 60 61 63 63 63 64 67
67 69 70 70 75 77 78 80 85 103

The P_{25} (Q1), P_{50} (Q2), and P_{75} (Q3) of this dataset are 51, 55, and 63 respectively. Therefore, the trimean is computed according to Equation 6 as (Equation 22):

$$TM = \frac{51 + (2 \times 55) + 63}{4} = 56 \qquad \textbf{Equation 22.}$$

Case 9 - calculation of the geometric mean

The geometric mean of two numbers, say 2 and 8, would be according to Equation 8 just the square root (i.e., the second root) of their product (Equation 23):

$$GM = \sqrt{2 \times 8} = \sqrt{16} = 4 \qquad \textbf{Equation 23.}$$

As another example, the geometric mean of **1**, **½**, and **¼** is the cube root (i.e., the third root) of their product (0.125), which is **½** (16) (Equation 24).

$$GM = \sqrt[3]{1 \times 0.5 \times 0.25} = \sqrt[3]{0.125} = 0.5 \qquad \textbf{Equation 24.}$$

The geometric mean of the scores: 1, 2, 3, and 10 is the fourth root of their product (Equation 25):

$$GM = \sqrt[4]{1 \times 2 \times 3 \times 10} = \sqrt[4]{60} = 2.78 \qquad \textbf{Equation 25.}$$

CASE STUDY 2: MEASURES OF DISPERSION

Case 1 – variability of the data

Consider two sets of data of mother's ages at first birth. The first sample of 10 cases with age values is:

18, 21, 23, 23, 25, 27, 27, 28, 30, 33.

The second sample of 10 cases with age values is:

23, 23, 24, 25, 26, 26, 27, 27, 27, 27.

The means for the two samples are 25.5, the medians - 26, and the modes – 27.

However, it is obvious that the samples are quite different. The second sample seems more homogeneous that the first one.

It is unquestionable, from this simple example, that knowledge of a single summary figure for describing the characteristics of a sample or population is not enough without a measure of the extent of variability or spread of the measurements around this summary index.

Case 2 – the range

Consider two sets of data from Case study 2, Case 1, again:

Set No 1: 18 21 23 23 25 27 27 28 30 33

Set No 2: 23 23 24 25 26 26 27 27 27 27

The means for the two samples are 25.5, the medians - 26, and the modes – 27 while the range is very different. The range for the first sample equals according to Equation 9 to 15, and for the second sample to 4. So, from sample to sample, drawn from the same population, the range tends to fluctuate considerably.

Case 3 – variance and standard deviation

Consider two samples from Case study 2, Case 1, again. In Table 5, elements for calculation of standard deviation are presented.

Table 5. Elements of calculation of standard deviation in two samples of equal arithmetic mean.

Sample 1				Sample 2			
Age (years)	**Mean**	**Diff.**	**Square of diff.**	**Age (years)**	**Mean**	**Diff.**	**Square of diff.**
(x)	$(\bar{x})$	$(x-\bar{x})$	$(x-\bar{x})^2$	(x)	$(\bar{x})$	$(x-\bar{x})$	$(x-\bar{x})^2$
18	25.5	-7,50	56,25	23	25.5	-2,50	6,25
21	25.5	-4,50	20,25	23	25.5	-2,50	6,25
23	25.5	-2,50	6,25	24	25.5	-1,50	2,25
23	25.5	-2,50	6,25	25	25.5	-0,50	0,25
25	25.5	-0,50	0,25	26	25.5	0,50	0,25
27	25.5	1,50	2,25	26	25.5	0,50	0,25
27	25.5	1,50	2,25	27	25.5	1,50	2,25
28	25.5	2,50	6,25	27	25.5	1,50	2,25
30	25.5	4,50	20,25	27	25.5	1,50	2,25
33	25.5	7,50	56,25	27	25.5	1,50	2,25
			$\Sigma(x-\bar{x})^2 =$ 176,5				$\Sigma(x-\bar{x})^2 =$ 24,5

Variance and standard deviation for the first sample are according to Equations 10 and 11 (Equations 26 and 27):

$$s^2_{sample\,1} = \frac{176.5}{9} = 19.6 \Rightarrow s_{sample\,1} = \sqrt{19.6} = 4.43$$ **Equation 26.**

Variance and standard deviation for the second sample are:

$$s^2_{sample\,2} = \frac{24.5}{9} = 2.7 \Rightarrow s_{sample\,2} = \sqrt{2.7} = 1.65$$ **Equation 27.**

These calculations confirm that the variation is much smaller in the second sample which means that the mean for the second sample is more accurate measure of a central tendency.

Case 4 – calculation of interquartile range

Let's g back to the example with two samples of 10 mothers at first birth:

Set No 1: 18 21 23 23 25 27 27 28 30 33

For the first dataset $Q_1 = 23$ and $Q_3 = 28$, so the IQR is 5 years.

Set No 2: 23 23 24 25 26 26 27 27 27 27

For the second data set $Q_1 = 24$ and $Q_3 = 27$, so the IQR is 3 years.

Case 5 – calculation of coefficient of variation

Let's g back to the example with two samples of 10 mothers at first birth again:

Set No 1: 18 21 23 23 25 27 27 28 30 33

Set No 2: 23 23 24 25 26 26 27 27 27 27

We have already calculated the mean and the standard deviation for both sets of data (Case study 2, case 3). The mean was in both cases the same (25.5 years), while the standard deviations were different (s_1 = 4.43 years; s_2 = 1.65 years). According to Equation 12, coefficient of variation could be calculated as (Equations 28 and 29):

$$CV_1 = \frac{4.45}{25.5} \times 100 = 17.45\%$$ **Equation 28.**

$$CV_2 = \frac{1.65}{25.5} \times 100 = 6.47\%$$ **Equation 29.**

We can conclude that the first sample has much higher variability than the second one. The second sample could be described as homogenous, while in the first sample variation is moderate.

EXERCISE

Task 1

For the following questions choose between "true" or "false":

1. With nominal data, the mean should be used as a measure of central tendency.	T	F
2. The mode represents the most frequently occurring score in a distribution.	T	F
3. With ordinal data we can use both the mode and the mean as a measure of central tendency.	T	F
4. When the data are interval or ratio, we can use the mean as a measure of central tendency.	T	F
5. If a continuous distribution is highly skewed, the median might be the appropriate measure of central tendency.	T	F
6. When a frequency distribution is positively skewed, the mean is greater than the median or the mode.	T	F
7. Given a normal distribution, the three measures of central tendency are equivalent.	T	F
8. The range is the simplest indicator of variability.	T	F
9. The range is calculated by adding the lowest score to the highest score in a distribution.	T	F
10. The square root of the variance is called the standard deviation.	T	F
11. Standard deviation indicates the extent to which scores are distributed around the mean.	T	F
12. When a distribution consists of very different scores, standard deviation will be relatively large.	T	F
13 The median is less affected than the mean by extreme scores of a distribution.	T	F
14. Central tendency describes the 'typical' value of a set of scores.	T	F

15. If the number of raw scores is odd, the median is the score in the middle position. T F
16. The mean must have a value equal to one of the scores in the distribution. T F
17. About 10% of scores fall 3 standard deviations above the mean. T F

Answers: 1. F; 2. T; 3. F; 4. T; 5. T; 6. T; 7. T; 8. T; 9. F; 10. T; 11. T; 12. T; 13 T; 14 T; 15 T; 16 F; 17F

Task 2

For the following two questions choose only one right answer:

1. Given a set of nominally scaled scores, the most appropriate measure of central tendency is the:
   A. Mean
   B. Mode
   C. Standard deviation
   D. Range

2. Which of the following statements is true?
   A. The mode is the most useful measure of central tendency.
   B. The variance is the square root of the standard deviation.
   C. The median and the 50^{th} percentile rank have different values.
   D. The mean is more affected by extreme scores than the median.

Answers: 1. B; 2. D

Task 3

Following four questions refer to the following set of data:

2 2 3 4 6 6 7.

1. Σx is equal to:
   A. 30
   B. 40
   C. 50
   D. None of the above A, B or C.

2. $(\Sigma x)^2$ is equal to:
   A. 124
   B. 128
   C. 130
   D. 900

3. The median is equal to:
   A. 6
   B. 5
   C. 4
   D. 3

4. The range for the above set of scores is:
   A. 7
   B. 5
   C. 2
   D. 1

Answers: 1. A; 2. D; 3. C; 4. B

Task 4

A clinic had 50 patients attending in a month. The number of times each patient visited the clinic is given below in the form of frequency distribution (Table 6):

Table 6. Elements of calculation of measures of location in a distribution of patients attending a clinic.

No of visits (*x*)	**No of patients (*f*)**
7	3
6	6
5	6
4	10
3	21
2	0
1	4

The following three questions refer to the example:

1. The total number of visits by the patients was:
   A. 194
   B. 28
   C. 50
   D. None of the above

2. The mean number of visits per patient was:
   A. 3.89
   B. 3.50
   C. 1.00
   D. 3.84

3. The median number of visits was:
    A. 3.88
    B. 3.50
    C. 3.00
    D. 4.00

Answers: 27. A; 28. D; 29. D

Task 5

Answer the following questions:

1. The more dispersed, or spread out, a set of scores is:
    A. The greater the difference between the mean and the median
    B. The greater the value of the mode
    C. The greater the standard deviation
    D. The smaller the interquartile range

2. The mean height of a student group is 167 cm. Assuming height is normally distributed this enables us to deduce that:
    A. Approximately half of all students are taller than 167 cm
    B. Being a student stunts your growth
    C. Approximately half of all students are shorter than 167 cm
    D. A and C
    E. None of the above

3. If we subtract the value of the mean from every score in a set of scores the sum of the remaining values will be:
    A. Impossible to determine
    B. Equal to the mean
    C. A measure of the dispersion around the mean
    D. Zero
    E. None of the above

4. Given a normally distributed continuous variable the best measure of central tendency is the:
    A. Mode
    B. Median
    C. Mean
    D. standard deviation
    E. None of the above

5. If a distribution is negatively skewed, then:
    A. The median is greater than the mean
    B. The mode is greater than the median
    C. The mean is greater than the median
    D. Both A and B are true
    E. None of the above are true

6. In a normal distribution, the mean, the median and the mode:
    A. Always have the same value
    B. The mean has the higher value
    C. The mean has the lower value
    D. Have no particular relationship
    E. Cannot take the same value

7. Given the group of scores 1, 4, 4, 4, 7, it can be said of the mean, the median, and the mode that:
    A. The mean is larger than either the median or the mode
    B. All are the same
    C. The median is larger than either the mean or the mode
    D. All are different
    E. The mode is larger than either the median or the mode

Answers: 1. C; 2. D; 3. D; 4. C; 5. D; 6. A; 7. B

Task 6

The following four questions refer to the example:
A nurse recorded the number of analgesic preparations taken by patients in a surgical ward as the following:

5 2 8 2 3 2 4 12

1. The mode for this distribution is:
    A. 2
    B. 3
    C. 8
    D. There is no mode

2. The median is:
    A. 2.00
    B. 3.50
    C. 3.00
    D. 3.25

3. The mean is:
    A. 3.52
    B. 5.43
    C. 4.75
    D. 4.15

4. The range is:
    A. 9
    B. 10
    C. 12
    D. 2

Answers: 1. A; 2. B; 3. C; 4. B

Task 7

The following two questions refer to this data:

3 3 4 5 6 7 8 9 9 10 38 60.

1. The interquartile range is:
    A. 5.0
    B. 4.5
    C. 6.0
    D. 9.0

2. The percentage of cases falling between SD = -1 and SD = +1 is:
    A. 16.8%
    B. 33.6%
    C. 34.1%
    D. 68.3%

Answers: 1. A; 2. D

REFERENCES

1. Cramer MS. Clinical epidemiology and biostatistics. Springer-Verlag Berlin Heidelberg, 1988.
2. Polit DF, Hungler BP. Nursing research. Principles and methods. 6th rev. ed. Lippincot Williams & Wilkins; 1999.
3. Grancharova G, Hristova P. Medical statistics (in Bulgarian). Pleven: Publishing House of the Medical University of Pleven; 2004.
4. Kirkwood BR, Sterne JAC. Essentials of medical statistics. 2nd rev. ed. Oxford: Blackwell Publishing Company, 2003.
5. Leaverton PE. A review of biostatistics. A program for self-instruction. 2nd rev. ed. Boston: Little Brown and Company; 1981.
6. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd rev. ed. Geneva: World Health Organization; 1999.
7. Munro BH, Visintainer MA, Page EB. Statistical methods for health care research. Philadelphia: J.B. Lippincott Company; 1986.

8. Palmer C. Lectures in biostatistics. Teaching Course in Modern Epidemiology. Varna: Medical University of Varna; 1994.
9. Campbell MJ, Machin D, Walters SJ. Medical statistics: A textbook for the health sciences. 4th rev. ed. Chichester West Sussex: John Wiley & Sons LTD; 2007.
10. Hassard TH. Understanding biostatistics. St. Louis, Missouri: Mosby Year Book; 1991.
11. Kalinov K. Statistical methods in behavioural and social sciences (in Bulgarian). Sofia: NBU Publishing House; 2001.
12. Wonnacott TH, Wonnacott RJ. Introductory statistics. 3rd rev. ed. New York: John Wiley & Sons; 1977.
13. Maxwell DL, Satake E. Research and statistical methods in communication disorders. Baltimore, Maryland: Williams & Wilkins; 1997.
14. Weiss NA. Elementary statistics. Instuctor's Manual. Reading, Massachusetts: Addison-Wesley Publishing Company; 1989.
15. Lane DM. HyperStat Online Statistics Textbook. Trimmed mean. Available from URL: http://davidmlane.com/hyperstat/example_1.html. Accessed: April 20, 2009.
16. Forbes Digital Company. Investopedia. Geometric mean. Available from URL: http://www.investopedia.com/terms/g/geometricmean.asp. Accessed: April 20, 2009.
17. Ranchov G. Biostatistics and biomathematics - concepts, methods, applications (in Bulgarian). Sofia: Eco Print; 2008.
18. Statistics.com. Trimean. Available from URL: http://www.statistics.com/resources/glossary/t/trimmean.php. Accessed: April 20, 2009.
19. Morton RF, Hebel JR. A study guide to epidemiology and biostatistics (including 100 multiple-choice questions. Baltimore: University Park Press; 1983.
20. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd rev.ed. Singapore: Churchill Livingstone; 1991.
21. Bailar JC, Mosteller F. Medical uses of statistics. Waltham, Massachusetts: NEJM Books; 1986.
22. Abramson JH. Survey methods in community medicine. 4th rev. ed. Edinburgh London Melbourne and New York: Churchill Livingstone; 1990.
23. Ingelfinger JA, Mosteller F, Thibodeau LA, Ware JH. Biostatistics in clinical medicine. New York: Macmillan Publishing CO, Inc.; 1987.
24. Weiss NA. Elementary statistics. Reading, MA: Addison-Wesley Publishing Company; 1989.

RECOMMENDED READINGS

1. Campbell MJ, Machin D, Walters SJ. Medical statistics: A textbook for the health sciences. 4th rev. ed. Chichester West Sussex: John Wiley & Sons LTD; 2007.
2. Kirkwood BR, Sterne JAC. Essentials of Medical statistics. 2nd rev. ed. Oxford: Blackwell Publishing Company; 2003.
3. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd rev. ed. Geneva: World Health Organization; 1999.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASURES OF LOCATION: QUANTILES**
Module: 1.2.5	**ECTS (suggested): 0.10**
Author(s), degrees, institution(s)	**Gena Grancharova**, MD, PhD, Associate Professor Faculty of Public Health, Medical University of Pleven, Bulgaria **Silviya Aleksandrova**, MD, MB, Assistant professor Faculty of Public Health, Medical University of Pleven, Bulgaria
Address for correspondence	**Gena Grancharova** Medical University of Pleven 1, Kliment Ochridski Str. Pleven 5800, Bulgaria e-mails: dean-ph@mu-pleven.bg; gegran@optisprint.net
Keywords	Quantiles, median, quartiles, percentiles
Learning objectives	After completing this module students and public health professionals should be able to: • understand the concept of quantiles; • define and compute different types of quantiles; • explain the use of quartiles and percentiles to summarize health data;
Abstract	Measures of central tendency and dispersion are essential for summarizing any data set of individual scores. This process is based on two main characteristics of quantitave data – its variability and its tendency to some typical level. This section is devoted to quantiles as measures of location in the numerical approach of data summarising with the objective to present identification as well as methods of calculation of different types of quantiles, especially percentiles and quartiles.
Teaching methods	A half ann hour lecture in the form of Power Point presentation should introduce the students to the main concepts of quantiles, types of quantiles and their use. After the lecture students should read and discuss in groups all the material presented in this section and individually answer the multiple choice questionnaire at the end of this lecture.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1-2 student), LCD projector, access to the Internet and statistical package software; • training materials: recommended readings or other related readings; • target audience: bachelor and master students in public health.
Assessment of students	Multiple choice questionnaire (MCQ) - minimum 70% success.

MEASURES OF LOCATION: QUANTILES

Gena Grancharova, Silviya Aleksandrova

THEORETICAL BACKGROUND

Quantiles

Quantiles are special measures of location that could be described as the points that divide the ordered series of data (from the lowest to the highest value) into subgroups of equal size regarding the number of units in a subgroup. In other words, quantiles are the data values marking the boundaries between consecutive subgroups ordered series of data (an array).

Types of quantiles

There exist several types of quantiles, dividing an array in different number of subgroups: e.g., terciles divide the distribution into three equal subgroups (called thirds), quartiles - into four subgroups (quarters), quintiles -into five (fifths), deciles - into ten (tenths), and centiles – into hundred (hundredths) (1).

Most frequently used quantiles are presented in Table 1, including their number and number of subgroups they are dividing the distribution.

Table 1. Most frequently used quantiles.

Quantile/s	Number of equal parts of a distribution	Number of quantiles of this type
Median	2	1
Quartiles	4	3
Deciles	10	9
Percentiles	100	99

As shown in Table 1, in all types of presented quantiles their number is one less then the number of corresponding equal parts of a distribution they are dividing.

Estimation of quantiles

Quantiles are measures of location that are usually not calculated but identified or determined. The procedure of identifying quantiles is as follows:

- first we need to rearrange all observations according to the magnitude of a value of a variable we are observing in an ordered series of data from the lowest to the highest value (NOTE: one must consider all observations even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the location of the very first quantile dividing the distribution in two parts equal in number of observations (the median) is simple - the median is simply the value of the middle observation in the ordered series of data. When the number of observations is even, we locate the central two observations. Afterwards we sum values of these two units and divide the sum by 2 - the median is just a halfway of values of the two middle observations;

- then we repeat exactly the same procedure in the lower half and in the upper half of the ordered series to locate the quantiles dividing the ordered series in four equal parts regarding the number of units or observations (quartiles);
- we repeat the procedure until we divide the distribution in wanted number of equal parts.

For the estimation of quantiles there exist also procedures to calculate them. They are described later in this module.

Median

Median (Me) is a quantile, dividing an ordered series to two halfs equal regarding number of observations (Table 1). Usually it is identified from the ordered using following procedure:

- first we need to rearrange all observations according to the magnitude of a value of a variable we are observing in an ordered series of data from the lowest to the highest value (NOTE: one must consider all observations even though some values may repeat more than once);
- then we must determine whether the number of cases is odd or even;
- when the number of observations is odd, the median is simply the value of the middle observation (unit) in the ordered series of data;
- when the number of observations is even, the median is just a halfway of values of the two middle observations.

This simple procedure of dentification of the median in practice is presented in Case study, Example 1.

To identify the median in a long ordered series of data we can use the formula for calculating the position (rank) of the unit carrying median value (Equation 1).

$$Me_{Rank} = \frac{(n+1)}{2} \qquad \textbf{Equation 1.}$$

Me = rank of the unit carrying median value
n = number of cases

This procedure is also presented in Case study, Example 1.

Calculation of median

Median is usually identified but it could also be calculated if we have groupped data. For a grouped frequency distribution, the calculation of the median might be a little more complicated. If we assume that the variable is continuous (for example time, height, weight, or level of pain) we can use a formula for calculating the mean. This

formula can be applied to ordinal data, provided the variable being measured has an underlying continuity (Equation 2) (2).

$$Me = X_L + \left(i \times \frac{(n/2) - cumf_L}{f_i} \right)$$ **Equation 2.**

Me = median
X_L = number of observations below a given observation
i = width of the class interval
n = number of cases
$cumf_L$ = cumulative frequency at the real lower limit of the interval
f_i = frequency of cases in the interval containing the median

Calculation of the median in practice is presented in Case study, Example 2.

Use of median

It is used as a measure of central tendency. It is described in details in separate module in this book.

Percentiles

Percentiles (also called centiles (1)) are points that divide an array into 100 equal parts (3-6). Thus, there are 99 percentiles. They are denoted as P_1, P_2, ... P_{25},...P_{50},...., P_{75}, ...P_{99}.

Some of characterristics of percentiles are as follows:

- a percentile tells us the relative position of a given observation,
- it allows us to compare scores on tests that have different means and standard deviations (7,8),
- the tenth percentile, for example, exceeds 10% and is exceeded by 90% of the observations. The seventy-fifth percentile exceeds 75% of the data, etc.

Calculation of percentiles

Percentiles are usually identified but they could also be calculated.

In ungrouped data set a percentile (P_i) can be calculated by the formula (Equation 3):

$$P_i = \frac{n_{below\ obs\ x}}{n_{all}} \times 100$$ **Equation 3.**

P_i = percentile i
$n_{below\ obs\ x}$ = number of observations below a given observation
n_{all} = number of all observations

Calculation of a percentile in practice in this case is presented in Case study, Example 3.

Calculation of percentiles in an interval scale is much more complicated. The following formula can be used (Equation 4) (3):

$$P_i = L_{pi} + \frac{c}{f} \times e \qquad \textbf{Equation 4.}$$

P_i = the value of the percentile i
L_{pi} = the lower limit of the interval where the percentile is situated
c = the difference between the percentile rank and the cumulative frequency in the previous interval
f = the number of cases in the percentile interval
e = the width of the percentile interval

An intermediate step in this procedure is calculation of rank of a given percentile (Equation 5):

$$Percentile\ rank = \frac{\sum f}{100} \times i \qquad \textbf{Equation 5.}$$

Σf = the total number of cases
i = the number, corresponding to the percentile rank (for P_{10} i = 10, for P_{25} - 25, etc.).

Calculation of a percentile in practice in this case is presented in Case study, Example 4.

Use of percentiles

Percentiles are practically used to establish the reference limits of normality in many clinical and other areas of investigation. For this purpose usually seven main percentiles are used - P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90} and P_{97} – to form the upper and lower limits of seven reference groups of population.

Percentiles are widely used for the establishment of "normal ranges" of values for health data that permits the selection of appropriate action in medical practice or allows for accurate estimate of many clinical and laboratory indicators. Percentiles have an advantage as compared to the other methods of determining "normal" values as they are applicable to any form of variables/distribution (not only to normal distribution). When the investigator prefers to use seven reference groups the limits of "normal" values are determined by P_{25} and P_{75} whereas P_{50} corresponds to the mean.

Quartiles

These are observations in an array that divide the distribution into four equal parts. Therefore, there 3 quartiles, denoted as Q_1, Q_2 and Q_3. If we have an array of 23 cases, the first quartile Q_1 is the 6^{th} observation with its corresponding value; the second quartile Q_2 is the 12^{th} observation, and the third quartile Q_3 is equal to the 18^{th} observation.

Identification of the quartiles in practice is presented in Case study, Example 5.

Comparison between different types of quantiles

It is worth mentioning that:

- P_{25} corresponds to Q_1,
- P_{50} corersponds to Q_2 and to the median, and
- P_{75} is equal to Q_3,
- taking into account that the sample median (Me) is the second quartile (Q_2), then the median of the lower half of the data gives the first quartile (Q_1), and similarly, the median of the upper half of the data gives the third quartile (Q_3).

Use of quantiles

Quantiles are used in description of both, central tendency and dispersion of a distribution they are describing. Median is used as a measure of central tendency while quartiles are used for quick estimation of the degree of dispersion in an array.

CASE STUDY

Example 1 - identification of the median in an ungroupped data

Suppose the age at first birth for a sample of 10 mothers is (9):

18 21 23 23 25 27 27 28 30 33

In this set of data where the number of units in the sample is even (ten), the median age at first birth is 26 – the mean of the two middle numbers 25 and 27.

18 21 23 23 25 27 27 28 30 33

In continuation we add to the observed set of data one case more:

18 21 23 23 25 27 27 28 30 33 33

In this set of data where the number of units in the sample is now odd (eleven), the median age at first birth is 27 – simply the value of the middle observation in the ordered series of data:

18 21 23 23 25 27 27 28 30 33 33

Now let us take two very simple examples to present the procedure of identification of the median by calculating the rank of the unit carrying median value first (Equation 1). Suppose we have the values of a variable in a group of 7 units (n = 7) are as follows:

5 8 9 10 15 18 28

Using the Equation 1 we find that in this ordered array the median will be the 4th score (Equation 6):

$$Me_{Rank} = \frac{(7+1)}{2} = 4 \qquad \textbf{Equation 6.}$$

So, the median is equal to 10. Now we have odd number of units. Suppose we have the the values of a variable a group of 8 units (n = 8) are as follows:

6 12 17 19 20 21 24 27

Using the Equation 1 we find that in this ordered array the median will be situated between 4th and 5th scores (Equation 7):

$$Me_{Rank} = \frac{(8+1)}{2} = 4.5 \qquad \textbf{Equation 7.}$$

So, the median is equal to 19.5.

Example 2 - calculation of the median in a groupped data

In a study of pain measurement the following data were obtained for 17 units (n=17) (2):

1 1 2 2 2 2 2 3 3 3 3 4 4 4 5 5 5

By inspection, we can see that the median will fall into the category "3" (the 9th score).

Assuming underlying continuity, we can present the data into real class intervals (Table 2):

Table 2. Data on a measurement of pain among 17 subjects.

Score	Real class interval	Frequency	Cumulative frequency
1	0.5 – 1.4	2	2
2	1.5 – 2.4	5	7
3	2.5 – 3.4	4	11*
4	3.5 – 4.4	3	14
5	4.5 – 5.4	3	17
		N = 17	* median

Now we can apply Equation 2 for calculation (Equation 8):

$$Me = 2.5 + \left(1 \times \frac{(17/2) - 7}{4} \right) = 2.5 + \left(\frac{1.5}{4} \right) = 2.5 + 0.375 = 2.875 \qquad \textbf{Equation 8.}$$

Example 3 – calculation of percentiles in an ungroupped data

Let's say that a student received a score of 90 points on a test given to a group of 50 examined students, 40 of them with scores less than 90 points (2). That means that his location in an ordered series is 41st place. The percentile rank for this student will be calculated according to Equation 3 (Equation 9):

$$P_i = \frac{40}{50} \times 100 = 80 \qquad \textbf{Equation 9.}$$

In other words, he achieved a higher score than 80% of the students who took the test and 20% of the students received better results than that particular student.

Example 4 – calculation of percentiles in an inteval scale and its use

In a representative sample of 120 urban male liveborns the height scores are presented in an interval array of 3 cm of width. Let's calculate the value of P_3, P_{10}, P_{25}, P_{50}, P_{75}, P_{90} and P_{97}. The values of the variable "weight" are presented in a table (Table 3), containing equal width intervals and their corresponding frequencies.

The calculation of percentiles includes the following steps (3,7,8):

1. Determining the cumulative frequences for each interval by adding up to the absolute frequences in a given interval the absolute frequences from the previous interval.
2. Determing the percentile ranks using the Equation 3 (Equations 10-16):

Table 3. The values of the variable "weight" ontaining equal width intervals and their corresponding frequencies

Height in cm ***x***	**Frequency** ***f***	**Cumulative frequency** ***f***	**Percentiles' rank**
41 - 43	3	-	
44 - 46	14	17	P_3, P_{10}
47 - 49	40	57	P_{25}
50 - 52	45	102	P_{50}, P_{75}
53 - 55	15	117	P_{90}, P_{97}
56 - 58	3	120	
	$\Sigma f = N = 120$		

$$P_3 = \frac{120}{100} \times 3 = 3.6$$ **Equation 10.**

$$P_{10} = \frac{120}{100} \times 10 = 12$$ **Equation 11.**

$$P_{25} = \frac{120}{100} \times 25 = 30$$ **Equation 12.**

$$P_{50} = \frac{120}{100} \times 50 = 60$$ **Equation 13.**

$$P_{75} = \frac{120}{100} \times 75 = 90$$ **Equation 14.**

$$P_{90} = \frac{120}{100} \times 90 = 108$$ **Equation 15.**

$$P_{97} = \frac{120}{100} \times 97 = 116.4$$ **Equation 16.**

3. According to the cumulative frequencies and the percentile ranks in step 2 we define for each percentile the interval where it should fall: in our example P_3

and P_{10} fall in the interval 44-46 cm, P_{25} - in the interval 47-49, P_{50} and P_{75} - in the interval 50-52 cm, P_{90} and P_{97} - in the interval 53-55 cm.

4. Using the Equation 2, we can calculate the value of each percentile. Here we need to take into account that value of "c" - the difference between the percentile rank and the cumulative frequency in the previous interval is:
    - for P_3 = 3.6 (percentile rank) – 0 (cumulative frequency in the previous interval) = 3.6,
    - for P_{10} = 10 (percentile rank) – 0 (cumulative frequency in the previous interval) = 10,
    - for P_{25} = 30 (percentile rank) – 17 (cumulative frequency in the previous interval)= 13,
    - for P_{50} = 60 (percentile rank) – 57 (cumulative frequency in the previous interval) = 3,
    - for P_{75} = 90 (percentile rank) – 57 (cumulative frequency in the previous interval) = 33,
    - for P_{90} = 108 (percentile rank) – 102 (cumulative frequency in the previous interval) = 6,
    - for P_{97} = 116.4 (percentile rank) – 102 (cumulative frequency in the previous interval) = 14.4.

    The values for each percentile are now calculated as follows (Equations 17-23):

$$P_3 = 44 + \frac{3.6}{14} \times 3 = 44.77$$ **Equation 17.**

$$P_{10} = 44 + \frac{10}{14} \times 3 = 46.14$$ **Equation 18.**

$$P_{25} = 47 + \frac{13}{40} \times 3 = 47.98$$ **Equation 19.**

$$P_{50} = 50 + \frac{3}{45} \times 3 = 50.20$$ **Equation 20.**

$$P_{75} = 50 + \frac{33}{45} \times 3 = 52.20$$ **Equation 21.**

$$P_{90} = 53 + \frac{6}{15} \times 3 = 54.20$$ **Equation 22.**

$$P_{97} = 53 + \frac{14.4}{15} \times 3 = 55.88$$ **Equation 23.**

We can use these data now to determine the reference groups for the weight of the male newborns in that urban area. The limits of the seven reference groups will be as presented in Table 4.

Table 4. Percentiles and their use for reference groups*.

Reference groups	Percentiles	Weight of the male newborns
Severely stunted	Below P_3	Below 44.77
Moderately below the norm	$P_3 - P_{10}$	44.77 – 46.14
Sligthly below the norm	$P_{10} - P_{25}$	46.14 – 47.98
Normal range	P_{25} - P_{75}	47.98 – 52.2
Sligthly above the norm	$P_{75} - P_{90}$	52.2 – 54.2
Moderately above the norm	$P_{90} - P_{97}$	54.2 – 55.88
Expressed accelaration	Above P_{97}	Above 55.88

* All the data in the table are an approximation

Example 5 – identification of quartiles in an ungroupped data

Say we have a sample where n = 16 and the values of the variable are the following (2):

1 5 7 7 8 9 9 10 11 12 13 15 19 20 20 20

By inspection of the data, we find:

- the first quartile Q_1 is located between 4^{th} and 5^{th} scores; thus $Q_1 = (7 + 8)/2 = 7.5$
- the second quartile Q_2 is located between the 8^{th} and 9^{th} scores; thus $Q_2 = (10 + 11)/2 = 10.5$
- the third quartile Q_3 is located between 13^{th} and 14^{th} scores; thus $Q_3 = (15 + 19)/2 = 17.0$

These values are located in the array as follows:

1 5 7 [7 8] 9 9 [10 11] 12 13 [15 19] 20 20 20

EXERCISE

Task 1

For the following questions choose between “true” or “false”:

1. The 50^{th} percentile score and the median will always be the same value. T F
2. Twenty five percent (25%) of the scores fall between Q1 and the median. T F
3. The distance between Q1 and the median is always different to the distance between Q3 and the median. T F
4. The median and the 50^{th} percentile rank have different values. T F

Answers: 1. T; 2. T; 3. F; 4. F

Task 2

Consider following set of data:

3, 3, 4, 5, 6, 7, 8, 9, 9, 10, 38, 60

1. The median is:
   A. 7.0
   B. 7.5
   C. 8.0
   D. 3 or 9

2. Q_1 is:
   A. 4.5
   B. 5.5
   C. 8.0
   D. 9.5

3. Q_3 is:
   A. 4.5
   B. 6.0
   C. 7.5
   D. 9.5

Answers: 1. B; 2. A, 3. D

REFERENCES

1. Last JM (editor). A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd ed. Singapore: Churchill Livingstone; 1991.

3. Grancharova G, Hristova P. Medical statistics (in Bulgarian). Pleven: Medical University of Pleven Publishing House, 2004.
4. Leaverton PE. A review of biostatistics. A program for self-instruction. 2nd edition. Boston: Little Brown and Company; 1981.
5. Munro BH, Visintainer MA, Page EB. Statistical methods for health care research. London: J. B. Lippincott Company; 1986.
6. Lwanga SK, Tye CY, Ayeni O. Teaching health statistics. Lessons and seminar outlines. 2nd ed. Geneva: World health organization; 1999.
7. Kalinov K. Statistical methods in behavioural and social sciences (in Bulgarian). Sofia: New Bulgarian University Publishing House; 2001.
8. Ranchov G. Biostatistics and Biomathematics – concepts, methods, applications (in Bulgarian). Sofia: Eco Print; 2008.
9. Palmer C. Lectures in biostatistics. Teaching Course in Modern Epidemiology. Varna: Medical University of Varna; 1994.

RECOMMENDED READINGS

1. Campbell MJ, Machin D, Walters SJ. Medical Statistics: A textbook for the health sciences. 4th rev. ed. Chichester: John Wiley & Sons LTD; 2007.
2. Kirkwood BR, Sterne JAC. Essentials of medical statistics. 2nd rev. ed. Chichester: Blackwell Publishing Company, 2003.
3. Polgar S, Thomas SA. Introduction to research in the health sciences. 2nd ed. Singapore: Churchill Livingstone; 1991.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**FREQUENCY MEASURES: ESTIMATING RISK**
Module: 1.2.6	**ECTS (suggested): 0.30**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Jadranka Božikov,** PhD, Professor Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Risk, cumulative incidence, simple cumulative method, actuarial method, density method
Learning objectives	After completing this module students should be: • familiar with differences between four different methods for estimation of cumulative risk, being simple cumulative, actuarial, density, and Kaplan Meier method; • able to estimate cumulative risk measures of different level of accuracy independently.
Abstract	Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period. In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and most widely used. Risk cannot be accurately estimated by this method unless all subjects in the observed candidate population are followed for the entire follow-up period or are known to develop the disease during the period (no censoring). Because of serious limitations of this method, several methods more or less susceptible to censoring were proposed. Considering the censoring of the data in estimating cumulative risk requires the use of special analytic methods. These methods are actuarial, density, and Kaplan Meier method.
Teaching methods	An introductory lecture gives the students first insight in four methods for calculation of cumulative risk. The theoretical knowledge is illustrated by case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards they on provided data set perform tasks on estimation of different types of measures. They are stimulated to compare results with other students and discuss the differences.
Specific recommendations for teachers	• work under teacher supervision/individual work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which detailed description of process of calculation is described.

FREQUENCY MEASURES: ESTIMATING RISK

Lijana Zaletel-Kragelj, Jadranka Božikov

THEORETICAL BACKGROUND

Introduction

In expressing relative incidence we are dealing with several measures. One of them is so called risk.

Risk is defined as the probability that a disease-free individual is developing a disease under observation over a specified period, conditional on that the same individual is not dying from any other disease during the period (1). Thus, risk is a conditional probability, with values varying between zero and one. It is dimensionless (1). It usually refers to the first occurrence of the disease for each initially disease-free individual, although it is possible to consider the risk of developing the disease under observation within a specified period more than once (1).

In practice, risk is estimated by using different methods. The simple cumulative method is the easiest and most widely used (1). For a cohort of subjects followed for a given period of time, risk is often estimated by calculating the proportion of candidate subjects who develop the disease during the observation period. This measure is usually referred as the cumulative incidence (CI) (1). Generally cumulative incidence is estimated only for first occurrence of the disease. If the durations of the individual follow-up periods for all non-cases are equal, the cumulative incidence is equivalent to the average risk for members of the cohort. This means thait under the condition of a fixed cohort cumulative incidence is good estimate of risk. This is the reason that cumulative incidence and risk are frequently equalized. But once again, because risk is, by its definition, a conditional probability, it cannot be accurately estimated by calculating cumulative incidence unless all subjects in the observed candidate population are followed for the entire follow-up period or are known to develop the disease (or other observed phenomenon) during the period (1).

The cumulative probability of the event during a given time interval is the proportion of new events during the interval in which the denominator is the initinal number of observed persons. The calculation of this measure is straightforward if no losses happen in the cohort during the interval (1-9). However, in real life the size of the cohort is more than likely to be decreased after a long period of follow-up as a result of different reasons. A situation in which the event and the time of individual is at risk for the event is unknown is usually called censoring (2,8-12).

There are usually three reasons why censoring occurs. The first is the termination of the observation because of the end of the study before the event occurs, the second is the termination because of some competing factors (death of other cause e.g. traffic accident), the third, the fourth simply the lost because of changing the domicile of the individual under observation, etc. In all cases the occurrence of observed phenomenon is unknown. The terms also used with this phenomenon are “withdrawals”, “losts-to-follow-up” and others (2,8-12). Considering the censoring of the data requires the use of special analytic methods.

The methods of risk estimation are the simple cumulative method, the actuarial method, the density method, and the Kaplan Meier product limit method (1,2,9-13).

Methods of risk estimation

Simple cumulative method

This method is the easiest for estimating risk (1,2,12). The risk calculated by this method is the most rough measure in this family of measures.

It is simply the proportion of new events during the interval in which the denominator is the initinal number of observed persons (Equation 1):

$$_{cum}R = \frac{N_{d+new\,cases(gp)}}{N_{all\ persons\,at\,risk\,(bgp)}} \qquad \textbf{Equation 1.}$$

$_{cum}R$ = *cumulative risk (risk of getting a disease during the entire period)*
$N_{d+\ new\ cases\ (gp)}$ = *number of new cases of the disease under observation during a given period*
$N_{all\ persons\ at\ risk\ (bgp)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*

Usually it is estimated only for the first occurrence of the disease. This is the reason that the population at risk (the denominator in the equation) consists of disease-free individuals at the beginning of the observational period. The observation period has to be clearly stated since the value of the measure is increasing with the prolongation of period of observation. This period could be based upon a callendar time or not (e.g. first year after the exposure, first year after surgery etc.). It is good estimate of the risk only in the case of fixed cohorts in which there are no withdrawals from the follow-up (1,12).

Estimation of cumulative risk over entire 5-year observational period in practice is presented in in Case study 1.

For avoiding the drawbacks of this rough direct method of estimation of cumulative risk over longer period, we could split this longer period first to shorter periods (i.e. 1-year periods) and obtain cumulative risk indirectly through calculating risks for these periods (partial risks). When partial risk refers to 1-year period it is known as annual risk (Equation 2):

$$_{ann}R = \frac{N_{d+new\,cases(1\text{-}year\ period)}}{N_{all\ persons\,at\,risk\,(beginning\,of\ 1\text{-}year\ period)}} \qquad \textbf{Equation 2.}$$

$_{ann}R$ = *annual risk (risk of getting a disease during the 1-year period)*
$N_{d+\ new\ cases\ (1\text{-}year\ period)}$ = *number of new cases of the disease under observation during1-year period*
$N_{all\ persons\ at\ risk\ (beginning\ of\ 1\text{-}year\ period)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given 1-year period*

The annual risk is annual probability of the event (12). The complement of this probability (the mirror image) is annual probability of survival without an event under observation (i.e. a breakout of a disease). Technically these probabilities are conditional probabilities. This means for example, that one has to survive through the first interval in order to be a part of the denominator for the calculation of the survival probability in the second interval. Similarly, the survival probability for the third interval is calculated only among those persons who survived first the first and then the second interval (12).

A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, 5-year interval, etc.) is obtained by multiplying the annual conditional survival probabilities over all intervals (12). Afterwards we calculate again complementary values (1 – cumulative survival) that are in fact cumulative risks over more than one interval.

By using this procedure the censoring is partially considered even when using simple method, as we need to define separately for every year the number of individuals under observation at risk, and all participants who terminated the observation because of extraneous factors (e.g. death because of traffic accident etc.) are not included.

Estimation of cumulative 5-year risk over observational period through calculation of annual risks is presented in in Case study 1.

Actuarial method

This is the first method in which the censoring is considered in calculation of risk estimate (1,8,11-13). It is tipically used to estimate the probability of death in survival analysis, but as mortality is a special case of incidence (12), it could be generalized to estimation of risk on general (2). It is refered also as interval-based life table or life table interval approach (12).

This method is working under the assumption that the censoring is occurring uniformly throughout the observed period (usually meaning that all withdrawals, i.e. censored observations, occur on average in the middle of the observational period) (1,2,11). If the periods are short (up to 1 year), or there is a small number of withdrawals this assumption does not affect the risk estimate seriously (1). However, one should be aware that this method still provides us more or less biased estimate of risk (1). The basic equation for calculating risk by using actuarial method directly is as follows (Equation 3):

$$ {}_{cum}R = \frac{N_{d+\,new\,cases(gp)}}{N_{all\ \ persons\,at\,risk\,(bgp)} - \frac{N_{w(gp)}}{2}} \qquad \textbf{Equation 3.} $$

${}_{cum}R$ = *cumulative risk (risk of getting a disease during the entire period)*
$N_{d+\,new\,cases\,(gp)}$ = *number of new cases of the disease under observation during a given period*
$N_{all\ persons\,at\,risk\,(bgp)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of a given period*
$N_{w\,(gp)}$ = *number of withdrawals during a given period*

For avoiding the drawbacks of this method we could again split longer period first to shorter periods (i.e. 1-year periods) and calculate risks for these periods (i.e. annual risks). Only afterwards, on the basis of risks of shorter periods as intermediate elements, the cumulative risk is calculated indirectly. Annual risks could be calculated as follows (Equation 4):

$$ {}_{ann}R = \frac{N_{d+\ new\ cases\ (1\text{-}year\ period)}}{N_{all\ \ persons\ at\ risk\ (beginning\ of\ 1\text{-}year\ period)} - \frac{N_{w\ (1\text{-}year\ period)}}{2}} $$ **Equation 4.**

${}_{ann}R$ = *annual risk (risk of getting a disease during the 1-year period)*
$N_{d+\ new\ cases\ (1\text{-}year\ period)}$ = *number of new cases of the disease under observation during the 1-year period*
$N_{all\ persons\ at\ risk\ (beginning\ of\ 1\text{-}year\ period)}$ = *number of all persons at risk for getting ill with the disease under observation at the beginning of the 1-year period*
$N_{w\ (1\text{-}year\ period)}$ = *number of withdrawals during the 1-year period*

Estimation of this measure in practice is presented in Case study 2.

Again, a cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, 5-year interval, etc.) is obtained by multiplying the partial conditional survival probabilities over all intervals (Equation 5) (12):

$$ {}_{cum}R = 1 - \left[\left(1 - {}_{ann}R_{(year1)}\right) \times \left(1 - {}_{ann}R_{(year2)}\right) \times \ldots \times \left(1 - {}_{ann}R_{(yearn)}\right)\right] $$ **Equation 5.**

${}_{cum}R$ = *cumulative risk (risk of getting a disease during the entire period of observation)*
${}_{ann}R_{(year\ 1)}$ = *annual risk (risk of getting a disease) during the 1st year*
${}_{ann}R_{(year\ 2)}$ = *annual risk (risk of getting a disease) during the 2nd year*
${}_{ann}R_{(year\ n)}$ = *annual risk (risk of getting a disease) during the nth year*

Estimation of this measure in practice is presented in Case study 2.

Because of serious limitations of this method, other methods were proposed (1).

Density method

Actuarial method is working under the assumption that all withdrawals occur on average in the middle of the observational period (1,2,11). If the periods are short, or there is a small number of withdrawals this assumption does not affect the risk estimate seriously (1). However, it is better to consider exact times of being at risk of developing a disease under observation. Another interval-based method based on the estimation of average incidence rates (person-time rate or incidence density) was proposed (1,3,4,11,12). This method depends on the functional relationship between a risk and an incidence rate (estimated through incidence density) (1).

Risk depends on incidence density and on the duration of the period of observation. Under the assumption that the cohort under observation is fixed (with no censored observations), and that the incidence density is constant over the period of observation, the risk estimate could be directly calculated as follows (Equation 6) (1,3):

$$_{cum}R = 1 - e^{\left(-ID \times t_{(gp)}\right)} \qquad \textbf{Equation 6.}$$

$_{cum}R$ = *cumulative risk (risk of getting a disease during the entire period)*
ID = *incidence density*
$t_{(gp)}$ = *duration of the given period of observation (period at risk)*

Incidence density, used in this equation was introduced in separate module in this book. It is the rate between the number of new cases which occur during the period under observation, and the quantity known under the term person-time (PT). It is calculated as (Equation 7):

$$ID = \frac{N_{d+new cases(gp)}}{PT} \qquad \textbf{Equation 7.}$$

ID = *incidence density*
$N_{d+ new cases (gp)}$ = *number of new cases of the disease under observation during a given period*
PT = *person-time*

However, usually the incidence density (as an estimate of incidence rate) does not remain constant during the entire follow-up period. Like in actuarial method, cumulative risk over a longer period also in this method is not calculated directly. We split this longer period first to shorter periods (i.e. 1-year periods) and calculate risks for these periods (partial risks), i.e. annual risks. They could be calculated as follows (Equation 8):

$$_{ann}R = 1 - e^{\left(-_{ann}ID \times 1\right)} \qquad \textbf{Equation 8.}$$

$_{ann}R$ = *annual risk (risk of getting a disease during the 1-year period)*
$_{ann}ID$ = *annual incidence density*

We can see that annual incidence densities need to be calculated prior calculation of annual risks (Equation 9):

$$_{ann}ID = \frac{N_{d+newcases(1-year period)}}{PT} \qquad \textbf{Equation 9.}$$

$_{ann}ID$ = *annual incidence density*
$N_{d+ new cases (gp)}$ = *number of new cases of the disease under observation during a 1-year period*
PT = *person-time*

Estimation of annual incidence densities and annual risks estimated by usig density method in practice is presented in Case study 3.

Only afterwards, on the basis of annual risks as intermediate elements, the cumulative risk is calculated as follows (Equation 10):

$$_{cum}R = 1 - e^{\left[\left(-_{ann}ID_{(year1)}\times 1\right)+\left(-_{ann}ID_{(year2)}\times 1\right)+\ldots+\left(-_{ann}ID_{(yearn)}\times 1\right)\right]} \qquad \textbf{Equation 10.}$$

$_{cum}R$ = *cumulative risk (risk of getting a disease during the entire period)*
$_{ann}ID_{(year\ 1)}$ = *annual incidence density during the 1st year*
$_{ann}ID_{(year\ 2)}$ = *annual incidence density during the 2nd year*
$_{ann}ID_{(year\ n)}$ = *annual incidence density during the nth year*

Estimation of this measure in practice is presented in Case study 3.

Kaplan Meier product limit method

Kaplan Meier product limit method (8,11,12) combines calculated probabilities of survival and estimates to allow censored observations, which are assumed to occur randomly. The intervals are defined as ending each time an event (i.e. disease, death, withdrawal) occurs and are therefore unequal (2,12). Again, these probabilities are conditional – they are conditioned on being at risk (present in the study without a disease under observation or censored) at each event time. The formula for calculation of conditional probability is simply (Equation 11):

$$p = \frac{N_{d+i}}{N_{persons\ at\ risk\ i}} \qquad \textbf{Equation 11.}$$

p = *conditional probability for an event in time i*
N_{d+i} = *number of events (new cases of a disease or death) occurring at time i*
$N_{persons\ at\ riski\ i}$ = *number of individuals still under observation (still at risk of the event under observation) at time i*

When time i is measured exactly, the number of events is usually 1.

The complement of this conditional probability of an event is probability of survival without an event under observation (i.e. a breakout of a disease) (12). A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, 5-year interval, etc.) is obtained by multiplying the annual conditional survival probabilities over all intervals (12).

Estimation of cumulative 5-year risk over observational period through calculation of conditional probabilities is presented in Case study 4.

CASE STUDIES

Data set

For the illustration of differences between the simple, the actuarial, the density, and the Kaplan Meier product limit method of calculation of cumulative risk an imaginary data-set is used. A cohort of 20 individuals initially without a disease under observation, were followed up for 5 years (Figure 1).

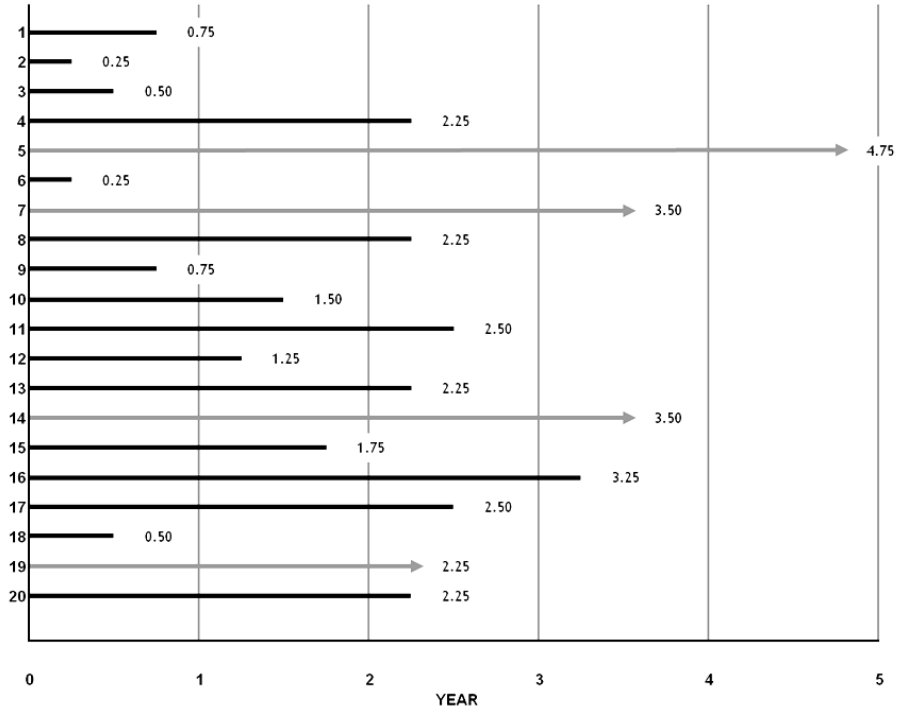


Figure 1. Graphic presentation of events in a cohort of 20 people. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

In this period, 16 individuals got a disease under observation (an event under observation) (Figure 1, persons with black lines of follow-up time), while 4 of them were lost to follow-up because of voluntarily withdrawal from the study or change of domicile (persons No. 5, 7, 14 and 19) (Figure 1, persons with gray lines of follow-up time). The lines with arrows indicate that individuals were alive at the time of the lost of follow-up.

In Figure 1 the members of a cohort are presented in order as they were numbered at the time of the entry into the study, while in Figure 2, the members are rearranged in

rank order regarding the time of an event or withdrawal. This presentation is useful in determination of times of being at risk fot the event under observation.

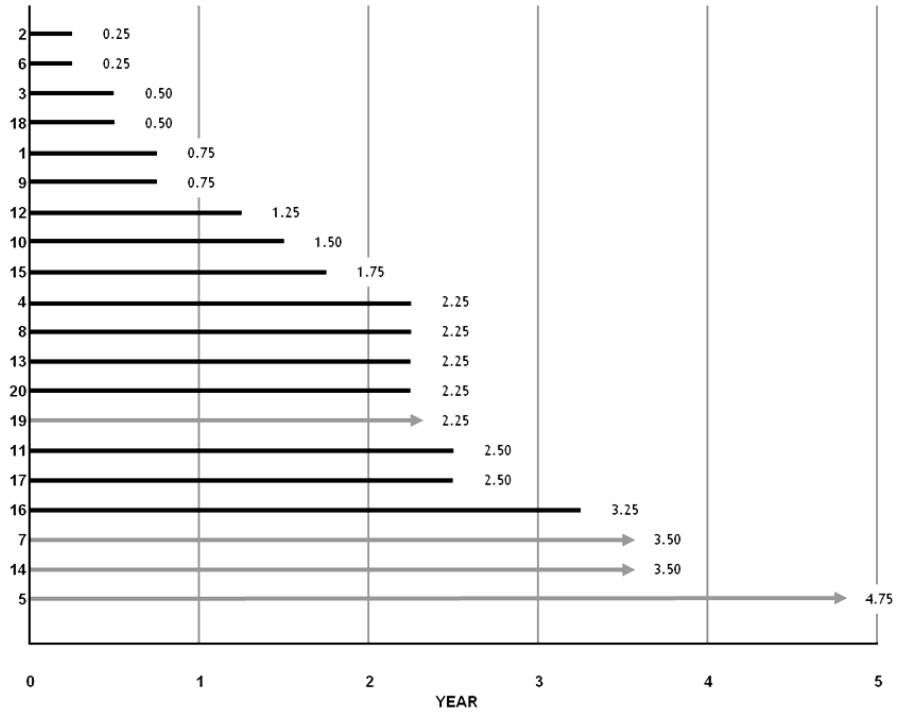


Figure 2. Ordered time of being at risk of developing a disease under observation in a cohort of 20 people from Figure 1. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individiuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individiuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Case study 1: Estimation of cumulative risk using simple cumulative method

Results of counting of cases of observed disease which broke out during the entire 5-year time of observation (Figure 1) show that the cumulative 5-year risk estimated by the simple cumulative method according to Equation 1 is (Equation 12):

$$_{cum}R = \frac{16}{20} = 0.8000 \qquad \textbf{Equation 12.}$$

But this estimate is unreliable as there are censorings in 4/20 individuals under observation (No. 5, 7, 14 and 19) (Figure 1). In these individuals the occurrence of the event of interst is uncertain because of the termination of the observation before the event occurred. To diminish the drawbacks of this method we can split 5-year interval to 5 1-year intervals, and for each 1-year interval we calculate the annual risk by following next steps:

- define the number of persons entered in the interval (Table 1, column 1), number of persons with the disease at the end of interval (Table 1, column 2), and the number of losts (withdrawals) (Table 1, column 3),
- by using Equation 2 calculate annual risks (Table 1, column 4).

From the Table 1 it could be seen that in case of calculation of annual risks, the censoring is partially considered even when using simple cumulative method, as we need to define separately for every year the number of individuals at risk, and all participants who terminated the observation because of extraneous factors (e.g. death because of traffic accident etc.) are not included.

Table 1. Elements for calculation and calculation of annual risks using simple cumulative method.

Year of observation	**1**	**2**	**3**	**4**
	Entered in the interval (N)	**With the disaese at the end of interval (d+)**	**Lost**	**d+/N (annual risk) ($_{ann}$R)**
1st	20	6	0	0.3000
2nd	14	3	0	0.2143
3rd	11	6	1	0.5455
4th	4	1	2	0.2500
5th	1	0	1	0.0000

The annual risk (Table 1, column 4) is annual probability of the event (12). The complement of this probability is annual probability of survival without an event under observation (i.e. a breakout of a disease) (Table 2, column 5). Technically these probabilities are annual conditional probabilities. A cumulative probability of survival without a disease under observation over more than one interval (2-, 3-, 4-, and 5-year interval) is obtained by multiplying the annual conditional survival probabilities over all intervals (Table 2, column 6) (12).

Table 2. Calculation of cumulative 5-year risk from annual risks using simple cumulative method.

Year of observation	**4**	**5**	**6**	**7**
	d+/N (annual risk) ($_{ann}$R)	**1 - $_{ann}$R**	**product (1 - $_{ann}$R) (Π)**	**1 – Π (cumulative risk) ($_{cum}$R)**
1st	0.3000	0.7000	0.7000	0.3000
2nd	0.2143	0.7857	0.5500	0.4500
3rd	0.5455	0.4545	0.2500	0.7500
4th	0.2500	0.7500	0.1875	0.8125
5th	0.0000	1.0000	0.1875	0.8125

The cumulative probability of having an event is the complement of joint probability of survival through every of five years of observation (Table 2, column 7) (12).

Case study 2: Estimation of cumulative risk using actuarial method

Simple cumulative method assumes no withdrawals during the period of observation. Since in our case (Figures 1 and 2) there were four individuals lost to observation, this must be considered. Their limited participation need to be considered in the denominator of the cumulative probability of an event. Actuarial method considers censored observations most roughly (Equation 3). Since we have at the end of the 5-year interval 16 individuals with a disease out of 20 persons at the beginning of the observation, and 4 persons were lost to follow up, we calculate cumulative 5-year risk directly as (Equation 13):

$$_{cum}R = \frac{16}{20 - \frac{4}{2}} = 0.8889 \qquad \textbf{Equation 13.}$$

Again, we can split 5-year interval first into five 1-year intervals and calculate first the annual risks and afterwards cumulative 5-year risk. For each 1-year interval we:

- define the number of persons entered in the interval (Table 3, column 1), number of persons with the disease at the end of interval (Table 3, column 2), and the number of withdrawals (Table 3, column 3),
- calculate the adjusted number of withdrawals (1,12),
- by using Equation 4 calculate annual actuarial risks (Table 3, column 6).

Table 3. Elements for calculation and calculation of annual risks using actuarial method.

Year of observation	1	2	3	4	5	6
	Entered in the interval (N)	With the disaese at the end of interval (d+)	Withdrawals (W)	W/2	N – (W/2)	d+/N–(W/2) (annual risk) ($_{ann}$R)
1st	20	6	0	0	20	0.3000
2nd	14	3	0	0	14	0.2143
3rd	11	6	1	0.5	10.5	0.5714
4th	4	1	2	1	3	0.3333
5th	1	0	1	0.5	0.5	0.0000

After annual risks are calculated we follow exactly the same priciples for calculation of 2-, 3-, 4- and 5-year cumulative risks as discussed in simple method. The results are presented in Table 4. Results of calculating the cumulative 5-year risk estimated by using the actuarial method (Table 4, column 9) show that its value is 0.8428, what is much higher than estimated by using the simple method.

Table 4. Calculation of cumulative 5-year risk from annual risks using actuarial method.

Year of observation	6	7	8	9
	d+/N – (W/2) (annual risk) ($_{ann}$R)	1 – $_{ann}$R	product (1 - $_{ann}$R) (II)	1 – II (cumulative risk) ($_{cum}$R)
1st	0.3000	0.7000	0.7000	0.3000
2nd	0.2143	0.7857	0.5500	0.4500
3rd	0.5714	0.4286	0.2357	0.7643
4th	0.3333	0.6667	0.1572	0.8428
5th	0.0000	1.0000	0.1572	0.8428

Case study 3: Estimation of cumulative risk using density method

The first method that consider exact times of being at risk of developing a disease under observation is density method.

Table 5. Data for calculation of person-years.

Id. number	Time of being at risk* (Years)	Status at the end of observation (1=with the disease, 0=cesored (cause of censoring))
2	0.25	1
6	0.25	1
3	0.50	1
18	0.50	1
1	0.75	1
9	0.75	1
12	1.25	1
10	1.50	1
15	1.75	1
4	2.25	1
8	2.25	1
13	2.25	1
20	2.25	1
19	2.25	0 – free of disease, change of domicile
11	2.50	1
17	2.50	1
16	3.25	1
7	3.50	0 – free of disease, voluntarily withdrawal
14	3.50	0 – free of disease, change of domicile
5	4.75	0 – free of disease, change of domicile
Total	38.75	Diseased = 16, Lost-to-follow-up = 4

* time in which an individual under observation is exposed to effect of noxious agent (is at risk of getting an event under observation)

In order to perform the procedure (Equation 6) we need first to calculate the person-years (PY) since we need this quantity in calculation of the incidence density. We use the information given in Figure 2. In Table 5 data for calculation of PY for the entire 5-year period are presented.

The incidence density for 5-year period could be now calculated using the Equation 7. The results are presented in following equation (Equation 14):

$$ID = \frac{16}{38.75} = 0.4129 \qquad \textbf{Equation 14.}$$

This quantity afterwards enters the equation for calculating the 5-year cumulative risk using the Equation 6. The results are presented in following equation (Equation 15):

$$_{cum}R = 1 - e^{(-0.4129 \times 5)} = 0.8731 \qquad \textbf{Equation 15.}$$

Again, we can split 5-year interval first into five 1-year intervals and calculate first the annual risk using the density method and afterwards cumulative 5-year risk. The steps are as follows

- first we summarize the events in each of 1-year intervals which are five as the duration of the longest observation is 4.75 let: entered in the interval (Table 6, column 1), with the disease at the end of interval (Table 6, column 2), lost to follow-up (Table 6, column 3), and present at the end of the period without a disease (Table 6, column 4),

Table 6. Summary of the events in each of 1-year intervals.

Year of observation	**1**	**2**	**3**	**4**
	Entered in the interval (N)	**With the disaese at the end of interval (d+)**	**Lost to follow-up**	**Present at the end of the period**
1st	20	6	0	14
2nd	14	3	0	11
3rd	11	6	1	4
4th	4	1	2	1
5th	1	0	1	0
Total		16	4	

- in following step we calculate the person-years (PY) for for each of 1-year periods (Table 7). We use the information given in Figure 2,

Table 7. The summary of calculation of person-years in each of 5 1-year intervals.

Year of observation	Contribution to person-years (PY) at the end of 1-year interval	PY Total
1st	(0.25 × 2) + (0.50 × 2) + (0.75 × 2) + (1.00 × 14)	17.00
2nd	(0.25 × 1) + (0.50 × 1) + (0.75 × 1) + (1.00 × 11)	12.50
3rd	(0.25 × 5) + (0.50 × 2) + (1.00 × 4)	6.25
4th	(0.25 × 1) + (0.50 × 2) + (1.00 × 1)	2.25
5th	(0.75 × 1)	0.75

- in following step the annual incidence density is calculated (Table 8). As the incidence density is not constant over 5-year period (the highest is in the third year of observation) this has to be considered in the calculation of cumulative risk,
- at the final step from incidence density the risk is calculated (Table 9).

Table 8. Calculation of incidence density in each of 5 1-year intervals.

Year of observation	2	5	6
	With the disaese at the end of interval (d+)	Annual person-years (PY)	Annual incidence density (d+/PY) ($_{ann}$ID)
1st	6	17.00	0.3529
2nd	3	12.50	0.2400
3rd	6	6.25	0.9600
4th	1	2.25	0.4444
5th	0	0.75	0.0000

Table 9. Calculation of the annual risk in each of 5 1-year intervals.

Year of observation	6	7	8
	Annual incidence density (d+/PY) ($_{ann}$ID)	$e^{(-IDann\times1)}$	$1-e^{(-annID\times1)}$ (annual risk) ($_{ann}$R)
1st	0.3529	0,.7027	0.2974
2nd	0.2400	0.7866	0.2134
3rd	0.9600	0.3829	0.6171
4th	0.4444	0.6412	0.3588
5th	0.0000	1.0000	0.0000

Results of calculating the cumulative 5-year risk estimated by using the density method (Figure 1) show that its value is 0.8643, what is much higher than estimated using the simple method, and also higher than estimated using the actuarial method. The elements for calculation and its results are presented in Table 10.

Table 10. Elements for calculation of cumulative risk using the density method.

	6	9	10	11
Year of observation	**Annual incidence density (d+/PY) ($_{ann}$ID)**	**$\sum(-_{ann}ID\times1)$**	**$e^{\sum(-annID\times1)}$**	**$1-e^{\sum(-annID\times1)}$ (cumulative risk) ($_{cum}$R)**
1st	0.3529	-0.3529	0.7026	0.2974
2nd	0.2400	-0.5929	0.5527	0.4473
3rd	0.9600	-1.5529	0.2116	0.7884
4th	0.4444	-1.9973	0.1357	0.8643
5th	0.0000	-1.9973	0.1357	0.8643

Case study 4: Estimation of cumulative risk using Kaplan Meier product limit method

This method also considers exact times of being at risk of developing a disease under observation (2,12). The intervals are defined as ending each time an event (i.e. disease, death, withdrawal) occurs. The procedure is as follows:

- first we determine the times when events or censoring occurred. We use the information given in Figure 2,
- define the number of persons entered in the interval (Table 11, column 1), number of persons with the event (occurrence of the disease or death) at time I (Table 11, column 2), and the number of censored cases (Table 11, column 3) at time i,

Table 11. Elements for calculation of cumulative risk by using the Kaplan Meier product limit method.

Time of the events/ censoring (years)	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	Entered in the interval (N)	**Occurrence of the event (d+)**	**Censored**	**d+/N (conditional probability of the event) (p)**	**1 – p (survival) (S)**	**Product (S) (Cumulative survival) ($_{cum}$S)**	**$1 - S_{cum}$ (cumulative conditional probability of an event) ($_{cum}$R)**
0.25	20	2	0	0.1000	0.9000	0.9000	0.1000
0.50	18	2	0	0.1111	0.8889	0.8000	0.2000
0.75	16	2	0	0.1250	0.8750	0.7000	0.3000
1.25	14	1	0	0.0714	0.9286	0.6500	0.3500
1.50	13	1	0	0.0769	0.9231	0.6000	0.4000
1.75	12	1	0	0.0833	0.9167	0.5500	0.4500
2.25	11	4	1	0.3636	0.6364	0.3500	0.6500
2.50	6	2	0	0.3333	0.6667	0.2333	0.7667
3.25	4	1	0	0.2500	0.7500	0.1750	0.8250
3.50	2	0	2	0.0000	1.0000	0.1750	0.8250
4.75	1	0	1	0.0000	1.0000	0.1750	0.8250

- by using Equation 11 calculate conditional probalities (Table 11, column 4),
- calculate the complement of conditional probalities of the event at every time of occurrence of the events or censoring – the conditional probability of survival without an event under observation up to the time i (Table 11, column 5),
- calculate cumulative probability of survival over more than one interval by multiplying the conditional survival probabilities over all intervals (Table 11, column 6),
- calculate the complement of cumulative probabilities of survival over more than one interval (Table 11, column 7).

Conclusion

In table 12 the summary over results of all four methods of estimation of cumulative risk is presented.

Table 12. Summary over results of estimating cumulative risk over 5-year period using four different methods of estimation.

Method	Direct 5-year cumulative risk	Indirect 5-year cumulative risk
Simple	0.8000	0.8125
Actuarial	0.8889	0.8429
Density	0.8731	0.8643
Kaplan Meier		0.8250

Since the most accurate measure is Kaplan Meier method we could compare all other results to this result. We could conclude that in this case study, the closest results to Kaplan Meier method are obtained by indirect simple method, and by actuarial indirect method, while the most far away were results obtained by direct actuarial method. One should be aware that this is not always so. The results depend on number of events and number of censored cases. When the events are rare and there is no censoring, the discrepancy tends to be smaller (12).

EXERCISE

Data set

In Figure 3, another imaginary data-set is presented. Again, a cohort of 20 individuals initially without a disease under observation, were followed up for 5 years.

Task 1

For the data set presented in Figure 3, calculate cumulative risk using simple method:

- directly,
- indirectly by calculating annual risks first.

Task 2

For the data set presented in Figure 3, calculate cumulative risk using actuarial method:

- directly,
- indirectly by calculating annual risks first.

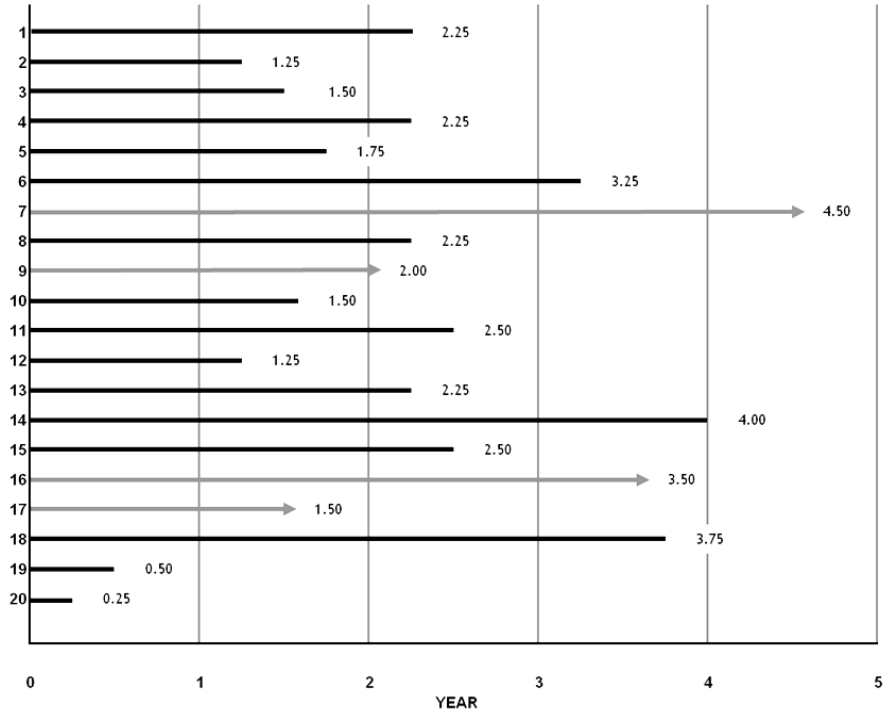


Figure 3. Graphic presentation of events in a cohort of 20 people. LEGEND: ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before an event occurred) in individuals that developed the disease under observation; ▬ the period of exposure to the effect of the noxious agent (being at risk of developing a disease under observation before censoring occurred) in individuals that were lost to follow-up (voluntarily withdrawal from the study or change of domicile).

Task 3

For the data set presented in Figure 3, calculate cumulative risk using density method:

- directly,
- indirectly by calculating annual risks first.

Task 4

For the data set presented in Figure 3, calculate cumulative risk using Kaplan Meier method.

REFERENCES

1. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
2. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
3. Dos Santos Silva I. Cancer epidemiology: principles and methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
5. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
6. Friis RH, Sellers TA. Epidemiology for public health practice. Fourth edition. Sudbury, MA: Jones&Bartlett Publishers; 2009.
7. Kue Youtng T. Population health: concepts and methods. Second edition. Oxford: Oxford University Press; 2005.
8. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
9. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
10. Kleinbaum DG, Klein M. Survival analysis. A Self-Learning Text. Second edition. New York: Springer; 2005.
11. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.
12. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
13. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.

RECOMMENDED READINGS

1. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
2. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.
3. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
4. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASURES OF ASSOCIATION AND POTENTIAL IMPACT**
Module: 1.3.1	**ECTS (suggested): 0.30**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Measures of association, measures of potential impact
Learning objectives	After completing this module students should: • know types of measures of association, and know how to calculate them; • know types of measures of potential impact, and know how to calculate them.
Abstract	Measures of association are quantities that express the strength of association between variables. Mathematically they are differences or ratios between different kinds of other measures (e.g. frequency measures). Measures of potential impact are quantities that express potential impact of one phenomenon on the frequency of the other. They are tightly interrelated. The module is describing basic features and types of them.
Teaching methods	An introductory lecture gives the students first insight in features and types of measures of association, and measures of potential impact. The theoretical knowledge is illustrated by two case studies. After introductory lectures students first carefully read the theoretical background of this module and complement their knowledge with recommended readings. Afterwards they on provided data set in pairs perform two extensive tasks on calculation of different types of measures. They are stimulated to compare results with results of other pairs and discuss the differences.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Written report on calculated measures in which detailed description of process of calculation is described.

MEASURES OF ASSOCIATION AND POTENTIAL IMPACT

Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Introduction to analysis of association and potential impact

Globally, measures of association, or measures of effect, are quantities that express the strength of association between variables (1-15). Mathematically they are differences or ratios between different kinds of other measures (e.g. means, or frequency measures). Measures of potential impact are quantities that express potential impact of one phenomenon on the frequency of another (2,7,8,11). On one hand they could express potential impact of risk factor on occurrence of observed health phenomenon (unfavourable or favourable) in population or among exposed persons. These measures are common in public health. On the other hand they can express potential impact of an intervention on disease occurrence (benefit or harm) (16,17). These measures are much more commonly used in clinical epidemiology than in public health.

Before introducing some important measures of association and measures of potential impact, we need to give, again, a warning about terminology. As we have already emphasized on several places, common problem in epidemiology is existence of multiple terms for the same concept. Also, there are instances where a single term is applied to different concepts. The confusion is aggravated by the multitude of terms that have been introduced, with usages that differ from one author to another (2).

Some important concepts in analysis of association

To make all considerations about analysis of association between phenomena under observation easier some of concepts in relationship analysis should be clarified:

1. Observed phenomenon.
   Observed phenomenon is a disease or other health-related condition under observation, frequently called also an outcome. In analysis of association is assigned the role of "effect".
2. Risk factor.
   Risk factor is defined as a phenomenon or characteristic (an aspect of behaviour or lifestyle, an environmental exposure, an inborn or inherited characteristic) which is on the basis of epidemiologic evidence known to be associated with health-related condition(s) considered important to prevent (1). Frequently is called also an exposure. In analysis of association is assigned the role of "cause".
3. Dependency.
   In epidemiology, the dependent variable is the manifestation or outcome whose variation is to be explained by risk factors (1).

   As the relationship is not analyzed only by the means of epidemiologic methods but statistical as well, it should be clarified that in statistics the dependent variable is the one predicted by a regression equation.

4. Causality.
   By the definition, causality is relating of causes to the effects they produce (1).
   Mostly epidemiology is concerned of causality. In analysis of association the phenomenon under observation is usually assigned the role of "effect" and the risk factor thereof the role of »cause«. Actually causality of the relationship is not as simple. It should be clearly stated that epidemiologic evidence of causality by itself is insufficient to establish the causality in nature. The later is usually possible to be proved only by studies in different places which could be carried out for more decades. However, for the analysis of association to be understandable, such setting up is very adequate. A picture (Figure 1) can offer additional help.

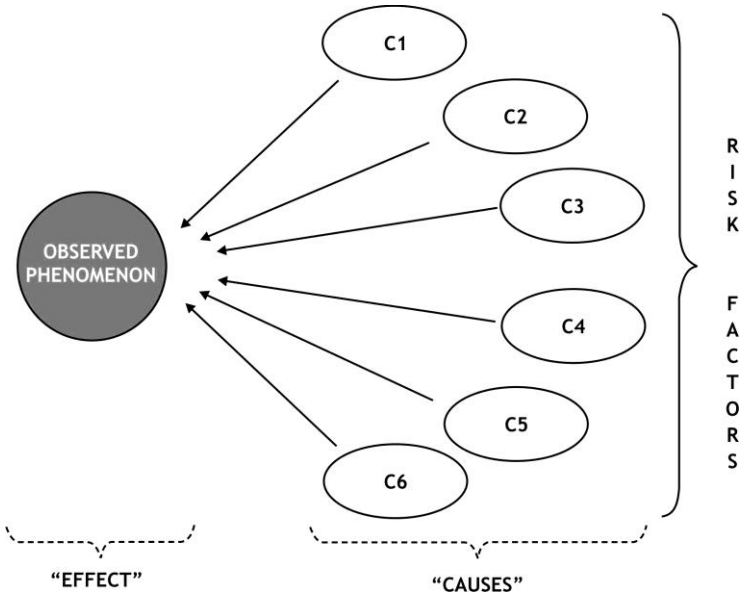


Figure 1. Graphical representation of the setting up of the observed disease or other health related phenomenon (observed phenomenon or "effect") and risk factors ("causes") into a relationship, as a help in analysis of association.

Some important considerations in analysis of association

Epidemiologic versus statistical measures

Distinguishing and separating between epidemiologic and statistical measures is rather difficult as they are interlaced, what is also discussed in one of chapters of this book. In fact, to some extent, statistical measures of association, used in epidemiology could be considered as epidemiologic and vice versa. Maybe the most appropriate attempt to distinguish between these two terms is as follows: when we are using the term »epidemiologic measure« this measure is applied to health phenomena of different kind, while »statistical measure« is more general term. »Epidemiologic

measures« of association are also characterized by that they are based only upon frequency measures (e.g. relative risk, or odds ratio), while »statistical measures« are based upon frequency measures (e.g. chi-square test statistics), or mean and variance (e.g. test statistics in analysis of variance, regression and correlation coefficients). Only »epidemiologic measures« will be discussed here.

Simple versus complex analysis of association

In analysing the relationship between observed outcome and risk factors two kinds of methods could be used:

1. Univariate methods.
   Univariate methods of analysis of association are concerned about the analysis of relationship of a single risk factor with a single outcome.
2. Complex or multivariate methods.
   Multivariate methods are concerned about the analysis of relationship of several risk factors usually with a single outcome at a time. The methods for applying such kind of analysis are usually classified in statistical methods, although it is very difficult to draw strict line of separation between statistical and epidemiologic methods. In fact they all strive to assess the relationship between different phenomena.

   As outcomes are associated not only with one risk factor, it would be most appropriate to think about using multivariate methods as an analytical tool starting from the designing phase of a study. In fact in epidemiology multivariate methods are very useful in controlling the effect of confounders, in the phase of data analysis, since the methods of controlling them in the designing phase (randomization, matching) are not applicable in several types of epidemiologic studies (e.g. cross-sectional studies).

Interrelation between measures of association and measures of potential impact

Both, epidemiologic measures of association and measures of potential impact are based upon measures of frequency (e.g. incidence measures). They are tightly interrelated. As we will see later on, there is a case that the same measure is at the same time a measure of association, and a measure of potential impact. Only the interpretation is different. Additionally, measures of potential impact could be based upon measures of association as well.

Like in measures of frequency, also these two families of measures could be classified according to various characteristics, what will be discussed later. The process of explanation of measures of both groups will be based upon example data in case studies after theoretical background.

Measures of association

Epidemiologic measures of association are quantities that express the strength of association between phenomena, related to health (1-15). Mathematically they are differences or ratios between different kinds of frequency measures.

Measures of association could be classified according to various characteristics. Four different classifications are summarized in Figure 2.

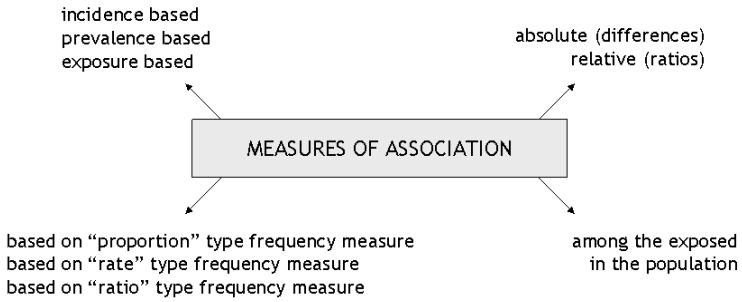


Figure 2. Classifications of measures of association according to various characteristics.

1. Classification of measures of association to incidence based, prevalence based, and exposure based.
   According to that if the measure is incidence or prevalence based, the measures of association are classified as incidence based and prevalence based. All measures have been already introduced, so in this place they are only listed by this criterion.
2. Classification of measures of association to absolute and relative.
   Absolute measures are differences between frequency measures of different kinds between two observed groups (e.g. between exposed and unexposed), while relative measures are ratios between frequency measures of different kinds between two observed groups (e.g. between exposed and unexposed).
3. Classification of measures of association according to the type of relative frequency measure upon which are based.
   Measures of association could be based on different types of frequency measures that were in details presented in one of previous modules in this book. According to this classification could be based on proportion (measure in which numerator is included in denominator in basic frequency measure), rate or ratio in a narrow sense frequency measures (measures in which numerator is not included in denominator in basic frequency measure) (4).
4. Classification of absolute measures of association to those related to exposed and those related to population.
   Absolute measures of association could be classified to those related to exposed, and those related to population. This classification is not so important in measures of association than in measures of potential impact. Consecutively, we will not pay attention to this classification at the moment.

Detailed interrelation between different measures of association are out of the scope of this module and are discussed in several textbooks of modern epidemiology (2,4,8,18).

In continuation, measures of association frequently used in epidemiologic research will be presented (Figure 3).

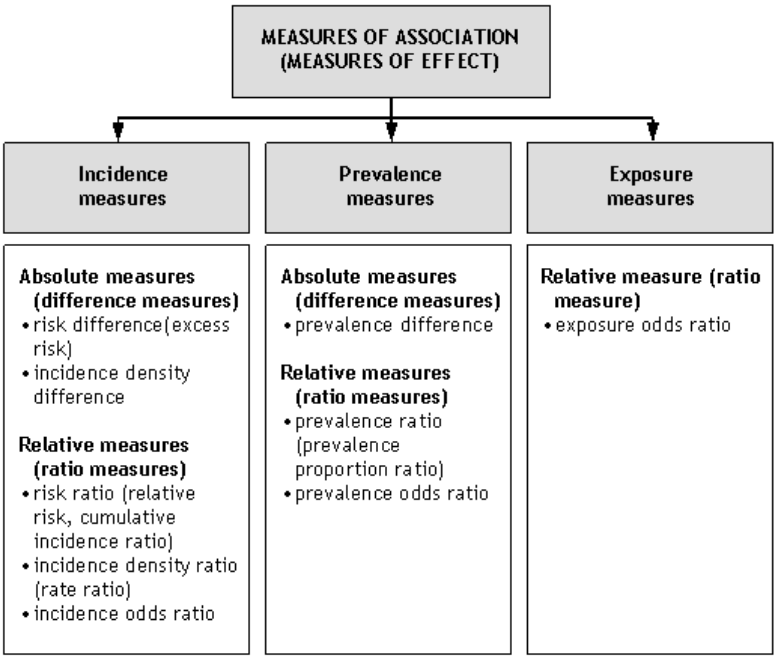


Figure 3. Measures of association frequently used in epidemiologic research.

Incidence comparisons

In incidence (cohort) studies, incidence measures are generated, being absolute (difference based) or relative (ratio based).

Difference based measures

Two difference based measures are commonly used in incidence comparisons: risk difference and incidence density difference.

1. Risk difference or excess risk.
   This measure is the absolute difference between two risks (1,5,12). When the risk difference is observed between exposed and unexposed (excess risk among exposed), the measure is calculated as follows (1,5,12) (Equation 1):

$$RD = R_{E+} - R_{E-} \qquad \textbf{Equation 1.}$$

RD = risk difference among exposed
R_{E+} = risk among exposed
R_{E-} = risk among unexposed

The risk difference can be estimated directly from the cumulative incidence difference (2,4) (Equation 2):

$$CID = CI_{E+} - CI_{E-}$$ **Equation 2.**

CID = cumulative incidence difference among exposed
CI_{E+} = cumulative incidence among exposed
CI_{E-} = cumulative incidence among unexposed

Calculation of this measure in practice is presented in Case study 1.

2. Incidence density difference (2,4,12).
   It is absolute difference between two incidence densities (1). The measure is calculated as follows (Equation 3):

$$IDD = ID_{E+} - ID_{E-}$$ **Equation 3.**

IDD = incidence density difference
ID_{E+} = incidence density among exposed
ID_{E-} = incidence density among unexposed

Calculation of this measure in practice is presented in Case study 1.

Ratio based measures

Following ratio based measures are commonly used in incidence comparisons: relative risk or risk ratio, incidence density ratio, and risk odds ratio.

1. Risk ratio or relative risk.
   Relative risk is a ratio of risk of observed outcome (e.g. disease, death) among the exposed to the risk among the unexposed (2,4,5,7,8,12). The measure is calculated as follows (Equation 4):

$$RR = \frac{R_{E+}}{R_{E-}}$$ **Equation 4.**

RR = relative risk
R_{E+} = risk among exposed
R_{E-} = risk among unexposed

Calculation of this measure is presented in Case study 1.

2. Incidence density ratio.
   Incidence density ratio or rate ratio is the ratio of two incidence densities (2,4,7,12). The measure is calculated as follows (Equation 5):

$$IDR = \frac{ID_{E+}}{ID_{E-}}$$ **Equation 5.**

IDR = incidence density ratio
ID_{E+} = incidence density among exposed
ID_{E-} = incidence density among unexposed

Calculation of this measure is presented in Case study 1.

In this place we need to introduce a new ration effect measure – a measure known under the term »hazard ratio« (2,19,20). This measure is tightly related to the incidence density ratio (19). It is very important measure since it is the measure of effect in survival analysis (19,20), which is closely related to epidemiologic concept of risk, what is in details presented in a separate module in this book. To understand relationship between incidence density ratio and hazard ratio, we need to recall the measure known as »hazard rate«. This measure, that measures the instantaneous potential for change in disease status (4,20,21), was in details presented in a separate module in this book. Incidence density is an average rate for estimating average of instantaneous incidence rates. Under certain conditions we could use the terms »incidence density«, and »hazard rate« as synonyms (19). The incidence density ratio, just presented, is the ratio of two incidence densities (two incidence rates). Conceptually this ratio is identical to a hazard ratio usually denoted as HR (19). Theoretically, the hazard ratio at a given point in time is the limiting value of the incidence density ratio as the time around the pint becomes very short, just as the hazard is the limiting quantity for incidence density (19).

3. Incidence odds ratio.

Incidence or risk odds ratio is a ratio of risk odds of observed outcome (e.g. disease, death) among the exposed to the odds among the unexposed (2,4,5,7,8,12). The measure is calculated as follows (Equation 6):

$$ROR = \frac{{}_{risk}O_{E+}}{{}_{risk}O_{E-}}$$ **Equation 6.**

ROR= risk odds ratio
${}_{risk}O_{E+}$ = risk odds among exposed
${}_{risk}O_{E-}$ = risk odds among unexposed

Calculation of this measure is presented in Case study 1.

Odds ratio is one of extremely useful measures in multivariate analysis as well. It is one of possible results of logistic regression method (22,23).

In this place we need to expose the relationship between relative risk and odds ratio. In medical literature odds ratio is often misinterpreted as estimate of relative risk. Odds ratio is good estimate of relative risk only under specific condition - only when the phenomenon under observation is rare. In this case only, denominators in calculation process of frequency measures in odds ratio (e.g. number

of people under observation without observed phenomenon among exposed) and in risk ratio (e.g. total number of exposed), are similar. Consecutively values of both measures, the odds ratio and relative risk, are very close to each other.

The ratio measures ranges in value from 0 to infinity. Values close to 1.0 indicate no relationship between the exposure and the outcome. Values less than 1.0 suggest a protective effect, while values greater than 1.0 suggest an adverse effect of exposure. When comparing all three ratio measures it can be shown that numerically, the odds ratio falls the furthest from the null, and the risk ratio the closest, with the rate ratio (incidence density ratio) in between (19).

Prevalence comparisons

In prevalence (cross-sectional) studies, prevalence measures are generated, being, similarly as in incidence studies, one absolute (difference based) or relative (ratio based).

Difference based measure

Only one difference based measure is commonly used in prevalence comparisons.

1. Prevalence proportion difference.
   Prevalence proportion difference, also known as prevalence rate difference (2,4,19), is calculated as follows (4) (Equation 7):

$$PD = P_{E+} - P_{E-} \qquad \textbf{Equation 7.}$$

PD = prevalence difference among exposed
P_{E+} = prevalence proportion among exposed
P_{E-} = prevalence proportion among unexposed

Calculation of this measure is presented in Case study 1.

Ratio based measures

Following two ratio based measures are commonly used in prevalence comparisons: prevalence ratio, and prevalence odds ratio.

1. Prevalence ratio.
   Prevalence ratio is a ratio of point prevalence proportion of observed outcome among the exposed to the point prevalence proportion among the unexposed (4,8,12,15). The measure is calculated as follows (Equation 8):

$$PR = \frac{P_{E+}}{P_{E-}} \qquad \textbf{Equation 8.}$$

PR = prevalence ratio
P_{E+} = prevalence (proportion) among exposed
P_{E-} = prevalence (proportion) among unexposed

Calculation of this measure is presented in Case study 1.

2. Prevalence odds ratio.
   Prevalence odds ratio is a ratio of prevalence odds of observed outcome among the exposed to the prevalence odds among the unexposed (2,4,8,12,15). The measure is calculated as follows (Equation 9):

$$POR = \frac{{}_{prevalence}O_{E+}}{{}_{prevalence}O_{E-}} \qquad \textbf{Equation 9.}$$

POR= prevalence odds ratio
${}_{prevalence}O_{E+}$ = prevalence odds among exposed
${}_{prevalence}O_{E-}$ = prevalence odds among unexposed

Calculation of this measure is presented in Case study 1.

Exposure comparisons

In case-control studies neither incidence nor prevalence measures could be generated. The only measures that can be generated are exposure related measures. The exposure odds ratio is the ratio of the odds in favour of exposure among the cases to the odds in favour of exposure among non-cases (1,2,4,8). The measure is calculated as follows (Equation 10):

$$EOR = \frac{{}_{exposure}O_{cases}}{{}_{exposure}O_{non-cases}} \qquad \textbf{Equation 10.}$$

EOR= exposure odds ratio
${}_{exposure}O_{cases}$ = exposure odds among cases
${}_{exposure}O_{non\text{-}cases}$= exposure odds among non-cases

Calculation of this measure is presented in Case study 1.

Major advance is that exposure odds ratio (EOR) and disease odds ratio (DOR) in case-control studies are mathematically equivalent. Consecutively, the exposure odds ratio can be used to estimate the relative risk, especially when the probability of positive response is small (2,4,8). The equivalence is presented in Case study 1.

Measures of potential impact

This family of epidemiologic measures quantifies potential impact of various exposures on observed phenomena. They could express potential impact of risk factor on occurrence of observed health phenomenon or potential impact of an intervention on disease occurrence change (1,2,4,5,13,16,17,24,25)

Measures of potential impact could be classified according to various characteristics. Five different classifications are summarized in Figure 4.

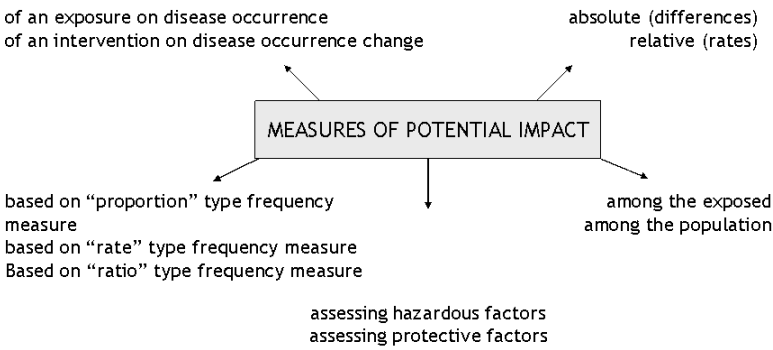


Figure 4. Classifications of measures of potential impact according to various characteristics.

1. Classification of measures of potential impact to those assessing impact of an exposure on disease occurrence, and those assessing impact of an intervention on disease occurrence change
   In this group of measures there are two subgroups according to what they are measuring. Measures of the first subgroup express potential impact of risk factor on occurrence of observed health phenomenon among exposed persons or in population. These measures are common in public health. Measures of the second subgroup express potential impact of an intervention on disease occurrence reduction. They are much more common in clinical epidemiology than in public health.
2. Classification of measures of potential impact to absolute and relative
   According to that if the measure is expressed as difference or ratio the measures of potential impact are classified as absolute (differences) and relative (ratios).
3. Classification of measures of potential impact to measures referring to the population or to exposed
   Measures of potential impact could be classified also to those related to exposed, and those related to population.
4. Classification of measures of potential impact to measures of potential impact according to the type of relative frequency measure upon which are based.
   Measures of potential impact could be based on different types of frequency measures that were in details presented in one of previous modules in this book. According to this classification could be based on proportion (measure in which numerator is included in denominator in basic frequency measure), rate or ratio in a narrow sense frequency measures (measures in which numerator is not included in denominator in basic frequency measure) (4).

5. Classification of measures of potential impact to measures assessing hazardous and measures assessing protective factors.
   Measures dealing with hazardous factors are known under the common term »etiologic fraction« (4). The other group of measures, those measures that deal with protective risk factors, is known under common term »prevented fraction« (4). The difference is that they express the proportion of potential new cases of the disease under observation that would have occurred in absence of exposure to a protective factor but did not occur (4), or the proportion of the hypothetical total load of disease that has been prevented by exposure to a protective factor (1).

Detailed interrelation between different measures of potential impact are out of the scope of this module and are discussed in several textbooks of modern epidemiology (2,4,8,18).

In continuation, measures of association frequently used in epidemiologic research will be presented (Figure 5).

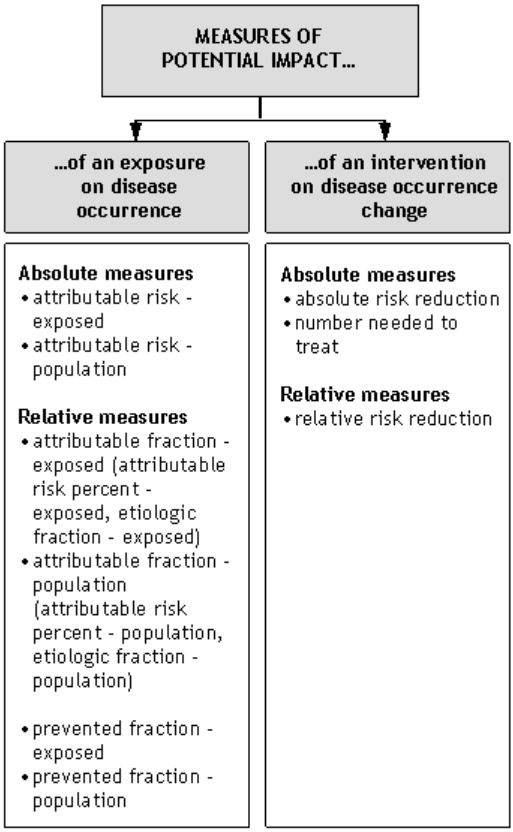


Figure 5. Measures of potential impact frequently used in epidemiologic research.

Measures assessing impact of an exposure on disease occurrence

Difference based measures

Difference based measures of potential impact are attributable risk in exposed and attributable risk in population.

1. Attributable risk in exposed.
   This measure is the absolute difference between two risks (1,2,5,8). The measure is calculated as follows (Equation 11):

$$AR = R_{E+} - R_{E-} \qquad \textbf{Equation 11.}$$

AR = attributable risk among exposed
R_{E+} = risk among exposed
R_{E-} = risk among unexposed

   In fact this measure has been already introduced among measures of association under the name »risk difference«. It is the portion of the risk in the exposed that is due to the exposure (could be attributed to exposure). In other terms, it is the risk of a disease in the exposed that could be eliminated if exposure were eliminated.

   Calculation of this measure, which is technically exactly the same as calculation of risk difference from in Case study 1, but the interpretation is different, is presented in Case study 2.

   In calculation process it is assumed that other risk factors than the one under observation have equal effects on the exposed and unexposed.

2. Attributable risk in population.
   Attributable risk could be expressed also for the population (population attributable risk or population attributable risk). It is a measure of the amount of disease attributed to a putative cause of the disease in the population (1,2,5,8). Mathematically it is the difference between the risk of disease in the entire population and among unexposed (Equation 12):

$$PAR = R_{pop} - R_{E-} \qquad \textbf{Equation 12.}$$

PAR = population attributable risk
R_{pop} = risk in population
R_{E-} = risk among unexposed

   Population attributable risk is the portion of the risk of a disease in population (exposed and non-exposed) that is due to exposure. It is the risk of a disease in the population that could be eliminated if exposure were eliminated.

   Calculation of this measure is presented Case study 2.

Ratio based measures – assessment of hazardous factors

Ratio based measures of potential impact of hazardous factors are attributable fraction in exposed and attributable fraction in population.

1. Attributable fraction in exposed.
   It is a fraction of people with the disease under observation which could be attributed to exposure to a risk factor under observation (1,2,5,7,8). It is known also under the terms attributable risk proportion (2), and etiologic fraction (4). The measure is calculated as follows (Equation 13):

$$AF = \frac{R_{E+} - R_{E-}}{R_{E+}} \qquad \textbf{Equation 13.}$$

AF = attributable fraction among exposed
R_{E+} = risk among exposed
R_{E-} = risk among unexposed

In calculation process it is assumed that other risk factors than the one under observation have equal effects on the exposed and unexposed.

When this measure is multiplied by a multiplier 100 it is known as attributable risk percent (5).

Attributable fraction or attributable risk percent is the fraction/percent of the risk of a disease in the exposed that is due to the exposure. It is the fraction/percent of the risk of a disease in the exposed that could be eliminated if exposure were eliminated.

The same measure could be calculated using another procedure (1) (Equation 14):

$$AF = \frac{RR - 1}{RR} \qquad \textbf{Equation 14.}$$

AF = attributable fraction among exposed
RR = relative risk

Calculation of this measure is presented Case study 2 as well.

2. Attributable fraction in population.
   When attributable fraction refers to the population (attributable fraction – population or population attributable fraction), the measure is calculated as follows (1) (Equation 15):

$$PAF = \frac{R_{pop} - R_{E-}}{R_{pop}} \qquad \textbf{Equation 15.}$$

PAF = attributable fraction in population
R_{pop} = risk in population
R_{E-} = risk among unexposed

When this measure is multiplied by a multiplier 100 it is known as population attributable risk percent (5).

Population attributable fraction or percent is the fraction/percent of the risk of a disease in the population (exposed and non-exposed) that is due to exposure. It is the fraction/percent of the risk of a disease in the population that would be eliminated if exposure were eliminated.

The same measure could be calculated using another procedure (Equation 16):

$$PAF = \frac{P_{E+} \times (RR - 1)}{1 + [P_{E+} \times (RR - 1)]}$$ **Equation 16.**

PAF = attributable fraction in population
RR = relative risk
P_{E+} = proportion of exposed in population

Calculation of this measure is presented Case study 2.

Ratio based measures – assessment of protective factors

When considering a protective exposure an appropriate alternative measures are prevented fraction in exposed and prevented fraction in population (1,19,26).

1. Prevented fraction in exposed.
   The reference point in this measure is the risk of disease if nobody is exposed (26). The fraction indicates the amount of disease that would be prevented by exposure to a protective factor (26). The measure is calculated as follows (Equation 17):

$$PF = \frac{R_{E-} - R_{E+}}{R_{E-}}$$ **Equation 17.**

PF = prevented fraction among exposed
R_{E-} = risk among unexposed
R_{E+} = risk among exposed

In calculation process it is assumed that other risk factors than the one under observation have equal effects on the exposed and unexposed.

Calculation of this measure is presented Case study 2.

The same measure could be calculated using another procedure (26) (Equation 18):

$$PF = 1 - RR$$ **Equation 18.**

PF = prevented fraction among exposed
RR = relative risk

Calculation using this procedure is presented Case study 2 as well.

This measure is identical to the measure, known in clinical epidemiology as »relative risk reduction« (26). It will be discussed later in this module. It is also identical to the measure, known in communicable diseases epidemiology as »vaccine efficacy« (26).

2. Prevented fraction in population.
   When the prevented fraction refers to the population (prevented fraction - population), the measure is calculated as follows (26) (Equation 19):

$$PPF = \frac{R_{E-} - R_{pop}}{R_{E-}}$$ **Equation 19.**

PPF = prevented fraction in population
R_{E-} = risk among unexposed
R_{pop} = risk in population

Calculation of this measure is presented Case study 2.

The same measure could be calculated using another procedure (26) (Equation 20):

$$PPF = P_{E+} \times (1 - RR)$$ **Equation 20.**

PPF = prevented fraction in population
RR = relative risk
P_{E+} = proportion of exposed

Calculation using this procedure is presented Case study 2 as well.

Beside these two measures, there exists also a measure called »preventable fraction«. This measure is less commonly used and will not be discussed here.

Measures assessing impact of an intervention on disease occurrence change

As already mentioned, this group of measures is much more commonly used in clinical epidemiology than in public health since they are measuring the effectiveness of clinical intervention trials. Nevertheless they could be effectively used in assessment of public health interventions as well.

There exist several subgroups of these measures, depending on if they are assessing undesirable (bad outcomes) or desirable events (good outcomes) (16,17). In this module we are presenting only three measures used most frequently: absolute risk reduction, relative risk reduction, and number needed to treat.

1. Absolute risk reduction.

Absolute risk reduction is defined as the difference in risk between the control (untreated) group and the intervention (treated) group (25). The measure is calculated as follows (Equation 21):

$$ARR = R_C - R_T \qquad \textbf{Equation 21.}$$

ARR = absolute risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

This measure expresses the proportion of individuals spared from the unfavourable outcome if they receive the intervention in comparison to not receive it (25).

In fact this measure is opposite to attributable risk in exposed (Equation 11). This calculation is opposite since clinical interventions on general reduce risk (26).

Calculation of this measure is presented Case study 2.

In this place we need to emphasize that in clinical epidemiology frequently another notation of basic frequency measures is used. Risk in the control group is denoted by CER (the control group event rate), while risk in the intervention (treated) group is denoted by EER (experimental group event rate). This notation is suitable since the studies conducted in clinical epidemiology are mostly clinical trials (experimental studies with the experimental and the control group).

2. Relative risk reduction.

This measure is the most commonly reported measure in this family of measures (25). It can be obtained by dividing the absolute risk reduction by the risk in the control group (25) (Equation 22):

$$RRR = \frac{R_C - R_T}{R_C} \qquad \textbf{Equation 22.}$$

RRR = relative risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

In calculation process it is assumed that other risk factors than the one under observation have equal effects on the exposed and unexposed.

Relative risk reduction measures how much of the risk is reduced in the experimental (treated) group compared to a control group. Here would be worthy to mention, that treatments with very large relative risk reductions may have a small effect in conditions where the control group has a very low bad outcome. On the other hand, modest relative risk reduction can mean major clinical importance if the risk in a control group is large.

Calculation of this measure is presented Case study 2.

The same measure could be obtained easily from the relative risk (the ratio of risk in the intervention group to the risk in the control group) using the following equation (25) (Equation 23):

$$RRR = 1 - RR = 1 - \frac{R_T}{R_C}$$ **Equation 23.**

RRR = relative risk reduction
RR = relative risk
R_C = risk in the control group
R_T = risk in the intervention (treated) group

Calculation using this procedure is presented Case study 2 as well.

This measure expresses how much the risk is reduced in the treated group compared to a control group. The greater the relative risk reduction, the more efficacious is the intervention (25).

3. Number needed to treat.

The last measure in this group we are presenting is the number needed to treat. It is in fact only another way to express the absolute risk reduction, since it is defined as inverse of the absolute risk reduction (25) (Equation 24):

$$NNT = \frac{1}{ARR} = \frac{1}{R_C - R_T}$$ **Equation 24.**

NNT = number needed to treat
ARR = absolute risk reduction
R_C = risk in the control group
R_T = risk in the intervention (treated) group

Calculation of this measure is presented Case study 2.

This measure is very popular because of its simplicity to compute and its ease to interpret – it is interpreted as the number of patients that would need to be treated to prevent one additional bad outcome.

CASE STUDIES

Case study 1: Measures of association

Incidence comparisons

An incidence study of impact of hyperirritable uterus on a premature delivery

In a ambidirectional cohort study, basing on Perinatal Informational System of Slovenia (PISS)[6] (27), data of 800 mothers and their newborns were analyzed.

[6] In Slovenia for already several years for the purpose of teaching epidemiologic methods in public health, comprising also statistical methods, data collection which enables learning such

The observed outcome was a preterm delivery (a delivery prior the completed 36th week of pregnancy). We related this outcome to a condition (»exposure«), known as hyperirritable uterus (abnormal contractility of the uterus during earlier stages of pregnancy). For analyzing this relationship we used different methods:

1. First we were interested only if preterm delivery occurred or not. The results of this simple cumulative method analysis are presented in Table 1.

Table 1. Frequency of preterm delivery in two groups of mothers according to absence or presence of hyperirritable uterus during pregnancy based on Slovene PISS data (27).

	Presence of hyperirritable uterus during pregnancy		**Total**
Preterm delivery	**No**	**Yes**	
No	719	30	749
Yes	44	7	51
Total	763	37	800

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 1 – are presented in Table 2.

methods in much comprehensive way has been created. These are the data collected in the frame of the Perinatal Informational System of Slovenia (PISS) (27), which is considered to be one of the permanent of health data-bases of the highest quality with many years' tradition in the country. It was started in 1987, when collection of perinatal data started according to a common protocol in all fourteen Slovene maternity hospitals.

In PISS the data on total period of pregnancy and on delivery are joined. Data concerning the pregnancy period are collected at the time of admission to the maternity. They are objective and subjective. Objective data come from a so-called »Maternal Booklet«. This is a booklet every pregnant woman in Slovenia gets at the time of establishing pregnancy. It holds the most important up-to-date information regarding pregnancy, determined by the obstetrician, responsible for a pregnant woman. This information is important in case of an urgent treatment or intervention and at the time of delivery. Subjective data come from a questionnaire filled-in at the time of admission to the maternity. They are collected by the maternity staff (partially by a midwife and partially by an obstetrician). After admission, a period from the admission to the discharge from a maternity is prospectively followed-up and data collected in PISS.

The basic data material for all epidemiologic and statistical activities is composed of 6,356 statistical units, representing the model of a population. For teaching different epidemiologic and statistical methods, samples of various sizes are then randomly selected from the population database. The data set used in this example is composed of 800 observational units.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal data have been removed, and moreover, the data are selected from the whole collection which shall be used only for the teaching purpose.

Table 2. The intermediate measures (frequency measures) - necessary elements for calculation of incidence risk and odds based measures of association calculated on the basis of data from Table 1.

Frequency measure	Notation	Calculation and value
Risk in the non-exposed	R_{E-}	$R_{E-} = \frac{44}{763} = 0.0577$
Risk in the exposed	R_{E+}	$R_{E+} = \frac{7}{37} = 0.1892$
Risk (incidence) odds in the non-exposed	${}_{risk}O_{E-}$	${}_{risk}O_{E-} = \frac{44}{719} = 0.0612$
Risk (incidence) odds in the exposed	${}_{risk}O_{E+}$	${}_{risk}O_{E+} = \frac{7}{30} = 0.2333$

Table 3. Duration of pregnancy and person-weeks quantity in two groups of mothers according to absence or presence of hyperirritable uterus during pregnancy based on Slovene PISS data (27).

Hyperirritable uterus during pregnancy no			Hyperirritable uterus during pregnancy yes		
Duration of pregnancy (weeks of gestation)	**n**	**Person-weeks (PW)**	**Duration of pregnancy (weeks of gestation)**	**n**	**Person-weeks (PW)**
22	1	22	28	1	28
23	1	23	29	2	58
25	1	25	32	1	32
26	1	26	33	1	33
29	2	58	34	1	34
30	2	60	35	1	35
31	2	62	37	5	185
32	7	224	38	3	114
33	2	66	39	9	351
34	5	170	40	11	440
35	12	420	41	2	82
36	8	288	Total	37	1392
37	33	1221			
38	97	3686			
39	182	7098			
40	295	11800			
41	108	4428			
42	4	168			
Total	763	29845			

2. Afterwards we were interested also in when preterm delivery occurred. We used the data on duration of pregnancy (assessed through gestational age of a newborn). These more detailed data are presented in Table 3.

   The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 3 – are presented in Table 4.

Table 4. The intermediate measures (frequency measures) - necessary elements for calculation of incidence density based measures of association calculated on the basis of data presented in Tables 1 and 3.

Frequency measure	Notation	Calculation and value
Incidence density in the non-exposed	ID_{E-}	$ID_{E-} = \frac{44}{29{,}845} = 0.00147$
Incidence density in the exposed	ID_{E+}	$ID_{E+} = \frac{7}{1{,}392} = 0.00503$

Calculation of measures of association in incidence comparisons

The procedures of calculation of risk difference, incidence density difference, relative risk, incidence density ratio (also known as rate ratio), and incidence odds ratio (also known as risk odds ratio), and interpretation of results is as follows:

1. Risk difference or excess risk.
   The risk difference calculated according to Equation 1 based on data presented in Tables 1 and 2 is (Equation 25):

$$RD = 0.1892 - 0.0577 = 0.1315 \qquad \textbf{Equation 25.}$$

   Mothers with hyperirritability of uterus have risk of preterm delivery for 0.1315 higher than mothers without it.

   We could interpret this result also in terms of a percentage. If we multiply the risk difference by 100, we get 13.15%, and the interpretation is that mothers with hyperirritability of uterus have risk of preterm delivery for 13.1% higher than mothers without it.
2. Incidence density difference.
   The incidence density difference calculated according to Equation 3 based on data presented in Tables 3 and 4 is (Equation 26):

$$IDD = 0.00503 - 0.00147 = 0.00355 \qquad \textbf{Equation 26.}$$

Mothers with hyperirritability of uterus have incidence density of preterm delivery for 0.00355 higher than mothers without it.

If we than multiply the incidence density difference by 1,000 we get the value 3.55, which could be interpreted as: on average in 42-week interval 3.55 more pregnant women with hyperirritability of the uterus experience preterm delivery per 1,000 population than mothers without it.

3. Risk ratio or relative risk.
   The relative risk calculated according to Equation 4 based on data presented in Tables 1 and 2 is (Equation 27):

$$RR = \frac{0.1892}{0.0577} = 3.2807 \qquad \textbf{Equation 27.}$$

   The risk of preterm delivery in mothers with hyperirritability of uterus is 3.28-times higher than in mothers without it.
4. Incidence density ratio.
   The incidence density ratio calculated according to Equation 5 based on data presented in Tables 3 and 4 is (Equation 28):

$$IDR = \frac{0.00503}{0.00147} = 3.4110 \qquad \textbf{Equation 28.}$$

   The »risk« of preterm delivery in mothers with hyperirritability of uterus, assessed through incidence density procedure, is 3.41-times higher than in mothers without it.
5. Incidence or risk odds ratio.
   The risk odds ratio calculated according to Equation 6 based on data presented in Tables 1 and 2 is (Equation 29):

$$ROR = \frac{0.2333}{0.0612} = 3.8129 \qquad \textbf{Equation 29.}$$

   The odds of preterm delivery in mothers with hyperirritability of uterus, is 3.81-times higher than in mothers without it.

When comparing all three presented ratio measures for the same set of data we can see that the odds ratio falls the furthest from the null (ROR = 3.81), and the relative risk the closest (RR = 3.28), with the incidence density ratio in between (IDR = 3.41).

Prevalence comparisons

A cross-sectional study of smoking in adults in Slovenia

In a cross-sectional study, basing on CINDI Health Monitor survey 2001, that was aiming at assessing the prevalence of health behaviours (28)[7], data of 9,034 adults were analyzed.

The observed outcome was smoking at the time (point of observation) of the survey. We related this outcome to a gender as a risk factor for unfavourable health behaviour (i.e. in this context in a role of »exposure«; since in general males are at highest risk for smoking than females, they are considered as »exposed«, while females are considered as »non-exposed«).

Out of 9,034 respondents 8,904 adults reported their smoking status. This outcome was related to a gender. The results of this analysis are presented in Table 5.

Table 5. Prevalence of smoking in both gender groups in adult population in Slovenia in 2001, based on CINDI Health Monitor Survey, Slovenia 2001 (28).

	Gender		Total
Smoking	**Females**	**Males**	
No	3,859	2,931	6,790
Yes	971	1,143	2,114
Total	4,830	4,074	8,904

[7] In Slovenia in 2001 the survey aiming at assessing the prevalence of health behaviours (stress perception, smoking habits, nutrition habits, alcohol consumption habits, physical activity habits, and traffic safety habits) was performed for the first time. This survey is conceptually a part of a wider international project in the frame of the Countrywide Integrated Non-communicable Diseases Intervention (CINDI) program, entitled CINDI Health Monitor Surveys, supported by the World Health Organization, CINDI Health Monitor. For the second and the third time this type of a survey was performed in 2004 and in 2008.

In the 2001 survey, the stratified random sample was drawn from the Central Population Registry of the Republic of Slovenia. The sample size was 15,379 with the age range 25-64 years. The sampling was performed by the Statistical Office of the Republic of Slovenia.

The data were collected in late spring 2001 by using a self-administered postal questionnaire, based on the CHM Core Questionnaire (28).

Out of 15,379 inhabitants included in the sample 15,153 were contacted (226 were excluded because of changed domicile, severe illness or death). The response rate was 63.8% (9,666 responses). The respondents did not differ statistically from non-respondents in age distribution or distribution of size of settlements of permanent residence, but the response to the survey was slightly lower among men (47.0%) than among women (53.0%) at a ratio 1:1.1 (according to population data in 2001 the ratio was 1:1). The questionnaires of 9,034 respondents were eligible for analysis (eligibility criteria: sex and age provided by Statistical Office of the Republic of Slovenia).

For the purpose of this module, we have chosen observation of smoking behaviour.

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 5 – are presented in Table 6.

Table 6. The intermediate measures (frequency measures) - necessary elements for calculation of prevalence measures of association calculated on the basis of data presented in Table 5.

Frequency measure	Notation	Calculation and value
Prevalence proportion in the non-exposed (females)	P_{E-}	$P_{E-} = \frac{971}{4{,}830} = 0.2010$
Prevalence proportion in the exposed (males)	P_{E+}	$P_{E+} = \frac{1{,}143}{4{,}074} = 0.2806$
Prevalence odds in the non-exposed (females)	${}_{prevalence}O_{E-}$	${}_{prevalence}O_{E-} = \frac{971}{3{,}859} = 0.3900$
Prevalence odds in the exposed (males)	${}_{prevalence}O_{E+}$	${}_{prevalence}O_{E+} = \frac{1{,}143}{2{,}931} = 0.2333$

Calculation of measures of association in prevalence comparisons

The procedures of calculation of prevalence proportion difference, prevalence ratio (also known as prevalence rate ratio), and prevalence odds ratio, and interpretation of results is as follows:

1. Prevalence proportion difference.
   The prevalence proportion difference calculated according to Equation 7 based on data presented in Tables 5 and 6 is (Equation 30):

$$PD = 0.2806 - 0.2010 = 0.0796 \qquad \textbf{Equation 30.}$$

   In Slovene adult males, prevalence of smoking is higher for 0.0796 than in Slovene adult females.

   We could interpret this result also in terms of a percentage. If we multiply the prevalence proportion difference by 100, we get 7.96%, and the interpretation is that in Slovene adult males, prevalence of smoking is higher for 7.96% than in Slovene adult females.
2. Prevalence ratio.
   The prevalence ratio calculated according to Equation 8 based on data presented in Tables 5 and 6 is (Equation 31):

$$PR = \frac{0.2806}{0.2010} = 1.3956 \qquad \textbf{Equation 31.}$$

The prevalence of smoking in Slovene adult males is 1.40-times higher than in Slovene adult females.

3. Prevalence odds ratio.
   The prevalence odds ratio calculated according to Equation 9 based on data presented in Tables 5 and 6 is (Equation 32):

$$POR = \frac{0.3900}{0.2516} = 1.5498 \qquad \textbf{Equation 32.}$$

   The (prevalence) odds of smoking in Slovene adult males, is 1.55-times higher than in Slovene adult females.

Exposure comparisons

A case-control study of impact of hyperirritable uterus on a premature delivery

On the basis of the Slovene PISS data (27), we could simulate a case-control study. Let's suppose that we have followed-up outcomes of deliveries in a given period of time in a selected maternity. In this period we registered 51 preterm deliveries. For each case we selected two controls out of mothers without preterm delivery (i.e. on-time delivery). We were interested if frequency of hyperirritability of uterus (exposure) was different in cases (mothers with preterm delivery) in comparison to controls (mothers without preterm delivery). The results of this analysis are presented in Table 7.

Table 7. Frequency of preterm delivery in two groups of mothers according to absence (controls) or presence (cases) of hyperirritable uterus during pregnancy based on Slovene PISS data (27).

	Presence of hyperirritable uterus during pregnancy		**Total**
Preterm delivery	**No (Non-exposed)**	**Yes (Exposed)**	
No (Controls)	95	7	102
Yes (Cases)	44	7	51
Total	139	14	153

The intermediate measures (frequency measures) - necessary elements for calculation of exposure odds ratio as a measure of association in exposure comparisons calculated on the basis of data presented in Table 7 – are presented in Table 8. For demonstrating the equality of exposure odds ratio to disease odds ratio, both, exposure odds in cases and controls, as well as disease odds in exposed and non-exposed are presented.

Table 8. The intermediate measures (frequency measures) - necessary elements for calculation of measures of association in exposure comparisons calculated on the basis of data presented in Table 7.

Frequency measure	Notation	Calculation and value
Exposure odds in controls	${}_{exposure}O_{controls}$	$_{exposure}O_{controls} = \frac{7}{95} = 0.0737$
Exposure odds in cases	${}_{exposure}O_{cases}$	$_{exposure}O_{cases} = \frac{7}{44} = 0.1591$
Disease odds in the non-exposed	${}_{disease}O_{E-}$	$_{disease}O_{E-} = \frac{7}{7} = 1.000$
Disease odds in the exposed	${}_{disease}O_{E+}$	$_{disease}O_{E+} = \frac{44}{95} = 0.4632$

Calculation of measures of association in exposure comparisons

As just mentioned, in this group of measures we will present only one measure, being exposure odds ratio.

1. Exposure odds ratio.
   The exposure odds ratio calculated according to Equation 10 based on data presented in Tables 7 and 8 is (Equation 33):

$$EOR = \frac{\frac{7}{44}}{\frac{7}{95}} = \frac{0.1591}{0.0737} = \frac{7 \times 95}{44 \times 7} = \frac{95}{44} = 2.1591 = \quad \textbf{Equation 33.}$$

$$= \frac{\frac{7}{7}}{\frac{44}{95}} = \frac{1}{0.4632} = \frac{7 \times 95}{44 \times 7} = \frac{95}{44} = 2.1591 = DOR$$

The odds of being exposed to hyperirritable uterus is in group of mothers that experienced preterm delivery (cases) 2.16-times higher than in group of mothers that did not experienced preterm delivery (controls), indicating that association between preterm delivery and preterm delivery is strong.

Case study 2: Measures of potential impact

Measures assessing impact of an exposure on disease occurrence

An incidence study of impact of hyperirritable uterus on a premature delivery

For demonstrating the calculation process in measures assessing the impact of hazardous factors we will use the same data set as presented in demonstration of calculation of measures of association in incidence comparisons - »An incidence study of impact of hyperirritable uterus on a premature delivery« (Table 1). We will use intermediate measures (frequency measures), presented in Table 2, as well as additional intermediate measures for assessment of population based measures, presented in Table 9.

Table 9. The intermediate measures (frequency measures) - necessary elements for calculation of population based measures for assessing the impact of exposure to a hazardous factor on disease occurrence.

Frequency measure	Notation	Calculation and value
Risk in population	R_{pop}	$R_{pop} = \frac{51}{800} = 0.0638$
Proportion of exposed in population	P_{E+}	$P_{E+} = \frac{37}{800} = 0.0463$

Calculation of measures of assessment of impact of hazardous factors

The procedures of calculation of attributable risks in exposed and in population, and attributable fractions in exposed and in population are as follows:

1. Attributable risk in exposed.
   The attributable risk in exposed calculated according to Equation 11 based on data presented in Tables 1, 2 and 9 is (Equation 34):

$$AR = 0.1892 - 0.0577 = 0.1315 \qquad \textbf{Equation 34.}$$

   The risk for preterm delivery in mothers suffering from the hyperirritable uterus is 0.1892 (18.92%). In absolute terms, 0.1315 (13.15%) of this risk could be attributed to the hyperirritable uterus. This is the portion of the risk of a preterm delivery in the exposed that could be eliminated if exposure were eliminated.
2. Attributable risk in population.
   The population attributable risk calculated according to Equation 12 based on data presented in Tables 1, 2 and 9 is (Equation 35):

$$PAR = 0.0638 - 0.0577 = 0.0061 \qquad \textbf{Equation 35.}$$

In absolute terms, out of 0.0638 (6.36%) of risk of preterm delivery in a population, 0.0061 (0.6%) could be attributed to hyperirritable uterus. This is the risk of a preterm delivery in the population that could be eliminated if exposure were eliminated.

3. Attributable fraction/percent in exposed.
   The attributable fraction in exposed calculated according to Equations 13 and 14 based on data presented in Tables 1, 2 and 9 is (Equations 36 and 37):

$$AF = \frac{0.1892 - 0.0577}{0.1892} = \frac{0.1315}{0.1892} = 0.6952 = 69.52\% \qquad \textbf{Equation 36.}$$

$$AF = \frac{3.2807 - 1}{3.2807} = 0.6952 = 69.52\% \qquad \textbf{Equation 37.}$$

   The second procedure involves the use of relative risk (Equation 27).

   This result indicates that 69.52% of risk of preterm delivery in exposed group could be attributed to hyperirritable uterus. In other words, if those mothers suffering from hyperirritable uterus are spared from this unfavourable phenomenon (e.g. pharmacological treatment, rest, etc.), their risk of preterm delivery would decrease by 0.1315, what would represent 69.52% reduction of their preterm delivery incidence.

4. Attributable fraction/percent in population.
   The population attributable fraction calculated according to Equations 15 and 16 based on data presented in Tables 1, 2 and 9 is (Equations 38 and 39):

$$PAF = \frac{0.0638 - 0.0577}{0.0638} = 0.0954 = 9.54\% \qquad \textbf{Equation 38.}$$

$$PAF = \frac{0.0463 \times (3.2807 - 1)}{1 + [0.0463 \times (3.2807 - 1)]} = 0.0954 = 9.54\% \qquad \textbf{Equation 39.}$$

   The second procedure involves the use of relative risk (Equation 27).

   This result indicates that 9.54% of total risk of preterm delivery in population (in total 6.38%) (exposed and non-exposed) could be attributed to hyperirritable uterus. In other words, 9.45% of preterm deliveries in the population could be prevented if all exposure to hyperirritable uterus is eliminated (e.g. pharmacological treatment, rest, etc.), or a reduction of 0.6 new cases of preterm delivery per 100 population (exposed and non-exposed) is expected if none of pregnant women is suffering from hyperirritable uterus. Such a reduction represents a 9.45% reduction of the incidence of preterm delivery in the population.

An incidence study of impact of medication with iron supplements during pregnancy on a low birth weight of a newborn

In a ambidirectional cohort study, basing on Perinatal Informational System of Slovenia (PISS) (27), data of 800 mothers and their newborns were analyzed again. The observed outcome this time was low birth weight of a newborn (in this analysis 2500 g or less). Let us suppose that in Slovenia we have a preventive programme for reducing iron deficiency during pregnancy. On an individual level with iron deficiency treatment during pregnancy better oxygenation of a fetus could be attained and consecutively better growth (»exposure« to a protective factor) what could result in higher birth weight of a newborn. On a population level with iron deficiency reduction programme during pregnancy lower frequency of low birth weight at birth could be attained. Thus, we related observed outcome to a medication with iron supplements during pregnancy. The results of this analysis are presented in Table 10.

Table 10. Frequency of preterm delivery in two groups of mothers according to medication with iron supplements during pregnancy based on Slovene PISS data (27).

	Presence of hyperirritable uterus during pregnancy		**Total**
Preterm delivery	**No**	**Yes**	
No	401	348	749
Yes	39	12	51
Total	440	360	800

The intermediate measures (frequency measures) - necessary elements for calculation of measures of association calculated on the basis of data presented in Table 10 – are presented in Table 11.

Table 11. The intermediate measures (frequency measures) - necessary elements for calculation of measures for assessing the impact of exposure to a protective factor on disease occurrence calculated on the basis of data from Table 10.

Frequency measure	**Notation**	**Calculation and value**
Risk in the non-exposed	R_{E-}	$R_{E-} = \frac{40}{440} = 0.0886$
Risk in the exposed	R_{E+}	$R_{E+} = \frac{11}{360} = 0.0333$
Relative risk	RR	$RR = \frac{0.0333}{0.0886} = 0.3761$
Risk in population	R_{pop}	$R_{pop} = \frac{51}{800} = 0.0638$
Proportion of exposed in population	P_{E+}	$P_{E+} = \frac{360}{800} = 0.4500$

Calculation of measures of assessment of impact of protective factors

The procedures of calculation of prevented fractions in exposed and in population are as follows:

1. Prevented fraction/percent in exposed.
   The prevented fraction in exposed calculated according to Equations 17 and 18 based on data presented in Tables 10 and 11 is (Equations 40 and 41):

$$PF = \frac{0.0886 - 0.0333}{0.0886} = \frac{0.0553}{0.0886} = 0.6239 = 62.39\% \qquad \textbf{Equation 40.}$$

$$PF = 1 - 0.3761 = 0.6239 = 62.39\% \qquad \textbf{Equation 41.}$$

   The results indicate that medication with iron supplements during pregnancy has reduced the risk of low birth weight of a newborn by 62.39% among treated pregnant women. The iron supplements were 62.39% efficacious.
2. Prevented fraction/percent in population.
   The population prevented fraction calculated according to Equations 19 and 20 based on data presented in Tables 10 and 11 is (Equations 42 and 43):

$$PPF = \frac{0.0886 - 0.0638}{0.0886} = \frac{0.0249}{0.0886} = 0.2808 = 28.08\% \qquad \textbf{Equation 42.}$$

$$PPF = 0.4500 \times (1 - 0.3761) = 0.2808 = 28.08\% \qquad \textbf{Equation 43.}$$

   The results indicate that preventive programme for reducing iron deficiency has reduced the risk of low birth weight of a newborn by 28.08% in the population of pregnant women as a whole. We might argue that the iron deficiency reduction programme during pregnancy was 28.08% effective.

Measures assessing impact of an intervention on disease occurrence change

An incidence study of impact of medication with iron supplements during pregnancy on a low birth weight of a newborn

For demonstrating the calculation process in measures assessing the impact of an intervention on disease occurrence change we will use the same data set as just - »An incidence study of impact of medication with iron supplement during pregnancy on a low birth weight of a newborn« (Table 10). The only difference is that we suppose now that our data originate from a preventive trial in which one group of pregnant women were treated with iron supplements and the other was a control group. In calculations we will in fact use the same intermediate measures (frequency measures)

as presented in Table 11, but since the notation is now slightly different, these intermediate measures are presented Table 12 again according to this different notation.

Table 12. The intermediate measures (frequency measures) - necessary elements for calculation of measures for assessing the impact of an intervention on disease occurrence change calculated on the basis of data from Table 10.

Frequency measure	Notation	Calculation and value
Risk in the control group (non-exposed)	R_C	$R_{E-} = \frac{40}{440} = 0.0886$
Risk in the treated (experimental) group (exposed)	R_T	$R_{E+} = \frac{11}{360} = 0.0333$
Relative risk	RR	$RR = \frac{0.0333}{0.0886} = 0.3761$

Calculation of measures of assessment of impact of an intervention on disease occurrence change

The procedures of calculation of absolute risk reduction, relative risk reduction and number needed to treat are as follows:

1. Absolute risk reduction.
   The absolute risk reduction calculated according to Equation 21 based on data presented in Tables 10 and 12 is (Equation 44):

$$ARR = 0.0886 - 0.0333 = 0.0553 \qquad \textbf{Equation 44.}$$

   Absolute risk reduction is just the absolute difference in risks for observed outcome between the control and the treatment group. It is less intuitive measure to interpret than relative risk reduction is, and its main role is to be used in calculation of number needed to treat. However, if we multiply the result in our case, and we get absolute risk reduction 5.53%, we can interpret it as follows: for every 100 pregnant women enrolled in the treatment group, about 5.5 bad outcomes (deliveries of a low birth weight newborn) would be averted.
2. Relative risk reduction.
   The relative risk reduction calculated according to Equations 22 and 23 based on data presented in Tables 10 and 12 is (Equations 45 and 46):

$$RRR = \frac{0.0886 - 0.0333}{0.0886} = \frac{0.0553}{0.0886} = 0.6239 = 62.39\% \qquad \textbf{Equation 45.}$$

$$RRR = 1 - 0.3761 = 0.6239 = 62.39\% \qquad \textbf{Equation 46.}$$

Relative risk reduction measures how much of the risk is reduced in the experimental (treated) group compared to a control group. In our example the result is 62.39%. This means that low birth weight of a newborn was reduced by 62.4% in the treatment group compared with the control group without treatment with iron supplements.

3. Number needed to treat.
   The number needed to treat calculated according to Equation 24 based on data presented in Tables 10 and 12 is (Equation 47):

$$NNT = \frac{1}{0.0553} = 18.08 \qquad \textbf{Equation 47.}$$

The result of calculating this measure indicates that for every 18 pregnant women treated with iron supplements one newborn with low birth weight would be prevented.

EXERCISE

Task 1

In a maternity hospital data on successive 800 deliveries were collected in an ambidirectional cohort study. The observed outcome was low birth weight of a newborn, which is defined as birth weight 2500 g or less. The exposure under observation is smoking of mother during pregnancy. The results of this study are presented in Table 13.

Table 13. Frequency of low birth weight in newborns in two groups according to smoking of mother during pregnancy based on PISS data (27).

Low birth weight of a newborn	Exposure to smoking of mother during pregnancy		Total
	No	**Yes**	
No	558	191	749
Yes	35	16	51
Total	593	207	800

Please:

- make small groups of students (maximum three students in a group),
- with other students in your group discuss what kind of a comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13,

- calculate and make interpretation of all these measures,
- make a short presentation,
- present results to other groups of students,
- discuss your results to the results of other students.

Task 2

In the same study, the same observed outcome was related to a different hazardous factor, this time being elevated blood pressure of mother during pregnancy. The results of this study are presented in Table 14.

Table 14. Frequency of low birth weight in newborns in two groups according to elevated blood pressure in mother during pregnancy based on PISS data (27).

Low birth weight of a newborn	Elevated blood pressure in mother during pregnancy		Total
	No	Yes	
No	707	42	749
Yes	46	5	51
Total	753	47	800

Please:

- make small groups of students (maximum three students in a group),
- with other students in your group discuss what kind of a comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13,
- calculate and make interpretation of all these measures,
- make a short presentation,
- present results to other groups of students,
- discuss your results to the results of other students.

Task 3

In a cross-sectional study, basing on CINDI Health Monitor survey 2001, that was aiming at assessing the prevalence of health behaviours (28), data of 9,034 adults were analyzed. The observed outcome was frequent perception of stress without or with poor coping mechanisms. We related this outcome to a gender as a risk factor for unfavourable health behaviour (i.e. in this context in a role of »exposure«). The results of this study are presented in Table 15.

Table 15. Prevalence of stres in both gender groups in adult population in Slovenia in 2001, based on CINDI Health Monitor Survey, Slovenia 2001 (28).

	Gender		Total
Smoking	**Males**	**Females**	
No	3,235	3,570	6,805
Yes	861	1,321	2,182
Total	4,096	4,891	8,987

Please:

- make small groups of students (maximum three students in a group),
- with other students in your group discuss what kind of a comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13,
- calculate and make interpretation of all these measures,
- make a short presentation,
- present results to other groups of students,
- discuss your results to the results of other students.

Task 4

In a case-control study outcomes of successive deliveries were followed-up in a given period of time in a selected maternity. In this period 36 deliveries of babies with low birth weight were registered (cases). For each case approximately three controls out of other deliveries (i.e. with normal birth weights) were selected. The research question was if elevated blood pressure of mother during pregnancy is associated with low birth weight of newborns. Data of 153 mothers and their newborns were analyzed. The results of this analysis are presented in Table 16.

Table 16. Frequency of preterm delivery in two groups of mothers according to absence (controls) or presence (cases) of hyperirritable uterus during pregnancy based on Slovene PISS data (27).

	Presence of elevated blood pressure in mother during pregnancy		Total
Low birth weight	**No (Non-exposed)**	**Yes (Exposed)**	
No (Controls)	69	48	117
Yes (Cases)	30	6	36
Total	99	54	153

Please:

- make small groups of students (maximum three students in a group),
- with other students in your group discuss what kind of a comparisons you can perform according to basic frequency measures that can be computed on the basis of data presented in Table 13,

- calculate and make interpretation of all these measures,
- make a short presentation,
- present results to other groups of students,
- discuss your results to the results of other students.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Schoenbach VJ, Rosamond WD. Understanding the fundamentals of epidemiology - an evolving text [Internet]. Chapel Hill: University of North Carolina, Epidemiology.net; 2000. http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: June 25, 2009.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
5. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
6. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
7. Centers for Diasese Control and Prevention. Principles of epidemiology in public health practice. 3rd edition. An introduction to applied epidemiology and biostatistics [Internet]. Atlanta: Centers for Diasese Control and Prevention; [cited 2009 Jun 25]. Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: June 24, 2009.
8. Szklo M, Nieto FJ. Epidemiology: Beyond the basics. Gaithersburg, MA: Aspen Publishers; 2005.
9. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maidenhead: Open University Press, McGraw-Hill Education; 2005.
10. Babuš V. Epidemiologic methods (in Croatian). Zagreb: Medicinska naklada; 2000.
11. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden MA: Blackwell Science, Inc.; 2005.
12. Pearce N. A Short Introduction to Epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003. Available from URL: www.massey.ac.nz/~wwwcphr/publications/introepi.pdf. Accessed: June 25, 2009.
13. Walter SD. Choice of effect measure for epidemiological data. J Clin Epidemiol. 2000;53:931-9. Available from URL: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6T84-417WDDC-5&_user=4776866&_rdoc=1&_fmt=&_orig=search&_sort=d&_docanchor=&view=c&_searchStrId=942536967&_rerunOrigin=google&_acct=C000033658&_version=1&_urlVersion=0&_userid=4776866&md5=426ecb1b2e927d8e1bc77e845dee1941. Accessed: June 24, 2009.
14. Symons MJ, Moore DT. Hazard rate ratio and prospective epidemiological studies. J Clin Epimiol. 2002;55:893-9. Available from URL: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6T84-470TK8G-8&_user=4776866&_rdoc=1&_fmt=&_orig=search&_sort=d&_docanchor=&view=c&_searchStrId=942540184&_rerunOrigin=google&_acct=C000033658&_version=1&_u

rlVersion=0&_userid=4776866&md5=ae285644eb3d7f2500700b3c6b2288b5. Accessed: June 24, 2009.
15. Pearce N. Effect measures in prevalence studies. Environ Health Perspect. 2004;112:1047-50. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1247374&blobtype=pdf. Accessed: June 24, 2009.
16. Jenicek M. Foundations of evidence-based medicine. London: Parthenon publishing; 2003.
17. University of Toronto, Centre for Evidence-Based Medicine. Glossary of EBM terms [homepage on the Internet]. Available from URL: http://www.cebm.utoronto.ca/glossary/. Accessed: June 30, 2009.
18. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
19. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
20. Kleinbaum DG, Klein M. Survival analysis. A self-learning text. Second edition. New York: Springer; 2005.
21. Parmar M, Machin D. Survival analysis. A practical approach. Chichester: John Wiley&Sons; 1995.
22. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.
23. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
24. Kleinbaum DG, Sullivan KM, Barker ND. A pocket guide to epidemiology. New York: Springer; 2007.
25. Schunemann HJ, Guyatt GH. Clinical epidemiology. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.1170-223.
26. Spasoff RA. Epidemiologic methods for health policy. New York: Oxford University Press, I999.
27. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia. Available from URL: http://www.ivz.si/index.php?akcija=podkategorija&p=44. (Accessed July 24, 2008).
28. Zaletel-Kragelj L, Fras Z, Maučec Zakotnik J, editors. Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001. II. Risky behaviours (in Slovene). Ljubljana: CINDI Slovenija, 2004. p.107-564.

RECOMMENDED READINGS

1. Benichou J, Palta M. Rates, risks, measures of association and impact. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.91-156.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
4. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sons, cop.; 1982.
5. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
6. Schunemann HJ, Guyatt GH. Clinical epidemiology. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.1170-223.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**CLUSTER ANALYSIS**
Module: 1.3.2	**ECTS (suggested): 0.30**
Author(s), degrees, institution(s)	**Anca Vitcu,** PhD, Assistant Professor Faculty of Computer Science, University "Al. I. Cuza", Iasi, Romania
Address for Correspondence	**Anca Vitcu** Faculty of Computer Science, University "Al. I. Cuza", Iasi Str. General Berthelot, nr. 16, Iasi, Code 700483, Romania Fax: +40 232 201490 Email: avitcu@yahoo.com, avitcu@info.uaic.ro
Keywords	Hierarchical clustering, k-means clustering, probabilistic clustering
Learning objectives	After completing this module students should: • understand and apply the basic clustering methods; • be aware of cluster analysis limitations; • gain expertise in the interpretation of clustering solutions • improve the skills needed for the use of statistical software packages • be able to undertake investigations into geographical pattern of disease; • identifie and investigate disease risks in certain areas; • increase knowledge regarding the efficient distribution of resources for prevention and treatment of disease.
Abstract	From the huge amount of data available in health field today, we have to find out what the emerging problems are and recommend the best scenario we can get to solve them. In light of this, we have to learn to use the most appropriate tools, which can help us to mine the "mountains" of data around us with the aim to create knowledge. The chapter below is a brief presentation of clustering analysis methodology which includes a number of different algorithms and methods accustomed to organize huge amount of observed data into meaningful structures.
Teaching methods	Teaching methods will include combination of lectures, exercises, individual work, interactive methods such as small group discussions. Before the introductory lecture a case study could be presented to increase students' motivation. An introductory lecture gives the students the basic theoretical knowledge on cluster analysis. After the introductory lecture students will work individual and in teams of 2-3 students, study the recommended readings and discuss the characteristics and pitfalls of clustering algorithms. Work will be followed by an individual case problem presentation and overall discussion.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and statistical programme; • training materials: recommended readings or other related documents; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment could be based on case problem presentations.

CLUSTER ANALYSIS

Anca Vitcu

THERORETICAL BACKGROUND

About clustering

Definition and main features

Cluster analysis is a data analysis tool which comprises a variety of goals, all of them related with grouping or segmenting objects into meaningful structures (clusters), without explaining why they exist, based only on the data that describes them and their relationships.

It includes a number of different algorithms and methods and is used when we do not have any a priori hypothesis about the data.

The components of a cluster are more closely related to one another than those assigned to different clusters. In other word, the degree of association between objects is maximal if they belong to the same cluster and minimal otherwise.

Aims

The main objective of cluster analysis is to identify homogeneous groups or clusters in a data set. Cluster analysis has the potential to generate new knowledge: may help formulate hypotheses concerning the origin of the sample, describe a sample in terms of a typology, or predict the future behaviour of population types. It also proved to be a useful starting point for other analysis procedures (e.g. based on geographical analysis methods and spatial scan statistics localised clustering can be assessed and the risk of disease inside and out side the study area can be compared).

Applications

In medical science there are different types of clustering such as:

- general clustering – involves the analysis of the overall clustering tendency of the disease incidence in a study area without searching the exact location of the clusters (1,2);
- specific clustering – involves specific disease clustering methods which are designed to examine the exact location of the clusters (3,4).

Some of the most important applications of cluster analysis are: pattern recognition, spatial data analysis, image processing. In epidemiology clustering is frequently used:

- to identify diseases and their stages, and by examining their characteristics to discover whether there are different subtypes of diseases grouped togheter under a single diagnosis (5);
- to detect patterns in the spatial and temporal distribution of a disease (6);
- to find if there are several distinct groups of patients with different symptoms and similar behavior habits which have been diagnosed with a certain disease (7);

- to group patients into nonoverlapping activity/inactivity clusters and then uses the outcomes in models of prevalent and incident overweight;
- to measure the different effects of treatments on classes within the population;
- to assist the disease surveillance.

Description

Types of clustering

In a cluster analysis we can distinguish different types of clustering (8,9):

- exclusive or non-exclusive;
- complete or partial;
- hierarchical or partitional.

In the following paragraphes we will shortly describe each of these clustering types:

1. Exclusive vs. non-exclusive clusters.
   Exclusive clusters assign each case to a single cluster such as each case is closer to all of the cases in its cluster than to any case in another cluster (Figure 1). In this situation a case can only belong to one cluster (e.g. case 1 belongs to cluster "*c*", case 2 to cluster "*a*", case 3 to cluster "*c*", case 4 to cluster "*a*" and so on).

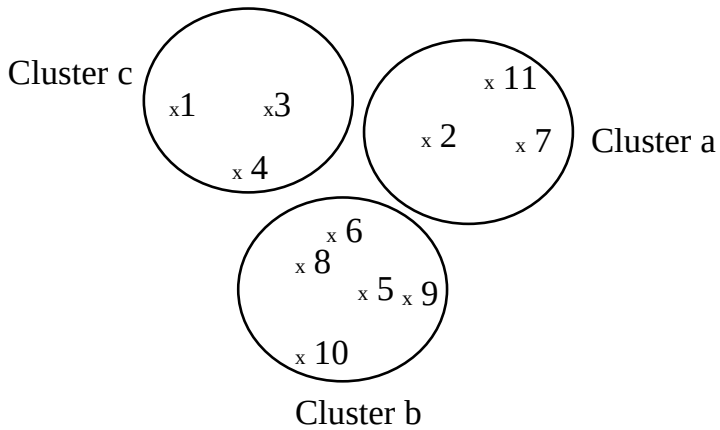


Figure 1. Exclusive clusters.

On the other hand, there are situations in which a case can logically be placed in more than one cluster, e.g. case 2 belongs to cluster "*a*" and "*c*", case 4 belongs to cluster "*b*" and "*c*" (Figure 2). These situations in which a case can simultaneously belong to more than one cluster or is "located between" two or more clusters, and can be assigned to any of them are addressed by *non-exclusive clustering*.

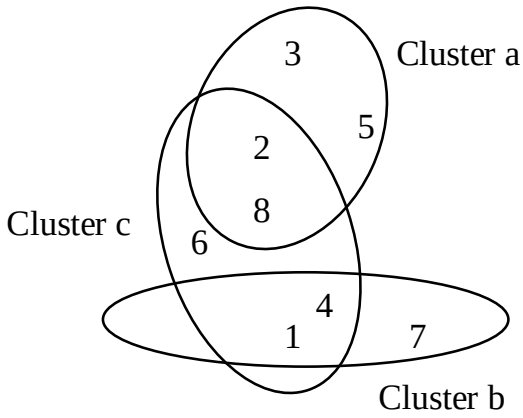


Figure 2. Non-exclusive clusters.

An example of non-exclusive clustering is fuzzy clustering, where every case belongs to every cluster with a membership weight which takes values between 0 (absolutely doesn't belong) and 1 (absolutely belongs). In fuzzy clustering it is often added the constraint that the sum of weights for each case must equal 1. In a similar way, probabilistic clustering techniques compute the probability with which a case belongs to each cluster, and these probabilities must also sum to 1. These approaches are most appropriate for avoiding the arbitrariness of assigning a case to only one cluster when logically it may belong to several. In practice, a fuzzy or probabilistic clustering is often converted to an exclusive clustering by assigning each case to the cluster in which its membership weight or probability is the highest (10,11).

2. Complete vs. partial clusters.
   A *complete clustering* assigns every case of a data set to a cluster while a *partial clustering* is used when some cases in a data set may not belong to well-defined groups (e.g. they may represent outliers, or uninteresting issues).
3. Hierarchical vs. partitional clusters.
   The goal of a cluster analysis is sometimes to arrange the cases or clusters themselves into a natural hierarchy. This involves successively grouping them so that at each level of the hierarchy, clusters (resp. cases) within the same group (resp. cluster) are more similar to each other than those in different groups (resp. clusters).

   In figure 3, we can see that the two subclusters with components (a,e) and (c) belong to cluster 1 while the other two subclusters composed of cases (f,b,d) and (g) respectively belong to cluster 2. In other words, if a cluster has subclusters we obtain a hierarchical clustering. We will say more about hierarchical clustering later in this chapter.

The opposite of hierarchical clustering is *partitional clustering* which is a division of the set of data into exclusive clusters.

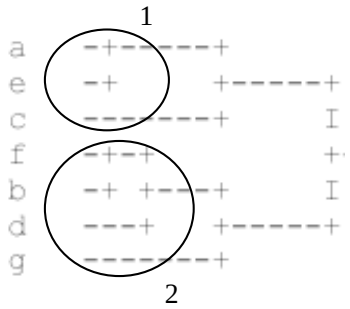


Figure 3. Hierarchical clusters.

Fundamental steps in cluster analysis

In principle, the basic steps in a clustering analysis are the following (8,12):

1. Selection of the appropriate variables on which cluster configuration will be based.
   Note: The initial choice of variables determines the characteristics that can be used to identify clusters; if important variables are excluded, poor or misleading findings may result.
2. Selection of the appropriate cases to be included in the analysis.
   Note: The clusters' structure could be quite different among selected groups: women vs. men, socioeconomic groups, life style, regions or countries.
3. Selection of the appropriate measures for calculation the distance or similarity between objects.
   Note: *Distance* is a measure of how different two objects are, while *similarity* is a measure of their closeness. Objects which are alike share a low distance and a high similarity. Selection of a distance/similarity measure should be based on both the properties of the measure and on the algorithm for cluster formation.
4. Selection of the appropriate methods to be used for combining objects into clusters.
   Note: Selection of a method depends on the data set to be analyzed, definition of measure supplied to it and the characteristics of the various methods available. Different methods will result in different cluster patterns.
5. Application of these methods.
6. Evaluation of the results.
   Note: Basically this refers to:
   - set up the number of clusters (if it is not known);
   - test the clusters stability;
   - test the validity of clusters (internal, relative and external validity).

7. Interpretation of the results (substantive interpretation of clusters).
   In the following sections we will give a brief description to the main features of the steps mentioned above.

Assumptions regarding variables

In cluster analysis variables should be commensurable, which means that they must have equal scales. Frequently in a data set we have to deal with the following situations:

- the variables are quantitative but have different scales (e.g. age, income, human development index);
- the variables included in the analysis are mixed, in other words they have different measurement levels: *quantitative* (e.g. age, weight, blood pressure), *ordinal* (e.g. attitude scale: 1=strongly agree, 2=agree, 3=disagree, 4=strongly disagree), *nominal* (e.g. types of health services, development areas);
- the occurrence of one variable depends on another variable (e.g. occupational level depends on occupational status).

If variables are quantitative or ordinal and have different scales a transformation to an equal scale is required. The most popular transformations are (12):

1. z - standardization ((a) theoretical (Equation 1) or (b) empirical (Equation 2)) – transforms the values so that they have a mean of 0 and a standard deviation of 1:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \qquad \textbf{Equation 1.}$$

x_{ij} = *values of case* i *in variable* j

μ_j = *mean value derived from the attributes of the scale and*

σ_j = *standard deviation derived from the attributes of the scale*

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \qquad \textbf{Equation 2.}$$

x_{ij} = *values of case* i *in variable* j

$\bar{x}_j$ = *empirical mean of variable* j

s_j = *empirical standard deviation of variable* j .

2. [0,1] – transformation ((a) theoretical (Equation 3) or (b) empirical (Equation 4)) - the procedure subtracts the minimum value from each item being standardized and then divides by the range:

$$z_{ij} = \frac{x_{ij} - \alpha_j}{\beta_j - \alpha_j} \qquad \textbf{Equation 3.}$$

x_{ij} = *values of case* i *in variable* j

α_j = *theoretical minimum of the scale*

β_j = *theoretical maximum of the scale*

$$z_{ij} = \frac{x_{ij} - a_j}{b_j - a_j}$$ **Equation 4.**

x_{ij} = *values of case* i *in variable* j

a_j = *empirical minimum of the scale*

b_j = *empirical maximum of the scale*

So, if the variables are quantitative or ordinal, the types of scales we have to deal with take account of: continuous values, and resp. discret equidistant values or discret non-equidistant values. For each of these scales we can compute the variables' theoretical mean and standard deviation according to the subsequent formula (13):

- for continuous values (Equations 5 and 6):

$$\mu_j = (\beta_j + a_j)/2$$ **Equation 5.**

and

$$\sigma_j = (\beta_j + \alpha_j)/2\sqrt{3}$$ **Equation 6.**

- for discret equidistant values (Equations 7 and 8):

$$\mu_j = (\beta_j + a_j)/2$$ **Equation 7.**

and

$$\sigma_j = \sqrt{(n_j - 1)(n_j + 1)/12}$$ **Equation 8.**

n_j = *number of categories of variable* j

- for discret non-equidistant values (Equations 9 and 10):

$$\mu_j = (\sum v_{jk})/n_j$$ **Equation 9.**

and

$$\sigma_j = \sum (v_{jk} - \mu_j)^2 / n_j$$ **Equation 10.**

n_j = *the number of categories of variable* j

v_{jk} = *scale value of category* k *of variable* j

The following alternatives are also available for transforming values:

- range -1 to 1 - each value for the item being standardized is divided by the range of the values,
- maximum magnitude of 1 - the procedure divides each value for the item being standardized by the maximum of the values,
- mean of 1 - the procedure divides each value for the item being standardized by the mean of the values,
- standard deviation of 1 - the procedure divides each value for the variable or case being standardized by the standard deviation of the values.
  Note: Standardization can be done by variables or by case.

Another situation mentioned above concerns the case of mixed variables. In this circumstance the following transformations can be approched:

- for binary variables (Equation 11):

$$x'_{ij} = x_{ij} \qquad \textbf{Equation 11.}$$

- for nominal variables (Equation 12):

$$x'_{ij}(n) = \begin{cases} 1/\sqrt{2} & if \quad x_{ij} = n \\ 0 & otherwise \end{cases} \qquad \textbf{Equation 12.}$$

- for ordinal and quantitative variables (Equations 13 and 14):

$$x'_{ij} = x_{ij} / r \qquad \textbf{Equation 13.}$$

or

$$x'_{ij} = \left(x_{ij} - min(x_j)\right) / r \qquad \textbf{Equation 14.}$$

r = *the range*

The resulting variables have values between 0 and 1.

If transformation $1/r$ causes problems, a priory standardization of the variables can overcome them.

Assumptions regarding data-missing values

In almost all databases we can find observations which have missing values in one or more of the variables. The most common method of including missing values in dissimilarity calculation is to omit each observation pair having at least one missing value (listwise deletion) or a case is eliminated only if the number of missing values

exceeds a certain threshold (pairwise deletion) (14). These methods can fail in the circumstance when both observations have no measured values in common.

An alternative is to estimate missing values with imputation techniques. Both observations could be imputed using the mean or median of each variable over the non-missing data.

For categorical variables, the value "missing" can be considered as another categorical value, if both objects have missing values on the same variables.

Distance and proximity measures

Fundamental to all clustering techniques is the choice of distance or dissimilarity measure between two objects. We first discuss distance measures before describing a variety of algorithms for clustering.

There are various measures that can be used to quantify similarity or dissimilarity between objects (13). These measures can be classified in four main groups:

1. correlation coefficients;
2. distance measures;
3. derived measured based on correlation coefficients or distances;
4. other similarity or dissimilarity measures developed for special purposes, mainly for binary variables.

Note: Correlation coefficients and derived measures based on correlation coefficients are *mostly used for clustering variables* while, distance measures and derived measures based on distances are *mostly used for clustering cases.*

In most cases in cluster analysis we work with raw data, but sometimes we don't have them. However, the data available is characterized in terms of proximity between pairs of cases (objects) (Example 1).

Example 1.

Patients from the same type of hospitals located in ten different counties are asked to judge by how much certain medical services differ from one another. The patients are selected randomly from the hospital database and asked to indicate their attitude toward each of the 7 items that refers to the quality of health services. The attitude is measured on a seven point Likert scale with the following response categories: 1=very favorable, 2=moderately favorable, 3=slightly favorable, 4=neither favorable nor unfavorable, 5=slightly unfavorable, 6=moderately unfavorable, 7=very unfavorable. Dissimilarities can then be computed by averaging over the collection of such judgments. The results can be represented by a symmetric matrix $(d_{ij})_{n\times n}$ *with nonnegative entries and zero diagonal elements, where* n *represent the number of objects and* d_{ij} *the proximity between cases* i^{th} *and* j^{th}*. This matrix is then provided as input to clustering algorithm. Table 1 envisages the proximity matrix of the imaginary situation presented above. Together with this cluster analysis certain statistics are also recommended to be computed: mean, percent of favorable, percent of unfavorable, percent of neutral, standard deviation, etc.*

Table 1. Proximity matrix of patients attitude: values are average pairwise dissimilarities of counties from a questionnaire given to sampled patients.

	County1	**County2**	**County3**	**County4**	**County5**	**County6**	**County7**	**County8**	**County9**	**County10**
County1	0									
County2	5.27	0								
County3	6.34	7.12	0							
County4	3.42	3.82	4.92	0						
County5	2.25	4.67	6.17	3.67	0					
County6	6.17	6.92	4.50	5.67	4.23	0				
County7	4.95	3.57	3.92	5.00	5.25	3.92	0			
County8	6.33	2.67	4.81	2.24	3.00	4.52	5.83	0		
County9	4.75	4.51	6.83	4.25	8.12	6.00	2.85	3.77	0	
County10	7.00	5.00	6.08	3.02	4.58	5.02	5.76	6.42	6.92	0

In most of the studies, data are measurements for n cases on k variables. As we know the popular clustering algorithms require the specification of a dissimilarity matrix as their input. For this reason, before deciding the type of clustering algorithm we have to construct the dissimilarity matrix (Equation 15):

$$D(x_i, x_{i'}) = \sum_{j=1}^{k} d_j(x_{ij}, x_{i'j}) \qquad \textbf{Equation 15.}$$

$d_j(x_{ij}, x_{i'j})$ = *the distance between values of* j^{th} *variable for cases* i *and* i'.

These measures have to be calculated according to the measurement level of variables: *quantitative*, *ordinal*, or *categorical*.

Quantitative variables

First we would like to remember that measurements of this type of variables are represented by continuous real-valued numbers. Among the distance measures the following ones are the most popular:

1. Euclidian distance (Equation 16):

$$d(x_i, x_{i'}) = \sqrt{\sum_j (x_{ij} - x_{i'j})^2} \qquad \textbf{Equation 16.}$$

2. Squared Euclidian distance (Equation 17):

$$d(x_i, x_{i'}) = \sum_j (x_{ij} - x_{i'j})^2$$ **Equation 17.**

Squared Euclidian distance places more weight on larger differences than smaller ones.

Note: (1) Euclidian distances are usually computed from raw data and not from standardized ones. (2) They can be affected by differences in scale among the dimensions from which the distances are computed.

3. City-block (Manhattan) distance (Equation 18):

$$d(x_i, x_{i'}) = \sum_j |x_{ij} - x_{i'j}|$$ **Equation 18.**

4. Chebychev distance (Equation 19):

$$d(x_i, x_{i'}) = \max_j |x_{ij} - x_{i'j}|$$ **Equation 19.**

Alternatively, for these types of variables clustering can be based on correlation measures.

5. Pearson's r can be used as a similarity or correlation measure (Equation 20):

$$\rho(x_i, x_{i'}) = \frac{\sum_j (x_{ij} - \bar{x}_i)(x_{i'j} - \bar{x}_{i'})}{\sqrt{\sum_j (x_{ij} - \bar{x}_i)^2 \sum_j (x_{i'j} - \bar{x}_{i'})^2}}$$ **Equation 20.**

x_{ij} = *value of case* i *in variable* j,

$x_{i'j}$ = *value of case* i' *in variable* j,

$\bar{x}_i = \sum_j x_{ij} / p$,

$\bar{x}_{i'} = \sum_j x_{i'j} / p$.

If inputs are first standardized, then we can write (Equation 21):

$$\sum_j (x_{ij} - x_{i'j})^2 = 2(1 - \rho(x_i, x_{i'}))$$ **Equation 21.**

which imply that clustering based on correlation is equivalent to that based on squared distance.

Ordinal variables

The values of this type of variables are represented as adjacent integers, and the possible values are considered to be an ordered set (e.g. degrees of preferences or agreement). Rank data are special kind of ordinal data.

The most common measures for ordinal data are:

- city block metric,
- coefficient kappa for ordinal variables, and
- correlation coefficients (Kendal's tau or Gamma).

Categorical variables

In this case, the degree of difference between pairs of values must be delineated explicitly.

For clustering variables the following measures can be used:

- Cramer's V,
- Phi,
- Lambda, and
- other association coefficients.

If cases are clustered the following measures can be applied for nominal variables:

- simple matching coefficient,
- coefficient kappa for nominal variables,
- city block metric, and
- squared Euclidian distances.

Binary variables

The measures for binary variables differ in the importance they attach to the different cells of a 2x2 table as the one presented in Figure 4.

		Case i	
		Present (1 or +)	Absent (0 or -)
Case i'	Present (1 or +)	***a***	***b***
	Absent (0 or -)	***c***	***d***

Figure 4. Elements of 2×2 table.

Reading the table we find out that for "*a*" variables both cases have value "present", for "*d*" variables both cases have value "absent", for "*b*" and "*c*" variables the cases have different values.

There are situations when we may want to weight the positive-positive cell more or less than the negative-negative cell, we may want to weight equal mismatches and matches cells or we may want to ignore one of them, etc. according to the information they provide. In this context, for binary variables following measures can be used (12,15):

1. Jaccard's coefficient I (Equation 22):

$$\rho(x_i, x_{i'}) = d/(2d + b + c) \qquad \textbf{Equation 22.}$$

2. Dice's coefficient (Equation 23):

$$\rho(x_i, x_{i'}) = 2d/(2d + b + c) \qquad \textbf{Equation 23.}$$

3. Sokal&Sneath's coefficient I (Equation 24):

$$\rho(x_i, x_{i'}) = d/(d + 2(b + c)) \qquad \textbf{Equation 24.}$$

4. Russel&Rao's coefficient (Equation 25):

$$\rho(x_i, x_{i'}) = d/(a + b + c + d) \qquad \textbf{Equation 25.}$$

5. Euclidian distance (Equation 26):

$$d(x_i, x_{i'}) = \sqrt{b + c} \qquad \textbf{Equation 26.}$$

Euclidian distance has a minimum value of 0 and no upper limit

6. Squared Euclidian distance (Equation 27):

$$d(x_i, x_{i'}) = b + c \qquad \textbf{Equation 27.}$$

Squared Euclidian distance has a minimum value of 0 and no upper limit.

7. Lance and Williams (non-metric dissimilarity measure) (Equation 28):

$$d(x_i, x_{i'}) = (b + c)/(2a + b + c) \qquad \textbf{Equation 28.}$$

Lance and Williams (non-metric dissimilarity measure) has a range of 0 to 1.

8. Pattern distance (Equation 29):

$$d(x_i, x_{i'}) = bc/(a+b+c+d)^2 \qquad \textbf{Equation 29.}$$

Pattern distance has a range of 0 to 1.

Mixed levels

Various methods have been developed to work with mixed measurement levels. One of these methods is based on Gower's dissimilarity coefficient (Equation 30):

$$d_{ii'} = \frac{\sum_{i'} w_{ii'j} \cdot d_{ii'j}}{\sum_{i'} w_{ii'j}} \qquad \textbf{Equation 30.}$$

$d_{ii'}$ = *dissimilarity between cases* i *and* i'

$w_{ii'j}$ = *weight for variable* j

$d_{ii'j}$ = *dissimilarity between cases* i *and* i' *in variable* j

or Gower's similarity coefficient (Equation 31):

$$s_{ii'} = \frac{\sum_{i'} w_{ii'j} \cdot s_{ii'j}}{\sum_{i'} w_{ii'j}} \qquad \textbf{Equation 31.}$$

$s_{ii'}$ = *similarity between cases* i *and* i'

$w_{ii'j}$ = *weight for variable* j

$s_{ii'j}$ = *similarity between cases* i *and* i' *in variable* j

While squared Euclidian distance and city block metric can be computed for all measurement levels they are often used as dissimilarity measures for Gower's coefficient. In each situation the weight is defined as the inverse value of the maximum distance such as its highest value is 1.

Clustering methods

Different clustering methods with different properties have been developed. A general answer to the question "which technique should be used?", cannot be given. The answer depends on the data used and the analysed question.

An important distinction is the question whether "cases or variables should be clustered?" and in this situation an answer can be provided.

Broadly speaking, we can identify three different general types of cluster analysis algorithms (16):

- those based on a hierarchical attempt to discover cluster structure,
- those based on an attempt to find the optimal partition into a specified number of clusters,
- those based on a probabilistic model for the underlying clusters.

In the following paragraphs we will give a brief discription to each of these algorithms.

Different methods have been developed to cluster cases or variables which result in two major types of assignments (17,18):

1. Deterministic:
    - hierarchical algorithms (divisive hierarchical algorithms, agglomerative hierarchical algorithms) – according to this algorithm we construct step by step a hierarchy or tree-like structure to see the relationship among objects,
    - non-hierarchical algorithms (K-means algorithms) – consistent with this method a position in the measurement is taken as central location, and distance between cases is measured from such central point (19).
2. Probabilistic:
    - probabilistic algorithms – according to this method there is a probability or a degree of membership with which the case belongs to each of the cluster, such as in the example presented in Table 2 (e.g. case 1 belongs with a probability of 0.2 to cluster 1, with a probability of 0.3 to cluster 2 and with a probability of 0.5 to cluster 3, we will see later in this chapter that ussualy, exclusive clusters are required and for this reason the case is assigned to the cluster which has the highest probability) (Example 2).

Table 2. Deterministic vs. probabilistic assignment of objects in cluster analysis.

Objects	Cluster membership	Deterministic assignment			Probabilistic assignment		
		Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
1	3	0	0	1	0.2	0.3	0.5
2	3	0	0	1	0.3	0.0	0.7
3	1	1	0	0	0.8	0.2	0.0
4	2	0	1	0	0.3	0.5	0.2
5	1	1	0	0	0.7	0.1	0.2
6	2	0	1	0	0.2	0.6	0.2
7	2	0	1	0	0.0	0.9	0.1

Example 2.

Periodically, the management team of hospital X conducts interview surveys with persons who benefit of the hospital assistance. The respondents are patients selected randomly from the registration list. As part of the survey each patient is asked to indicate her/his perception toward a battery of Likert-type items regarding the quality and safety of health services provided by the hospital (staff behavior and attitude, hospital facilities, etc.). If the management team wants to analyze what services are alike, they have to use methods for clustering variables (e.g. agglomerative hierarchical algorithms). If they want to find out if patients differ in their attitude and dissimilar patterns of preference can be recognized, cases have to be clustered. Theoretically all three algorithms mentioned above (hierarchical, non-hierarchical, and probabilistic) can be used for this purpose.

Note: (1) Hierarchical techniques may be used to cluster cases or variables while K-means and probabilistic methods only to cluster cases. (2) Hierarchical algorithms can be applied for small and moderate sample sizes. K-means algorithms require at least a moderate sample size while probabilistic algorithms require large sample size.

Hierarchical algorithms

Hierarchical algorithms involve a concept of ordering motivated by the number of observations that can be combined at a time or the assumption that the distance between two observations or clusters is not statistically different from 0.

Hierarchical algorithms can be classified into *divisive algorithms* and *agglomerative algorithms* (15,20,21).

Divisive methods start with the assumption that all objects are part of a single cluster. The algorithm splits this large cluster step by step until each object is a separate cluster.

Agglomerative methods start inversely, are bottom-up procedures, and ended when all observations are combined in one cluster. Most common agglomerative methods are:

- single linkage (nearest neighbor approach) – The method works in the following way: At the first step each cluster consists of one object. At next step we agglomerate those two observations that have the shortest distance. A third observation, which has the next least distance, is added to the two observation cluster to create a third observation cluster or is combined with another observation to form a two observation cluster. The clusters are combined step by step. In each step those two clusters with the smallest dissimilarity or the highest similarity are merged. Iteration continues until all objects are in one single cluster. Single linkage leads to chaining and may result in too few large and heterogeneous clusters,
- complete linkage (furthest neighbor approach) – The method is similar to single linkage except that this is based on maximum distance. This result in a very strong definition of the homogeneity of clusters: The largest dissimilarity between all objects of one cluster should be less than a certain value. The farthest neighbour of

each object should have a distance less than a certain value. Complete linkage results in dilatation and may produce too many clusters.

Note: Complete linkage and single linkage are extreme procedures with completely different properties.

- *average linkage within groups* (within-groups linkage) is the mean distance between all possible inter- or intra-cluster pairs. The method is based on the fact that the average distance between all pairs in the resulting cluster is made to be as small as possibile. This method is therefore appropriate when the research purpose is homogeneity within clusters. This method try to avoid the effects mentioned above (chaining and dilatation).
- *centroid method.* In this situation the distance between two clusters is determined as the difference between centroids. Cluster to be merged is the one with the smallest sum of Euclidean distances between cluster means for all variables,
- *median method* (incremental sum of squares method). This method is similar with the centroid method but included weighting to control the differences in clusters sizes. It also uses Euclidean distance as the proximity measure,
- *Ward's method*. This one is a method distinct from all other methods presented above because it uses an analysis of variance approach to evaluate the distance between clusters (it attempts to minimize the sum of squares of any two clusters).

All these methods differ in the way similarities or dissimilarities are re-computed after two clusters are merged. The last three methods (centroid, median and Ward) have a different approach based on two important assumptions:

1. there is a data file to work with (all other methods require only a dissimilarity or similarity matrix that can be computed from a data file but that can be observed directly too),
2. clusters can be described by their centres (means in the variables).

In these methods the centres are computed step by step based on minimization or maximization of a certain criteria. Ward´s method minimises the within sum of squares, the centroid method and median method select in each step those clusters whose centres are closest. They differ in the way the centres are calculated. The methods are primarily designed for clustering cases. The methods require squared Euclidean distances, in other words are applied to interval-scaled variables or variables that can be treated as interval-scaled and the distances are weighted implicitly. A larger distance in one variable has a higher weight than small distances in many variables.

Evaluation of the results

The answer to the question "how many clusters we found?" is quite difficult regardless the clustering method applied. No one of the methods usually do not result in a unique solution.

In most of the cases the number of clusters is determined on the basis of a *dendrogram* (called also hierarchical tree diagram) by counting the number of clusters that combine objects at a convenient distance level (12).

Another frequently used method is the *scree test*. In this case a graphic is constructed such as the x-axis contains the number of clusters, and the y-axis the agglomeration levels. A sharp increase in the agglomeration schedule results in an elbow knick.

Another important issue regarding the evaluation of the outcomes of a cluster analysis concerns the stability of a cluster solution. Different methods can provide different clusters. A cluster solution is said to be stable, if a small modification of the method specified and the data used does not change the results too much. A stability analysis of methods is usualy based on modifications of clustering techniques and dissimilarity (resp. similarity) measure (22) (Example 3).

Example 3.

A common interest in epidemiological practice is to look for dissimilarities among geographical or economical regions regarding a certain illness. Based on the data provided by World Health Organization (WHO) database we apply hierarchical cluster analysis to group countries for which standardized death rate on diabetes is available. The data we are working with are from 2005, for males of 0-64 ages. In this framework we apply three of the methods discussed above (Ward method, single linkage and complete linkage). Squared Euclidian distance is used as dissimilarity measure. The outcomes are presented in Figures 5, 6 and 7.

Analysing the dendograms (Figures 5, 6 and 7) we see that all three methods involved provide the same number of clusters (seven) but moderatly different solutions.

The problem became more complex if beside standardized death rate on diabetes we include in the analysis other variables such as: total health expenditure as % of gross domestic product (GDP), health at current prices (% of total household consumption expenditure). In this situation standardization is requested (23).

An essential topic which should also be considered refers to the assessment of clusters validity, more specific: internal, relative and external validity of the clusters.

The assumptions that has to be checked in the analysis of internal validity are:

1. The clusters should be homogenous.
2. The clusters should be different in structure.
3. The classification should be able to explain the variation in the data.

The analysis of the relative validity implies the examination of the following constraints (24,25):

4. The classification should be better than the null model that assumes no clusters are present.
5. The classification should be better than other possible classifications.

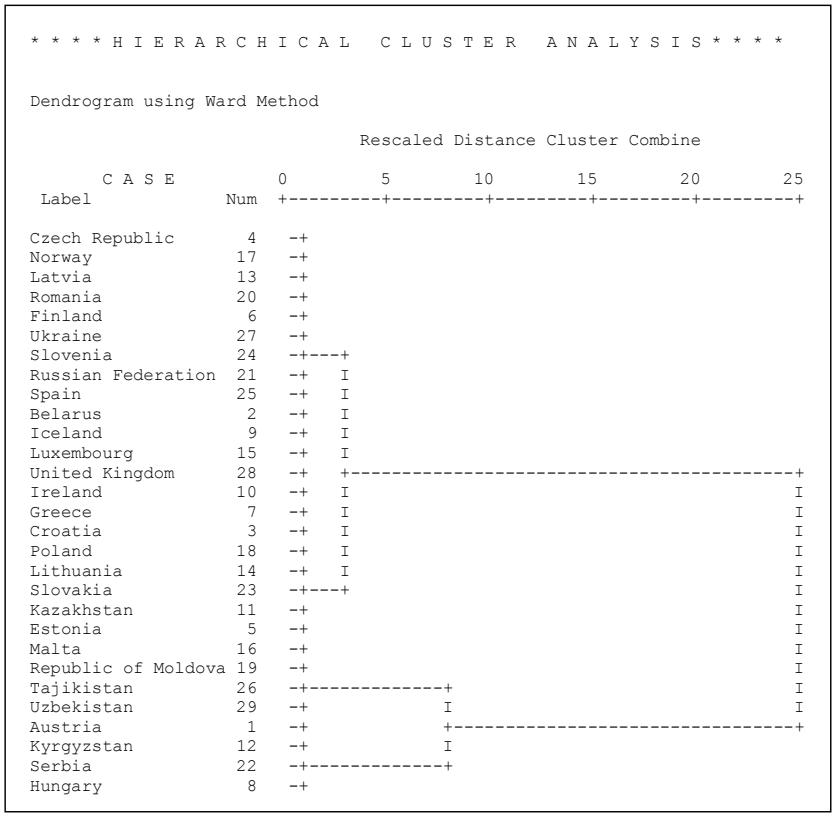


Figure 5. Hierarchical cluster analysis based on Ward method.

Example 3. Cont.

According to Figure 5, using Ward method the solution is composed of the following seven clusters:

1. *Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation*
2. *Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia*
3. *Cluster III: Kazakhstan, Lithuania, Slovakia, Poland, Croatia*
4. *Cluster IV: Malta, Estonia, Republic of Moldova*
5. *Cluster V: Austria, Kyrgyzstan, Serbia*
6. *Cluster VI: Hungary*
7. *Cluster VII: Tajikistan, Uzbekistan*

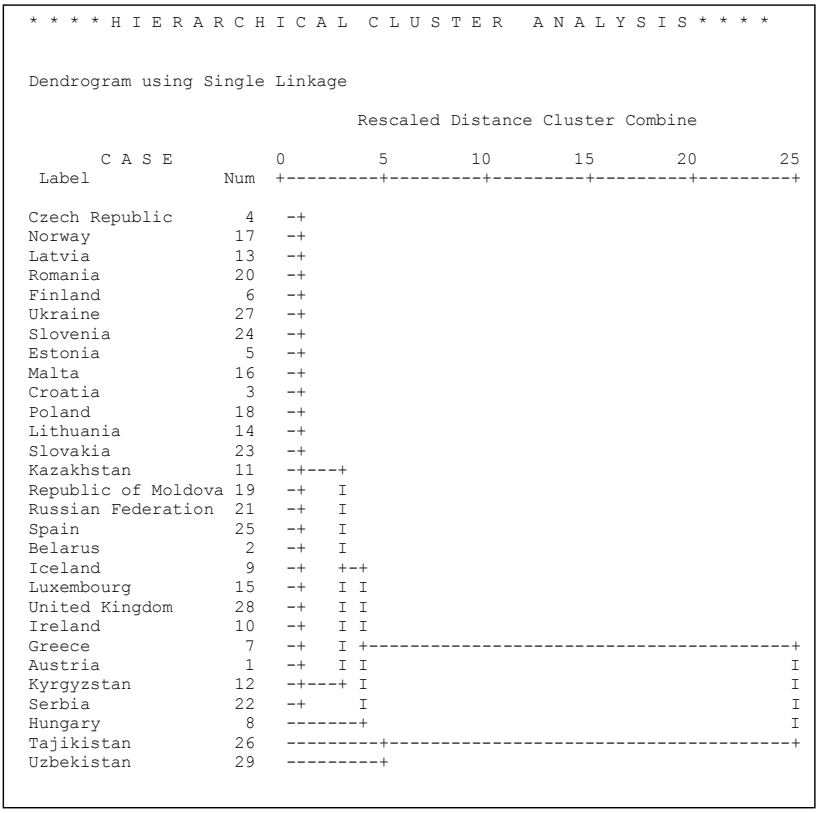


Figure 6. Hierarchical cluster analysis based on single linkage method.

Example 3. Cont.

Analyzing the Figure 6 it can be seen that the single linkage method provided the folowing seven clusters solution:

1. *Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation*
2. *Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia, Kazakhstan, Lithuania, Slovakia, Poland, Croatia, Malta, Estonia*
3. *Cluster III: Republic of Moldova*
4. *Cluster IV: Austria, Kyrgyzstan, Serbia*
5. *Cluster V: Hungary*
6. *Cluster VI: Tajikistan*
7. *Cluster VII:Uzbekistan*

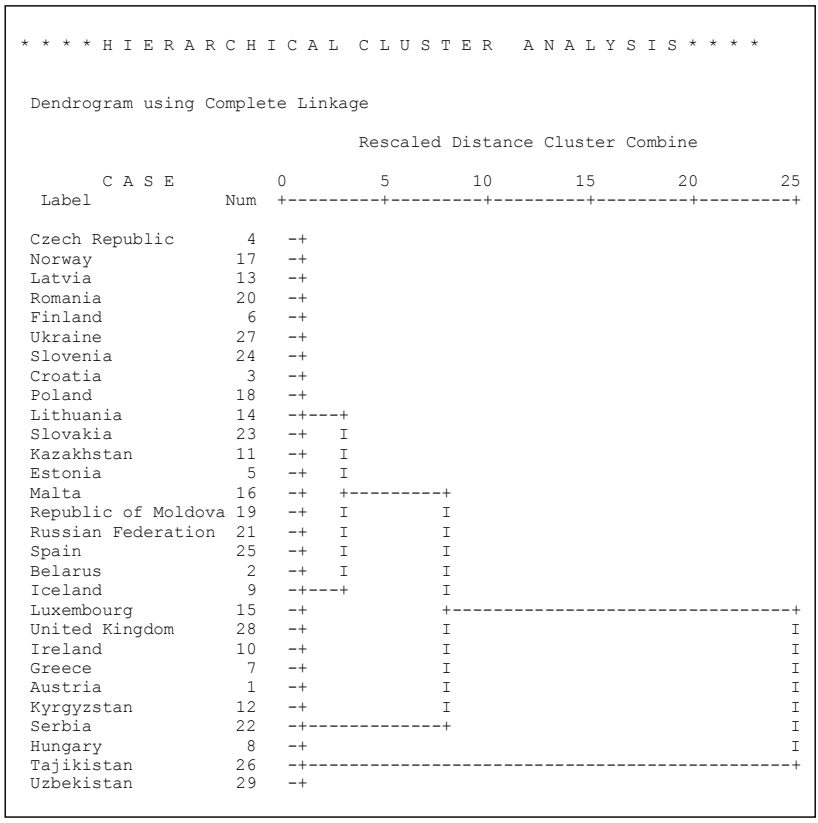


Figure 7. Hierarchical cluster analysis based on single linkage method.

Example 3. Cont.

Complete linkage method provided the folowing seven clusters solution:

1. *Cluster I: Greece, UK, Luxembourg, Ireland, Belarus, Iceland, Spain, Russian Federation*
2. *Cluster II: Latvia, Romania, Norway, Czech Republic, Finland, Ukraine, Slovenia, Kazakhstan, Lithuania, Slovakia, Poland, Croatia*
3. *Cluster III: Malta, Estonia, Republic of Moldova*
4. *Cluster IV: Austria, Kyrgyzstan, Serbia*
5. *Cluster V: Hungary*
6. *Cluster VI: Tajikistan*
7. *Cluster VII:Uzbekistan*

In the case of external validity it has to be analysed if the clusters correlate with the external variables that are known to be correlated with the classification and that are not used for clustering.

The last step but not less important in clustering analysis refers to the clusters interpretability and in regard of this it has to considered that for a succeful analysis clusters need to have a substantive interpretation (26).

K-means clustering

K-means clustering is a top-down procedure which belongs to a more general group of clustering techniques known as partitioning or optimization methods.

The job in this type of clustering is to partition a data set into *k* disjoint sets of objects such that the objects within each set are as homogeneous as possible. Homogeneity is captured by an appropriate *score function* such as minimizing the distance between each object and the centroid (average) of the cluster to which it is assigned. Often the *centroid* belonging to a cluster is considered to be a representative point for that cluster, and there is no explicit statement of what sort of shape of cluster is being sought. Maximizing (or minimizing) the score function is a computationally *search* problem. Iterative algorithms based on local search are particularly common for cluster analysis (27).

The general idea in this type of clustering is to start with a randomly chosen clustering of the objects, then to reassign objects so as to give the greatest increase (or decrease) in the score function, then to recalculate the updated cluster centers, to reassign points again, and so forth until there is no change in the score function or in the cluster memberships.

This approach has the advantage of being simple and guaranteeing at least a local maximum (minimum) of the score function. The major drawback of the search algorithms is that we do not know how good the clustering that it converges to is relative to the best possible clustering of the data (the global optimum for the score function being used).

The classical iterative algorithm used to find the optimal partition is called *k* – means clustering, which has close connection to the EM algorithm, about which we will talk later. In this case, we have to fix in advance the number of clusters we require (this is typical problem of many clustering algorithms). The number of clusters is represented by parameter *k*.

There are several variants of the *k* -means algorithm (19). The first step of the basic version involves the chosen at random of *k* objects to represent initial cluster centers. In the next step all objects are assigned to the nearest cluster center according to Euclidean distance, the mean value of the objects in each cluster is computed to form its new cluster center, and iteration continues until there are no changes in the clusters.

When the number of clusters cannot be specified in advance, we can apply an incremental clustering method based on a hierarchical grouping of objects which use a measure of cluster quality or, a statistical clustering method based on a mixture model of different probability distributions, one for each cluster (28,29).

Suppose we are using *k* -means but do not know the number of clusters in advance. A solution is to test different possibilities and see which is best, which one minimizes the total squared distance of all objects to their cluster center. In this

context, a possible strategy is to start from a given minimum, *k=1*, and worked up to a small fixed maximum, using cross-validation to find the best value. Another strategy is to begin by finding few clusters and determining whether it is worth splitting them. Suppose *k=2*, perform *k* -means clustering until it terminates, and then consider splitting each cluster.

One way to split a cluster is to make two new seeds: a seed one standard deviation away from the clusters center in the direction of its greatest variation, and the other seed the same distance in the opposite direction. Next apply *k* -means to the points in the cluster with these two new seeds. If the split can be retained, try splitting each cluster further. The process continues until no splits remain.

Partition-based methods of cluster analysis begin with a specified number of clusters and search through possible allocations of objects to clusters to find an allocation that optimizes some clustering score function. A large variety of score functions can be used to determine the quality of clustering and a wide range of algorithms has been developed to search for a good partition.

In the following section we will present some basic features of score functions.

Score functions

In order to define the clustering score function we need to look at *within cluster variation* and *between cluster variation* of a clustering *C*.

The *within cluster variation* measures how compact the clusters are, while the *between cluster variation* looks at the distances between different clusters.

Suppose that we have selected *cluster centers*, noted r_k, from each cluster. This can be a designated representative data point (object) that is defined to be "central" in some manner. If the input objects belong to a space where means have sense, we can use the centroid of the objects in the cluster C_k as the cluster center and r_k, will be defined by formula (Equation 32):

$$r_k = \frac{1}{n_k} \sum_{x \in C_k} x \qquad \textbf{Equation 32.}$$

n_k *= the number of objects in the* k^{th} *cluster.*

Within cluster variation can be defined as the sum of square of distances from each point to the center of the cluster (Equation 33):

$$within\_c(C) = \sum_{k=1}^{K} within\_c(C_k) = \sum_{k=1}^{K} \sum_{x(i) \in C_k} d(x, r_k)^2 \qquad \textbf{Equation 33.}$$

$d(x, r_k)$ *= Euclidian distance*

$within\_c(C)$ *= the within-cluster sum-of-squares*

Between clusters variation can be defined by the distance between cluster centers (Equation 34):

$$between\_c(C) = \sum_{1 \le j < k \le K} d\left(r_j, r_k\right)^2 \qquad \textbf{Equation 34.}$$

In this framework, the score function of a clustering C can then be defined as a monotone combination of the factors $within\_c(C)$ and $between\_c(C)$, such as the ratio $between\_c(C) / within\_c(C)$.

The k - means algorithm, uses the means within each group as cluster centers and Euclidean distance for d to search for the clustering that minimizes the within cluster variation.

If we are given a candidate clustering, it is important to know how difficult is it to evaluate $within\_c(C)$ and $between\_c(C)$. Computing $within\_c(C)$ takes $O(n)$ operations, while $between\_c(C)$ can be computed in $O(k^2)$ operations. Hence, computing a score function for a single clustering requires a pass through the whole data.

Stability analysis

After we get the clusters based on the algortithm described above we have to proceed with a stability analysis. In *K*-means clustering the idea of stability analysis is to check, whether *modifications of methods* or *data* have a negative effect on the results.

In contrast to hierarchical methods, the technique and the distance measure (squared Euclidean distances) are fixed for k-means, in other words they cannot be modified. In this case, the only thing which is not fixed is the starting partition. Therefore, the stability of the results can be tested by modifying the starting partitions. Having in view this aim, the following strategies can be applied:

- generate different random starting partitions, if random starting values are used,
- re-order the cases,
- change the starting values, if centers are entered or computed using a hierarchical technique,
- use different starting procedures (e.g. randomly generated starting values),
- if a classification is stable, the starting procedure should have no influence.

Regarding the data stability, both cases and variables can be analyzed. In this framework cases stability can be verified considering the subsequent steps:

- divide the data set in *M* sub-datasets,
- run for each sub-dataset *k*-means and save the cluster centres, and
- compare the cluster centres.

M is usually set equal to 2. This method only allows comparing cluster centers. If classifications are to be compared, the strategy has to be modified in the following way:

- divide the population in *M* (usually *M* = 2) subpopulations,
- use one subpopulation as reference population and compute the cluster centres for this reference population,
- run two cluster analysis for the other subpopulation: an 'ordinary' *k*-means analysis and a 'confirmatory' *k*- means analysis with fixed centres. Use the centres of the solution of the previous step, and
- compute an index to compare the classifications within each subpupolation analysed above (e.g. Rand index).

Variables stability can be tested by adding randomly distributed variables.

Instead of testing stability attempts have been made to change the algorithm in order to find a more robust classification, a best solution or to compute a partition of partitions.

Experience proves that a more robust classification can be obtained using the following methods:

- eliminating outliers in a first stage (the outliers may be assigned to clusters in the next stage) (30,31),
- using the city block metric instead of squared Euclidean distances, because it is less sensitive towards outliers, and
- weighting variables automatically according to their contribution to separate the clusters.

In this way, variables with a high proportion of random noise should be eliminated from analysis.

Note: *k*-means cluster analysis is very sensitive to outliers and is recommended to remove them before starting the analysis.

Two-step clustering

Two-step clustering is a method preferred for *large data sets* and when *categorical variables with three or more levels* are involved. The algorithm follows the subsequent steps:

1. Pre-clusters are identified.
2. The pre-clusters identified in the first step are treated as single cases and clustered hierarchical.

Note: *k*-means cluster analysis and two-stage cluster analysis usually generate different solutions.

Probabilistic clustering

In the case of probabilistic clustering the problems of *k*-means clustering are avoided:

- the problem of incommensurability does not occur – variables of different measurement levels and different scale units can be analysed without any transformation of the variables.

- each case is assigned probabilistic to a cluster — usually to the cluster from which it is most likely to have come.
- the model has a statistical basis (the typically score function is likelihood of the data).

From a probabilistic perspective the goal of clustering is to find the most likely set of clusters given the data. The foundation for statistical clustering is a statistical model called finite mixture.

Finite mixture

A mixture is a set of k probability distributions, representing k clusters that govern the variables values for members of those clusters.

Each distribution gives the probability that a particular case would have a certain set of variables values if it were known to be a member of that cluster. Each cluster has a different distribution. Any particular case belongs to one and only one of the clusters, but it is not known which one. In the end the clusters are not equally likely (32).

The simplest finite mixture situation occurs when there is only one numeric variable, which has a normal distribution for each cluster, with different means and variances. The clustering problem is to take a set of objects and a specified number of clusters and work out each cluster mean and variance and the population distribution between the clusters. The mixture model combines several normal distributions and its probability density function looks like a "mountain range with a peak for each component" (29).

Suppose we have three clusters A, B and C and each has a normal distribution with means and standard deviations: μ_A and σ_A for cluster A, μ_B and σ_B for cluster B, μ_C and σ_C for cluster C. Samples are taken from these normal distributions using cluster A with probability p_A, cluster B with probability p_B and, cluster C with probability p_C such as $p_A + p_B + p_C = 1$. We also suppose that we have a set of objects and want to determine the parameters that characterize the model. If we knew which of the three distributions each object (case) came from, the finding of the parameters is easy, the only thing we have to do is to estimate the mean and standard deviation for the three clusters separately, using the clasical formulas (Equations 35 and 36):

- for the mean:

$$\mu = \frac{x_1 + x_2 + \ldots + x_n}{n} \qquad \textbf{Equation 35.}$$

- for the standard deviation:

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_n - \mu)^2}{n-1} \qquad \textbf{Equation 36.}$$

If we knew the parameters finding the probabilities that a given object comes from each distribution would be easy. Given an object x the probability that it belongs to cluster A is given by formula (Equation 37):

$$Pr[A|x] = \frac{Pr[x|A] \cdot Pr[A]}{Pr[x]} \qquad \textbf{Equation 37.}$$

The final result of the method is not a particular cluster but rather the probabilities with which case x belongs to cluster A, B and C.

Expectation maximization algorithm

The general objective of the expectation maximization (EM) algorithm is to discover clusters in observations (or variables) and to assign those observations to the clusters. An application for this type of analysis is marketing segmentation - a marketing research study in which a number of consumer behaviors related variables are measured for a large sample of respondents. The purpose of the study is to detect groups of customers that are similar to each other when compared to respondents that "belong to" other clusters. In addition to identifying such clusters, it is also interesting to determine how the clusters are different, to determine the specific variables or dimensions that vary and how they vary in regard to members in different clusters (32,33).

While the *k*-means clustering refers to the fact that given a fixed number of *k* clusters observations are assigned to those clusters so that the means across clusters (for all variables) are as different from each other as possible, the *EM* algorithm extends this approach to clustering in two significant ways:

1. Instead of assigning cases or observations to clusters to maximize the differences in means for continuous variables, the *EM* clustering algorithm computes probabilities of cluster memberships based on one or more probability distributions. In this situation the goal of the clustering algorithm is to maximize the overall probability of the data, given the (final) clusters.
2. Unlike the classic implementation of *k*-means clustering, the general *EM* algorithm can be applied to both continuous and categorical variables .

The basic approach and logic of this clustering method is as follows: suppose we measure a single continuous variable in a large sample of observations and that the sample consists of two clusters of observations with different means and standard deviations. The goal of *EM* clustering is to estimate the means and standard deviations for each cluster so as to maximize the likelihood of the observed data.

The results of *EM* clustering are different from those computed by *k*-means clustering. The latter will assign observations to clusters to maximize the distances between clusters. The *EM* algorithm does not compute actual assignments of observations to clusters, but classification *probabilities*.

Finding the right number of clusters in *k*-Means and EM clustering: v-fold cross-validation

An important question that needs to be answered before applying the *k*-means or *EM* clustering algorithms is how many clusters there are in the data. This is not known *a priori* and there is no unique answer.

An estimate of *k* can be obtained from the data using the method of cross-validation. Remember that the *k*-means and *EM* methods will determine cluster solutions for a particular user-defined number of clusters. The *k*-means and *EM* clustering techniques can be optimized and enhanced for typical applications in data mining (13).

To determine *k* we can use v-fold cross-validation algorithm for automatically determining the number of clusters in the data (34).

This algorithm is useful in all general "pattern-recognition" tasks - to determine the number of market segments in a marketing research study, the number of distinct spending patterns in studies of consumer behavior, the number of clusters of different medical symptoms, the number of weather patterns in meteorological research, etc.

The general idea of the method is to divide the overall sample into a number of *v* folds. The same type of analysis is then successively applied to the observations belonging to the training sample, and the results of the analyses are applied to sample *v* to compute some index of predictive validity. The results for the *v* replications are aggregated to yield a single measure of the stability of the respective model.

Software Products

Through a large variaty of software packages available for developing cluster analysis, the most popular are: SPSS, CLUSTAN, ALMO, WEKA.

EXERCISE

Task 1

What role does variable and data selection play in cluster analysis? Please give a solid argumentation of your answer. Provide examples.

Task 2

Is standardization needed in cluster analysis? Why? Please give a solid argumentation of your answer. Provide examples.

Task 3

How can we use clustering methods to identifying the outliers in a dataset?

Task 4

What happens if in cluster analysis, distance measure is replaced by similarity measure?

Task 5

What is the most important problem with non-hierarchical algorithms?

Task 6

List the factors that affect the stability of clustering solution. Select three of them and propose a method that drives to the right solution.

REFERENCES

1. Ritter G, Gallegos MT. Outliers in statistical pattern recognition and an application to automatic chromosome classification. Pattern recognition letters. 1997;18:525-39.
2. Waller LA, Jacquez GM. Disease models implicit in statistical tests of disease clustering. Epidemiology. 1995;6:584-90.
3. Rezaeian M, Dunn G, Leger St, Appleby L. Geographical epidemiology, special analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007;61:98-102.
4. Sjurdur FO, Martuzzi M., Elliott P. Cluster analysis and disease mapping- why, when, and how? A step by step guide. Br Med J. 1996;313:863-66.
5. Knox EG. Detection of clusters. In: Elliot P, editors. Methodology of enquiries into disease clustering. London: Small Area Health Statistic Unit; 1989. p.17-20.
6. Elliott P, Martuzzi M, Shaddick G. Spatial statistical methods in environmental epidemiology. Stat Meth Med Res. 1995;4:139-61.
7. Jyun-de Wu, Milton DK, Hammond SK, Spear RC. Hierarchical cluster analysis applied to workers exposures in fiberglass insulation manufacturing. Ann Occupat Hygiene. 1999;43:43-55.
8. Kaufman L, Rousseeuw PJ. Finding groups in data: An introduction to cluster analysis. New York: J. Wiley & Sons; 1990.
9. Kiers HAL, Rasson JP, Groenen PJF, Schader M., editors. Data analysis, classification, and related methods. Proceedings of the 7th Conference on the International Federation of Classification Societies (IFCS-2000), University of Namur, Belgium; 2000 11-14 July; Berlin, Heidelberg: Springer; 2000.
10. Bezdek JC, Keller J, Krisnapuram R, Pal NR, editors. Fuzzy models and algorithms for pattern recognition and image processing. Vol 5 Dordrecht: Kluwer Acad. Publ.; 1999.
11. Jain AK, Dubes RC. Algorithms for clustering data. Englewood Cliffs, NJ: Prentice-Hall; 1988.
12. Aldenderfer M, Blashfield R. Cluster analysis. San Francisco: Sage Publications; 1984.
13. Witten IH, Frank E. Data Mining: Practical machine learning tools and techniques. 2nd ed. San Francisco: Elsevier; 2005.
14. Little R, Rubin D. Statistical analysis with missing data. New York: J. Wiley & Sons; 1987.
15. Hartigan JA. Clustering algorithms. New York: J. Wiley & Sons; 1975.
16. Everitt B, Landau S, Leese M. Cluster analysis. 4th ed. London: Hdder Arnold; 2001.
17. Kendall M. Multivariate Analysis. 2nd ed. London: Charles Griffin & Company; 1980.
18. Mardia K, Kent J, Bibby J. Multivariate analysis. London: Academic Press; 1979.

19. Massart DL, Plastria F, Kaufman L. Non-hierarchical clustering with MASLOC. J Pattern Recognat Soc. 1983;16:507-16.
20. Banks D, House L, McMorris FR, Arabie P, Gaul W, editors. Classification, clustering and data mining applications. Proceedings of the Meeting of the International Federation of Classification Societies (IFCS), Illinois Institute of Technology; 2004 15-18 July; Chicago. New York: Springer; 2004.
21. Cherkassky V, Mulier F. Learning from data: concepts, theory and methods. New York, NY: J. Wiley & Sons; 1998.
22. Michie D, Spiegelhalter D, Taylor C, editors. Machine learning, neural and statistical classifications. New York: Ellis Horwood; 1994.
23. Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. J Classificat. 1988;5:181-204.
24. Gordon AD, Boca Raton FL. Classification. London: Chapman and Hall; 1999.
25. Hand D. Discrimination and classification. Chichester: J. Wiley & Sons; 1981.
26. Forgy E. Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications. Biometrics. 1965;21:768-9.
27. McLachlan GJ. Discriminant analysis and statistical pattern recognation. New York: J. Wiley & Sons; 1992.
28. Hartigan JA, Wong MA. A k-means clustering algorithm. Applied Statistics. 1979;28:100-8.
29. Coleman DA, Woodruff DL. Cluster analysis for large datasets: An effective algorithm for maximizing the mixture likelihood; 2000. J Comput Graph Statist. 2000;9:672-88.
30. Barnett V, Lewis T. Outliers in statistical data. 3rd ed. Chichester: J. Wiley & Sons; 1994.
31. Wild CJ, Seber GAF. Chance encounters: A first course in data analysis and inference. New York: J. Wiley & Sons; 2000.
32. Jordan M, Jacobs R. Hierarchical mixtures of experts and the EM algorithm. Neural Comput. 1994;6:181-214.
33. Neal R, Hinton GE. A view of the EM algorithm that justifies incremental, sparse, and other variants. In: Jordan M, editor. Learning in Graphical Models. Dordrecht: Kluwer Acad. Publ.; 1998. p.355-68.
34. Zhang P. Model selection via multifold cross-validation. Ann Statist. 1993;21:299-311.

RECOMMENDED READINGS

1. Everitt B, Landau S, Leese M. Cluster analysis. 4th ed. London: Hdder Arnold; 2001.
2. Jyun-de Wu, Milton DK, Hammond SK, Spear RC. Hierarchical cluster analysis applied to workers exposures in fiberglass insulation manufacturing. Ann Occupat Hygiene. 1999;43:43-55.
3. Rezaeian M, Dunn G, Leger St, Appleby L. Geographical epidemiology, special analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007;61:98-102.
4. Sjurdur FO, Martuzzi M., Elliott P. Cluster analysis and disease mapping- why, when, and how? A step by step guide. Br Med J. 1996; 313: 863-66.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**TOTAL RISK ASSESSMENT**
Module: 1.3.3	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Mariana Dyakova**, MD, PhD, Associate Professor Faculty of Public Health, Medical University Sofia, Bulgaria **Emilia Karaslavova,** MD, PhD, Assistant Professor Department of social medicine and health management, Medical University Plovdiv, Bulgaria **Hristo Mateev,** MD University Hospital "Lozenetz", Sofia, Bulgaria
Address for correspondence	**Mariana Dyakova** Faculty of Public Health, Medical University Sofia 8 "Bialo more" str., 1527 Sofia, Bulgaria E-mail: dr_dyakova@abv.bg
Keywords	Risk factor, risk assessment, total risk, absolute risk, prevention, chronic diseases, cardiovascular diseases, primary prevention
Learning objectives	After completing this module students should: • know the characteristics and use of global (total/absolute) risk; • be familiar with different methods of total risk assessment (TRA); • understand the advantages and disadvantages of TRA; • be familiar with the specific implementation of TRA for chronic non-communicable disease prevention and treatment; • be aware of European populations' specificities and recommendations for TRA for one case, for example of cardiovascular diseases; • be able to analyse the results from TRA and consider it in health planning and management.
Abstract	During the last decades the field of risk analysis has grown rapidly, focusing on the identification and quantification of threats to human health called risk assessment. Different health risks act jointly and various risk factors interact to cause certain outcome (i.e. death). The causal web model of disease causation has been developed, reflecting the fact that risk factors often increase not only the risk of disease, but also levels of other risk factors. It is the combination and interaction of multiple causes that determines the absolute risk of an individual. The concept of TRA has been introduced, as a measure. TRA for a certain disease is essential for clinical practice as well as for public health policy. The individual TRA and the following risk stratification of a population appear to be relatively simple and cost-effective.
Teaching methods	An introductory lecture gives the students first insight in characteristics of TRA. Case studies are presented. Students are asked to find other published materials, analyze and compare them through individual and group work.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 20%/80%; • facilities: a lecture room, a computer room; • equipment: computers, LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Design and presentation of a Model for epidemiological study, involving TRA.

TOTAL RISK ASSESSMENT

Mariana Dyakova, Emilia Karaslavova, Hristo Mateev

THEORETICAL BACKGROUND

Assessing health risks

Preventing death, disease, and injury requires systematic assessment and reduction of their causes e.g. focusing on risks to health. During the last several decades the interest in health determinants and their risk factors has intensified, included new perspectives and faced new challenges. The field of risk analysis has grown rapidly, focusing on the identification, quantification and characterization of threats to human health and the environment – a set of activities broadly called *risk assessment* (1). While a very long interest in comparing risks posed by different threats to health has existed, formal frameworks have been developed relatively recently. Governments, ensuring overall population health, need information from risk assessments that are comprehensive, reliable, relevant and timely. However, such information, which is crucial to prioritization, is typically limited. Many aspects are relevant in prioritizing strategies to prevent / reduce risks to health (1): the extent of the threat posed by different risk factors, the availability of cost-effective interventions, and societal values and preferences are particularly important. Still most of the scientific efforts and healthcare resources are directed towards treating disease - the "rule of rescue" still dominates (2), while assessments of risk factors estimate the potential of prevention.

An effective risk assessment should have a well-defined scope, which depends on the purpose of the analysis, as the range of risks to health is almost limitless. In order to get the necessary comprehensive, consistent and applicable data from a risk assessment study and provide policy-relevant information, a number of issues should be considered (1):

1. Risks need to be identified and studied comprehensively irrespective of factors such as their place in a causal chain or the methods used for their analysis.
2. Risks to health do not occur in isolation. The chain of events leading to an adverse health outcome includes both proximal and distal causes. Research into the different levels of risks should be seen as complementary.
3. Understanding risk factors requires contributions from different areas of health impact: environmental, communicable, noncommunicable, injury etc as well as different intellectual and scientific tools and methods.
4. Many risks to health are widely distributed in the population, with individuals differing in the extent of their risk rather than whether they are at risk or not. Binary categorization into "exposed" and "unexposed" substantially underestimate the importance of continuous risk factor - disease relationships.
5. The impact of each risk factor should be assessed in terms of a "common currency" that incorporates loss of quality of life as well as loss of life years.
6. Risk factor does have a negative connotation, but ideally a risk assessment should include a range of protective as well as hazardous risk factors.
7. In recent years, a life-course approach to the study of health and illness has been widely accepted, which suggests that exposure to disadvantageous

experiences and environments accumulates throughout life and increases the risk of illness and premature death (3).

8. It is important in any risk assessment to review quantitatively the best available evidence for both "definite" and "probable" risks. Estimation of the potential impact of a health hazard can never wait until perfect data are available, since that is unlikely to occur.
9. Timeliness is essential. We should not rely on assertions of uncertainty or certainty when, in fact, there are different degrees of uncertainty and disagreement about tolerable thresholds.
10. Risk assessments to date have typically used only attributable risk estimates, basically addressing the question "what proportion of current burden is caused by the accumulated effects of all prior exposure?" However, often a more policy - relevant question is "what are the likely future effects of partial removal of current exposure?" Two key developments are therefore needed: an explicit focus on future effects and on less-than-complete risk factor changes.

Multifactorial causality and total risk assessment

Different risks to health act jointly and various risk factors interact to cause certain outcome (death, disease or injury) – *"a multiple causality of disease"* (4). The chain of events, leading to a specific (adverse) outcome, includes proximal (direct) and distal (indirect) factors, as well as a number of intermediate ones. The impact of a single risk factor is often summarized as the proportion of disease caused by, or attributable to, that risk factor. When several risks (risk factors) affect the same disease or injury outcome, then the net effect can be less or more than the sum of their separate effects (i.e. more than 100%). The size of these joint effects depends principally on the amount of prevalence overlap and the biological effects of joint exposures (4). Separate estimation of the effects of individual risk factors does not typically take into account the effect of changes on the levels of other risk factors. Thus the *"causal web model of disease causation"* has been developed, reflecting the fact that risk factors often increase not only the risk of disease, but also levels of other risk factors.

It is the combination and interaction of multiple causes that determines the *absolute risk of an individual,* i.e. the probability of event (disease) occurrence in a defined period of time. The absolute risk approach (5):

- is counter-intuitive: e.g. benefit from lowering blood pressure does not depend (mainly) on level of blood pressure;
- differs from 'threshold' approaches of the past. It recommends that terms implying thresholds, such as 'hypertension' and 'hypercholesterolaemia' should be abandoned.

By a simple process of aggregation the same logic can be applied at a population level: in populations at high (average) absolute risk of a disease, all reversible risk factors should be lowered, and not because their risk factor levels are higher, but because the benefits of lowering them will be greater (6). Thus the absolute risk approach is to be more extensively applied at both an individual (clinical) and population (public health) level.

The concept of *"global/total risk" assessment* has been introduced, as a measure for individual, respectively population absolute risk. Total (global, absolute) risk assessment for a certain disease is essential for clinical practice as well as for public health policy, as multiple risk factors confer greater risk than the sum of their components (5,6).

Multicausality and total (absolute) risk assessment have important implications for prevention. The individual total (global) risk assessment and the following risk stratification of a population (in low, moderate and high risk groups) appear to be relatively simple and cost-effective. On individual level it could be very effective for outcome prognosis and setting treatment objectives. On population level it offers opportunities to tailor prevention. The key message of multicausality is that different sets of interventions can produce the same goal, with the choice of intervention being determined by such considerations as cost, availability and preferences. Thus, prevention should not wait until further causes are elucidated as in the close future we will not know all the causes of disease, or how to avoid the entire disease burden attributable to genetic causes. Nonetheless, multicausality means that in many cases considerable gains can be achieved by reducing the risks to health that are already known (1).

Total risk assessment of chronic non-communicable diseases

It is already clear that all known to men diseases have a multifactorial causation. For an infectious disease development, not only the contact with a contagious agent (bacteria, virus etc) is important, but also its quantity and distribution in the surrounding environment, the susceptibility of the organism, the duration of exposure etc. However, usually the most effective and efficient way for communicable disease prevention is interruption of the multicausal chain at the level of agent transmission, e.g. treatment of the infected individuals (fig 1). Obviously this is not the case when non-communicable diseases are concerned. Example for a causal web model of a chronic non-infectious disease can be illustared also by the model, given in fig 1: distal socioeconomic causes include income, education, occupation, all of which affect levels of proximal factors such as inactivity, diet, tobacco use and alcohol intake; these interact with physiological and pathophysiological causes, such as blood pressure, cholesterol levels and glucose metabolism, to cause cardiovascular disease (CVD) such as coronary heart disease (CHD) or stroke. The sequelae include death or disability, such as angina or hemiplegia. Thus, prevention of non-communicable diseases appears to be effective when:

- as much as possible risk factors have been identified and modified (reduced);
- it is timely, e.g. realized as early as possible, as usually these diseases (or their preceding stages) have silent, long-term development before any measurable symptoms appear; and
- risk factor reduction is as complete as possible (smoking cessation, blood pressure control etc) and as long in life-course as possible (weight loss, everyday physical activity etc).

The proportion of a chronic disease risk that is attributable to a given exposure or risk factor is the proportion that would be avoided if lifetime exposure had been at a more favourable specified level (5,6). For binary harmful exposures such as

cigarette smoking, the attributable disease burden in the population is estimated by comparing the actual burden with what would be expected in an 'unexposed' (non-smoking) population. However, for many important risk factors, zero exposure is not meaningful – e.g. blood pressure, blood cholesterol concentration, adiposity etc. In these cases it is necessary to be explicit about the level (or distribution) of the risk factors with which the effects of the current levels (or distributions) are to be compared. That comparison level is referred to as the counterfactual level. Counterfactual exposure distributions can be more or less extreme in the contrast they provide with the actual distribution of exposure. The most extreme in the Burden of Disease approach is that distribution of a risk factor which confers theoretical minimum risk, i.e. **"**theoretical-minimum-risk exposure distribution" (7)**.**

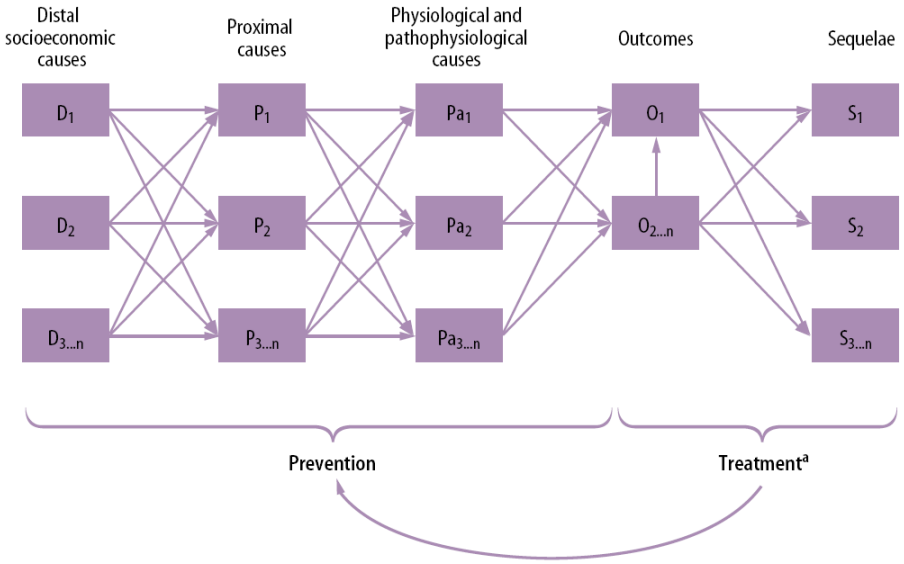


Figure 1. Causal chains of exposure leading to disease. Source: World Health Organization. World Health Report, 2002 (1). LEGEND: a = treatment of infectious disease can lead to prevention of further cases if it interrupts transmission.

The case of cardiovascular diseases total risk assessment

Estimating the global (absolute) risk for cardiovascular diseases (CVD) is a typical and most widely discussed case of total risk assessment of non-communicable diseases. They are also perfect example for multifactorial diseases, where a vast number of known and still obscure risk factors act together. The CVD multicausal web has been extensively studied through the years and lots of the risk factors' interactions and attributable burdens have been identified. The contribution of Law and Wald to understanding of how absolute risks of vascular disease should be managed in individuals is presented in the following statements (8,9):

1. Risk on a logarithmic (proportional) scale tends to be related linearly to the levels of continuously distributed risk factors (such as blood pressure and blood cholesterol concentration). This means, for example, that constant absolute differences in the risk factor produce constant proportional changes in risk, e.g. a 5 mmHg reduction in systolic blood pressure makes a constant percentage reduction in the risk of heart attack irrespective of the level of the blood pressure.
2. Optimal values for vascular risk factors are well below current levels, so virtually everyone could benefit from lower values.
3. Many of the most important determinants of absolute risk are not modifiable, e.g. age, sex and past history of a heart attack or stroke.

The implications of 1, 2 and 3 are that the intensity of efforts to reduce risk should depend mainly on the absolute level of risk and not on the level of particular risk factors. In individuals at high absolute risk (from whatever cause) all reversible risk factors should be lowered (9) – *that is the absolute risk approach.* This again can be extrapolated to population level: in populations at high (average) absolute risk of CVD, all reversible risk factors should be lowered, and not because their risk factor levels are high, but because the benefits of lowering them will be greater.

Historically, patients have been assessed and treated, according to the presence or absence of a particular risk factor, such as hypertension, obesity, diabetes etc. However logic this approach seems to be, it eventually results in two unfavourable consequences: first, patients with low absolute (total) cardiovascular risk are subjected to long-term (sometimes life-long) drug treatment, and second, patients with high absolute (total) cardiovascular risk (but no explicitly high level of any single risk factor) are left untreated. Another very important disadvantage of the *single-risk approach* to CVD is that is, in most cases, not cost-effective and thus it is inappropriate for low- and middle-resource settings.

As it was mentioned above, any government is aiming at achieving the best possible health results for most of the population. In this aspect, the total (absolute) risk approach for CVD prevention, especially applied on population level appears to be most relevant. So, current recommendations on the prevention of coronary heart disease in clinical practice stress the need to base intervention on an assessment of the individual's total burden of risk rather than on the level of any particular risk factor.

Models (scores) of total risk assessment for CVD

The first model of a global (total) individual cardiovascular risk and its estimation is published in 1991. This coronary prediction algorithm is based on the famous Framingham Heart Study (10,11), a long-term cohort study, started in 1948. It followed 5,209 adults (ages 30–62) from Framingham, Massachusetts, USA who had not yet developed overt symptoms of cardiovascular disease and who had not suffered a heart attack or stroke. In 1971 the study enrolled a second-generation group to participate in similar examinations. It consisted of 5,124 of the original participants' adult children and their spouses. The second study was named the Framingham Offspring Study. In 2002, recruitment began for a third generation of participants – the children of the 1971 study group. The first phase of this Third Generation Study

was completed in 2005, with 4,095 participants. The initially proposed formula has been developed and periodically innovated, using the continuously coming data from the Framingham Heart Study and its extension - the Framingham Offspring Study (12). The proposed risk function estimates the probability for development of coronary heart disease (CHD) - myocardial infarction, angina pectoris, coronary insufficiency and coronary death over the course of 10 years. Separate score sheets are used for men and women and the factors used to estimate risk include age, blood cholesterol (or LDL cholesterol), HDL cholesterol, blood pressure, cigarette smoking, and diabetes mellitus. Relative risk for CHD is estimated by comparison to low risk Framingham participants. In addition to score sheets for men and women, a sample score sheet is provided to illustrate how they can be used. The high risk is considered over 20% (10).

European populations have also been affected by the CVD epidemic, mostly Scandinavian countries (13) as well as Central and Eastern European countries (14), especially in the last decades (15,16). Initially the Framingham Heart Study data have been used as a source for several European risk prediction systems (17-19). They have been incorporated into numerous guidelines and served as a model for the development of Recommendations for coronary heart disease prevention in clinical practice by the European Societies on coronary prevention - the First and the Second Joint Task Force (20-22). However, a number of European studies revealed unambiguously that risk calculators, based on the Framingham formula overestimate the total cardiovascular risk in most European populations. That was firstly found for France, Italy and other Mediterranean countries, considered to have lower coronary heart disease rates and mortality (23,24). Afterwords it was also confirmed for "high coronary heart disease rate" countries, such as Germany, UK and even Scandinavian countries (25-27). Two major German cohort studies - MONICA Augsburg and PROCAM (Prospective Cardiovascular Münster Study) were especially conclusive for that (25). The specificities and limitations of the Framingham total risk calculator and its European variation as well as its disadvantages for implementation in principally all European populations may be summarized as follow (28-31):

1. The risk estimating Framingham score sheets are only for persons without known heart disease.
2. The Framingham Heart Study risk algorithm encompasses only coronary heart disease, not other heart and vascular diseases.
3. The Framingham Heart Study population is a relatively small cohort, almost all Caucasian. The Framingham risk algorithm may not fit other populations.
4. For some of the sex-age groups in Framingham, the numbers of events are quite small. Therefore, the estimates of risk for those groups may lack precision.
5. The Framingham risk score estimates the risk of developing CHD within a 10-year time period. This risk score may not adequately reflect the long-term or lifetime CHD risk of young adults, which is: one in two for men and one in three for women.
6. The presence of any CHD risk factor requires appropriate attention because a single risk factor may confer a high risk for CHD in the long run, even if the 10-year risk does not appear to be high.

7. The Framingham algorithm does not include important risk factors as family (genetic) predisposition, body mass weight (BMI), and social status.
8. Since age is a prominent determinant of the CHD risk score, the 10-year hazards of CHD are, on average, high in older persons. This may over-identify candidates for aggressive interventions. Relative risk estimates (risk in comparison with low risk individuals) may be more useful than absolute risk estimates in the elderly.
9. The applicability of a risk function derived from US data to European populations is highly questionable, as it clearly overestimates absolute risk in countries with lower coronary heart disease rates as well as in those with high coronary heart disease risk.
10. The definition of nonfatal end-points used in the Framingham Study (10) differs from definitions used in most other cohort studies, and from endpoints used in clinical trials. It includes, in addition to non-fatal myocardial infarction, new onset angina and coronary insufficiency (i.e. unstable angina), making it difficult to validate the function with data from other cohort studies, and difficult to relate to the results of therapeutic trials.
11. There is a considerable difficulty in using local data to adjust the model for use in individual European countries.
12. The score derived from this algorithm should not be used in place of a medical examination.
13. The statistical estimations, based on the Framingham cohort, represent and are predictive for the high cardiovascular risk in USA during the 70-ies and 80-ies of the 20th century. As at present CVD is declining in most developed economies, this calculator tends to overestimate the up-to-date CVD total risk.

In the course of these findings a rigorous research started out and a number of absolute risk prediction scores, derived from nationally or regionally conducted epidemiological studies, were developed. One of the first and most widely used European based total risk prediction algorithms was PROCAM, named after the extensive German "PROspective CArdiovascular Münster (PROCAM) study", started in 1978 (32,33). The results from the following PRIME study (Prospective Epidemiological Study of Myocardial Infarction) showed that both Framingham and PROCSM risk calculators overestimate the absulure cardiovascular risk in some European populations (e.g. in UK and France) and recommended the elaboration of a separate national risk prediction charts, taking into consideration the specific cardiovascular risk factor burden in the targeted population (34). Such algorithms (scores) were published for the UK: separately for England - QRISK (Cardiovascular disease risk score for the United Kingdom: prospective open cohort study ((35) and Schtland - ASSIGN, based on the SHHEC study (Scottish Heart Health Extended Cohort) (36), and West of Scotland Coronary Prevention Study (37). The last one included also the impact of socio-economic status in the total risk prediction equation. Accordingly, the European Society of Cardiology and the Second Joint Task Force instigated the development of a risk estimation system based on a large pool of representative European data sets that would capture the regional variation in risk. This led to the establishment of the SCORE (Systematic COronary Risk Evaluation)

project as a European Concerted Action project funded under the European Union BIOMED programme (31).

The SCORE project was initiated to develop a risk scoring system for use in the clinical management of cardiovascular risk in European clinical practice, in liaison with the Third Joint Task Force (38). This is being done in three phases (31):

1. First, the development of simple paper-based risk charts for high-risk and low-risk European populations;
2. Second, the development of methods for creating national or regional risk charts based on published mortality data; and
3. Third, the integration of risk estimation into a computer-based risk factor management application.

Characteristics, new aspects and advantages of the SCORE calculator

The SCORE project pooled data from cohort studies from 12 European countries (31,39). Most cohorts were population-based, though some occupational cohorts were included to increase representation of regions of lower risk. Subjects were excluded from the development of the risk chart if they had a previous history of heart attack.

Features of the SCORE method

Features of the SCORE method are as follows:

1. First and most important feature of the SCORE method is that it is estimating total cardiovascular risk rather than risk of coronary heart disease. This represents a shift from the traditional epidemiological concern with the causes of specific diseases to a public health perspective which focuses on the consequences of risk factors. By calculating total cardiovascular risk, a better estimate of individual CVD risk is acheived and also a better reflection of the health service implications of cardiovascular risk factors. Non-coronary cardiovascular disease is important because it represents a greater proportion of all cardiovascular risk in European regions with low rates of coronary heart disease
2. The SCORE project shift the emphasis in risk estimation to fatal cardiovascular disease events only, instead of combined fatal and non-fatal events. There is no doubt that both patients and physicians are as interested in non-fatal as in fatal cardiovascular disease events, and furthermore morbidity and incapacity caused by non-fatal cardiovascular disease events is the major economic burden for the health care system and the society. Non-fatal CVD pose, however, a number of problems for the development of risk estimation systems, because they are critically dependent on definitions and methods used in their diagnosis. The SCORE project considered the use of ‘hard’ coronary heart disease end-points (coronary death and non-fatal myocardial infarction) and ‘hard’ cardiovascular disease end-points (cardiovascular death and non-fatal cardiovascular disease events). An important reason for this decision was also that the ultimate aim of the SCORE project is to develop cardiovascular disease risk estimation systems applicable at national level in different European countries representing different rates of cardiovascular disease and

different mixes of coronary and noncoronary cardiovascular disease. Many European countries do not have cohort studies of cardiovascular disease, but all countries have national causespecific mortality data. These data can be used to estimate the baseline risk of the population.

3. Changing thresholds for high risk to 5%. A shift in the risk estimation from the risk of any coronary heart disease event to the risk of fatal cardiovascular disease will also mean a redefinition of the threshold for the 10-year absolute risk considered to signal the need for intensified risk modification efforts. Such decisions have to be made by international and national expert bodies formulating recommendations on cardiovascular disease prevention on the basis of scientific evidence and considering constraints related to practical and economic factors. The First and Second Joint Task Force of the European Societies (20, 21) recommended as a threshold for intensified risk factor intervention a 10-year absolute risk of 20% or more of developing any manifestation of coronary heart disease based on the risk chart derived using the Framingham risk function. This recommendation focused the attention on the importance of absolute risk as the basis of multi-factorial assessment of cardiovascular disease risk, but oversimplified a complex issue. To emphasise that there is no single level of absolute risk that defines an optimal threshold for risk factor intervention, regardless of the persons' age, sex or nationality, the SCORE risk charts display the 10-year risk of cardiovascular death both as figures as well as categories. Health economic research has suggested that the risk threshold for cost effectiveness of risk factor interventions, such as cholesterol lowering drug therapy, is not a simple function of absolute risk but also varies with age and sex.
4. Versions for total cholesterol and cholesterol/HDL ratio Persons with multiple risk factors tend to have lower HDL cholesterol levels and there is therefore a concern that failing to take HDL cholesterol into account will underestimate risk in those most at risk. A number of clinicians therefore, have expressed interest in a risk estimation system based on cholesterol/HDL ratio. Accordingly, two parallel systems were develpped.
5. Versions for low and high cardiovascular risk European populations. Two SCORE charts were developed – one for low CVD risk, including France, Belgium, Italy etc and one for high CVD risk, including Scandinavian countries, Germany etc.
6. Change in the ages for which the risk is displayed. The SCORE risk charts are providing more detail in the age group 50 to 65, which is the period during which risk changes most rapidly. Risk for age 30 has been suppressed. Persons aged 30 are essentially risk free within the next 10 years, and in many of the SCORE datasets there were no events in this age group. As was pointed out earlier, showing the 10-year risks for them would give a wrong message about the long-term risk of the young people with high risk factor levels.
7. SCORE risk charts are intended for risk stratification in the primary prevention of cardiovascular disease. There are no provided risk estimates for persons with established coronary heart disease, as there is now a widely accepted consensus that all persons with clinically established coronary heart disease or other atherosclerotic disease should be treated as high risk cases, recognising,

however, that the same major risk factors which are important in primary prevention remain important also in secondary prevention (38). Life expectancy model analyses suggest that the relative benefits of risk factor modification are almost similar for both low-risk and high-risk groups of patients with cardiovascular disease.

Limitations of the SCORE study and risk calculator

The underlying risk functions are based on single risk factor measurements, not on the persons "usual" levels. The charts also consider only the principal risk factors. In practice, the impact of other risk factors modulating disease risk needs to be considered also. These factors include a strong family history of early-onset cardiovascular disease, milder degrees of impaired glucose regulation, triglycerides, and fibrinogen. Future risk estimation systems may incorporate at least some of these factors. However, as yet their impact on the overall accuracy of risk estimation is uncertain, as a statistically significant association is no guarantee of a material gain in predictive power.

Data and populations, included in the SCORE project, were gathered during the 70^{ies} and 80^{ies} of the 20^{th} century (1974-78). From the beginning of 1970 the death rate from CVD is diceasing with 30% to 50% on average in Western Europe (40). The lifestyle and constitution of most European populations also is changing in Western Europe. It is highly probable that SCORE calculator predicts higer absolute risk than real, e.g. a total CVD risk of 5% in 1985 could be equal to 2,5% in 2003 (41).

Applying cardiovascular risk assessment to low-resource settings

The most common application of risk stratification is as a tool to cost-effective health policy decisions (42). It is more rational to choose a treatment programme that will produce the greatest benefit to patients. After a decision has been made regarding the general level of cost-effectiveness that is acceptable within a health system, risk stratification can help identify the appropriate patients that require treatment. To apply this method, however, it is necessary to have evidence from epidemiological studies on the absolute risk that patients in various categories would experience. Unfortunately, these data do not exist in virtually all developing country settings. Under such conditions, risk stratification can serve only a more limited purpose. Given a level of expenditure that represents the current reality, risk stratification can help identify the subset of patients most in need of treatment. It is important in this context, however, not to confuse this triage application with a cost-effectiveness analysis. As a result of a formal cost-effectiveness study, the group of patients that would benefit from treatment might be much larger than those whom current budgets can accommodate. It is imperative therefore, that the necessary epidemiological data be obtained so that future decisions can be based on evidence. Feasible risk-assessment methods need to be devised that use simple clinical indicators, such as age, sex, smoking habits, history of premature CVD in the family, presence or absence of diabetes, and presence or absence of hypertension – that are measurable in less well-resourced settings. With this

information, it is possible to develop a pragmatic risk-stratification system to rank people with mild hypertension into low-risk and high-risk groups in order to make treatment decisions. Such systems, although they may be less accurate, are likely to be the only feasible option in such contexts (42).

Explicit absolute risk assessment is an essential starting point when considering primary preventive treatment for CHD. However, uncritical application of any risk skore may mislead patients and health professionals and ongoing studies are needed to ensure CHD risk assessment is as accurate as possible for the group of patients to which it is applied (43).

CASE STUDY: THE CASE OF SCORE IMPLEMENTATION IN BULGARIA

Total (absolute) cardiovascural risk assessment in Bulgarian urban population

Introduction - justification for the study:

Cardiovascular diseases are major problem for the Bulgarian population, accounting for more than 60% of the all-cause mortality and considerable percent of the disability in the last two decades (44).

The total (absolute) risk assessment for CVD and its application in everyday clinical practice as well as in public health policy has been widely discussed recently among Bulgarian epidemiologists and clinicians. Several research studies were conducted and attempts were made to adapt the SCORE calculation charts to the Bulgarian population, especially in females. However, a large-scale population-representative epidemiologic study, assessing the absolute individual and population cardiovascular risk is still lacking.

Apart from the mentioned above, ineffective and virtually lacking primary prevention of CVD is considerd to be one of the major causes for the continuously increasing cardiovascular morbidity and mortality.

Under these circumstances, the necessity for extensive, up-to-date population research, aimed at the absolute CVD risk assessment and development of nationally relevant cardiovascular prevention guidelines for clinical practice appears to be indispensable.

Herewith, a short summary of methods and results from a cross-sectional observation study is presented as an example of the SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

Study aim and methods used

A cross-sectional observation study was conducted during the period 2005-2007.

The main objectives

The main objectives of the study were:

1. Overview of the main risk factor burden for CVD in Bulgaria (healthy population).

2. Evaluation of the absolute (total) cardiovascular risk, using the European SCORE calculator and population stratification for clinical and health policy recommendations.

Study population

A representative sample for the Bulgarian urban population, 3,810 subjects in total, age: 25–74, men and women, was investigated.

The study was conducted in five biggest Bulgarian cities with the help of the general practionioners (GPs): Sofia (48 GPs), Varna (20 GPs), Plovdiv (21 GPs), Stara Zagora (14 GPs) and Bourgas (3 GPs) - 106 GPs altogether.

The sample population was randomly selected, with no history of hard cardiovascular event (no heart attack or stroke) or serious CVD (CHD, angina pectoris etc).

Basic risk factores studied

Basic risk factores studied were age, sex, BMI, blood pressure, waist and hip circumference, serum lipids, glucose, profession, education, diet, physical activity, behavioural characteristics etc.

Main outcome

The absolute 10-year risk of a fatal cardiovascular event was estimated using the SCORE (HeartScore®) method (31), high risk chart, and in complaiance with the European guidelines, 2003 (38). A risk level categorization of the population sample was done, defining groups with low, intermediate and high risk. The results were compared to other similar European studies. A relevant threshold for a high-risk prevention strategy was proposed.

The total risk assessment was performed through a specifically developed software application, which consisted of a CVD Programme for global risk estimation ("CardioDB, Ver 1.0.0") and "Scorecard Ver 4.0.0.15" (electorinic version of SCORE charts). Each general practitioner received this computer application, where he/she could fill in the required data as well as calculate the actual cardiovascular absolute risk of the individual and create a database for future medical check-ups.

The population stratification was done, according to the European guidelines, 2003 (38) (Box 1):

Box 1. Stratification according to the European guidelines in SCORE implementation for absolute risk estimation in Bulgarian population.

Low risk group - SCORE ≤ 1%
Intermediate risk group - SCORE between 2% and 4%
High risk group - SCORE ≥ 5%

Results[8]

CVD risk factor and SCORE distribution in the studied population are presented on Tables 1 and 2. The presented data show that the mean and median values of most of the studies CVD risk factors are slightly increased or borderline, compared to clinical thresholds. This is in contrast with the considerably high total risk score, which is over 5% for the total sample, studied and have very high mean value in men - 9.12%. However the results for SCORE correspond to the high CVD death rate in the Bulgarian population, which once again is conclusive that not individual risk factor levels are to be examined and treated, but the total cardiovascular risk. Another interesting result is the considerably higher average risk score for men then for women, which also cannot be explained by the individual risk factor levels.

Table 1. CVD risk factors' and SCORE median, mean and standard deviation in the sample, in SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

CVD risk factors	**Median**	***X* (Mean)**	**σ (SD)**
Age	59	58.31	9.25
Body mass index - BMI	27.02	27.67	4.49
Systolic blood pressure - SBP	140	149.84	21.03
Diastolic blood pressure - DBP	82	90.30	11.80
Total Cholesterol	5.10	5.67	1.47
HDL cholesterol	1.10	1.12	0.36
LDL cholesterol	3.40	3.46	1.06
SCORE	4.00	6.43	6.62

Table 2. CVD risk factors' and SCORE mean and standard deviation, according to sex, in SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

Risk factors	**MEN**		**WOMEN**	
	Mean	**SD**	**Mean**	**SD**
Age	56.90	9.81	59.47	8.59
BMI	27.99	3.87	27.41	4.93
SBP	152.41	21.36	147.72	20.51
DBP	92.16	12.11	88.87	11.65
Total cholesterol	5.63	1.38	5.70	1.54
LDL	3.43	1.07	3.50	1.06
HDL	1.13	0.38	1.11	0.96
SCORE	9.12	8.08	4.21	3.89

[8] Here, only a limited data is presented for illustration of the theoretical background

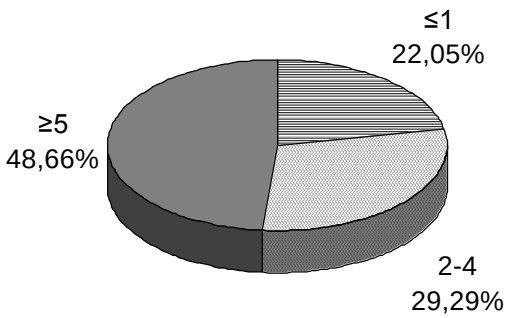


Figure 2. SCORE stratification of the studied population, according to ESC, 2003, in SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

Figure 2 presents a considerably high percent of the studied population (almost 50%) with SCORE>5%, i.e. nearly half of the sample belongs to the "high risk group" for fatal cardiovascular event in the next 10 years.

In relation to the above results and considering the lack of a national-specific absolute cardiovascular risk calculator, it was of interst to make a further cathegorization of the "high risk group" into "relatively high", "very high" and "excessively high risk" (Box 2):

Box 2. Cathegorization of the SCORE "high risk group" into "relatively high", "very high" and "excessively high risk" for absolute risk estimation in Bulgarian population.

Relatively high risk for fatal CVD - SCORE ≥ 5%
Very high risk for fatal CVD - SCORE ≥ 10%
Excessive risk for fatal CVD - SCORE ≥ 15%

Table 3. "High risk" distribution of the sample population in SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

SCORE	Number	%
5-9	927	24,31
10-14	512	13,46
≥15	415	10,89
Total	1854	48,66

From table 3 is seen that nearly one quarter of the studied sample has an absolute CVD risk, according to SCORE of over 10% (13,46 + 10,89 = 24,35%). Even this is an economically not cost-effective and not applicable to Bulgarian healthcare system and resources threshold for high-risk prevention strategy group.

The people with SCORE>15% are nearly 10% of the studied population, which appears to be already a reasonable group for high-risk strategy for CVD primary prevention.

The age and sex distribution of the sample population with SCORE over 5% is presented on figures 3 and 4. Most of the high risk subjects are found in the age groups over 45 years and especially over 55. The percent of men with SCORE>5% is significantly higher than that of women ($p<0.05$).

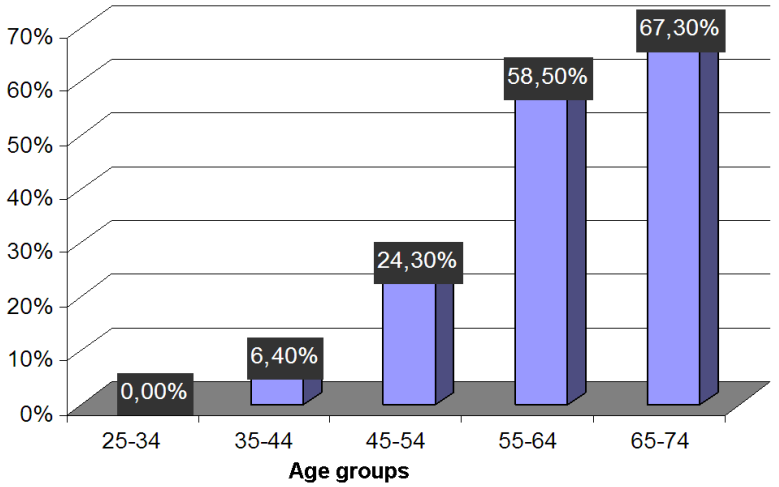


Figure 3. Age distribution of "high risk group" (SCORE>5%) in SCORE implementation for individual and population absolute risk estimation in Bulgarian population.

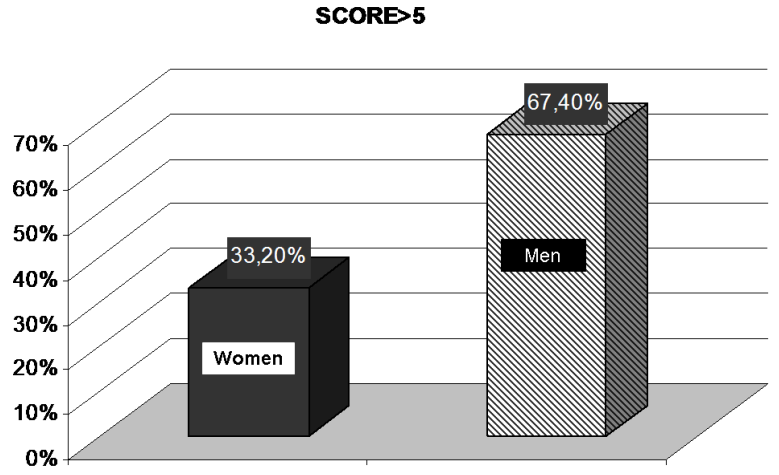


Figure 4. Sex distribution of "high risk group" (SCORE>5%) in SCORE implementation for individual and population absolute risk estimation in Bulgarian population

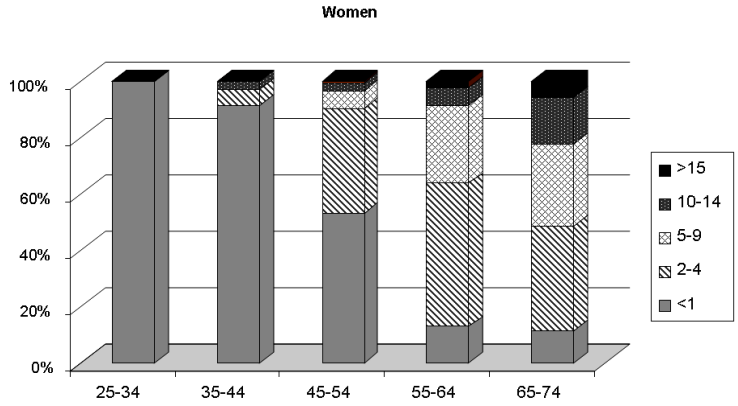


Figure 5. SCORE groups distribution according to age – women in SCORE implementation for individual and population absolute risk estimation in Bulgarian population

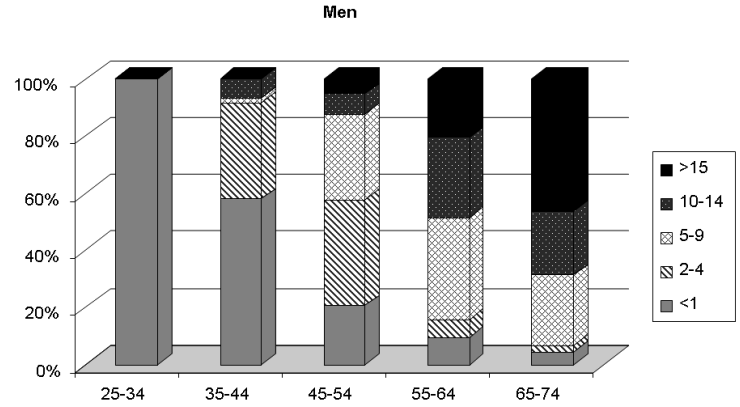


Figure 6. SCORE groups distribution according to age – men in SCORE implementation for individual and population absolute risk estimation in Bulgarian population

The SCORE stratification, according to age and sex, shown on figures 5 and 6, is characterized by:

1. Low absolute cardiovascular risk in the age group 25-34 in both men and women - 100% of the sample has SCORE<1%. However this could be explained by the small number of examined subjects in this group.
2. The SCORE (total risk) difference between men and women in the age groups 35-44 and 45-54 is significant ($p<0.05$). Women maintain higher percent with SCORE<1% in both age groups, while the percent of men with SCORE over 2 is rising. Nearly half of the men aged 45-54 years (42,4%) have SCORE>5%.
3. The gender differences in the age groups 55-64 and 65-74 are increasing.

4. The global CVD risk is increasing steeply in the age groups over 45. Only 6,6% of the men, aged 65-74 have low and average absolute risk (SCORE<5%), while the percentage of excessively high risk (SCORE>15%) reaches 46,6%, i.e. nearly half of the men over 65 yers of age have extremely high absolute risk of developing a fatal cardiovascular event in the next 10 years.

Discussions and conclusions

The high percentage of the studied population with SCORE>5% is consistent with the high CVD death rate in the Bulgarian population. It can be also considered as an indicator for a present "hidden morbidity", the so-called "morbidity iceberg", as most of the examined subjects are free of serious CVD, have relatively low levels of the main studied cardiovascular risk factors and most of them do not receive any medication treatment. Unfortunately, it is a common fact in Bulgaria, that the first cardiovascular event in such people is fatal (acute myocardial infarction or stroke).

One of the main advantages of the SCORE algorithm and the ESC guidelines, 2003 is the possibility for national-specific adaptation of the risk calculator and fixing a national relevant threshold for CVD high risk primary prevention strategy. However, the SCORE project considers an absolute risk of over 5% as high, but that is not the only element, on which health policy decision for CVD prevention targets should be taken. It has to take into account the health care resources available in the system as well as the general socio-economic, environmental and cultural background. Thus, according to our data, the most possibly cost-effective threshold for intensive risk factor modification strategy could be 15%.

According to the 2003 European guidelines on cardiovascular disease prevention in clinical practice (38), the SCORE method is leading in assessing global (absolute) cardiovascular risk. They set *three specific objectives, concerning CVD prevention:*

1. Adaptation of the proposed guidelines and recommendation at national level;
2. A tool for prioritizing patients (population stratification; and
3. A tool for councelling in clinical practice.

Our findings can be discussed with regard to these three explicit objectives of the guidelines:

1. Adaptation to national specificity.
   The regional and local differences in morbidity and mortality as well as different population risk profiles require risk evaluation and scoring systems development against epidemiological data from the target population to be screened before implementation in clinical practice. Whether the high risk chart applies to Bulgaria still needs more investigation and comprehensive population-representative studies.
2. Tool for prioritising patients.
   Our study found that the guidelines are unlikely to serve as an effective tool for prioritising Bulgarian population, as they classify an unreasonable number of people as at high risk (over 5%).
3. Tool for counselling in clinical practice.

> The European guidelines for CVD prevention are clearly recommended for direct use in counselling in clinical practice (38). Several ethical dilemmas arise from the likelihood of overestimating someone's true risk for cardiovascular disease (45). The systematic coronary risk evaluation project does not discuss the problem of retrospective risk bias. We question whether it was scientifically justifiable to include the risk charts of the systematic coronary risk evaluation project in guidelines intended for implementation in a clinical setting before validation in a contemporary context. Any overestimation of a person's risk for cardiovascular disease can have important implications. Apart from causing unnecessary concern, it undermines the patient's informed choice for intervention. It is also likely to increase prescribing costs and affect life insurance premiums. As yet little scientific knowledge is available on how the communication of this kind of risk affects people's understanding of themselves, their bodies, and their lives.

Methods for the development of guidelines for prevention of disease should be scientifically consistent so as to ensure that concordance with guidelines is practically feasible and likely to result in the desired outcomes. Despite the contribution of numerous experts and professional societies, it seems that authoritative clinical guidelines on the basis of the systematic coronary risk evaluation project may be an example of premature application of medical technology in routine clinical practice. On the other side, the CVD epidemic in Bulgaria requires an immediate concentrated action and systematic and effective solution. The insufficiency of enough local (national) data and evidence for development of a specific population targeted risk assessment tool, should not be an excuse for lack of public and health policy activity, addressing the problem.

EXERCISE

Task 1

Students are asked to search for risk factors for a certain socially-important) chronic non-communicable disease (e.g. diabetes, breast cancer, osteoporosis etc) and to build a multicausal web for it.

Task 2

Students are asked to find published studies of total risk assessment through different methods and compare them.

Task 3

Students should design a total risk assessment study (cross-sectional, cohort etc) for their own country, for a nationally important chronic disease.

REFERENCES

1. World Health Organization. Reducing risk, promoting healthy life. In World Health Organization. The World health report 2002. Geneva: World Health Organization; 2002. p.9-26. Available from: http://www.who.int/whr/2002/en/index.html. Accessed: June 03, 2008.
2. Hope T. Rationing and life-saving treatments: should identifiable patients have higher priority? J Med Ethics. 2001;3:179-85.
3. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. Int J Epidemiol. 2002;31:285-93.
4. Rothman K, Greenland S. Modern epidemiology. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1998.
5. Miettinen O. Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol. 1974;99:325–32.
6. Walter S. Prevention for multifactorial diseases. Am J Epidemiol. 1980;112(3):409-16.
7. Murray C, Lopez A. On the comparable quantification of health risks: lessons from the Global Burden of Disease Study. Epidemiology. 1999;10:594-605.
8. Wald N, Law M. A strategy to reduce cardiovascular disease by more than 80%. Br Med J. 2003;326:1419.
9. Law M, Wald N. Risk factor thresholds: their existence under scrutiny. Br Med J. 2002;324:1570-6.
10. Wang T, Massaro J, Levy D, Vasan R, Wolf P, D'Agostino R, et al. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community. The Framingham heart study. JAMA. 2003;290:1049-56.
11. Kannel W. Some lessons in cardiovascular epidemiology from Framingham. Am J Cardiol. 1976; 37:269-82.
12. Kannel W, Feinleib M, McNamara P, Garrison R, Castelli W. An investigation of coronary heart disease in families. The Framingham offspring study. Am J Epidemiol. 1979;110:281–90.
13. Kastarinen M, Saloma V, Vartiainen E, Jousilahti P, Tuomilehto J, Puska P, Nissinen A. Trend in blood pressure levels and control of hypertension in Finland from 1982 to 1997. J Hypertens. 1998;16:1379-87.
14. Shapo L, Pomerleau J, McKee M. Epidemiology of hypertension and associated cardiovascular risk factors in a country in transition: a population based survey in Tirana City, Albania. J Epidemiol Community Health. 2003;57:734-9.
15. Hirte L, Nolte E, Mossialos E, McKee M. The changing regional pattern of ischaemic heart disease mortality in southern Europe: still healthy but uneven progress. J Epidemiol Community Health. 2008;62:e4.
16. Powles J, Zatonski W, Vander Hoom S, Ezzati M. The contribution of leading diseases and risk factors to excess losses of healthy life in Eastern Europe: burden of disease study. BMC Public Health. 2005;5:116.
17. Assmann G, Schulte H. Modelling the Helsinki Heart Study by means of risk equations obtained from the PROCAM Study and the Framingham Heart Study. Drugs. 1990;40 Suppl 1:13-8.
18. Haq I, Ramsay L, Yeo W, Jackson P, Walliset E. Is the Framingham risk function valid for northern European populations? A comparison of methods for estimating absolute coronary risk in high risk men. Heart. 1999;81:40–6.
19. Marrugat J, D'Agostino R, Sullivan L, Elosua1 R, Wilson P, Ordovas J, et al. An adaptation of the Framingham coronary heart disease risk function to European Mediterranean areas. J Epidemiol Community Health. 2003;57:634–8.
20. Pyorala K, De Baker G, Graham I, Wilson P, Wood D. Prevention of coronary heart disease in clinical practice. Recommendations of the Task Force of the European

Society of Cardiology, European Atherosclerosis Society and European Society of Hypertension. Eur Heart J. 1994;15:1300–31.
21. Pyorala K, Wood D. Prevention of coronary heart disease in clinical practice. European recommendations revised and reinforced. Eur Heart J. 1998;19:1413-5.
22. Wood D, De Backer G, Faergeman O, Graham I, Mancia G, Pyörälä K. Prevention of coronary heart disease in clinical practice. Summary of recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. Blood Press. 1998;7:262-9.
23. Ducimetiere P, Richard J, Cambien F, Rakotovao R, Claude J. Coronary heart disease in middle-aged Frenchmen. Comparisons between Paris Prospective Study, Seven Countries Study, and Pooling Project. Lancet. 1980;1:1346-50.
24. Menotti A, Puddu P, Lanti M. Comparison of the Framingham risk function-based coronary chart with risk function from an Italian population study. Eur Heart J. 2000;21:365-70.
25. Hense H, Schulte H, Löwel H, Assmann G, Keil U. Framingham risk function overestimates risk of coronary heart disease in men and women from Germany-results from the MONICA Augsburg and the PROCAM cohorts. Eur Heart J. 2003;24:937-45.
26. Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, et al. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. BMJ. 2003;327:1267-72.
27. Van Steenkiste B, van der Weijden T, Stoffers H, Grol R. Barriers to implementing cardiovascular risk tables in routine general practice. Scand J Prim Health Care. 2004;22:32-7.
28. Brindle P, Beswick A, Fahey T, Ebrahim S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. Heart. 2006;92:1752-9.
29. Jones A, Walker J, Jewkes C, Game F, Bartlett W, Marshall T, Bayly G. Comparative accuracy of cardiovascular risk prediction methods in primary care patients. Heart. 2001;85:37–43.
30. Wilson P, D'Agostino R, Levy D, Belanger A, Silbershatz H, Kannel W. Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97(18):1837-47.
31. Conroy R, Pyorala K, Fitzgerald A, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J. 2003;24:987-1003.
32. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. Circulation. 2002;105:310-5.
33. Assmann G, Cullen P, Schulte H. The Münster Heart Study (PROCAM). Results of follow-up at 8 years. Eur Heart J. 1998;19(Suppl A):A2-A11.
34. Empana J, Ducimetière P, Arveiler D, Ferrières J, Evans A, Ruidavets J, et al. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. Eur Heart J. 2003;24:1903-11.
35. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK. A new cardiovascular disease risk score for the UK. Br Med J. 2007;335:136.
36. Haq I, Jackson P, Yeo W, Ramsay L. Sheffield risk and treatment table for cholesterol lowering for primary prevention of coronary heart disease. Lancet. 1995; 346:1467-71.
37. The West of Scotland Coronary Prevention Study Group. West of Scotland Coronary Prevention Study. Baseline risk factors and their association with outcome in the West of Scotland Coronary Prevention Study. Am J Cardiol. 1997;79:759-62.
38. De Backer G, Ambrosioni E, Borch-Johnsen K, Brotons C, Cifkova R, Dallongeville J, et al. European guidelines on cardiovascular disease prevention in clinical practice.

Third Joint Task Force of European and other Societies on Cardiovascular Disease Prevention in Clinical Practice. Eur J Cardiovasc Prev Rehabil. 2003;10:S1-S10.
39. Cowie M. Coronary risk - time for a more sophisticated approach? Eur Heart J. 2002; 23:589–91.
40. Kuulasmaa K, Tunstall-Pedoe H, Dobson A, Fortmann S, Sans S, Tolonen H, et al. Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the World Health Organization MONICA Project populations. Lancet. 2000;355:675-87.
41. Holmen J, Midthjell K, Kroger O, Langhammer A, Holmen T, Bratberg G, et al. The Nord-Trondelag health study 1995-7 (HUNT-2): objectives, contents, methods and participation. Norsk Epidemiol. 2003;13:19-32.
42. World Health Organization. Integrated management of cardiovascular risk: Report of a World Health Organization meeting. Geneva: World Health Organization; 2002.
43. Madhok V, Fahey T. Cardiovascular risk estimation: important but may be inaccurate. Br Med J. 2006;332:1422. Available from: http://bmj.com/cgi/content/full/332/7555/1422. Accessed: August 20, 2007.
44. National Statistical Institute. Sofia. Deaths by causes in 2006. Available from: http://www.nsi.bg/SocialActivities_e/Health.htm. Accessed: August 22, 2008.
45. Getz L, Kirkengen AL, Hetlevik I, Romundstad S, Sigurdsson J. Ethical dilemmas arising from implementation of the European guidelines on cardiovascular disease prevention in clinical practice: descriptive epidemiological study. Scand J Prim Health Care. 2004;22:202-8.

RECOMMENDED READINGS

1. Ezzati M, Lopez A, Rodgers A, Murray C. Comparative Quantification of Health Risks. Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Geneva: World Health Organization; 2004.
2. Lopez A, Mathers C, Ezzati M, et al. Measuring the Global Burden of Disease and Risk Factors, 1990-2001. Global Burden of Disease and Risk Factors. New York: Oxford University Press; 2006.
3. EUROASPIRE Study Group. EUROASPIRE: A European Society of Cardiology survey of secondary prevention of coronary heart disease: principal results. Eur Heart J. 1997;18:1569-82.
4. Reynolds T, Twomey P, Wierzbicki A. Concordance evaluation of coronary risk scores: implications for cardiovascular risk screening. Curr Med Res Opin. 2004;20:811-8.
5. Marteau T, Kinmonth A. Screening for cardiovascular risk: public health imperative or matter for individual informed choice? Br Med J. 2002;325:78-80.
6. Shekelle P, Woolf S, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. Br Med J. 1999;318:593-6.
7. Getz L, Kirkengen A, Hetlevik I, Sigurdsson J. Individually based preventive medical recommendations - are they sustainable and responsible? A call for ethical reflection. Scand J Prim Health Care. 2005;23:65-7.
8. Saltman R, Figueras J. European Health Care Reform: Analysis of Current Strategies. Copenhagen: World Health Organization, Regional Office for Europe; 1997.
9. Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. Br Med J. 1991;303:276-282.
10. MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, et al. Blood pressure, stroke, and coronary heart disease. Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. Lancet. 1990;335:765-74.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**INTRODUCTION TO EPIDEMIOLOGICAL STUDIES**
Module: 1.4.1	**ECTS (suggested): 0.10**
Author(s), degrees, institution(s)	**Enver Roshi**, MD, MPH Faculty of Medicine, University of Tirana, Albania **Genc Burazeri**, MD, PhD Department of International Health, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
Address for correspondence	**Enver Roshi** Faculty of Medicine, University of Tirana, Rr. "Dibres", N.370, Tirana, Albania Tel. +3554371859 E-mail: gburazeri@yahoo.com E-mail: e_roshi@yahoo.com
Keywords	Descriptive surveys, analytic surveys, observational studies, experiments, cohort studies, case-control studies, cross-sectional studies.
Learning objectives	After completing this module students should: • be familiar with the major classifications of epidemiological studies; • be familiar with the major types of epidemiological studies: cohort, case-control and cross-sectional approaches; • be aware of the difference between observational studies and experiments; • be familiar with the major epidemiological study designs: applications, advantages and disadvantages.
Abstract	Epidemiological studies may deal merely with the distribution of diseases/conditions in human populations, and/or with the factors influencing the distribution and the frequency of diseases. Three major types of epidemiological studies are: cohort studies, case-control studies, and cross-sectional studies
Teaching methods	Introductory lecture, illustrative examples of main epidemiological study designs.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Take-home exercise on the applications, advantages and limitations of the main epidemiological studies.

INTRODUCTION TO EPIDEMIOLOGICAL STUDIES

Enver Roshi, Genc Burazeri

THEORETICAL BACKGROUND

Some classifications of epidemiological studies

Epidemiological studies may deal merely with the distribution of diseases/conditions in human populations (descriptive surveys), and/or with the factors influencing the distribution and the frequency of diseases (analytic surveys: cross-sectional, case-control, cohort, quasi-experiment and experimental studies) (1-6).

Description versus analysis

This is the most classical way of classifying the epidemiological inquiries. However, this classification has little practical value in itself, since the same study could be both descriptive and analytic or, in principle, all studies could be regarded as analytic (e.g. the distribution of a certain condition by sex or age could be regarded as a sort of implicit analysis rather than just description of the observed facts).

1. *Descriptive survey*
   Describes a situation, e.g. distribution of a disease/condition in a certain population in relation to sex, age, or other characteristics.
2. Analytical survey (explanatory study).
   Tests hypotheses and looks for associations based on:
   - groups (ecological/correlation studies, trend studies);
   - based on individuals (cross-sectional studies, case-control studies, cohort studies, experiments and quasi-experiments).

Observation versus experiment

The most important way of classifying the epidemiological studies is the one, which accounts for the role/control of the researcher over the study. According to this classification, a major distinction is being made of observational studies, and experiments.

1. Observational studies.
   Characteristics of observational studies are:
   - the investigator observes the occurrence of the condition/disease in population groups that have assigned themselves to a certain exposure.
   - often most practical and feasible to conduct.
   - carried out in more natural settings – representative of the target population.
   - often, there is little control over the study situation – results are susceptible to distorting influences.
   - Experimental approach.

   Characteristics of experimental studies are:

- this design is the most powerful study design for testing ethiological hypothesis;
- the investigator exercises control over the allocation of exposure, its associated factors and observation of the outcome, and
- for obvious ethical and practical reasons, the possibilities of conducting experiments in human populations are very limited.

Direction of study question

According to the direction of inquiry, epidemiological studies are classified as prospective, retrospective, non-directional, and ambispective studies:

1. Prospective (forward-looking) design.
   Prospective (forward-looking) study is the study in which (disease free) people who are exposed and non-exposed are followed up and compared with respect to the subsequent development of the disease/outcome under study.
2. Retrospective (backward-looking) design.
   Retrospective (backward-looking) study is study in which people with the disease are compared with people without the disease, to determine whether they differ in their past exposure to the (hypothesized) causative factor.
3. Non-directional design.
   Non-directional design is the design in which the investigator observes simultaneously the exposure and disease status in the study population.
4. Ambispective design.
   Ambispective design is the design in which one primary variable/factor is measured prospectively and the other one retrospectively, or one primary variable/factor is measured both prospectively and retrospectively.

Direction of data collection

According to this classification, epidemiological studies are classified as retrolective and prolective:

1. Retrolective design.
   Retrolective studies are epidemiological studies in which data are collected before the study design (not necessarily for the purpose of the actual study).
2. Prolective design.
   Prolective studies are epidemiological studies in which data are collected after study design (for the purpose of the actual study).

Three major types of epidemiological studies

Three major types of epidemiological studies are: cohort studies, case-control studies, and cross-sectional studies.

Cohort studies

Characteristics of cohort studies

The essence of cohort studies is the identification of a group of subjects about whom certain exposure information is collected. The group is then followed up in time to ascertain the occurrence of disease/condition of interest.

For each individual prior exposure can be related to subsequent disease experience. Since the first requirement of such studies is the identification of the individuals forming the study group – a cohort, prospective or longitudinal studies are usually referred to as cohort studies.

The design of a cohort study is sketched in Figure 1.

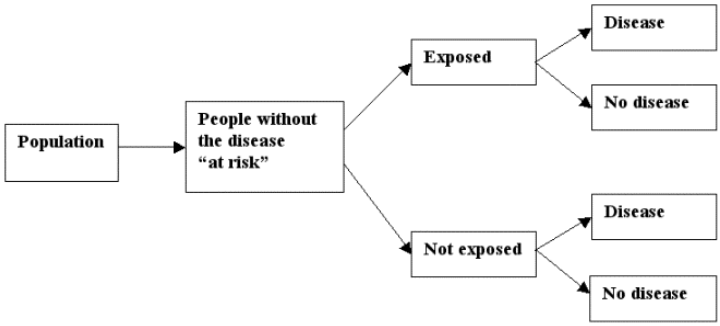


Figure 1. Schematic presentation of a cohort study.

Advantages of cohort studies

Advantages of cohort studies are as follows:

- they measure incidence and thus permit direct estimation of risk of disease;
- they do not rely on memory for information about exposure status, hence avoid bias due to selective recall;
- since cohort studies begin with people free of disease, potential bias due to selective survival is eliminated;
- they provide a logical approach to studies of causation or effects of treatment, and
- cohort approach can yield information on associations of exposure with several diseases.

Disadvantages of cohort studies

Disadvantages of cohort studies are as follows:

- they require large samples to yield the same number of cases that could be studied more efficiently in a case-control study;
- they are particularly inefficient for studies of rare diseases;
- they are logistically difficult – long follow-up period, often serious attrition to study subjects;
- direct observation of participants may cause changes in health behaviour;

- possible bias in ascertainment of disease due to changes over time in criteria and methods, and
- they are very costly.

Prospective versus historical cohort studies

An alternative strategy to the costly and time consuming prospective cohort design is the historical cohort study. A cohort is identified (enumerated) as of some historical point in time and is then followed over past time to the present. Disease rates and relative risk (RR) can be derived from this type of study as well. The retrospective (historical) cohort approach is particularly amenable to the study of exposure-disease relationship for which the exposure group is unusual in some way, e.g. many occupational or environment exposure-disease relationships.

Case-control studies

Characteristics of case-control studies

The case-control study begins with a group of cases of a specific disease. This (the disease) is the starting point of the study, unlike cohort studies where the interest is in drawing a contrast between exposed and non-exposed subjects. So, the case-control approach is directed at the prior exposures, which caused the disease and thus proceeds from effect (outcome) to cause (exposure).

The design of a case-control study is sketched in Figure 2.

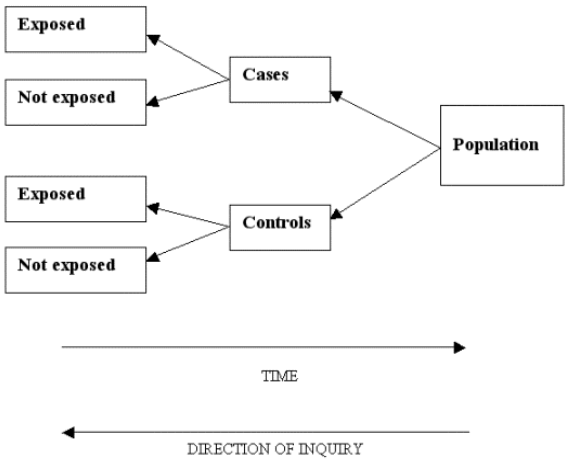


Figure 2. Schematic presentation of a case-control study.

Advantages of case-control studies

Advantages of case-control studies are as follows:

- they are highly informative in comparison to other designs: several exposures or potential causal agents can be examined;

- their design is efficient (low cost per study) primarily because few subjects are needed to obtain stable estimate of RR, and
- they are particularly appropriate for studies of rare diseases (e.g. a case-control study with 100 cases of a disease having an annual incidence of 1/1,000. A cohort design for this disease would require 1,000 persons to be followed up for 100 years or 10,000 persons to be followed up for 10 years in order to yield the same number of cases).

Disadvantages of case-control studies

Disadvantages of case-control studies are as follows:

- the absolute frequency of a disease cannot be determined. No counts are made of population at-risk, thus, there are no denominators available to obtain the incidence rates. Lacking absolute risks, it's not possible to compare disease rates among different studies, nor is it possible to estimate the attributable risk;
- they are particularly subject to two types of bias: selection bias in choosing controls, and recall bias (cases may recall better prior exposures than controls), and
- "philosophically" they are difficult to interpret: the antecedent-consequent relationship (exposure-outcome) is subject to considerable uncertainty.

Cross-sectional studies

Characteristics of cross-sectional studies

Cross-sectional design is referred to as non-directional or one point in time survey, where data is collected on both outcome and exposure status of the individuals under study. Such studies are useful to describe characteristics of study population and can generate new aetiological hypothesis. This study design involves disease prevalence. Cross-sectional studies can evaluate the impact of changes in health services during an intervening period. This can be accomplished by conducting a cross-sectional study twice: before and after the intervention.

The design of a cross-sectional study is sketched in Figure 3.

Advantages of cross-sectional studies

Advantages of cross-sectional studies are as follows:

- they describe the distribution of both exposure and outcome in a population (particularly useful for studying frequent outcomes of long duration);
- they provide estimates of the magnitude of a disease problem in a community, which might be very important for planning of health services;
- in comparison to other studies they are relatively quick and inexpensive. Often, involve only one-time survey, and
- they are largely applicable: provision of health care services as well as generation of aetiological hypothesis.

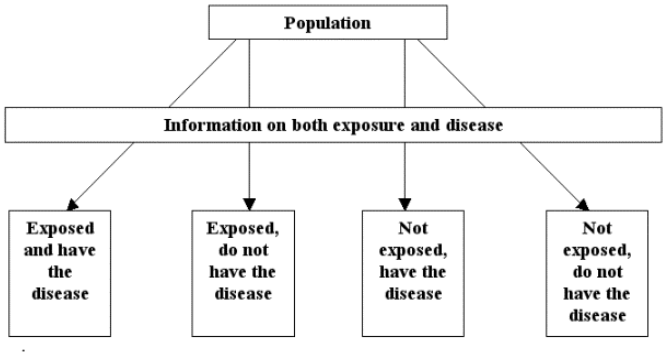


Figure 3. Schematic presentation of a cross-sectional study.

Disadvantages of cross-sectional studies

Disadvantages of cross-sectional studies are as follows:

- they do not measure incidence risk, because this would require incidence data (but they measure prevalence risk);
- they can not determine cause-effect relationship;
- current exposure status may be due to changes that have occurred as a result of the disease rather than having led to the disease, and
- diseases of short duration may be missed. Thus, cross-sectional studies are best applied to the study of chronic or persistent conditions.

EXERCISE

Task 1

Using international bibliographic data-bases (e.g. Medline, PubMed) find at least 3 cohort, 3 case-control, and 3 cross-sectional epidemiological studies.

Task 2

Carefully read them and find their characteristics.

Task 3

Discuss the characteristics, strengths and limitations of selected studies with your colleagues.

REFERENCES

1. Gordis L. Epidemiology (3rd edition). Philadelphia: W.B. Sounders Company; 2004.
2. Abramson JH, Abramson ZH. Survey methods in community medicine. New York: Churchill Livingstone; 1999.
3. Lilienfeld AM, Lilienfeld DE. Foundations of epidemiology. New York: Oxford University Press; 1980.
4. MacMahon B, Pugh TF. Epidemiology: principles and methods. Boston, Massachusetts: Little, Brown; 1970.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. Belmont, California: Lifetime Learning Publications; 1992.
6. Abramson JH, Feinleib M, Detels R, Greenberg RS, Ibrahim MA. Oxford Textbook of Public Health. Vol. 3: Investigative methods in public health. Oxford: Oxford University Press; 1985.

RECOMMENDED READINGS

1. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
2. Detels R, McEwen J, Beaglehole R, Tanaka H (editors). Oxford textbook of public health. Oxford: Oxford University Press; 2002.
3. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. Lifetime Learning Publications, Belmont, California; 1992.
4. Lilienfeld AM, Lilienfeld DE. Foundations of epidemiology. Oxford University Press, New York; 1980.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**FEATURES OF EPIDEMIOLOGICAL STUDIES**
Module: 1.4.2	**ECTS (suggested): 0.15**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Doncho Donev**, MD, PhD, Professor Institute of Social Medicine, Medical Faculty, University of Skopje, Macedonia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Epidemiological studies, features of epidemiological studies, classifications of epidemiological studies
Learning objectives	After completing this module students should: • be familiar with the major classifications of epidemiological studies. • understand differences between various types of studies and can interpret them.
Abstract	Rarely, prior introducing the readership main study designs, textbooks on epidemiological methods are discussing about different features/ characteristics of epidemiological studies. Since one study design has several of them, this could be for the audience sometimes rather confusing. To clarify this aspect of epidemiological studies, we decided to describe these features in more details
Teaching methods	An introductory lecture gives the students insight in different features of epidemiological studies and distinction between them. After introductory lectures students discuss in small groups these features and confront them.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Essay type exam.

FEATURES OF EPIDEMIOLOGICAL STUDIES

Lijana Zaletel-Kragelj, Ivan Eržen, Doncho Donev

THEORETICAL BACKGROUND

Introduction

There exist a pleiad of textbooks, manuals and other types of books, describing epidemiological concept and its methods in one or another way. Most of them are focused first on different epidemiological measures, and then from one or another perspective on different study designs (ecological studies, cross-sectional studies, case-control studies, cohort studies, clinical trials, community trials, etc.) (1-12). A brief overview on three major types (designs) of epidemiological studies, being cross-sectional, case-control, and cohort studies, is presented also in a previous module of this book, while some points of view or more detailed description of individual design will be given in following modules of this chapter.

Rarely, prior introducing the readership main study designs, these books are discussing about different features/characteristics of epidemiological studies. Since one study design has several of them, this could be for the audience sometimes rather confusing. To clarify this aspect of epidemiological studies, we decided to describe these features in more details.

Overview of different features of epidemiological studies

Out of numerous different features, two of them could be classified as most important, being:

- is the study observational or experimental, and
- is the study descriptive or analytical.

But in fact, epidemiological studies have many different features (1-8) and not only just mentioned. These features are used as criteria for classification of epidemiological studies as well. Thus, epidemiological studies could be classified in many different ways. Some of these groups of features are as follows (Figure 1):

1. is the study observational, or experimental (interventional),
2. is the study descriptive, or analytical,
3. is the study cross-sectional, or follow-up,
4. is the study incidence, or prevalence study,
5. is the study question directed into the future, or into the past,
6. is the collection of study data directed into the future, or into the past,
7. is the study using individual, or population/aggregated level data,
8. is the study for generating hypotheses, or testing it/them,
9. is the study using permanent data sources, or focused study data,
10. is the study focused on effect, or on exposure,
11. is the study for planning or evaluating an intervention,
12. is the study focused on solving problem on individual or on the population level.

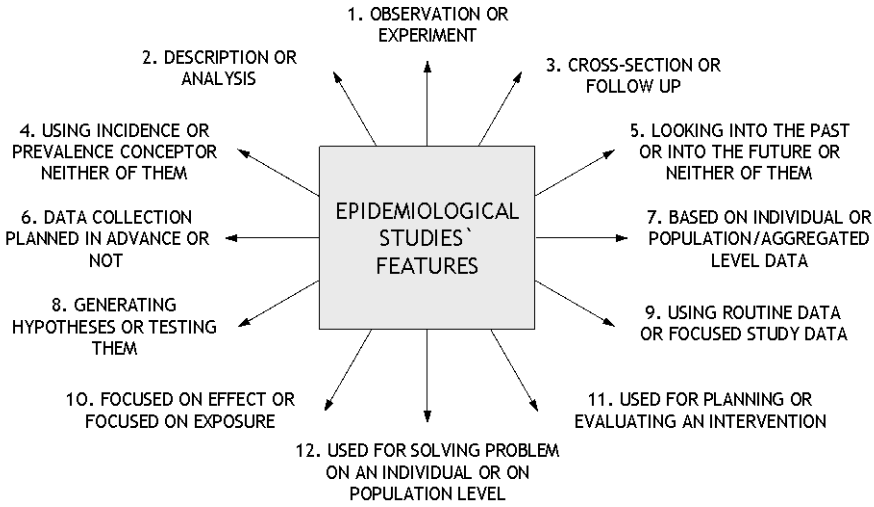


Figure 1. Some groups of features of epidemiological studies.

Beside the first two, also classification listed under the number #3-5 could be met frequently then the rest of them. In continuation we will describe all of them in more details.

Classifying epidemiological studies by different features

Experimental versus observational studies

Experimental studies are those characterized by assignment of exposure by a researcher. This means that a researcher gains the mastery over the situation (1-8).

Another important characteristic is that experimental studies use well known method of randomization for controlling confounding. The importance of randomization is that it leads to a balance of confounders in exposed and non-exposed study groups, providing theoretically unbiased evaluation of exposure-outcome associations. The study and control groups are comparable except in exposure under observation.

Because of their characteristics, for experimental studies holds, that this design is the most powerful design of all.

But in spite their importance, experimental studies have an important disadvantage in human research – they could be extremely unethical and therefore not possible to conduct. Here observational studies take advantage over experimental studies.

In contrast, observational or non-experimental studies are epidemiological studies that do not involve any intervention, experimental or otherwise. Exposure in this type of studies is not assigned by a researcher. In such a study nature is allowed to take its course, with changes in one characteristic being studied in relation to changes in other characteristics.

Analytic epidemiological methods such as case-control and cohort study designs, are properly called observational epidemiological studies because the investigator is observing without intervention other than to record, classify, count, and statistically analyse results.

An important disadvantage of observational studies is a little control over the subject under study and confounding factors that may influence the results substantially. As a consequence are more susceptible to different types of bias.

Descriptive versus analytical studies

The distinction between "descriptive" and "analytic" studies is one of intent, objective, and approach, rather than one of design. In this respect, data obtained in public health research usually could be explored in a descriptive or analytical mode (8). Data obtained in an analytic study must be first described, and data obtained in a descriptive study can be analyzed to test hypotheses if indicated.

Epidemiological studies designed and concerned primarily to describe the existing distribution of health phenomena in the population, without regard to causal or other hypotheses are usually studies based on a routine data. Such studies describe the health conditions and health-related characteristics of populations, typically in terms of person, place, and time. Their results are usually presented in health statistics yearbooks, and similar publications. This information serves as the foundation for studying populations. It provides essential contextual information with which to develop hypotheses, design studies, and interpret results. Surveillance is a particular type of descriptive epidemiology, to monitor change over time.

In contrast, analytical studies are usually concerned with testing one or more specific hypotheses, typically whether an exposure is a risk factor for a disease or an intervention is effective in preventing or curing disease (or any other occurrence or condition of interest) (1-10).

In this place we need to emphasize that it is a stereotype that analytic studies (testing hypotheses) take advantage over descriptive studies. As a consequence, descriptive studies are labelled as "less important" than, or "inferior" to analytic studies. However, well designed descriptive study should be first step in a process of investigating health problems of the population, since they are source for generating sound hypotheses for more in-depth and usually more expensive analytical studies.

Cross-sectional or transversal versus follow-up or longitudinal studies

One of the most important features of epidemiological studies is, if we are making observation as a transversal cross-section through health status of a population or as a follow-up in time (Figure 2).

Cross-sectional or transversal studies are those studies that examine the relationship between diseases or other health related states at one particular time, being a moment or a period (1). In this design subjects are sampled without respect to disease status and are studied at a particular point in time (8), and the presence or absence of an outcome, as well as presence or absence of exposure is observed in the same point in time.

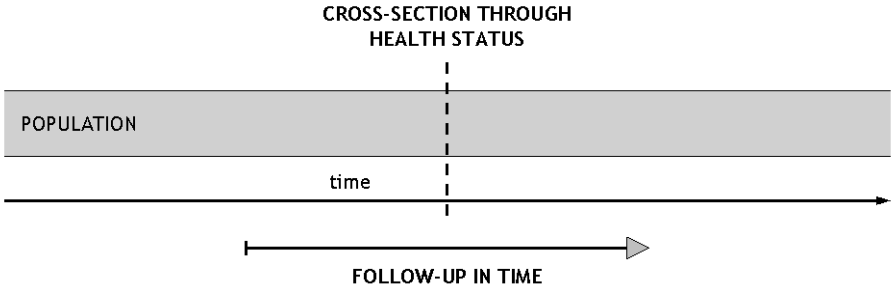


Figure 2. Schematic presentation of cross-sectional and follow-up studies.

The term "cross-sectional study" usually refers to studies at the individual data level, even though ecologic studies at aggregate level are typically (though not necessarily) cross-sectional, as well.

The target population is usually one whose identity is of some wider interest (e.g., a political or geographical entity, a profession or workforce, or a major organization, but may not necessarily be so) (8).

In cross-sectional studies, the current status of individuals could be examined in relation to some current or past exposure. When these studies are used with an analytical purpose, one should be cautious when interpreting the relationship between outcomes and exposures, especially the causal one, since temporal sequence of cause and effect cannot necessarily be determined in this study design (1,4).

These studies are most useful for conditions that are not rapidly fatal, not terribly rare, and/or not routinely brought to medical attention (e.g., elevated blood pressure, elevated blood cholesterol, etc) (8).

In contrast, in follow-up or longitudinal studies people without the disease at the beginning of the observation time (usually referred as »at risk«) are followed-up to see who develops the disease over time. If the population followed is a defined group of people (a »cohort«), then the study is referred to as a cohort study.

Special type of longitudinal studies are so-called ecological longitudinal studies (3) that are studies made on ongoing frequent cross-sectional studies (surveillance or monitoring) to measure trends in disease rates over many years in a defined population. By comparing the trends in disease rates over time with considering other changes in the population, it could be determined the impact of these changes on the disease rates.

Prevalence versus incidence studies

The term "prevalence studies" is referring to cross-sectional studies at the individual data level. The frequency of an outcome variable is in this type of studies measured in terms of prevalence. Prevalence is a common term for a group of measures which are quantifying the situation (state) of a given health phenomenon (e.g. a disease, a disorder, an unhealthy health behaviour etc.) at a designated time (at a specified moment, or at any time during a specified period). The detailed description of prevalence measures (e.g. prevalence risk, prevalence rate, prevalence odds), as well as that of characteristics of cross-sectional studies at the individual data level, are beyond the scope of this module, and are given in other modules of this book.

Similarly, the term "incidence studies" is referring to studies, also known as follow-up, longitudinal, or cohort studies. The frequency of an outcome variable is in this type of studies measured in terms of incidence. Incidence is a common term for a group of measures which are quantifying a breakout of new cases of a health phenomenon (e.g. a disease) under observation during a specified period in a specified group of persons. The detailed description of incidence measures (e.g. incidence risk, incidence rate, incidence odds, and incidence density), and characteristics of cohort studies, are beyond the scope of this module. They are given in other modules of this book.

In contrast, case-control studies are not incidence or prevalence studies, and are intended to observe past exposure.

Retrospective versus prospective versus non-directional studies

One very important feature concerns the timing of collection of exposure information. In this respect we distinguish between retrospective, prospective, and non-directional studies (Figure 3).

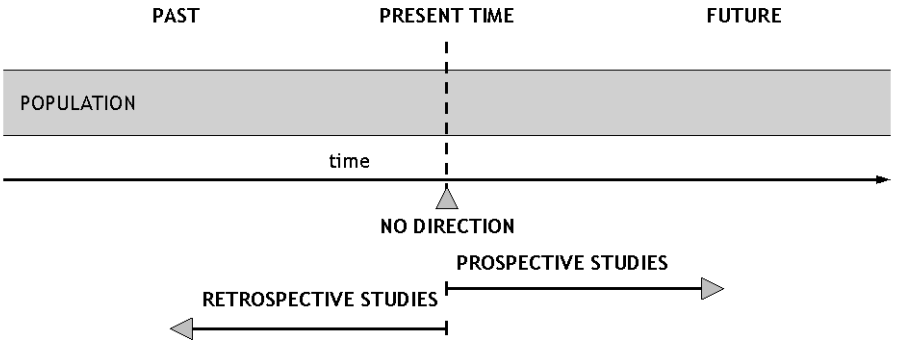


Figure 3. Schematic presentation of retrospective, prospective, and non-directional studies.

A retrospective study is a study that looks backwards in time and examines exposures to suspected risk or protection factors in relation to an outcome that is established at the start of the study. Thus, in a retrospective study design, the outcome of interest has already occurred at the time the study is initiated. Thus we find people that have a disease under observation and try to figure out why they got the disease.

An investigator conducting a retrospective study typically utilizes administrative databases, medical records, or interviews with patients who are already known to have a disease or condition.

The biggest problem in a retrospective study is that some of the information that we need may be hard to get, or it is subjected to a so-called recall bias. We have to rely on patients to recall things that may have happened many years ago. Memory

In contrast, a prospective study looks forward in time. In this study design, we select a group of subjects without a condition under observation and observe them over specified period if they develop the condition after they were exposed to a suspected risk or protection factor(s).

One of disadvantages of this study design is that in a case the outcome under observation has a long pre-clinical phase it could take a long time to accumulate sufficient data to get correct and strong conclusions. When we are studying a disease that takes a long time to appear, we usually need to use a retrospective study, and not a prospective one.

The outcome of interest also should be common; otherwise, the number of outcomes observed will be too small to be statistically meaningful (indistinguishable from those that may have arisen by chance).

All efforts should be made to avoid sources of bias such as the loss of individuals to follow up during the study. Prospective studies usually have fewer potential sources of bias and confounding than retrospective studies.

Despite these disadvantages, for prospective study design holds that it is the best design for establishing relationships between outcome of interest and exposure variables.

The third study design in this group is non-directional, in which outcome(s) and exposure(s) are observed at the same time. Transversal or cross-sectional studies are non-directional studies. We have already discussed some characteristics of this type of studies, while some more details will be discussed in a separate module in this book.

Retrolective versus prolective studies

Less frequently used feature and corresponding classification relates to the mode of gathering of data. In this respect we distinguish between retrolective and prolective studies (1) (Figure 4).

According to Feinstein, who coined both terms at the beginning of eighties (1), this classification describe more precisely the actions of researchers than more commonly used terms “retrospective” and “prospective” studies.

The term “retrolective studies” relates to data gathered from medical records or other sources when data collection took place without prior planning for the needs of a present study, while the term “prolective studies” relates to data collected by planning in advance.

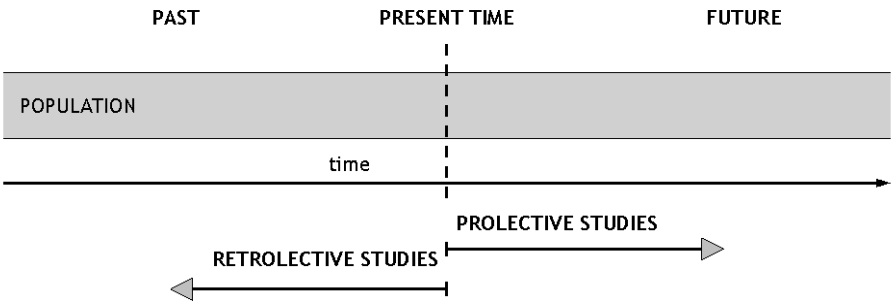


Figure 4. Schematic presentation of retrolective and prolective studies.

Studies based on individual level data versus studies based on population/aggregated level data

Another feature and corresponding way of classification of epidemiological studies involves the level of measurement. According to the level of measurement of variables that enter the studies, studies are classified as (11-21):

- individual level studies;
- aggregated level studies: measures in these studies are summaries of attributes calculated from data on individuals for whole populations, usually in well-defined geographic or administrative regions (e.g. countries, communities etc.). Examples for that kind of measurements would be: mean income; percentage of families below the poverty line or mean number of household members;
- group level studies: measures in these studies are estimates of (environmental) attributes that have individual analogues. Usually these measures are obtained from different environmental surveys. Examples for that kind of measurements would be: maximum daily exposure to ozone, mean annual exposures to radon gas; daily mean levels of environmental tobacco smoke in public buildings;
- population studies: measures in these studies are attributes that pertain to groups and do not have analogues at the individual level. Examples for that kind of measurements would be: total area of green space; number of private medical clinics; population density.

Studies for generating versus studies for testing hypotheses

This classification of epidemiological studies is one of less frequently used, since studies for testing hypotheses take advantage over studies for generating them.

Usually, epidemiological studies for generating hypotheses are basically descriptive, and as such are kept in the background in comparison to analytical studies. This is from historical point of view understandable to the certain extent,

since public health is in many respects tightly connected to biomedicine, where studies for testing hypotheses are taking huge advantage, especially experimental one.

In spite this fact, importance of studies for generating hypotheses lies in the fact that they provide essential contextual information for sound analytical studies for testing hypotheses which are mostly also more expensive (Figure 5).

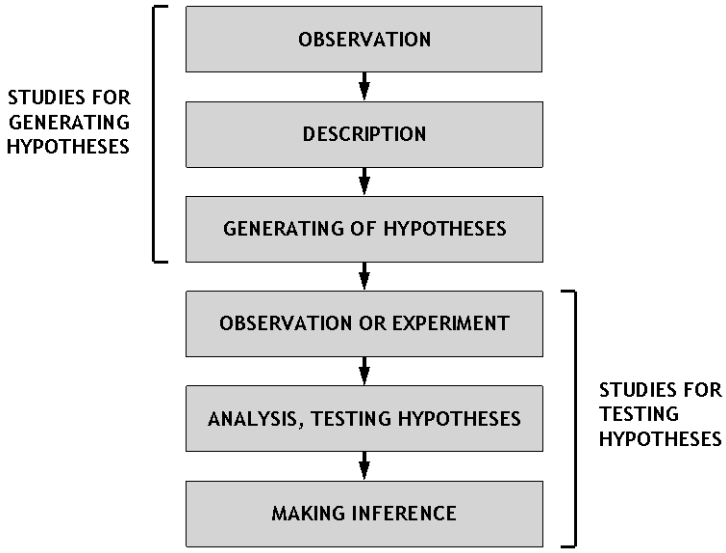


Figure 5. Schematic presentation of position of studies for generating, and studies for testing hypotheses in epidemiologic inference.

One should be aware that both types of epidemiological studies are part of the same process which could be more sound and effective when both types of studies complement each other.

Among others, very important studies for hypotheses generation are routine analyses of vital statistics and other notifiable events, periodic surveys of health status, knowledge, beliefs, attitudes, practices, behaviours, environmental exposures, and health care services performance, as well as ecological studies that compare information across geographical or administrative units.

In contrast, epidemiological studies for testing hypotheses are basically analytical. The most respected are experimental and cohort studies. Hypotheses could be tested also in case-control, and cross-sectional studies.

This distinction is more and more important, especially in the context of evidence based public health. Formulation of sound hypotheses on the basis of available data, e.g. on surveillance and monitoring data that are later verified by in-depth analyses is becoming an imperative.

Studies based on routine data versus studies based on data focused in analyzing a specific problem

A lot of data in public health is gathered in the frame of routine surveillance of major public health events. Basically, routine data are meant to describe major health conditions and health-related characteristics of populations, typically in terms of person, place, and time. As stated previously, their results are usually presented in health statistics yearbooks, and similar publications. Nevertheless, they could be logically and reasonably used also for other more in-depth analyses.

On the other hand, for less important or evolving problems, specific studies are more appropriate, since they are less expensive.

Studies focused on exposure rather than on effect

In the past, environmental epidemiology and occupationally epidemiology have been manly oriented in studying associations between disease and environmental agents. A broader approach is currently envisaged, which is primarily focused on exposure circumstances and which considers as dependent variables all possible health effects of environmental agents to which populations are exposed (22,23).

There are several reasons for the shift from disease- to exposure-centred environmental epidemiology:

1. Firstly, particularly in developed countries, degenerative, chronic diseases (such as cancer, lung emphysema, etc.) have become the prevailing pathology: the aetiology of many of these conditions is multifactorial, i.e. no specific hazard can be considered as a necessary cause. To further complicate the picture, many environmental hazards (e.g. excess dietary fat, asbestos, etc.) are causally associated with more than one disease.
2. Secondly, most environment-induced ill-effects are dose-related. For a given hazard, there may well be exposures either low enough, or of short enough duration, as to be negligible in terms of risk. It has also become obvious that ill-effects are frequently the result of interaction (addition, synergism, antagonism, etc.) between different hazards. For the same exposure to a given hazard, the risk may differ according to which other hazards are present or not.
3. Thirdly, analytical techniques for measuring pollutants in the environment have been used more and more, and their sensitivity has increased by several orders of magnitude. Consequently, there has been a dramatic increase in hazard-specific environmental data requiring risk evaluation.
4. Finally, health authorities, public opinion, and the scientific community have become increasingly concerned by the number of environmental contaminants for which potentially deleterious effects are unknown or poorly understood.

Studies for planning versus studies for evaluating public health interventions

Again, the distinction between “studies for planning public health interventions” and “studies for evaluating public health interventions” is one of intent, objective, and approach, rather than one of design. In this respect, data obtained in public health research usually could be explored for planning or evaluation.

Clinical epidemiologic studies versus public health epidemiologic studies

Clinical epidemiological studies are studies conducted in clinical settings, usually by clinicians, with patients as the subjects of the study. They apply epidemiological principles and methods to problems observed in clinical medicine. Their intention is to use the information from classical epidemiology to aid clinical decision making (1).

In contrast, classic (public health) epidemiological studies are intended to identify causes of diseases, and measure risk (1).

Some other features of epidemiologic studies

Along presented features of epidemiologic studies there exist other features as well. It could be worthy to mention at least two of them, being pragmatic and explanatory study:

- according to Last et al. (1), pragmatic study is a study aimed at providing a basis for decisions about health care, or evaluating previous action (interventions),
- according to Last et al. (1), explanatory study is a study aimed at explaining rather than merely describing the situation of a certain health problem by isolating the effects of specific variables and understanding the mechanisms of action.

EXERCISE

Task 1

Students carefully read the theoretical background of this module and recommended readings.

Task 2

Students make groups for crossover method of discussion. Every student is labeled with two labels being letters and figures: A1, A2, A3, B1, B2, B3, C1, C2, C3, etc.

For the first part of this task, all As, Bs, and Cs work together. The first part of the task is:

- GROUP A: discuss differences of experimental versus observational studies in relation to descriptive versus analytical studies;
- GROUP B: discuss differences of cross-sectional or transversal versus follow-up or longitudinal studies in relation to prevalence versus incidence studies;
- GROUP C: discuss differences of retrospective versus prospective versus non-directional studies in relation to retrolective versus prolective studies.

For the second part of this task, all 1s, 2s, and 3s work together. The second part of the task is the same as the first, instead groups are different:

- GROUP 1: discuss differences of experimental versus observational studies in relation to descriptive versus analytical studies;
- GROUP 2: discuss differences of cross-sectional or transversal versus follow-up or longitudinal studies in relation to prevalence versus incidence studies;
- GROUP 3: discuss differences of retrospective versus prospective versus non-directional studies in relation to retrolective versus prolective studies.

Task 3

In large group discuss other classifications.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
3. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia, W.B. Saunders company; 2001.
4. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
6. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
7. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Brown and Company; 1987.
8. Schoenbach, VJ, Rosamond WD. Understanding the fundamentals of epidemiology. An evolving text. Chapel Hill: University of North Carolina, School of Public Health, Department of Epidemiology, 2000. Available from URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: August 16, 2008.
9. Marušić M, Petrovečki M, Petrak J, Marušić A. Introduction to the scientific research in medicine (in Croatian). Zagreb: Medicinska naklada; 2000.
10. Gjorgov A, Gjorgov N. Research methodology and clinical epidemiology. Skopje: Grigor Prlicev Publishing House; 1999.
11. McMichael AJ. Setting environmental exposure standards: The role of the epidemiologist. Int. J Epidemiol 1989;18:10-16.
12. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993; 110:23-38.
13. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Monographs in Epidemiology and Biostatistics, New York Oxford University Press; 1994.
14. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994;84:825-9.
15. Susser M. The logic in ecological: II. The logic of design. Am J Public Health. 1994;84:830-5.
16. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16:61-81.

17. Susser M, Susser E. Choosing a future for epidemiology: I. Eras and paradigms. Am J Public Health. 1996;86:668-73.
18. Susser M, Susser E. Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. Am J Public Health. 1996;86:674-7.
19. Pearce N. Traditional epidemiology, modern epidemiology, and public health. Am J Public Health. 1996;86:678-83.
20. Künzli N, Tager IB. The semi-individual study in air pollution epidemiology: a valid design as compared to ecologic studies. Environ Health Perspect. 1997;105:1078-83.
21. Ribeiro MC, Pereira MJ, Soares A, Branquinho C, Augusto S, Llop E, et al. A study protocol to evaluate the relationship between outdoor air pollution and pregnancy outcomes. BMC Public Health. 2010;10:613.
22. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S, editors. Modern Epidemiology. Philadelphia: Lippincott-Raven; 1998. p.459-80.
23. Harvey T, Mahaffey KR, Velazquez S. Dourson M. Holistic risk assessment: an emerging process for environmental decisions. Regul Toxicol Pharmacol 1995; 22:110-7.

RECOMMENDED READINGS

1. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
2. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia, W.B. Saunders Company; 2001.
5. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
6. Schoenbach, VJ, Rosamond WD. Understanding the fundamentals of epidemiology. An evolving text. Chapel Hill: University of North Carolina, School of Public Health, Department of Epidemiology, 2000. Available at URL: http://www.epidemiolog.net/evolving/TableOfContents.htm. Accessed: August 16, 2008.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**ECOLOGICAL STUDIES: BASIC PRINCIPLES**
Module: 1.4.3	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Epidemiologic study, ecological study, ecological fallacy
Learning objectives	After completing this module students should: • know the definition and characteristics of epidemiological studies; • be familiar with strengths and weaknesses of epidemiological studies; • know how to get data for performing epidemiological studies, • know how to prepare data for an epidemiological study, and • know how to interpret results of an epidemiological study.
Abstract	Ecological study is an epidemiological study in which the units of analysis are populations or groups of people, rather than individuals. They are very applicable in situations in which data are not available on an individual level, but one should be careful when interpreting results of this type of epidemiological studies. The module is describing strengths and limitation of epidemiological studies, as well as procedure for analysing data in SPSS programme.
Teaching methods	An introductory lecture gives the students first insight in characteristics of ecological studies. The theoretical knowledge is illustrated by a case study. After introductory lectures students first carefully read the recommended readings. Afterwards they discuss the characteristics of ecological studies with other students. In continuation, they perform an ecological study analysis by themselves.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet, statistical programme and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire, and output of analysis of an ecological study with its interpretation.

ECOLOGICAL STUDIES: BASIC PRINCIPLES

Lijana Zaletel-Kragelj, Ivan Eržen

THEORETICAL BACKGROUND

Introduction

When we meet the term "ecological study" for the first time without knowing basic principles of epidemiology, the very first association is on ecology, which is defined as the study of the relationship among living organisms and their environment (1). A branch of ecology is human ecology, which is the study of human groups as influenced by environmental factors, including social and behavioural factors. But this association is not completely accurate. "Ecological studies" as a type of epidemiological studies are not related only to studying the influence of natural environment (mostly physical component of natural environment) on people. As a special type of epidemiological studies are useful in analysis in various situations not only in just described.

It would be logical and reasonable to think about this term in a broader, and in a narrower sense. Meant in their broad meaning, ecological studies are studies of influence of environmental factors on human health, while in a narrow, technical meaning, ecological studies are a special type of epidemiological studies. In this module, the narrow meaning will be discussed. However, this module describes only the most simple view of the ecological studies. In fact it describes only multiple-group studies design - the one the most frequently used in environmental epidemiology. More detailed view is far too complex to be presented here. All who wish to deepen their knowledge will find relevant literature in recommended readings list and in the module on principles and methods of environmental epidemiology in this book.

About ecological studies

Definition and description

According to Last et al. (1), ecological study is a study in which the units of analysis are populations or groups of people, rather than individuals. Similar definition is given by Bailey et al. (2): ecological studies are observational epidemiological studies that consider the characteristics of a disease and risk factors measured at the population rather than the individual level. They could be descriptive or analytical.

Thus, in the case of this type of epidemiological studies, if confronted to other types of epidemiological studies, the group of individuals, and not the individual person, is unit of observation and analysis. In this case, individual measurements are aggregated. Afterwards an aggregated measure is used (e.g. average or median value of individual values, or percentage of people with observed state) in ecological study. However, this is not the only type of data that can enter the ecological study.

Aggregation is usually carried out in a geographical region, or administrative region, as well as in different types of settings, e.g. health care settings, schools, etc.

Ecological studies have been conducted by social scientists for more than a century, and have been used extensively by epidemiologist in many research areas. Nevertheless, the distinction between individual level and group/population level (ecological) studies and the inferential implications are far more complicated and subtle than they first appear (3).

Characteristics

Level of measurement of entry data

According to the level of measurement of variables that enter the ecological studies, there are three different types of measures(4-6):

- aggregated level measures: summaries of attributes calculated from data on individuals for whole populations, usually in well-defined geographic or administrative regions (e.g. countries, communities etc.). Examples for that kind of measurements would be: mean income; percentage of families below the poverty line or mean number of household members;
- group level measures: estimates of (environmental) attributes that have individual analogues. Usually these measures are obtained from different environmental surveys. Examples for that kind of measurements would be: maximum daily exposure to ozone, mean annual exposures to radon gas; daily mean levels of environmental tobacco smoke in public buildings;
- population level measures (contextual): attributes that pertain to groups and do not have analogues at the individual level. Examples for that kind of measurements would be: total area of green space; number of private medical clinics; population density.

Studies may include also variables of different levels. The outcome variable in an ecological study could be measured on a quantitative scale (e.g. percentages, epidemiological rates), or qualitative (2,7).

Ecological studies purposes

The main purposes of using a study on a population level are to (2,7-10):

- study data that could be obtained only at group-level; health related data are sometimes available only at the group level (e.g. water or air pollution, percent of green areas in the community etc.);
- study group-specific effects; this is important since in public health interventions are performed at the group level rather on the individual level;
- assess very roughly a negative phenomenon which is perceived in a community level, and generate hypotheses for further investigation;
- investigate differences between populations – in some health phenomena differences are greater between populations than within them (e.g. due to differences in culture or health care system);
- describe patterns or trends on a geographic or administrative level;
- explore potential associations between community-level risk factors and disease.

Data source for ecological studies

Data for ecological studies are obtained from (2,7-9):

- most frequently ecological studies are performed on a routine data, since the valuable information about disease and exposure could often be abstracted from published statistics on international, national, regional, or local level. This means, that ecological studies usually do not require expensive or time consuming data collection. These routine data could be first obtained on individuals and than

aggregated, or are by nature data that are measured in natural environment, i.e. measurements of air pollution in vicinity of industry;
- data could also be obtained from periodical surveys, like health interview surveys about health behaviour.

Data could be obtained from one source only, or by combining different sources, and could be collected at different times for different purposes.

Advantages and disadvantages

Like other epidemiological studies, ecological studies have some advantages, and some disadvantages (2, 7-10). Some of them are presented in Table 1.

Table 1. Some advantages and disadvantages of ecological studies.

ADVANTAGES	DISADVANTAGES
1. quick and relatively inexpensive; 2. may be able to use readily available data; 3. useful in hypothesis generation; 4. allow estimation of effects not easily measurable for individuals; 5. permit exploratory analyses of potential factors in disease etiology; 6. appropriate when inferences are to be made about groups and not individuals; 7. useful for social scientists as well as epidemiologists; 8. useful in evaluation of new policies.	1. not able to analyze information on important factors that may be associated with the observed outcome because data are already collected for other purposes, thus difficult to control for confounders; 2. do not provide information about the relationship between risk factor levels and disease in individuals; 3. presence of so called "ecological fallacy" – association observed between variables on an aggregate/population level does not necessarily represent the association at an individual level; 4. exposures and outcomes are not measured on the same individuals; 5. in longitudinal ecological studies migration patterns over time could influence (e.g. diminish) the difference between observed groups.

Methods of analysis of ecological studies

Graphical presentation

Essential part of ecological study analysis is a graphical presentation. In fact, whenever observing the relationship between two quantitative variables, the first step is to plot the diagram. In the case of relating two variables, the diagram/chart is the joint distribution two-dimensional graph, called "scattergram", "scatter diagram", or "scatter plot" (9,11-13). The chart establishes the relationship of a dependent variable to an independent variable. The dependent variable is plotted on the vertical y-axis; the independent is plotted on the horizontal x-axis. Each point represents a place where the dependent and independent variables intersect. In Figure 1a some general examples of scatter plot are presented.

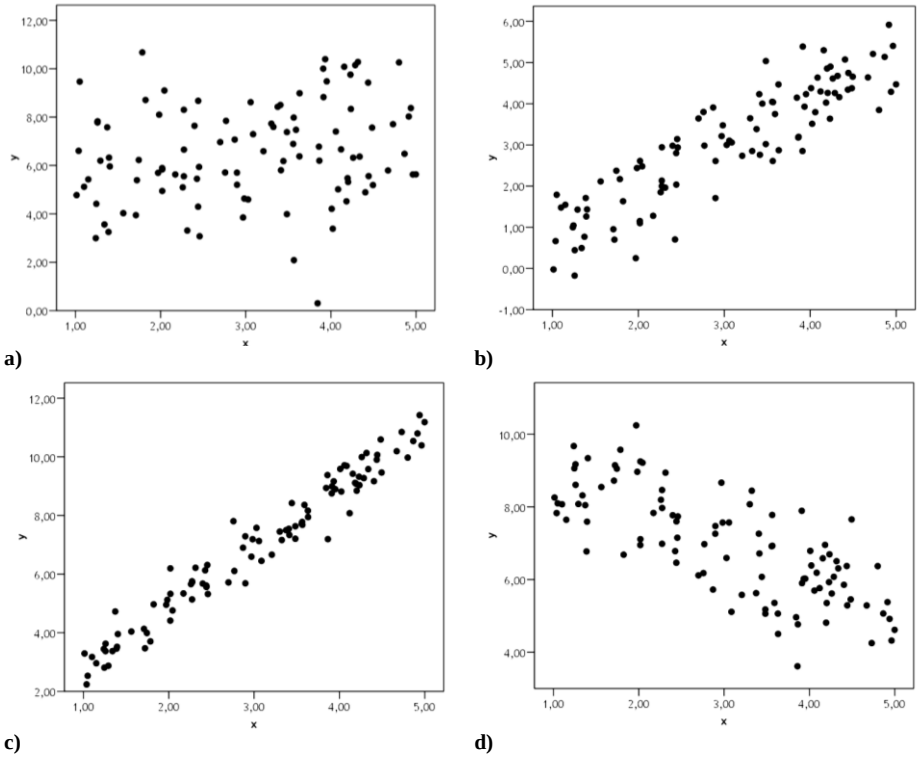


Figure 1. An example of a scatter plot with various degree of dispersion of intersection points (Figures 1a - c). In Figures 1a - c, relationship between variables X and Y is positive, while in Figure 1d is negative.

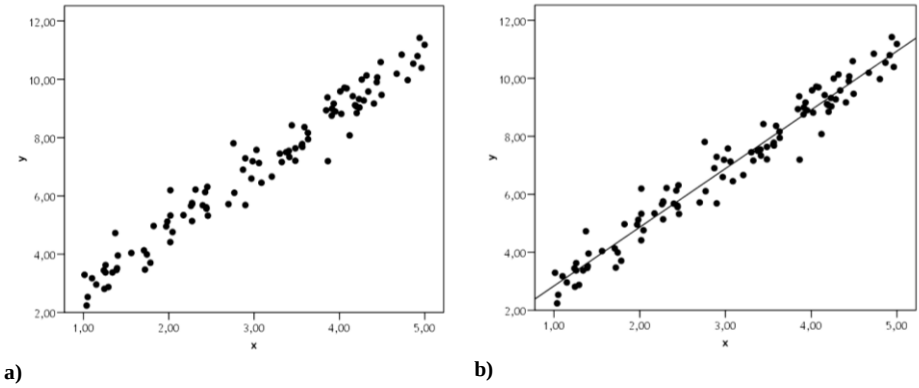


Figure 2. In Figure 2b, mathematical model (regression line) is added to a scatter plot from Figure 2a.

The dispersion (the scatter) of points of intersection of variable X and variable Y could express a pattern that could be summarized by a mathematical model. In the case presented in Figure 1, a straight line is the proper mathematical model (Figure 2b), so-called "regression line". The equation of this mathematical model on the sample level is (Equation 1):

$$y = a + bx \qquad \textbf{Equation 1.}$$

The presented relationship is only the most simple, being linear. There exist several others, but this issue is beyond the scope of this module.

Typical scatter plot in ecological studies has intersection points labelled. An example will be presented in case studies.

Correlation

Regarding the nature of the variables that enter the ecological studies, it is logical to use as an analytical method the statistical method called "correlation" (7, 9, 11-13), which measures the strengths of association between two variables, or in other words the grade of dispersion of intersection points around the mathematical model. The outcome measure is so-called "correlation coefficient" labelled at sample level as "r", if it is calculated using parametrical method (Pearson's correlation coefficient). The equation is as follows (Equation 2):

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \qquad \textbf{Equation 2.}$$

The value of Pearson's correlation coefficient lies between 0 and 1. The value 0 indicates that there is no linear relationship between variables, while the value 1 indicates the strongest relationship. In this case all intersection points lie on the regression line. In Figure 1a, the value of correlation coefficient is near 0, while in Figure 1c is rather close to 1.

The value of Pearson's correlation coefficient could have positive or negative sign. If with increasing of values of variable X values of variable Y are increasing (Figure 1b, 1c, 2a, 2b), the sign is positive, otherwise is negative (Figure 1d).

Pearson's correlation coefficient is only one correlation coefficient. The detailed description of characteristics of various types of correlation analysis is beyond the scope of this module. It is assumed that students are familiar with basic statistical methods, including correlation.

In the case of ecological studies the correlation is a special one, called "ecological correlation". According to Last et al. (1), ecological correlation is a correlation in which the units studied are populations rather than individuals. Correlations found in this manner may not hold true for the individual members of these populations.

Because the method of analysis is correlation, ecological studies are frequently called also "correlational studies" (8).

Interpretation of ecological studies

Although ecological studies are easily and inexpensively conducted, the results are often difficult to interpret.

In interpretation of results of epidemiologic studies one should be extremely cautious and careful. In fact, ecological studies may be useful pointer to further research, but conclusions derived from them must be interpreted wisely. Primarily, we need to be aware that research question in ecological study is about a population, and not about an individual (Example 1).

Example 1. *Does the overall occurrence of disease X in a population correlates with occurrence of the exposure in the population?*

It should be pointed out that ecological study design doesn't enable to draw any conclusion on the etiological factor of the observed phenomena otherwise there is a risk of so-called ecological fallacy.

Ecological fallacy

In ecological analysis, errors of inference may result because associations may be artifactually created or masked by the aggregation process.

The ecological fallacy, also known as aggregation bias or ecological bias (1), is the mistaken assumption that a statistical association observed between two ecologic (group-level) variables is equal to the association between the corresponding variables at the individual level. This assumption is often made implicitly or explicitly when using ecologic data to make inferences about the biologic (individual-level) effect of an exposure on the risk of a disease or other health outcome (1,2,7-9).

In extreme situation, an association at one level may disappear at another, or even be reversed. Suppose, for example, we observe a positive ecologic association between exposure prevalence and the rate of a disease across many regions (groups). The magnitude and direction of the association between exposure status and disease risk within regions (at the individual level) could be different from the ecologic association, even if there is no error in measuring either ecologic variable. Just because the disease rate is higher in regions with a larger exposure prevalence does not mean that exposed individuals are at greater risk of disease than are unexposed individuals. It is possible that the risk is particularly high for unexposed individuals living in regions with relatively high exposure prevalence. The underlying problem of the ecologic fallacy, therefore, is that each group is not entirely homogeneous with respect to the exposure. If every region were made up entirely of exposed individuals or unexposed individuals, then there would be no ecologic fallacy because information on the joint distribution of exposure and disease within groups would not be missing.

In conclusion, the aggregation of data that defines ecological studies results in an information loss that can lead to ecological bias, ecological fallacy respectively. It is due to the inability of ecological data to characterize within-

area variability in outcomes, exposures and confounders, when available. The only way to overcome such bias, while avoiding assumptions, concerning the missing information, that could not be checked, is to regard ecological studies as first step in analyzing the problem, and try to supplement the ecological-level information with individual-level.

Meaningless correlations

In ecological studies, meaningless correlations can occur. This kind of correlations sometimes occur when social, economic, or technological changes have the same trend over time as incidence or mortality rates.

Longitudinal (time trends) ecological studies

Usually, under term "ecological study" we understand the cross sectional ecological studies, that were already described. Both case studies presented in this module are also examples of cross-sectional ecological studies. But there exist also so-called longitudinal ecological studies (9).

Longitudinal ecological studies are studies made on ongoing frequent cross-sectional studies (surveillance or monitoring) to measure trends in disease rates over many years in a defined population.

By comparing the trends in disease rates over time with considering other changes in the population, it could be determined the impact of these changes on the disease rates. Important causal associations have been suggested by results of this type of studies. The detailed description is beyond the scope of this module.

Rationale for conducting ecological studies, and purpose of this type of studies in public health

There are several reasons for the widespread use of ecologic studies in epidemiology, despite frequent cautions about their methodological limitations (2,3,7-10):

1. Low cost and convenience.
   Ecologic studies are inexpensive and take little time because various secondary data sources, each involving different type of information needed for the analysis, can easily be linked at the aggregate/population level. For example, data obtained from population registries, vital statistics records, large sample surveys, and the census are often linked at the state, county or census-tract level.
2. Measurement limitations of individual-level studies.
   In environmental epidemiology and other research areas, we often cannot accurately measure relevant exposures of doses at the individual level for large numbers of subjects – at least not with available time and resources. Thus, the only practical way to measure the exposure may be ecologically. This advantage is especially true when investigating apparent clusters of disease in small areas (14). Sometimes individual-level exposures, such as dietary factors, cannot be measured accurately because of substantial within-person variability; yet ecologic measures might accurately reflect group averages.
3. Design limitations of individual-level studies.

Individual-level studies may not be practical for estimating exposure effects if the exposure varies little within the study area. Ecologic studies covering a much wider area, however, might be able to achieve substantial variation in mean exposure across groups.

4. Interest in ecologic effects and hypotheses generation.
   As noted above, the stated purpose of a study may be to assess an ecologic effect; i.e. the target level of inference may be ecologic rather than biologic – to understand differences in disease rates among populations. Ecologic effects are particularly relevant when evaluating the impacts of social processes or population interventions such as new programs, policies, or legislation. However, an interest in ecologic effects does not necessarily obviate the need for individual-level data.
5. Study group-specific effects.
   This is important since in public health interventions are performed at the group level rather on the individual level.
6. Simplicity of analysis and presentation.
   In large complex studies conducted at the individual level, it may be conceptually and statistically simpler to perform ecological analyses and to present ecological results than to do individual level analyses. For example, data from large periodic surveys are often analyzed ecologically by treating some combination of year, region, and demographic group as the unit of analysis.

Despite several practical advantages of ecologic studies, there are many methodological problems that severely limit causal inference (on the individual level), including ecologic and cross-level bias, problems of confounder control, within-group misclassification, and lack of adequate data, temporal ambiguity, co-linearity, and migration across groups (15).

CASE STUDIES

Case study 1: Standardized death rate in relation to Gross domestic product in countries of European Region of World Health Organization

Introduction

As an example of an ecological study could be observation of relationship between different measures/indicators, available on the level of countries, i.e. the indicators, available from Health for All data-Base of European Region of World Health Organization (16). For this module, we have chosen to present the relationship between standardized death rate (SDR) because of all causes in all age groups per 100,000 population, and Gross domestic product (GDP) per capita (in US$), in countries of European Region of World Health Organization. The data for the year 2001 were chosen since for this year they were available for most of the countries of the region.

Scatter plot

In Figure 3, the scatter plot of SDR because of all causes in all age groups per 100,000 population, and GDP per capita, in countries of European Region of World Health Organization, is presented.

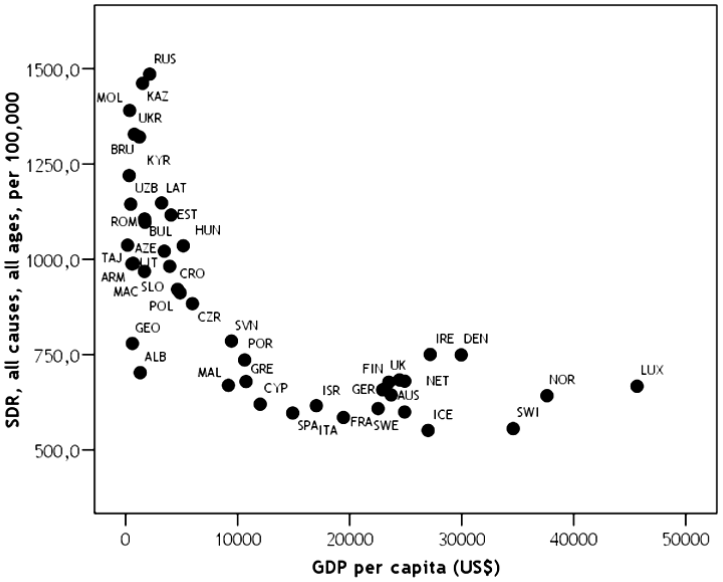


Figure 3. Standardized Death Rate (SDR) due to all causes in all age groups per 100,000 population in relation to Gross Domestic Product (GDP) (in US$), in countries of European Region of World Health Organization (data for 2001). LEGEND: ALB - Albania, AND - Andorra, ARM - Armenia, AUS - Austria, AZE - Azerbaijan, BRU - Belarus, BEL - Belgium, BH - Bosnia and Herzegovina, BUL - Bulgaria, CRO - Croatia, CYP - Cyprus, CZR - Czech Republic, DEN - Denmark, EST - Estonia, FIN - Finland, FRA - France, GEO - Georgia, GER - Germany, GRE - Greece, HUN - Hungary, ICE - Iceland, IRE - Ireland, ISR - Israel, ITA - Italy, KAZ - Kazakhstan, KYR - Kyrgyzstan, LAT - Latvia, LIT - Lithuania, LUX - Luxembourg, MAL Malta, MON - Monaco, MTN - Montenegro, NET - Netherlands, NOR - Norway, POL - Poland, POR - Portugal, MOL - Republic of Moldova, ROM - Romania, RUS - Russian Federation, SMA - San Marino, SER - Serbia, SLO - Slovakia, SVN - Slovenia, SPA - Spain, SWE - Sweden, SWI - Switzerland, TAJ - Tajikistan, MAC - TFYR Macedonia, TUR - Turkey, TUS - Turkmenistan, UKR Ukraine, UK - United Kingdom, UZB - Uzbekistan.

From this scatter plot could be seen that relationship between observed variables is not linear, so further analysis will not be just simple one, and the Pearson's correlation coefficient that assumes linearity of the relationship cannot be used.

Correlation analysis

One method to overcome the problem of non-linearity is that countries are classified in three groups according to GDP. It would be reasonable to make following groups:

- GDP up to 2999 US$ per capita;
- GDP 3000-7999 US$ per capita, and
- GDP 8000 US$ per capita or higher.

The boundaries were set arbitrary and only for the purposes of demonstration of one way of analysing ecological study data, and do not in any case mean that these boundaries are argumented.

In Table 2, values of correlation coefficients (r) in these groups are presented.

Table 2. Correlation coefficients in countries of European Region of World Health Organization, grouped according to GDP per capita.

Group according to GDP per capita	Correlation coefficient (r)
GROUP 1: GDP up to 2999 US$	0.182
GROUP 2: GDP 3000-7999 US$	-0.714
GROUP 3: GDP 8000 US$ or higher	-0.190

Interpretation of the results

From results of correlation analysis for three groups according to GDP could be concluded that GDP per capita is important factor in reducing general mortality in the population but only in a specified interval. Our results indicate that this interval is approximately between 3000 and 8000 US$. Only when a country attains certain level of GDP, it could expect that mortality could start to decrease. We could estimate that this threshold is about 3000 US$ per capita. It is interesting that also after a specified threshold increase in GDP per capita has no longer influence on reduction of general mortality of a population.

These results are valid only for the population level, and in any case not for an individual level.

Where were in 2001 PH-SEE Network Countries?

The scatter plot presented in Figure 3 could be supplemented with additional information. This could be very illustrative.

Since we are preparing this module for helping public health teachers from PH-SEE Network, it could be interesting to add on the scatter plot the information on participation of a county in PH-SEE Network (17).

Unfortunately, even we have chosen the year 2001 as the year with the most data available on both analyzed variables, data for both variables were not available

for Bosnia and Herzegovina, Montenegro, and Serbia for this year in Health for All database for the year 2007.

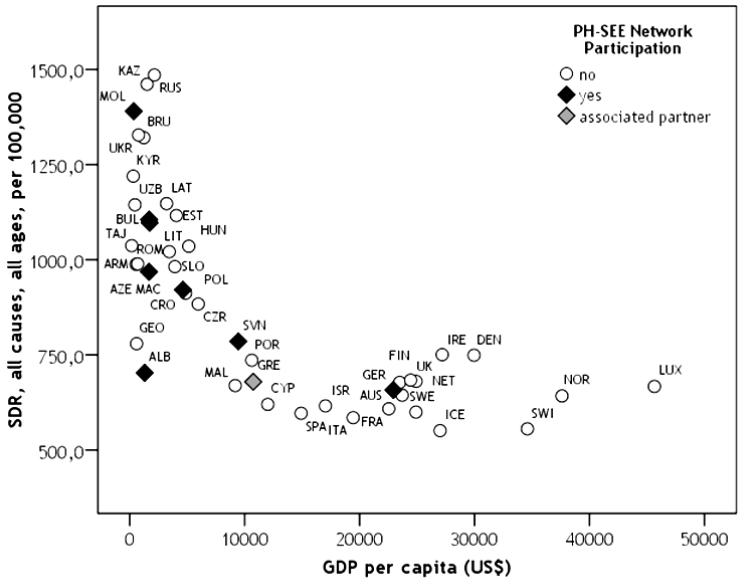


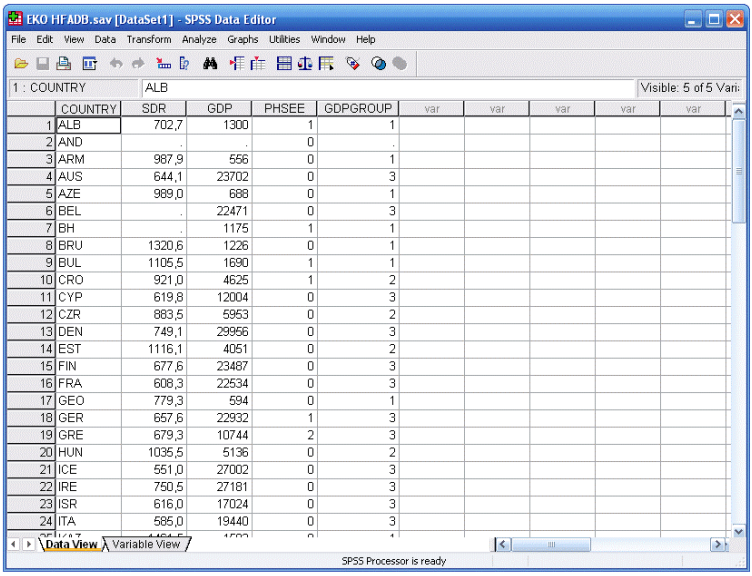
Figure 4. Standardized death rate (SDR) because of all causes in all age groups per 100,000 population in relation to Gross domestic product (GDP) (in US$), in countries of European Region of World Health Organization , by participation in PH-SEE Network (data for 2001). LEGEND: ALB - Albania, BUL - Bulgaria, CRO - Croatia, GER - Germany, GRE - Greece, MAC - Macedonia, MOL - Moldova, ROM - Romania, SVN - Slovenia.

Some tips for SPSS users

Performing basic analysis of simple ecological study in SPSS statistical programme is rather simple (18), nevertheless some tips could be welcome.

Data entry

Figure 5 presents the data matrix for analysis of ecological study data. We need to have five variables what mean five columns. Beside column with data on SDR and GDP, we at least have to have the column with country codes. Additionally we provided the information on participation in PH-SEE Network (codes: 0 - no, 1 - yes, 2 - associated partner), and group according to GDP (codes: 1 - GDP up to 2999 US$, 2 - GDP 3000-7999 US$, 3 - GDP 8000 US$ or higher)



	COUNTRY	SDR	GDP	PHSEE	GDPGROUP
1	ALB	702,7	1300	1	1
2	AND	.	.	0	.
3	ARM	987,9	556	0	1
4	AUS	644,1	23702	0	3
5	AZE	989,0	688	0	1
6	BEL	.	22471	0	3
7	BH	.	1175	1	1
8	BRU	1320,6	1226	0	1
9	BUL	1105,5	1690	1	1
10	CRO	921,0	4625	1	2
11	CYP	619,8	12004	0	3
12	CZR	883,5	5953	0	2
13	DEN	749,1	29956	0	3
14	EST	1116,1	4051	0	2
15	FIN	677,6	23487	0	3
16	FRA	608,3	22534	0	3
17	GEO	779,3	594	0	1
18	GER	657,6	22932	1	3
19	GRE	679,3	10744	2	3
20	HUN	1035,5	5136	0	2
21	ICE	551,0	27002	0	3
22	IRE	750,5	27181	0	3
23	ISR	616,0	17024	0	3
24	ITA	585,0	19440	0	3

Figure 5. SPSS Data Editor window with data properly prepared for making ecological study scatter plot in SPSS statistical programme.

Scatter plot

It is rather complicated to explain how to draw a scatter plot by using options offered by menu Graphs that is available for example from the SPSS Data Editor window. Instead of this we are rather providing SPSS Syntax Editor window with syntax for making ecological study scatter plots as presented in Figures 3 and 4. The syntax is provided in Figure 6.

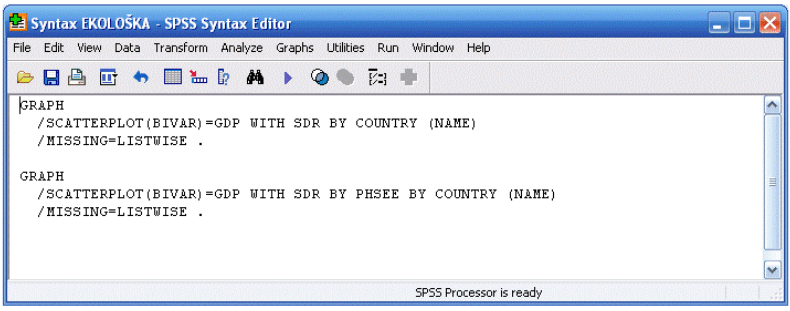


Figure 6. SPSS Syntax Editor window with syntax for making ecological study scatter plots as presented in Figures 3 and 4 in SPSS statistical programme.

The results of running the first syntax presented in Figure 6 is presented in Figure 7.

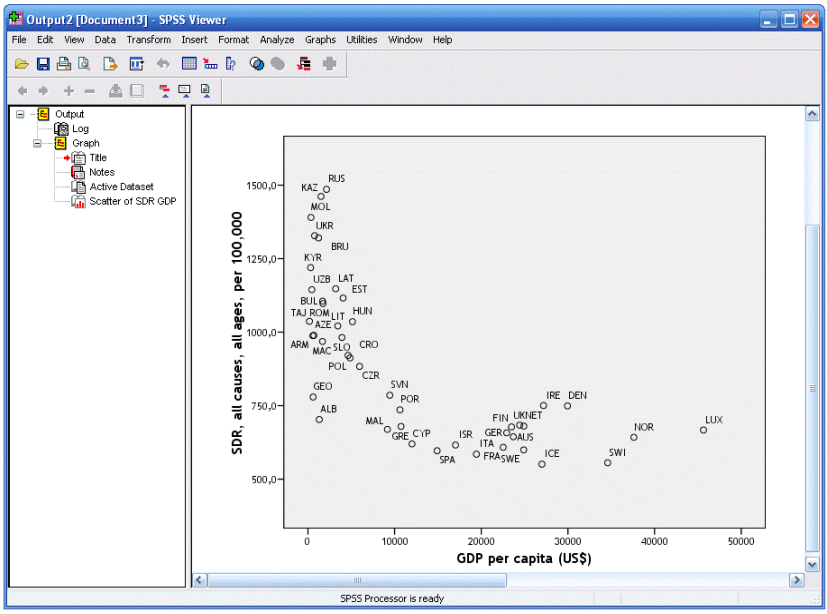


Figure 7. SPSS Viewer window with basic scatter plot in SPSS statistical programme.

Basic (default) scatter plot could be adapted according needs of the use using SPSS Chart Editor.

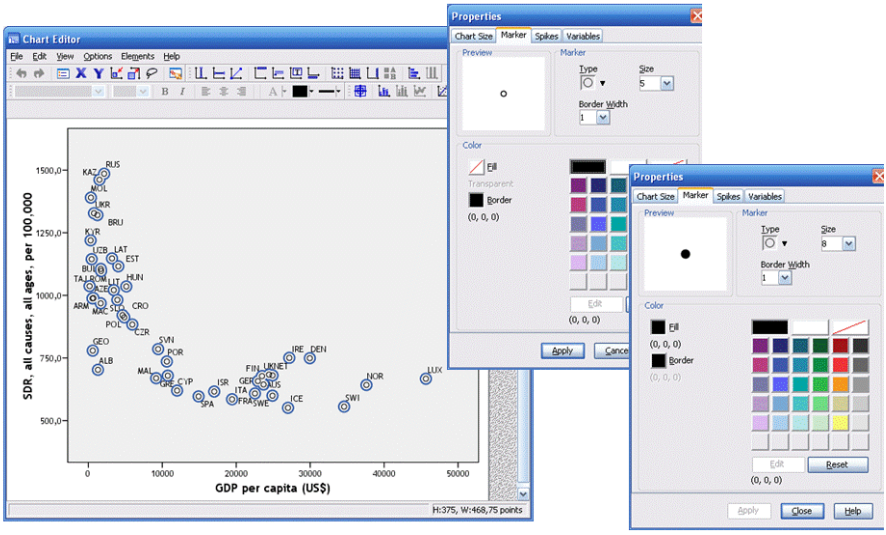


Figure 8. SPSS Chart Editor window with communication windows for adapting scatter plot according to user needs (e.g. as presented in Figures 2 and 3) in SPSS statistical programme.

Corerrelation analysis

Correlation analysis could be performed in SPSS using different procedures. The most simple is to use procedure Bivariate Correlation (Figure 9). The results of running this procedure are presented in Figure 10.

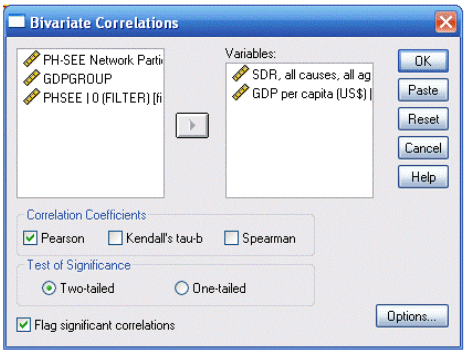


Figure 9. SPSS dialog box for running the Bivariate Correlation procedure in SPSS statistical programme.

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	,182
	Sig. (2-tailed)		,516
	N	15	15
GDP per capita (US$)	Pearson Correlation	,182	1
	Sig. (2-tailed)	,516	
	N	15	18

a. GDPGROUP = 1

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	-,714*
	Sig. (2-tailed)		,047
	N	8	8
GDP per capita (US$)	Pearson Correlation	-,714*	1
	Sig. (2-tailed)	,047	
	N	8	8

*. Correlation is significant at the 0.05 level (2-tailed).

a. GDPGROUP = 2

Correlations[a]

		SDR, all causes, all ages, per 100,000	GDP per capita (US$)
SDR, all causes, all ages, per 100,000	Pearson Correlation	1	-,190
	Sig. (2-tailed)		,410
	N	21	21
GDP per capita (US$)	Pearson Correlation	-,190	1
	Sig. (2-tailed)	,410	
	N	21	22

a. GDPGROUP = 3

Figure 10. SPSS Viewer window with results of running Bivariate Correlation procedure in SPSS statistical programme.

Case study 2: Hypertension (self-rated) in relation to average monthly gross earnings per person in paid employment in nine health region of Slovenia

Introduction

The second case study is basing on Slovene data. In this case we have used data from two different sources, being:

- data from the Statistical Yearbook of the Statistical Office of Republic of Slovenia (19) on gross earnings per person in paid employment on the community level. Since we needed the information on the level of nine health regions, we averaged values of communities, covered by corresponding Regional Institutes of Public Health;
- data on hypertension prevalence from CINDI Health Monitor Survey 2001 (20).

Thus we used data from one routine data source and data from one periodical survey.

Scatter plot

In Figure 11, the scatter plot of hypertension prevalence, and gross earnings per person in paid employment, is presented.

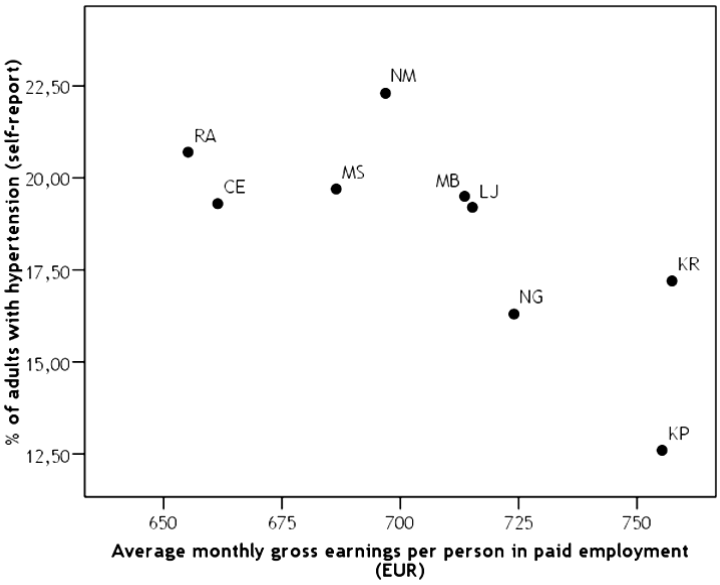


Figure 11. Percent of adults with hypertension (self-reported) in relation to average monthly gross earnings per person in paid employment in EUR, in nine health regions of Slovenia (data for 2001). LEGEND: CE - Celje Health Region, NG - Nova Gorica Health Region, KP - Koper Health Region, KR - Kranj Health Region, LJ - Ljubljana Health Region, MB - Maribor Health Region, MS - Murska Sobota Health Region, NM - Novo mesto Health Region, RA - Ravne Health Region.

From this chart we could assume that rather strong negative correlation is present between variables we put in relation. The relationship is very close to linear, thus Pearson's correlation coefficient could be calculated.

Correlation analysis

Correlation analysis confirmed our observation – value of Pearson's correlation coefficient is – 0.716, indicating rather strong negative correlation.

Interpretation of the results

From results of correlation analysis could be concluded that amount of average monthly gross earnings per person in paid employment is important factor in reducing hypertension in the population.

Again, these results are valid only for the population level, and in any case not for an individual level.

Is there some other pattern?

The scatter plot presented in Figure 11 could be supplemented with additional information on geographical location of health regions. This could be very illustrative (Figure 12).

Figure 12. Percent of adults with hypertension (self-reported) in relation to average monthly gross earnings per person in paid employment in EUR, in nine health regions of Slovenia by rough geographical location of health regions (data for 2001). LEGEND: CE - Celje Health Region, NG - Nova Gorica Health Region, KP - Koper Health Region, KR - Kranj Health Region, LJ - Ljubljana Health Region, MB - Maribor Health Region, MS - Murska Sobota Health Region, NM - Novo mesto Health Region, RA - Ravne Health Region.

It is interesting that situation is the best in the western part of Slovenia, and it is worsen in direction towards eastern part.

EXERCISE

Task 1

Carefully read the theoretical background on this module, and recommended readings.

Task 2

Using snowball technique, discuss the characteristics of ecological studies. Special attention pay on problems in interpretation of results of this type of epidemiological studies.

Task 3

In table 3, you will find data on obesity prevalence for 12 statistical regions of Slovenia for the year 2001. From the Web Page of Statistical Office of Republic of Slovenia (http://www.stat.si/eng/index.asp) find Statistical Yearbook with corresponding data (http://www.stat.si/eng/pub_letopis_prva.asp) on gross earnings per person in paid employment for 12 statistical regions of Slovenia (NOTE: in 2001 in Slovenia the currency was Slovenian tolar; the conversion rate to Euros is 1: 239.64). If available, make the scatter plot using SPSS statistical programme, otherwise make it manually.

Table 3. Data on obesity prevalence for 12 statistical regions of Slovenia for the year 2001. Data originate from CINDI Health Monitor Survey 2001 (20), and were prepared exclusively for this module.

Statistical region		Obesity prevalence (%)
1.	Pomurska	18.8
2.	Podravska	16.3
3.	Koroška	11.2
4.	Savinjska	16.1
5.	Zasavska	18.8
6.	Spodnjeposavska	21.6
7.	Jugovzhodna Slovenija	17.8
8.	Osrednjeslovenska	13.2
9.	Gorenjska	12.6
10.	Notranjsko-kraška	14.4
11.	Goriška	9.6
12.	Obalno-kraška	14.0

Task 4

In a group of three students prepare an example of ecological study using the World Health Organization, Regional Office for Europe Health for all Data Base

(http://www.euro.who.int/hfadb). Make a choice by yourselves. Provide short interpretation. The results are meant to be part of an assessment.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maindenhead: Open University Press; 2005.
3. Gail MH, Benichou J. Encyclopedia of epidemiologic methods. Chichester: John Wiley&Sons; 2000.
4. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993; 110:23-38.
5. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Monographs in Epidemiology and Biostatistics, New York Oxford Iniversity Press; 1994.
6. McMichael AJ. Setting environmental exposure standards: The role of epidemiologist. Int. J Epidemiol 1989;18:10-16.
7. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.231-262. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
8. Hennekens CH, Buring JE. Epidemiology in medicine. Boston, Little, Bown and Company; 1987.
9. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders company; 2001.
10. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing group, 1997. Available from URL: http://www.bmj.com/epidem/epid.html. Accessed: August 12, 2008.
11. Kirkwood B, Sterne JAC. Essential medical statistics. Second edition. Oxford: Blackwell Publishing Ltd.; 2005.
12. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1993.
13. Petrie A, Sabin C. Medical statistics at a glance. Malden, Mass: Blackwell; 2005.
14. Walter SD. The ecologic method in the study of environmental health. Overview of the method. Environ Health Perspect. 1992;94:61-65.
15. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annu Rev Public Health. 1995;16:61-81.
16. World Health Organization, Regional Office for Europe. European Health for all Database, HFA-DB. Copenhagen: World Health Organization, Regional Office for Europe, 2007. Available from URL: http://www.euro.who.int/hfadb. Accessed: August 12, 2008.
17. Forum for Public Health in South Eastern Europe, Programmes for Training and Reasearch in Public Health. PH-SEE Participating Centres Page [Internet]. Zagreb: Andrija Stampar School of Public Health, 2007 Nov 6. Available from URL: http://www.snz.hr/ph-see/centres.htm. Accessed: August 23, 2008.
18. SPSS 16.0 Brief guide. Chicago: SPSS Inc., 2007. Available from URL: www.hanken.fi/student/media/3618/spssbriefguide160.pdf. Accessed: August 12, 2008.
19. Statistični urad Republike Slovenije. Statistical yearbook 2006. Ljubljana: Statistični urad Republike Slovenije, 2006. Available from URL: http://www.stat.si. Accessed: 23 August, 2008.
20. Zaletel-Kragelj L, Fras Z, Maučec Zakotnik J, editors. Risky behaviours related to health and selected health conditions in adult population of Slovenia: results of Slovenia CINDI Health Monitor Survey 2001 (in Slovene). Ljubljana: CINDI Slovenia; 2004.

RECOMMENDED READINGS

1. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
2. Coggon D, Rose G, Barker DJP. Epidemiology for the uninitiated. Fourth edition. London: BMJ Publishing group; 1997. Available from URL: http://www.bmj.com/epidem/epid.html. Accessed: August 12, 2008.
3. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
4. Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health. 1994;84:819-24.
5. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994;84:825-9.
6. Susser M. The logic in ecological: II. The logic of design. Am J Public Health. 1994;84:830-5.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**CROSS-SECTIONAL STUDIES**
Module: 1.4.4	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Lijana Zaletel-Kragelj** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: lijana.kragelj@mf.uni-lj.si
Keywords	Epidemiologic study, cross-sectional study
Learning objectives	After completing this module students should: • know the definition and characteristics of cross-sectional studies; • be familiar with designing phase of cross-sectional studies; • be familiar with planning phase of cross-sectional studies.
Abstract	Cross-sectional studies are observational epidemiological studies of health status of the population in which a cross-section through frequency and characteristics of health outcomes and other health related events like exposures are studied and therefore provide prevalence data. They are very applicable in searching for general insight in health states and conditions that last a relatively long time as well as various for risk factors for them. They provide descriptive information for designing other types of epidemiological studies. The module is describing principles of cross-sectional surveys, especially their designing and planning phase.
Teaching methods	An introductory lecture gives the students first insight in characteristics of cross-sectional studies. The theoretical knowledge is illustrated by a case study. After introductory lectures students first carefully read the recommended readings. Afterwards they discuss the characteristics of cross-sectional studies with other students, especially the designing and planning phase of this type of epidemiological studies. In continuation, they need to find published materials (e.g. papers) on cross-sectional studies and present their findings to other students.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire.

CROSS-SECTIONAL STUDIES

Lijana Zaletel-Kragelj, Ivan Eržen

THEORETICAL BACKGROUND

About cross-sectional studies

Definition and description

There exist several similar definitions of cross-sectional studies (CSS):

- according to Last et al. (1), CSS are studies that examine the relationship between diseases or other health-related characteristics, and other phenomena of interest in a defined population at a particular time,
- a summary of several other definitions is that CSS are observational epidemiological studies of health status of the population in which a »snap-shot« of or a cross-section through frequency and characteristics of health outcomes and other health related events like exposures are studied (2-6). This characteristic also gave the name to this type of epidemiological studies,
- CSS are studies that measure the prevalence of health outcomes or determinants of health, or both, in a population at a specific point in time, or over a short period (Figure 1) (4).

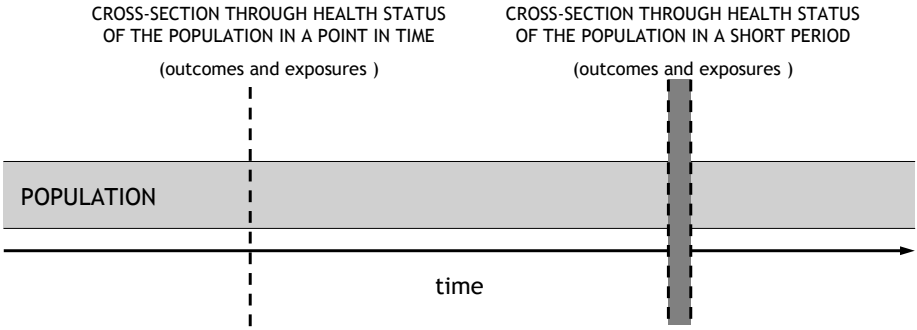


Figure 1. Schematic presentation of a cross-sectional study.

Health outcomes and other health related events could be measured in CSS on different measurement scale. In those CSS in which the outcome event is dichotomous the prevalence of this dichotomous event is recorded. This is the reason that CSS are also called prevalence studies (3,5,7,8). Prevalence studies thus could be on one hand regarded as a subgroup of cross-sectional studies, while on the other hand all CSS could be regarded as prevalence studies since we can dichotomize values of every observed outcome.

The selected specific point in time could be a time window within which data are collected (e.g. calendar week or month). It could also be a specific point in time in the course of events, differing in respect of each individual study subject with regard to the actual time (beginning of schooling, retirement, etc.) (3,9,10).

Frequently, CSS are designated as surveys.

Advantages and disadvantages

There exist several advantages and disadvantages of CSS (2,4,8,10). They are summarized in Table 1.

Table 1. Some advantages and disadvantages of cross-sectional studies.

ADVANTAGES	DISADVANTAGES
1. Relatively easy, quick and inexpensive, 2. Good study design for hypothesis generation, 3. Health outcome is measured at one point in time and the exposure may be measured from the same individuals at the same time and/or historical exposure information may be available, 4. Suitable for studying multiple exposures and/or multiple health outcomes, 5. Suitable for assessment of the prevalence of the events, 6. Suitable for estimating overall and specific health events prevalence, 7. Suitable for observation of frequent states of long duration, 8. Particularly suitable for observation of non-fatal diseases, degenerative diseases with no clear point of onset (e.g. Chronic bronchitis), or for examining effects on physiologic variables (e.g. blood pressure, serum glucose etc.), 9. Suitable for monitoring of the relationship between permanent and invariable exposures (risk factors) and health outcomes, 10. High generalizability, 11. Often good first step for new study issue in public health, 12. Good for public health programmes and health care service planning.	1. Study design not always appropriate, 2. Susceptible to selection bias (possibility of high proportion of long term survivors) i.e. individuals who either recover or die from a disease quickly have less of a chance of being included in the disease group, 3. Susceptible to misclassification (e.g. recall) 4. Not suitable for rare diseases/exposures, or diseases/exposures with short duration 5. Not a useful study design for establishing causal relationships because of problems with temporal sequence of data but this problem could be avoided in repeated cross-sectional studies.

Aims

The aims of the CSS are:

- to describe the frequency and characteristics of the observed health related phenomena at a certain point in time or within a time window (3,6),
- to analyse the relationship between two or more health related phenomena (2,4),

Since CSS are useful in description and in analysis of health phenomena they could be classified among descriptive (3,6), as well as among analytical epidemiological studies (2,4). Today they represent one of the most important tools of evidence based public health (11)

Methods

Sources of data

Theoretically, the source for CSS is a population. But since total population is usually hard to reach, a sample is drawn from the population. Not all members of a population or a sample under observation respond to the invitation and take part in CSS. These relations should be clear.

1. Population.
   All epidemiologic studies are based on a particular population. In this respect we need to distinguish between target and source population:
   - target population is the population which is to be subject to inference on the basis of the results of the CSS (1,4),
   - source population is the group of participants from whom we have collected data (4). It is also called the study population or base population.

   Source and target population could be the same.
2. Sample.
   Since CSS is usually not possible to be conducted on the total source population, usually a sample is drawn.

   There are several methods of sampling. In general, they could be classified into two major groups (9):
   - probability sampling – this type of sampling is also called random sampling. Types of random sampling procedures are simple random sampling, systematic sampling, multi-stage sampling, stratified random sampling and cluster sampling (4,9,12),
   - non-probability sampling - convenience sampling and purposive sampling are the types of non-probability sampling procedures (4,9,12).

   In epidemiology, probability sampling is preferred.

   The sample size depends on the characteristics of population under observation, on the purpose of the CSS, methods of data collection and data analysis methods.
3. Respondents.
   We usually cannot include all residents invited to take part in the CSS (selected in a sample) as some of them could simply not be found and the others refuse to cooperate. Those willing to take an active part in the CSS are representing only a sub-group of the randomly selected sample. They are called respondents or participants.

A lot of effort should be put to reach as high response rate as possible in order to avoid as much selection bias as possible (4,7,9).

Tools and methods of data collection

There exist several tools and methods for data collection in cross-sectional studies.

1. Tools for data collection.
   Collection of data in CSS could be carried out:
   - by the means of questionnaires which enables to pose the same questions in the same way to each respondent in the CSS. The questionnaire should be as short as possible and each question should be well considered (4,9),
   - by the means of health examination including diagnostic and laboratory tests.

   According to which tool is used to collect data there exist two main types of CSS (13):
   - health interview surveys or HIS surveys in which collection of data is carried out only by the means of questionnaires, and
   - health examination surveys or HES surveys which are usually a combination of questionnaires and health examination including diagnostic and laboratory tests.
2. Methods of data collection.
   In HIS, questionnaires may be communicated to the randomly selected study subjects in three ways: through mail, through personal interview or through telephone interview. Each of these methods has their own advantages and disadvantages (9), which are summarised in Table 2.

   In HES, the contact between participants and research personnel is personal since the health examination is a component part of the CSS. In this type of surveys, also questionnaires are usually communicated to the randomly selected study subjects through personal interview.

Preparing data for analysis and analysis

Getting data ready for analysis in CSS starts already at drafting the questionnaire where in respect of individual questions the codes for different answers are already predefined. The encoded data are then entered in the data matrix or data spreadsheet. For data entry the widely used spreadsheet programmes may be applied, however the analysis should be carried out by the means of one of the quality programmes specific for statistical analysis of data.

The basic analyses encompass the assessment of the prevalence[9] of phenomena under observation as a frequency measure in CSS studies.

Whenever we also wish to observe the relationship between a disease and a risk factor, the whole group of observed subjects should be divided with respect to exposure to the effect of the risk factor. By the term »risk factor« different

[9] NOTE: there exist several different types of prevalence measures: prevalence risk, prevalence rate and prevalence odds (14).

characteristics of respondents are indicated, which are already known to be related to their health status, and which, should be prevented or the extent of effect of which should be decreased (1,15). The assessments of prevalence of observed health outcome for each observed sub-group should be calculated, to be then compared.

Table 2. Some advantages and disadvantages of three different ways of intermediation of questionnaires to the respondents in cross-sectional studies.

ADVANTAGES	DISADVANTAGES
MAIL QUESTIONNAIRES	
1. low costs 2. rapid implementation 3. respondent fills-in the questionnaire at the most appropriate time for her/him 4. high level of anonymity 5. uniformity of posed questions 6. no bias resulting from subjectivity of the interviewer 7. respondent may check her/his answers 8. easy access to respondents	1. possibility of low response rate 2. small individual adaptability at posed questions (no help of interviewer) 3. not possible to use more complicated questions 4. possible influence of social environment (family members) 5. identity of the person who completed the questionnaire can not be controlled 6. order of precedence of answers to questions can not be supervised 7. many questions may not be answered to 8. spontaneity of answers can not be supervised 9. non-verbal messages of the respondent can not be observed 10. possibility of selection bias due to low response rate
FACE-TO-FACE INTERVIEWS	
1. high adaptability of the interviewer to the responder's understanding of questions 2. possibility of high response rate 3. non-verbal messages of respondent can be assessed 4. possibility of supervising the environment in which the respondent is completing the questionnaire 5. possibility of supervising the order of precedence of answering the questions 6. higher possibility of spontaneity of answers 7. possibility to control the identity of respondent 8. possibility of posing more complicated questions 9. supervision over completeness of answers	1. very high costs 2. conduction taking long time 3. respondent can not supervise and check her/his answers 4. possibility of not-suitable time for filling-in 5. low anonymity rate 6. less unified way of posing questions 7. possibility of difficult access to respondents 8. possibility of bias due to the influence of interviewer

Table 2. Cont.

ADVANTAGES	DISADVANTAGES
TELEPHONE INTERVIEWS	
1. possible adaptability of the interviewer to the responder's understanding of questions 2. possibility of supervising the order of precedence of answering the questions 3. possibility of spontaneity of answers 4. possibility of posing more complicated questions 5. supervision over completeness of answers	1. high costs 2. conduction taking a long time, however not as long as at personal interviews 3. respondent can not supervise and check her/his answers 4. high possibility of not-suitable time for filling-in 5. possibility of low response rate 6. low anonymity rate, however higher than at personal interviews 7. less unified way of posing questions 8. possibility of bias due to the influence of interviewer 9. the control of the identity of person who completed the questionnaire not possible 10. non-verbal messages of the respondent can not be observed

Strength of association may be assessed in different ways. We can observe the difference between two prevalence risks or rates, or the prevalence risk or rate ratio, as well as use the odds ratio (16).

Presentation and interpretation of the results

The results of CSS may be presented in different ways depending on for who the presentation is intended:

- to the expert public, results are usually presented in the form of articles published in scientific and other journals, or as presentations at different meetings,
- wider laic public is usually informed through the mass media; nowadays also web pages are frequently applied media.

Interpretation of results should be carried out with all due attention and impact of possible biases and errors, which were eventually done during the designing, planning or conducting of the CSS, should be taken into consideration. We should be aware, that the prevalence should be interpreted with a lot of caution, taking into account all potential influences on its value. Even if the errors are not the evident explanation for the observed relationship between the two phenomena, possibility of causality should be assessed very carefully.

Course

The course of a CSS should follow the general principles, which are common to most study designs, so to CSS as well.

Phases and periods of the course of cross-sectional studies

The whole course of CSS can be divided into different units.

The smallest unit of the CSS course is called phase of the CSS. Phases of the CSS are usually conducted according to the following scheme (Figure 2) (4,7,17,18):

1. Designing phase,
2. Planning phase,
3. Organizing phase (preparation for the implementation),
4. Data collection/analysis phase
5. Results dissemination phase (interpretation and presentation of the results).

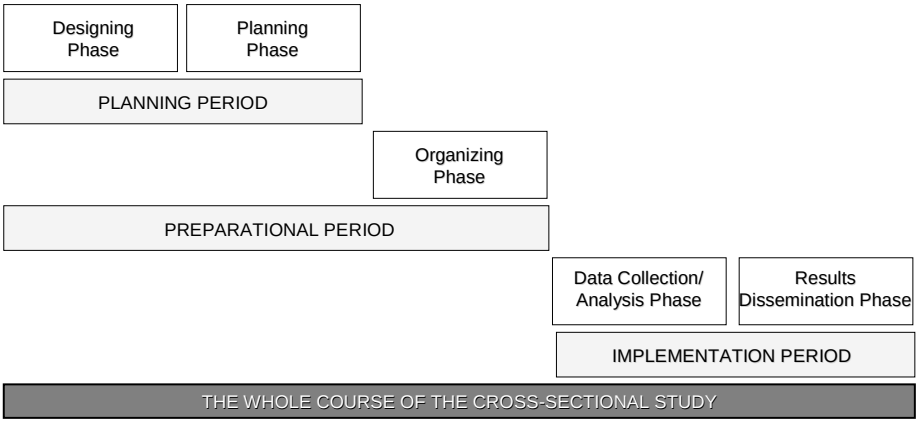


Figure 2. Phases and periods of the course of the cross-sectional study (NOTE: to make the presentation easier, at the picture all phases are represented as being of the same duration, which does not correspond to the facts of the practice; also overlapping of phases and periods is not taken into consideration).

As regards the duration, phases can vary significantly. With various circumstances taken into account, they can be conducted successive or parallel or may be even more or less intertwined among themselves. Intertwinement is usually more explicitly expressed with the CSS of rapid course. The first and the second, and the third and the fourth phase of the CSS are frequently intertwining.

Periods are wider units of CSS course. A single period may be consisted of one single phase or of more phases (Figure 2):

1. Planning period of the CSS.
   The first two phases of the CSS are included in the planning period – designing and planning phases.
2. Preparational period of the CSS.

Preparation period includes more phases than the planning period. In addition to both phases of the planning period also the organizing phase of the CSS is included in this period.

3. Implementation period of the CSS.

There are numerous textbooks on epidemiological methods containing recommendations with regard to designing and planning and management of each individual phase and/or period (4,17-19).

Although all phases/periods of CSS course are important, planning period is the most important and most sensitive period. If designing and planning the CSS in the wrong way, the whole CSS could be set on an inappropriate basis, and the deficiencies of this period are very difficult to be eliminated in the later phases of the CSS. In order to avoid as many faults as possible, the course of the CSS must be planned systematically and with all due care. Therefore, in continuation special attention is given in this module to recommendations for designing and planning a CSS.

Recommendations for designing a cross-sectional study

Designing of a CSS is a creative process. A precise management of individual CSS is very difficult to be advised. Nevertheless, common recommendations on actions in this phase of the CSS, do exist (4,10,17,19) and are very similar to those applied in biomedical studies in general. They can be summarised in the following steps (Figure 3):

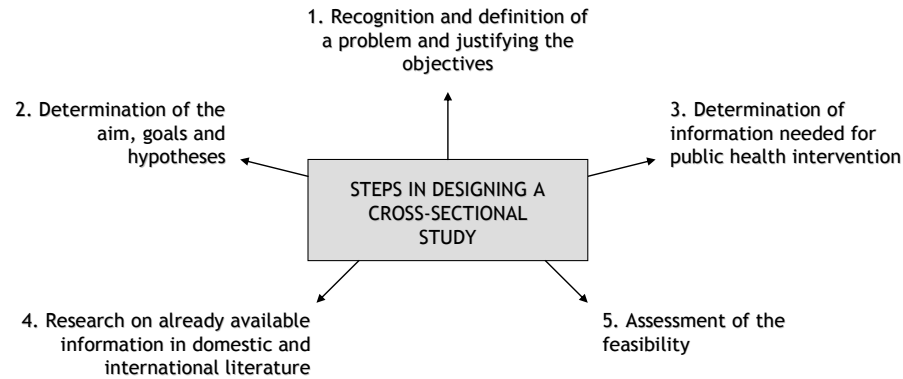


Figure 3. Steps in designing a cross-sectional study.

1. Recognition and definition of a problem and justifying of the objectives for the CSS.

   The cornerstone of a CSS is usually a public health event, which is recognised as a problem, which is so noticeable that we would like to investigate it and

search for the possible solutions. However this is not enough. Since the public health research field being so wide, we usually meet with problems already when focusing on the access to a certain problem. Our decision on being interested in something does not suffice for initiation of the CSS. This is only to define the study problem. The selected scope of the CSS must also be grounded. Since CSS are usually associated with big funds, these should be grounded by arguments.

The scope of the CSS is usually grounded on one or more objectives. However, the whole process of activities connected with the beginning of the CSS, is usually initiated by one of them.

2. Determination of the aim, goals and of the hypotheses of the CSS.
   To have most clear idea on the aims and goals it is one of the most important parts of the designing phase of CSS course. It could be helpful to know that:
   - the aim of the CSS is defined as that »what we shall strive for« in the CSS,
   - the goal as that »what should be attained« during the CSS in order to realise our striving to the maximal possible extent,
   - hypothesis or assumption is our proposal for understanding of the events and processes (our opinion on connectedness between the events under study).

   The aim and the goals should be set clearly. In the opposite case, it could be seemed that the CSS being conducted only as its own purpose.

   When determining hypothesis at CSS, problems are usual to arise. CSS are distinguishable from other studies in medicine for being first of all oriented in studying the size of health problems (its extension), rendering the hypothesis often possible to be set only upon the results obtained. CSS are primarily meant to set as a result the hypotheses which should be then tested in another study design.
3. Determination of information needed for public health intervention.
   With the aim to collect as many information on the problem as possible and to avoid collecting too much data, we should consider carefully which data to be of use for solving the problem. These are data, among others, for exact evaluation of the size and spread of the problem as well as data for forming the strategies for solving the problem.
4. Research on already available information in domestic and international literature.
   Aiming at determining the problem exactly, as well as the purpose and the goals of the CSS thereof, we have to carry out preliminary review on similar CSS at home or abroad, if available. Such a review could serve on one hand in choosing similar methods, but on the other hand it could result even in a decision that the CSS shall not have to be conducted in such a wide extent as it has been anticipated at the beginning. Besides it could also point out the possibilities and limitations at studying the selected problem.
5. Assessment of the feasibility of the planned CSS.
   When the planned CSS is established to be worth of conducting, it is to be estimated if:

- there are enough funds ensured at least for starting of the CSS,
- possibilities are existing for sustainability of the CSS,
- there are enough human sources for its execution,
- the CSS is feasible in the population,
- the CSS shall result in the information we are looking for,
- information, when collected and analysed, shall still be interested and useful.

Recommendations for planning a cross-sectional study

When we believe that the CSS is worth of conducting and that at least the main part of it is feasible, we shall set about its planning and assuring funds for the same (4,17-19) The steps in this phase are as follows (Figure 4).

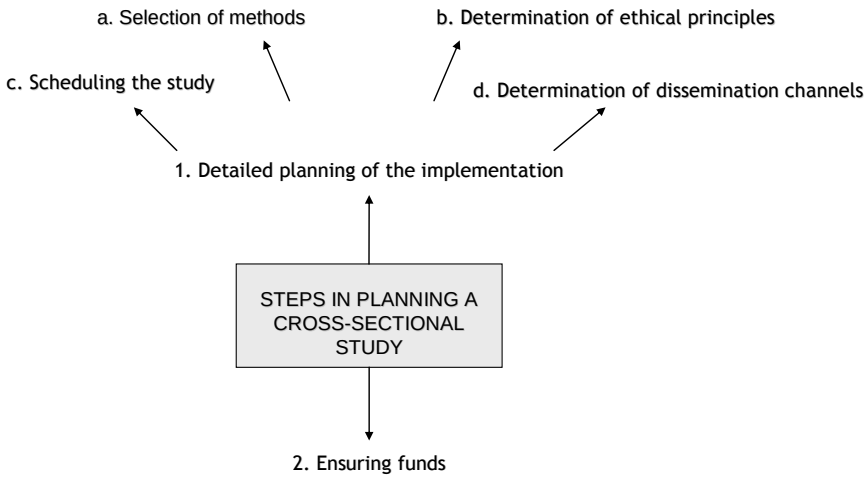


Figure 4. Steps in planning a cross-sectional study.

1. Detailed planning of the implementation of the CSS.
   Planning of the CSS is considered to be determination of supporting points in conducting of the CSS which shall enable us to make intended steps, and the way of their carrying out. Planning of the CSS should include the following elements:
   a. Selection of the method of approach.
      Different research methods were established by different sciences. The public health research is characterised by intertwining of biomedical and social sciences approaches (6,16). Regardless to this, following should be determined and/or argued:
      - what the target population is,

- how to select a sample from the source population and its size,
- tools and methods of collection of data,
- tools and methods of data analysis,
- results dissemination channels.

b. Determination of ethical principles.
The question whether to conduct a CSS is depending also in great deal on ethical problems. Therefore ethical instructions for the CSS s have been prepared by the World Health Organization (WHO) (20), basing these on the Helsinki Declaration (21) and other ethical principles (22).

In general, ethical principles of the CSS should keep to basic ethical principles in biomedicine, however differing from them. The CSS of such kind shall bear the responsibility on two levels – on the level of individuals and on the level of the community thereof. Such responsibility does not refer only to individuals and communities taking part in individual study, but also to other people, health of which could be protected or improved by the means of the results of the CSS. Therefore all respondents must be informed that with taking part therein something useful shall be done in respect of their own health as well in respect of the health of the community (5,22).

According to ethical principles, each individual should be acquainted with the purpose and course of the study. This should be paid special attention to when possible harmful effects are included. Each individual who decided to cooperate in the CSS shall have the right to withdraw there from at any time. If the person decides to cooperate we shall ask him/her for his/her consent. This could be in the form of a signed consent - informed consent, or a returned completed questionnaire shall be deemed a consent.

In CSS, the right to privacy of an individual and to confidentiality of information should be respected. However we are also obliged to inform the population on what we shall be doing and why and which shall be their benefits resulting thereof. Therefore all proposals for the CSS of this kind should be first addressed to the corresponding commissions (10).

c. Scheduling the CSS.
Duration of the CSS should be planned in advance. The timetable of the course as a whole and separate phases as well must be done. In order to be able to follow the timetable, each phase shall be set dates called milestones.

d. Determination of dissemination channels.
In the case of a CSS, also the dissemination channels i.e. how to present results to different public is recommended to be determined.

2. Searching for possibilities of ensuring funds for conduction of the CSS.
During the planning phase of the CSS, special attention should be put to ensuring the funds. Since epidemiological CSS are by the rule expensive, financing by the means of one big or more sources shall be needed. Therefore institutions should be found prior to start conducting the CSS, which shall have

interest in the problem and which shall also be willing to invest some funds in the implementation.

At the end of this phase considered also the end phase of the planning period of the CSS, a protocol of the CSS needs to be fully worked out.

The protocol – a plan of a cross-sectional study

Final product of the planning period is a study plan, composed of a conceptual and implementing part. Such plan is called protocol of the CSS (4,7).
The CSS protocol should be composed of the following parts:

1. Problem under observation and objectives of the CSS,
2. Aims and goals of the CSS,
3. Methods,
4. Ethical principles,
5. Time schedule of the course,
6. Results dissemination channels,
7. References.

CASE STUDY: CINDI HEALTH MONITOR SURVEY IN SLOVENIA, 2001

Introduction

The CSS entitled »Risk factors for non-communicable diseases in adults in Slovenia« (23) was the first CSS conducted in Slovenia to such large extent, aiming at studying health behaviour of the population (smoking, nutritional and exercise habits, habits connected to drinking alcoholic beverages, behaviour connected with the road-traffic safety, etc.) on the national and at the same time on the regional level.

By the means of this CSS, an attempt to establish the system of monitoring of health behaviour of Slovene population was intended to be made.

Conceptually, this CSS represented a part of a wider WHO project CINDI Health Monitor (CHM) (24). The project was conducted within the international programme for combating the non-communicable diseases of the WHO Countrywide Integrated Non-communicable Diseases Intervention (CINDI) Programme (24-28).

The ongoing WHO CHM project is mostly aiming at monitoring, assessing and comparing the trend of health behaviour in CINDI countries with different politically-economic systems. Owing to comparability, monitoring should be conducted under the uniform methodology.

The CHM methodology is based on the project Finbalt Health Monitor (24,29), which involves Finland and Baltic Republics Lithuania, Latvia and Estonia. The project is coordinated by Finland, due to its rich tradition and great experiences and successes in preventing non-communicable diseases. The most well-known project is the North Karelia Project (24). In this project, which is now lasting for more than 25 years, Finns succeeded to decline the coronary mortality rate among the male population aged 35-64 years, by about 75% (30). Besides, Finland is collecting data on health behaviour of its adult population on a random sample each year, starting in

1974. For this purpose they make use of mail CSS. Through the project Finbalt Health Monitor they started to spread their experiences also to the neighbouring countries. In 1990, Estonia was first to joint the project, to be then followed by Latvia in 1994 and Lithuania in 1998.

Since the Slovene survey of health behaviour represents a part of the CHM project, it can be also called CHM survey of Slovenia (CHMS-SI).

The activities associated with the CHMS-SI started at the end of February 2001. At that time, a first joint meeting was organised upon the initiative of the representatives of the Ministry of Health (MH). It was participated by representatives of the institutions, the knowledge and capacities of which enabled them to collaborate in the survey: CINDI Slovenia (CINDI-SI), Faculty of Medicine (FM), National Institute of Public Health of the Republic of Slovenia (NIPH) and regional health protection institutes (RHPI).

The first two tasks of invited public health experts were to communicate the idea of the survey to their institutions, and to assess the feasibility for carrying out such an extensive survey in Slovenia, which was to result in setting the basis for the system of monitoring of the health behaviour of the population. The estimation was not easy to be reached. The biggest problem represented an extremely short term, planned for the survey to be started. It was planned to be started in the middle of May of the same year. The third task was to find other public health experts, capable and willing to cooperate in such an extensive project outside the invited group. The meeting was aiming at forming a research group in the shortest possible time, which would be able to make the entire plan for the survey within the available period of time, and also to initiate the implementation thereof within the planned term.

All institutions which were invited to collaborate in the project, agreed to the collaboration. The research group was also formed quickly. Within the short time given, the group managed to make the plan for implementation of the survey and to start the survey within the planned term.

Designing phase

Recognition and definition of a problem and justification of the objectives of the survey

A problem

In Slovenia the problem of preterm mortality caused by non-communicable diseases has become so important that the need for studying and solving it has become evident (23,31).

Non-communicable diseases were in Slovenia, similar as in other European countries, one of the leading causes of death (31). Mortality rate resulting from such diseases was in Slovenia higher than for example in France, Germany, Italy, or in Finland, but lower than in Baltic republics (31).

Objectives

The need for studying and solving the problem of preterm mortality, caused by non-communicable diseases, was based on following objectives:

1. The first objective was related to the risk factors for non-communicable

diseases.

The problem of non-communicable diseases is mostly related to behavioural risk factors (12). When reviewing the existence of studies conducted within the period from 1991 on, the real estimation of the extension of these factors in Slovenia was shown not to be actually assessed before the CHMS-SI was started (32).

2. Second objective represented the guidelines of the Slovene health policy in the field of health care and health insurance development. These guidelines were determined in the National Health Care Programme of the Republic of Slovenia (NHCPRS) (33).

   In this programme, rising of the quality of health of the Slovene inhabitants was pointed out, as well as adjustment and improvement of the health system's operating in relation to the financial possibilities of Slovenia. In the programme, also the aims of the WHO were taken into consideration, which have been laid down in two documents: »Health for all« (34) and »Health 21« (35). The priority tasks of the NHCPRS were among others the following:

   - stimulating all holders of the health care to collaborate in forming and carrying out of the programmes on health promotion,
   - stimulating research of the population health by the means of interdisciplinary researches, first of all researches of life style and health behaviour patterns of people and the influence thereof on the health of the population,
   - taking part in the comprehensive programmes of health promotion of the WHO and EU, and
   - implementing the strategy of health protection of the population in compliance with the guidelines laid down in the documents of the WHO.

   All stated above was jointed in the CHMS-SI.

3. The third objective was the key aim of the international CINDI programme (23,25-28), based on the health promotion and prevention of the risk factors for non-communicable diseases. Slovenia has been taking part in the CINDI programme unofficially for over ten years, however officially from 1993. Also the aims thereof are in compliance with the strategic aims of the WHO.
4. The fourth objective was the new development policy of the CINDI programme.

   A decision was adopted in June 2000 at the headquarters of CINDI at the WHO Regional Office for Europe, to upgrade their activity (25-28). Surveys conducted up to that time, were limited only to demonstration areas of the participating countries. There were two surveys already conducted in Slovenia up to that time in 1991 and in 1996 (36,37), which were limited to the Ljubljana region only. With the stated date, the activities were supposed to be moved from community-regional level to the national level. The surveys were to be conducted each two years at the most instead of each five years. The activities were entitled CINDI Health Monitor (CHM) (24). This decision was based on the assessment that behavioural patterns linked with non-communicable diseases in wider Europe area, needed to be evaluated and improved. At the CINDI Winter School, held about one half of the year later in Helsinki, on 12 and 13 February 2001 (24), the representatives of

Slovenia, among which the representatives of the MH, the NIPH, RHPIs and CINDI-SI, estimated that a survey under the methodology CHM could be conducted in Slovenia already in the spring of 2001. Upon this decision, our survey was started.

Determination of the aim and goals

Aiming at contributing to detailed knowledge on health behaviour of adult population in Slovenia and consecutively contributing to realisation of the measures of the priorities determined in the NHCPRS (33), following goals were set as main (23):

- to investigate smoking habits on national and on regional level,
- to investigate nutritional habits on national and on regional level,
- to investigate alcohol consumption habits on national and on regional level,
- to investigate physical activity habits on national and on regional level,
- to investigate oral health habits on national and on regional level,
- to investigate road safety habits on national and on regional level,
- to investigate burden of stress on national and on regional level.

The hypotheses were not determined at this phase of the survey since they were to numerous. It was decided detailed hypotheses should be the issue of specific studies based on CHMS-SI data-base.

Definition of what information is needed

In the case of a CHMS-SI survey, we did not have to consider which information should be the result of the survey in order to be able to assess the size of the problem and to find the ways for the resolution thereof, since with entering the project CHM we accepted a common international questionnaire (24). The task of research group was to translate the questionnaire correctly into Slovene language and to adjust it to conditions in Slovenia.

Research on potential availability of the required information

The review of studies dealing with the prevalence of risk factors for non-communicable diseases in Slovenia (32) was carried out prior to the initiative event for the survey has been started. In the review, the results of studies were included, which were carried out in Slovenia after 1991 when our country became independent. These were the results of surveys of the programme CINDI-SI, of the Institute of Oncology, survey within the framework of the Slovenian public opinion (SPO), and of some other surveys (36-45). By the means of this review, the urgent need for exact evaluation of the spread of risk factors for non-communicable diseases in Slovenia as a whole and in each individual health region was established.

Assessment of feasibility of the planned survey

When assessing on feasibility of the planned survey, the funds showed to suffice at least for the beginning thereof. Human sources needed for the start of conducting of the survey were not questionable as well. Moreover, also the information, if possible to be collected and analysed, shall still be applicable and interesting.

When estimating feasibility of the survey, the response of the Slovene population to the survey and feasibility due to a short time available for the preparations thereto, turned out to be two major unknown items.

Regarding the response of the population, studies we were able to find in the short limited term in relation to the estimation of the response (46-48) indicated, that in Slovenia all kinds of response rates could be expected, from very good to very bad. Therefore we were also not able to estimate whether we shall acquire the searched information.

Nevertheless the preparations were carried on, since the CHMS-SI conducted in 2001 should represent also the study of feasibility of setting the monitoring system of the health behaviour of the adult population of Slovenia up.

Planning phase

Planning of the CHMS-SI survey

The planning of the CHMS-SI was in great extent directed by the WHO CHM international project. Recommendations within the framework of this project were the following (24):

1. Organisational recommendations.
   Organisational recommendations are as follows:
   - the survey should be conducted for the first time during the period 2001-2002, if possible,
   - the survey should be conducted in all countries at approximate the same season, recommended period being March-May,
   - it should be carried out, if possible, at the national level,
   - it should be carried out, if possible, every second year.
2. Methodological recommendations.
   Methodological recommendations are as follows:
   - the sample should be selected upon the principles of simple random sampling,
   - it should be of the size of at least 3000 units,
   - target population should be adults preferably of the age between 25-64 years (the range could also be wider with regard to the needs of individual country),
   - the survey should be based on a common core WHO CHM questionnaire,
   - a self-administered postal questionnaire is recommended; if not possible, face-to-face or telephone interviews can be used,
   - non-respondents should not be replaced with other individuals. They should be reminded by sending them a new invitation.

Exact planning of the course of the survey

Selection of methods

When planning survey methods in Slovenia, we mostly aimed at holding on the recommendations of the CHM, but this was not always possible. Whenever another methodology had to be applied, we tried to deviate from the recommended to minimal possible extent.

1. Selection of target population.
   We defined the target population in compliance with the recommendations of the CHM (24), meaning to have included in the study adult population aged 25-64 years.

   If comparing Slovenia with other countries in which similar studies were conducted (30), determination of the target population was characterised by regional approach. Since differences existed in mortality and certain diseases prevalence and incidence, and in some other socio-economic indicators between individual health regions in Slovenia (49-52), we decided to observe the target population separately in respect of each of 9 health regions.
2. Sampling method and sample size.
   According to the recommendations (24), the simple random sampling should be applied and the sample, if possible, it should be based on the population-based registry. The sample for studying the risk factors was selected on the basis of database of the Central Registry of Population (CRP) (53), what entirely corresponds to the recommendations. But we selected a slightly different sampling method. Since a single health region was selected as basic observational unit, the stratified random sampling was applied. Individual stratum of sample was represented by an individual health region of Slovenia. Such sampling method was allowed by the CHM.

   When determining the size of sample we were subject to the limitation of the smallest number of 3000 units, determined by the CHM. Due to the regional approach, the size of sample in Slovenia was differing from that in countries in which a similar study was conducted (30). As per the final estimations, the anticipated number of inhabitants included in the survey, was 15,426 for the whole Slovenia (from 578–4591 units in individual region). We reasoned such high number with the planned multivariate methods of data analysis on regional level and with the planned postal administration of questionnaires due to which also the drop-out at response had to be taken into consideration when estimating the size of sample (23).
3. Questionnaire.
   The questionnaire used in the CHMS-SI was entitled »The Health Behaviour Questionnaire«. It originated from the project Finbalt Health Monitor (24,30). The original CHM questionnaire was slightly adjusted to the circumstances in Slovenia.

   The content of »The Health Behaviour Questionnaire« was arranged into the following data groups:
   - basic demographic data on respondent,
   - habits of the use of some medical services and evaluation of the respondent′s health status,
   - smoking habits,
   - nutritional habits,
   - alcohol beverages consumption habits,
   - physical activity habits,
   - road safety habits.

   Altogether there were 73 questions.

4. Method and course of interviewing.
   Upon proper consideration we selected mail survey as the method of interviewing. Such method has also been recommended by the project CHM (24).

   To respondents who would not respond within 14 days, we decided to deliver a reminder accompanied once more by the questionnaire. If the first reminder would still not be responded to, we decided to deliver them another one within one week, consisting of only a letter asking them to answer the questionnaire.

   In the case if the results of simultaneous analysing of the returning of questionnaires in Slovenia and by its regions would indicate the response to be worse as planned at the beginning, the first reminder would be followed by the permitted methods for stimulation of cooperation. For such purpose, the rewarding with „healthy rewards“ was planned, such as visiting Slovene health resorts, healthy food such as fruit, etc.

   Our expectations were that the final response of the respondents in individual regions to be of at least 45%.
5. Statistical methods and tools.
   Besides basic analysis, among which survey of distribution of behavioural risk factors in Slovenia as a whole and by individual health regions, and basic analysis of association, also multivariate methods were planned (23).

   In respect of each of the planned analyses, methodological instructions should be prepared and, if necessary, education on use of the instructions in praxis, since analysis should not be carried out only in one institution, but widespread through all health regions.

   In respect of all kinds of analyses, the actual version of the statistical programme SPSS (Statistical Package for Social Sciences) for Windows was planned to be used.

   Database for analysis was planned to be prepared upon the recommendations of the WHO CHM group (24,54): the same names of variables and the same codes thereof should be applied as proposed by the WHO CHM group.

Determination of ethical principles

In the survey CHMS-SI, ethical principles and provisions of the Personal Data Protection Act (55) were implemented in the following manners:

1. Informing population about the survey (why, when and how it would be conducted and its benefits).
   The communication channels were:
   - mass media - planned to be started about fourteen days prior to sending the questionnaire, and through advertising material in the form of posters, aiming at getting familiar with the survey of as wide public as possible and not of only the respondents,
   - a cover letter to the questionnaire and the questionnaire itself.

2. Informed consent for cooperation in the survey.

As informed consent, a completed questionnaire returned by the respondent would be considered, since the respondent would be doing this voluntarily.

3. Respecting the right to privacy.
   We intended to respect this right as much as possible, therefore this was also one of the reasons for having decided for mail survey.
4. Protection of personal data.
   Due to the expected low response rate, we had to anticipate at least one reminder, on account of which we had to keep a record on the respondents already having returned the questionnaire. For that reason, the interviewing could not be conducted absolutely anonymously. Nevertheless, the anonymity was ensured to the maximal possible extent. At carrying out the analysis, names and surnames of individual respondents were replaced by special identification numbers, and only specified authorised persons were authorised to know connection between IDs and personal data.
5. Application for approval by authorised ethical commission.
   In Slovenia, a rule is in force, according to which each study from the field of medicine, including epidemiological study, which does otherwise not interfere with physical integrity of people, should be acquired a consent of the Ethical Committee of the Republic of Slovenia prior to the conduction thereof. The survey was therefore presented to this ethical body. It was approved at the beginning of April 2001.

Scheduling the survey

The anticipated duration of the survey from the start to the dissemination of majority of the results was 3 years. During this period, the activities and analyses should take place first, which would contribute to improving the course of the next survey of the same type. The analysis mentioned was, for example, the analysis of adequacy and success of the preparations for the survey and the analysis of efficiency of carrying out the interviewing.

When carrying out statistical analyses, we have first of all anticipated data description, which to be followed by the univariate and multivariate analyses. Within each phase of the statistical analysis, drafting of the recruitment of procedures was first planned.

Determination of data base management and results dissemination channels

Data were planned to be kept at the Institute of Social Medicine of the Faculty of Medicine, Ljubljana, which shall be responsible to prepare data in such form so as to be accessible for the wider public.

The survey on behavioural risk factors should be characterised significantly by accessibility of data for public, and by the results thereof to be communicated simultaneously as quickly as possible.

1. Accessibility of data.
   Data on the CHMS-SI are public, meaning that each Slovene citizen may have access thereto providing to respect the rules and under specified conditions.

   The most important rule was that no data referring to the identity of respondents in the survey shall be accessible for wider public.

In order to avoid disputes to arise between the potential users of data, a special group was anticipated to be established to keep the list of proposers and their proposals and to mutually adjust these.

2. Informing public on the results of the study and their applicability.
   Results of analysis, which shall be carried out by the members of the study group, were intended to be communicated to the inhabitants of Slovenia regularly and simultaneously. For this purpose, the media should be applied, and the more detailed information should be available on the web-side of the survey.

   Our intention was to inform the expert public on the results of the survey through a series of publications.

With the aim the database to be managed qualitatively, a body was anticipated to be established (project's council), to exercise control over operation of the database and realisation of its purpose.

Provision of funds for conduction of the survey

The survey was, in addition to the funds, which the CINDI-SI invested in the initiation of the survey, financed also by the Ministry of Health and the Ministry of Education and Sport of the Republic of Slovenia.

EXERCISE

Task 1

Carefully read the part on theoretical background of this module. Critically discuss the characteristics of cross-sectional surveys with your colleagues.

Task 2

From domestic (e.g. Biomedicina Slovenica, and COBISS-Cooperative Online Bibliographic System of Slovenia in Slovenia), and/or international bibliographic data-bases (e.g. Medline, PubMed) find out if any other cross-sectional survey has been already performed in your country.

Task 3

If yes, then try to find out its characteristics. If not, try to find an example from other countries (e.g. FINBALT Health Monitor Surveys).

Task 4

Discuss the characteristics, strengths and limitations of selected survey with your colleagues.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002. p.509-528.
3. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
4. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
5. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998.
6. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders Company; 2001.
7. Premik M. Introduction to epidemiology (in Slovene). Ljubljana: Faculty of Medicine; 1998.
8. Pearce N. A short introduction to epidemiology. Second Edition. Occasional Report Series No 2. Wellington, NZ: Massey University Wellington Campus, Centre for Public Health Research; 2003. Available from URL: www.massey.ac.nz/~wwwcphr/publications/introepi.pdf. Accessed: June 25, 2009.
9. Bailey KD. Methods of social research. New York: The Free Press (A Division of Macmillan, Inc.); 1987.
10. Vaughan JP, Morrow RH. Manual of epidemiology for district health management. Geneva: World Health Organization; 1989.
11. Brownson RC, Baker EA, Leet T, Gillespie KN. Evidence based public health. Oxford, New York: Oxford University Press; 2003.
12. Hansen MH, Hurwitz WN, Madow WG. Sample survey methods and theory. Volume I Methods and applications. New York: John Wiley Sons; 1993.
13. Aromaa A, Koponen P, Tafforeau J (uredniki).Status and future of health surveys in the European Union. Final report of Phase 2 of the project Health Surveys in the EU: HIS and HIS/HES Evaluations and Models. Helsinki, Bruxelles: National Health Institute KTL, Scientific Institute of Public Health; 2003.
14. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
15. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
16. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1993.
17. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 1. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan; 1992.
18. Lutz W, Chalmers J, Lockerbie L, Hepburn W. Health and community surveys. Volume 2. A practical manual for health and community workers. Geneva: IEA, World Health Organization, MacMillan; 1992.
19. Aday LA. Designing and conducting health surveys: a comprehensive guide. San Francisco: Jossey-Bass; 1996.
20. Bankowski Z, Bryant JH, Last JM, editors. Ethics and epidemiology: international guidelines. Proceedings of the XXVth CIOMS Conference. Geneva: Council for International Organizations of Medical Sciences; 1991.
21. World Medical Association. World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. Available from URL: http://www.wma.net/e/policy/b3.htm. Accessed: 17 April 2008.

22. Public Health Leadership Society. Principles of the ethical practice of public health. New Orleans: Public Health Leadership Society; 2002. Available from URL: http://209.9.235.208/CMSuploads/PHLSethicsbrochure-40103.pdf. Accessed April 23, 2008.
23. Zaletel-Kragelj L. Risk factors for noncommunicable diseases in adults in Slovenia Ljubljana. Survey protocol (in Slovene). Ljubljana: Institute of Public Health, Faculty of Medicine and CINDI-Slovenia; 2001.
24. Prättälä R, Helasoja V, Laaksonen M, Laatikainen T, Nikander P, Puska P. Cindi health monitor. Proposal for practical guidelines. Helsinki: National Public Health Institute; 2001.
25. World Health Organization, Regional Office for Europe. CINDI Countrywide Integrated Noncommunicable Diseases intervention Programme. Protocol and guidelines for monitoring and evaluation procedures. Berlin: Springer-Verlag; 1987.
26. World Health Organization, Regional Office for Europe. CINDI 2000. Positioning CINDI to meet the challlenges. Copenhagen, World Health Organization, Regional Office for Europe; 1992.
27. World Health Organization, Regional Office for Europe. CINDI 2000. Prevention in primary care. Recommendations for promoting good practice. Copenhagen: World Health Organization, Regional Office for Europe; 1994.
28. World Health Organization, Regional Office for Europe. CINDI 2000. Protocol and guidelines. Copenhagen: World Health Organization, Regional Office for Europe; 1996.
29. Finbalt Health Monitor. Feasibility of a Collaborative System for Monitoring Health Behaviour in Finland and the Baltic Countries. Helsinki: National Public Health Institute; 1999.
30. Puska P, Tuomilehto J. Nissinen A, Vartainen E. The North Karelia project. 20 year results and experiences. Helsinki: National Public Health Institute; 1995.
31. World Health Organization, Regional Office for Europe. The Health for all Statistical Database. Copenhagen: World Health Organization, Regional Office for Europe; 2007.
32. Zaletel-Kragelj L, Pokorn D. Risk factors for diseases in Slovenia (in Slovene). Ljubljana: University of Ljubljana, Faculty of Medicine; 2000.
33. National health care plan of Republic of Slovenia - Health for all to 2004 (in Slovene). Ur List RS. 2000;10:6650-77.
34. World Health Organization, Regional Office for Europe. Targets for health for all. Targets in support of the European strategy for health for all. Copenhagen: World Health Organization, Regional Office for Europe; 1989.
35. World Health Organization, Regional Office for Europe. Health 21. The health for all policy framework for the WHO European Region. Copenhagen: World Health Organization, Regional Office for Europe; 1999.
36. Čakš T. Relation of smoking to other risk factors for noncommunicable diseases in Ljubljana population (in Slovene). Master degree thesis. Zagreb: University oif Zagreb, School of medicine; 2001.
37. Gradišek A, Šoln D, Tršan V, Zakotnik-Maučec J, Prešeren N, Kovač M, Čakš T, Bulc M, Gabrovšek S, Milohnoja M. Study of risk factors for noncommunicable diseases in Ljubljana (in Slovene). Zdrav Var. 1992;31:71-7.
38. Toš N et al. Slovenia in transition II. Slovene public opinion 1993–1999 (in Slovene). Ljubljana: University of Ljubljana, Faculty for Social Sciences; 2000.
39. Koch V. Some characteristics of nutrition habits of adults in Slovenia (in Slovene). PhD Thesis. Ljubljana: University of Ljubljana, Faculty of Biotechnology; 1997.
40. Moravec Berger D, Ravnikar B, Jezeršek P, Lovše B, Pavlin-Klemenc M, Florjančič M, Mikulič C, Gale I. Spread of some known risk factors for cardiovascular diseases in Zgornja Ščavnica and Branik (in Slovene). Zdrav Var. 1992;31:63-70.

41. Jezeršek P, Accetto R, Cibic B, Dobovišek J, Kolšek B, Lapanja Z, Mihelič-Brčič M, Petrin J, Pirc-Čerček O, Zaletel-Kragelj L, Žemva A. Epidemiological assessment of arterial hypertension in Republic of Slovenia (in Slovene). Zdrav Vestn. 1988;57:403-5.
42. Premik M. Some results of project "Oral health and quality of life in Slovenia (in Slovene). Zdrav Var. 1994,33:85-9.
43. Primic-Žakelj M. Health enlighment of adult Slovenes about cancer (in Slovene). Specialistic thesis. Ljubljana: Institute of Oncology; 1991.
44. Pompe-Kirn V, Japelj B, Primic-Žakelj M. Trends in lung cancer and smoking habits in the male population of Slovenia (in Slovene). Zdrav Vestn. 2000;69:97-9.
45. Pompe-Kirn V, Japelj B. Trends in lung cancer and smoking habits in the female population of Slovenia (in Slovene). Zdrav Vestn. 2000;69:679-82.
46. Prevolnik Rupel V, Rebolj M. The Slovenian VAS tariff based on valutations of EQ-5D health states from the general population. Ljubljana: National Public Health Institute, 2000.
47. Avguštin C. An attempt of measuring of health among adults inhabitants of Celje (in Slovene). 2nd Congress of Slovene Preventive medicine. Book of abstracts. Ljubljana: National health Institute; 1998.
48. Kersnik J. Satisfaction of patients with family doctor (in Slovene). PhD thesis. University of Ljubljana, Faculty of Medicine; 2001.
49. Pompe-Kirm V, Primic Žakelj M, Ferligoj A, Škrk A. Maps of cancer incidence in Slovenia 1978–1987 (in Slovene). Ljubljana: Institute of Oncology; 1992.
50. Cancer incidence in Slovenia 1997 (in Slovene). Ljubljana: Institute of Oncology, Cancer Registry; 2000.
51. Cancer incidence in Slovenia 1998 (in Slovene). Ljubljana: Institute of Oncology, Cancer Registry; 2001.
52. Hanžek M, editor. Report on human development. Slovenia 1998. Ljubljana: Institute for Macroeconomic Analysis and Development; 1998.
53. Tršinar I. Central registry of population (in Slovene). Ur List RS. 1999;9:194.
54. National Public Health Institute, Finland. Finbalt Health Monitor. The data of the 1998 Surveys. Helsinki: National Public Health Institute; 1999.
55. Personal data protection act (in Slovene). Ur List RS. 1999;9:7534-39.

RECOMMENDED READINGS

1. Abramson JH. Cross-sectional studies. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press, 2002. p.509-528.
2. Cross-sectional surveys. In: Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999. p.213-229. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**CASE-CONTROL STUDIES**
Module: 1.4.5	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Slavenka Janković**, MD, PhD, Professor School of Medicine, University of Belgrade, Serbia
Address for correspondence	**Slavenka Janković** Institute of Epidemiology and Centre - School of Public Health, School of Medicine, University of Belgrade, Serbia Višegradska 26, 11000 Belgrade, +Serbia Tel: +381 11 3607 121; Fax: + 381 11 3607 062 Email: slavenka@eunet.rs
Keywords	Case-control studies, case-referent studies, case-comparison studies, nested case-control studies, cross-over studies, odds ratio, cross-product ratio.
Learning objectives	After completing this module students and public health professionals should: • discuss the traditional and modern views of case-control studies (CCS); • differentiate case-control studies from other epidemiologic study designs; • describe the key feature of conducting CCS, including the selection of cases and controls; • calculate and interpret an odds ratio; • discuss the strengths and limitations of CCS; • list the settings in which CCS are desirable.
Abstract	In CCS investigator selects two groups – a group of individuals with a disease of interest (or other outcome), called cases and a suitable group of people without that disease, called controls. Choice of the most appropriate control group is one of the most difficult and controversial aspects of study design The past history of exposure to suspected risk factors is then determined and compared retrospectively between "cases" and "controls." The odds ratio obtained from a CCS may be used as an estimate of the relative risk. CCS are useful for studying rare diseases or outcomes. When there is a long latent period between an exposure and the disease, CCS are the only feasible option. CCS are no longer seen as an inferior alternative to a cohort studies, but rather they are regarded as highly efficient design for learning about exposure-disease relationships. In recent years, case-control design has proven to be useful for evaluation of vaccine effectiveness, treatment efficacy, etc.
Teaching methods	Design issues are best taught by using a mixture of different teaching methods including lectures, exercises, individual work and interactive methods such as small group discussions, seminars etc. Published articles could be presented to the students and discussed with respect to the used design. Students at a more advanced level should be asked to write a protocol on a specific topic and have this protocol discussed in plenary.
Specific recommendations for teachers	The half of credit will be done under supervision, while the rest is individual student's work. Prerequisite for participants: basic knowledge of biostatistics.
Assessment of students	Oral examinations or written essays are preferable to multiple choice types of exams.

CASE-CONTROL STUDIES

Slavenka Janković

THEORETICAL BACKGROUND

Traditional versus modern view of case-control studies

The case-control study (synonyms: case referent study, case comparison study) is one of the two principal types of observational studies (the other one is the cohort study). The subjects are selected on the basis of their disease status. Investigator selects two groups – a group of individuals with a disease of interest (or other outcome), called cases and a suitable group of people without that disease, called controls. The past history of exposure to suspected risk factors is then determined and compared retrospectively between "cases" and "controls." (Figure 1).

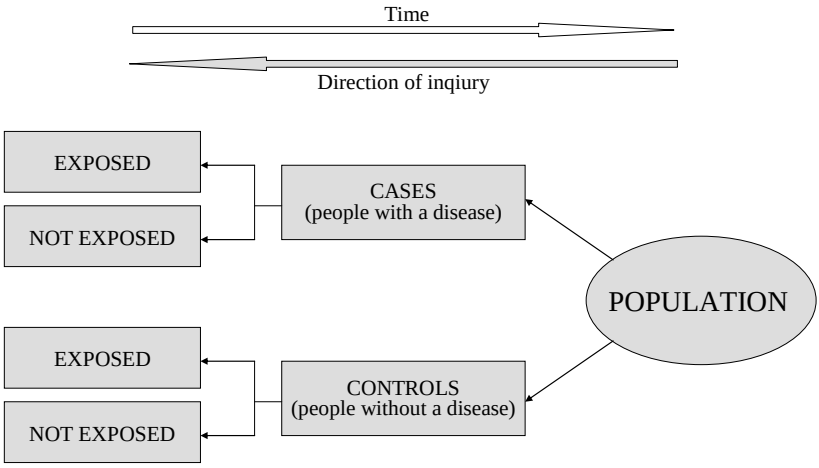


Figure 1. Design of case-control study. Modified from Beaglehole et al (1).

Traditionally, epidemiologists viewed case-control studies as an alternative to cohort studies and believed that their logic is backwards, hence the term "TROHOC" study (COHORT spelled backwards) proposed by Feinstein; its use is deprecated by most epidemiologists (2).

Epidemiologists' view of case-control studies began to change in the 1980s with the work of Miettinen who coined the term "TROHOC fallacy" to express his disagreement with the traditional conceptualization of case-control study (3). Miettinen declared that the case-control study is a method of sampling a population in which cases of disease are identified and a sample of the source population that gave rise to the cases (controls), so that the disease rates in exposed and non-exposed groups can be compared, like in a cohort study (3).

"The case-control study is no longer seen as an inferior alternative to a cohort study, but rather it is regarded as highly efficient design for learning about exposure – disease relationships" (4).

Situations in which a case-control study is desirable

Case control studies are often used to generate hypotheses that can then be studied by prospective cohort or other studies. They are useful for studying rare diseases or outcomes. When there is a long latent period between an exposure and the disease, case-control studies are the only feasible option.

In recent years, the case-control design has proven to be useful for evaluation of vaccine effectiveness, treatment efficacy, evaluation of screening programs and outbreak evaluation (4,5).

Selection of cases

There are several important issues in defining and selecting cases. The starting point of most case control studies is the identification of cases. This requires a suitable case definition.

1. Criteria for case definition should lead to accurate classification of diseased and non-diseased individuals.
2. Efficient and accurate sources should be used to identify cases (e.g. hospital databases, cancer registries).
3. For causal research incidence cases are preferable to prevalent ones.
4. It is not necessary to include all cases of disease occurring within the defined population in the study (4).

Selection of controls

Choice of the most appropriate control group is one of the most difficult and controversial aspects of study design (6). The guiding principles for the valid selection of controls are:

1. Controls must represent the source population.
2. Controls must be sampled independently of exposure status.

Several sources are available for identifying controls.

Population-based controls

When cases are identified from a well defined population (e.g. residents of a defined geographic area) population-based controls are selected. They can be identified using different sources: voter registration lists, telephone directories, national identity registries etc. The advantage of population controls is that they came from the same base population as the cases. However, it is time consuming and expensive to identify them. Also it may be difficult to find eligible controls willing to participate because they are not ill and usually they have not the same interest in participating as do cases and controls from other sources.

Hospital controls

When cases are selected from hospitals, it is appropriate to select hospital controls. It is difficult to determine which diseases are suitable for the control group, but the main principle is that it should be unrelated to the exposure under study. The advantages of hospital controls are easy identification and good participating rates. They recall of prior exposure is comparable to that of cases because they are also ill, and they are less expensive to identify than population controls.

Relatives, friends and neighbours

Selecting relatives, friends or neighbours is a good method to control for possible differences in socioeconomic status, education and other characteristics which are common for cases and controls.

The investigator can decide to use multiple control of the same type (e.g. two or three controls for each case, to increase the power of the study), or multiple controls of different types (e.g., population-based and hospital controls).

Matching

Matching is the process of selecting the controls so that they are similar to the cases in characteristics other than the one that has been studied, such as age, sex and occupation.

The controls may be a matched random sample from the unaffected population.

In individual matching, for each case person a control person is selected who is similar to the case person in terms of the specific variables (Example 1).

Example 1.

A case control study was conducted in order to assess possible relationships between potential risk factors and Graves' disease. The study included 100 newly diagnosed patients with Graves' disease and 100 controls matched with respect to sex, age (±2 years) and type of residence (rural, urban). All the subjects were interviewed by the same medical doctor. The findings indicated that stressful life events, lack of social support and family history of Graves' disease were significantly associated with the occurrence of Graves' disease (7).

Analysis of case-control studies

Calculation and interpretation of odds ratios

When investigators do not know the size of the total population that produced the cases (like in the most cases of case-control studies), they calculate a measure called an odds, a special type of rate (4).

From the two-by-two table (Figure 2) the odds of being a case among the exposed is a/b, and the odds of being a case among the non-exposed is c/d. The ratio of these two odds is known as the *disease-odds ratio* (Equation 1).

	DISEASE STATUS	
EXPOSURE STATUS	CASES	CONTROLS
YES	***a***	***b***
NO	***c***	***d***
TOTAL	***a + c***	***b + d***

Figure 2. Design of a case-control study

$$OR = \frac{a/b}{c/d} = \frac{ad}{bc}$$ **Equation 1.**

The disease-odds ratio is the ratio of the odds in favour of disease among the exposed to the odds in favour of disease among the unexposed.

The odds ratio can also be calculated in another way: the ratio of the odds of being exposed among the cases (a/c) divided by the odds of being exposed among the controls (b/d). This is known as the *exposure-odds ratio* (Equation 2).

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc}$$ **Equation 2.**

The exposure-odds ratio is the ratio of the odds in favour of exposure among the cases to the odds in favour of exposure among non-cases.

The exposure odds ratio is equivalent to disease odds ratio. The term *cross-product ratio* can be used for both of these odds ratios.

An odds ratio close or equal to 1 indicates that the odds of exposure are very similar in the two groups. If the odds ratio is greater than 1, it indicates that cases are more likely to be exposed to a particular factor than controls, and if the odds ratio is less than 1, the opposite is true. The odds ratio (obtained from a case-control study) may be used as an estimate of the relative risk (8).

Different types of case-control studies

Nested case control study

Nested case-control study is a case-control study "nested" within an ongoing cohort study. Assessment of exposure may be time-consuming and costly and instead to undertaking measurement on everyone in a cohort, it may be more efficient to construct a case-control study within the cohort, once a significant number of cases of the disease of interest at follow-up have emerged (the cases for the nested case-control study). Thereafter a control group could be selected among those from the cohort who had not developed the disease (Example 2). Compared to case-control study, nested case-control study can reduce the recall bias and temporal ambiguity. Compared to cohort it can reduce the cost and save time.

Example 2.

Mueller et al (9) conducted a nested case-control study to identify possible relationships between Epstein-Barr virus (EBV) and Hodgkin's disease. Cases on Hodgkin's disease were identified in a cohort of over 240,000 persons from whom blood had been drawn and stored in serum bank several years ago. Tests were carried out on the sera of cases (43 persons with Hodgkin's disease) for EBV antibody and results were compared with tests on coded sera from matched controls (96 persons) from cohort who did not develop Hodgkin's disease.

The case-crossover study

A case-crossover study design is a new type of the case-control study that is used to study the acute effects of transient exposure (4). In this type of case-control study cases serve as their own controls. The period of increased risk following a transient exposure is called the hazard period. The exposure frequency during the hazard period is compared to that during a control period (Example 3).

Example 3.

The case-crossover design was first used to study the risk of myocardial infarction following heavy physical exertion (10). The researches interviewed 1228 myocardial infarction patients. Patients were their own controls. The patients' frequency of physical exertion in the one-hour interval immediately before the start of the heart attack (hazard period) was compared to their usual frequency of physical exertion during the previous year (control period).

Strengths and weaknesses of case-control studies

Like all study designs, case-control studies have some strengths some weaknesses. They are presented in Box 1.

Box1. Strengths and weaknesses of case-control studies. Modified from Aschengrau, Seage (4).

Strengths

- *Efficient for rare diseases*
- *Efficient for diseases with long induction and latent period*
- *Can evaluate multiple exposures in relation to a disease*
- *Time required for study is relatively short*
- *Relatively inexpensive*

Weaknesses

- *Inefficient for rare exposure*
- *May have poor information on exposures*
- *Vulnerable to bias (selection bias, recall bias)*
- *Difficult to infer temporal relationship between exposure and diseases*

CASE STUDY: CIGARETTE SMOKING AND LUNG CANCER

This case study is modified from: A Disease Detectives Exercise adapted from a CDC case study: Cigarette Smoking and Lung Cancer. Teacher's Guide and Answer Key, by Dr. Natale A Carasali (11).

A causal relationship between cigarette smoking and lung cancer was first suspected in the 1920s on the basis of clinical observations. To test this apparent association, numerous epidemiologic studies were undertaken between 1930 and 1960. The case-control study conducted by Richard Doll and Austin Bradford Hill in Great Britain begun in 1948 comparing the smoking habits of lung cancer patients with the smoking habits of other patients (12,13).

Data for this case-control study were obtained from hospitalized patients in London and vicinity over a 4-year period (April 1948 – February 1952). Initially, 20 hospitals, and later more, were asked to notify the investigators of all patients admitted with a new diagnosis of lung cancer. These patients were then interviewed concerning smoking habits, as were controls selected from patients with other disorders (primarily non-malignant) who were hospitalized in the same hospitals at the same time.

Over 1,700 patients with lung cancer, all under age 75, were eligible for the case-control study. About 15% of these persons were not interviewed because of death, discharge, severity of illness, or inability to speak English. The final study group included 1465 cases (1,357 men and 108 women). Only men were included in the study.

Table 1 shows the relationship between cigarette smoking and lung cancer among male cases and controls.

Table 1. Smoking status before onset of the present illness, lung cancer cases and matched controls with other diseases

	Cases	**Controls**
Cigarette smoker	1,350	1,296
Non-smoker	7	61
Total	1,357	1,357

From this table we can calculate different ratios (Equations 3 to 8):

- proportion smoked, cases:

$$PS_{cases} = \frac{1{,}350}{1{,}357} \times 100 = 99.5\% \qquad \textbf{Equation 3.}$$

- proportion smoked, controls:

$$PS_{controls} = \frac{1{,}296}{1{,}357} \times 100 = 95.5\% \qquad \textbf{Equation 4.}$$

What can we infer from proportions? The answer is that the prevalence of smoking is similar in both groups and extremely high in both groups.

- odds of smoking among the cases:

$$OS_{cases} = \frac{1{,}350}{7} = \frac{192.86}{1} \qquad \textbf{Equation 5.}$$

- odds of smoking among the controls:

$$OS_{controls} = \frac{1{,}296}{61} = \frac{21.25}{1} \qquad \textbf{Equation 6.}$$

- disease odds ratio:

$$OR_{disease} = \frac{1{,}350/1{,}296}{7/61} = \frac{1{,}350 \times 61}{7 \times 1{,}296} = 9.1 \qquad \textbf{Equation 7.}$$

- exposure odds ratio:

$$OR_{exposure} = \frac{1{,}350/7}{1{,}296/61} = \frac{1{,}350 \times 61}{7 \times 1{,}296} = 9.1 \qquad \textbf{Equation 8.}$$

What can we infer from the odds ratios about the relationship between smoking and lung cancer? The answer is that the odds ratio is very high indicating that odds for getting the lung cancer are 9.1 times higher in smokers than in non-smokers.

Table 2 shows the frequency distribution of cases and controls by the average number of cigarette smoked per day.

Table 2. Number of lung cancer cases and controls according to daily number of cigarettes smoked.

Daily number of cigarettes smoked	Cases	Controls
0	7	61
1-14	565	706
15-24	445	408
25+	340	182
All smokers	1,350	1,296
Total	1,357	1,357

From Table 2 it is possible to calculate odds ratio (OR) for each smoking category (Equations 9-12):

- odds ratio in the group in which daily number of cigarettes smoked was 1-14 cigarettes (Equation 9):

$$OR_{disease} = \frac{565/7}{706/61} = \frac{565 \times 61}{706 \times 7} = 7.0 \qquad \textbf{Equation 9.}$$

- odds ratio in the group in which daily number of cigarettes smoked was 15-24 cigarettes (Equation 10):

$$OR_{disease} = \frac{445/7}{408/61} = \frac{445 \times 61}{408 \times 7} = 9.5 \qquad \textbf{Equation 10.}$$

- odds ratio in the group in which daily number of cigarettes smoked was 25+ cigarettes (Equation 11):

$$OR_{disease} = \frac{340/7}{182/61} = \frac{340 \times 61}{182 \times 7} = 16.3 \qquad \textbf{Equation 11.}$$

- odds ratio in the group smokers irrespective the daily number of cigarettes smoked (all smokers) (Equation 12):

$$OR_{disease} = \frac{1{,}350/7}{1{,}296/61} = \frac{1{,}350 \times 61}{1{,}269 \times 7} = 9.1$$ **Equation 12.**

The OR of acquiring lung cancer increases as the daily number of cigarettes smoked increases. This correlation is known as a dose-response relationship.

EXERCISE

This exercise and corresponding tasks are consisted first of individual work, and then of a group discussion.

Task 1

Suppose that the case-control study was conducted among men in the United States in order to find out whether a mother's use of hormones during pregnancy influenced her son's risk of developing testicular cancer later in life. Investigators selected 500 cases who were hospitalized for testicular cancer and 1,000 controls. The study found that 90 cases' mothers and 50 controls' mothers had used hormones during pregnancy (3).

Every student individually[10]:

1. sets up the two-by-two table for these data;
2. calculates the odds ratio;
3. interprets the odds ratio;
4. describes the most appropriate control group for this study and justifies his/her choice;

[10] **Answers:**

1. Two-by-two table:

Mother used hormones during pregnancy	**Testicular cancer**	
	Yes (cases)	**No (controls)**
Yes	(a) 90	(b) 50
No	(c) 410	(d) 950
Total	(a + c) 500	(b + d) 1,000

2. Odds ratio = ad/bc = (90 × 950)/(50 × 410) = 4.2.
3. Men whose mothers took hormones during their pregnancies have 4.2 times the risk of testicular cancer than do men whose mothers did not take hormones.
4. For example, one could select men who were hospitalized at the same facilities as the cases. The basic principle is that the controls should be comparable to the cases. Advantages: easy assessable, good cooperation. The difficulty lies in choosing suitable diagnoses for which the controls may be hospitalized. One should not pick diagnoses that are known to be associated with maternal hormone use.
5. Incident cases should be chosen because prevalent cases may represent a biased subset of all men who are diagnosed with testicular cancer. If prevalent cases are used, men who died from testicular cancer may be excluded.

5. answers to the question: Will you choose incident or prevalent cases of testicular cancer? He/she should be able to justify his/her answer.

Task 2

Give an example of a hypothesis that you could test by:

1. case-control study
2. nested case-control study
3. case-crossover study

REFERENCES

1. Beaglehole R, Bonita R, Kjellström T. Basic Epidemiology. Geneva: World Health Organization; 1993.
2. Last JM, editor. A dictionary of epidemiology. 4th.ed. New York: Oxford University Press; 2001.
3. Miettinen OS. The case-control study: valid selection of subjects. J Chron Dis. 1985;38:543-58.
4. Aschengrau A, Seage GR. Essentials of epidemiology in public health. 2nd ed. Sudbury MA: Jones and Barlett Publishers; 2008.
5. Friis RH, Sellers TA. Epidemiology for public health practice, 3rd ed. Sudbary MA: Jones and Bartllet Publishers; 2004.
6. Kelsey JL, Whittemore AS, Evans AS, Thompson DW. Methods in observational epidemiology. New York: Oxford University Press; 1996.
7. Janković S, Radosavljević V, Marinković J. Risk factors for Graves' disease. Eur J Epidemiol. 1997;13:15-8.
8. Gordis L. Epidemiology. 3rd ed. Philadelphia: Elsevier-Saunders; 2004.
9. Mueller N, Evans A, Harris NL, Comstock GW, Jellum E, Magnus K et al. Hodgkin's disease and Epstein-Barr virus. Altered antibody pattern before diagnosis. N Engl J Med. 1989;320:689-95.
10. Mittleman MA, Maclure M, Robins JM. Control sampling strategies for case-crossover studies: an assessment of relative efficiency. Am J Epidemiol. 1995;142:91-8.
11. Centers for Disease Control and Prevention, Epidemic Intelligence Service. Cigarette smoking and lung cancer. Teacher's guide and answer key. Available from URL: http://www.cdc.gov/excite/classroom/smoking_a.pdf. Accessed: March 24, 2009.
12. Doll R, Hill AB. Smoking and carcinoma of the lung. Br Med J. 1950;2:739-48. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2038856. Accessed: March 23, 2009.
13. Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. Br Med J. 1952;2:1271-86. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2022425. Accessed: March 23, 2009.

RECOMMENDED READINGS

1. Aschengrau A, Seage GR. Essentials of epidemiology in public health. Sudbury, MA: Jones and Barlett Publishers; 2008.

2. Centers for Disease Control and Prevention, Epidemic Intelligence Service. List of case studies. Available from URL: http://www.cdc.gov/eis/casestudies/casestudy-list.htm. Accessed: March 23, 2009.
3. Friis RH, Sellers TA. Epidemiology for public health practice, 3rd ed. Sudbury, MA: Jones and Bartllet Publishers; 2004.
4. Gordis L. Epidemiology. 3rd ed. Philadelphia: Elsevier-Saunders; 2004.
5. Janković S. Observational studies. In: Kirch W, editor. Encyclopedia of public health. Leipzig: Springer Verlag; 2008. p.1008-11.
6. Kelsey JL, Whittemore AS, Evans AS, Thompson DW. Methods in observational epidemiology. New York: Oxford University Press; 1996.
7. Last JM, editor. A dictionary of epidemiology. 4th.ed. New York: Oxford University Press; 2001.
8. Last JM, editor. A dictionary of public health. Oxford, New York: Oxford University Press, 2007.
9. Olsen J, Saracci R, Trichopoulos D. Teaching epidemiology. 2nd ed. New York: Oxford University Press; 2001.
10. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. PLoS Med. 2007;4:e297. doi:10.1371/journal.pmed. 0040297.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**COHORT STUDIES**
Module: 1.4.6	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Slavenka Janković** MD, PhD, Professor School of Medicine, University of Belgrade, Serbia
Address for correspondence	**Slavenka Janković** Institute of Epidemiology and Centre - School of Public Health, School of Medicine, University of Belgrade, Serbia Višegradska 26, 11000 Belgrade, Serbia Tel: +381 11 3607 121; Fax: +381 11 3607 062 Email: slavenka@eunet.rs
Keywords	Cohort studies, retrospective cohort studies, prospective cohort studies, follow-up studies, longitudinal studies, historical cohort studies, incidence, relative risk
Learning objectives	After completing this module students and public health professionals should: • understand the basic concepts of cohort studies (CS); • differentiate CS from other epidemiologic study designs; • distinguish between the various types of CS including prospective, retrospective and ambidirectional designs; • describe the main characteristics of conducting CS, including the selection of the exposed and nonexposed cohorts; • calculate and interpret a relative risk, attributable risk and population attributable risk; • discuss the strengths and limitations of CS.
Abstract	A classical CS examines one or more health effects of a single exposure. Subjects are defined according to their exposure status and followed over time to determine the incidence of the health outcome. CS is considered the strongest of all observational designs. Although the basic characteristic of CS is measurement of exposure and follow-up for outcome, there are several types in CS based on temporal differences in cohort design: In a prospective CS, the investigator collects information on the exposure status of the cohort members at the time the study begins and identifies new cases of disease from that time forward. In a retrospective CS the exposure status is established from information recorded at some time in the past, and disease incidence is determined from then until the present Historical prospective CS is a combination of both retrospective CS and prospective CS. Each type of design of CS has its strengths and weaknesses which tend to be complementary.
Teaching methods	Design issues are best taught by using a mixture of different teaching methods including lectures, exercises, individual work and interactive methods such as small group discussions, seminars etc. Published articles could be presented to the students and discussed with respect to the used design. Students at a more advanced level should be asked to write a protocol on a specific topic and have this protocol discussed in plenary.
Specific recommendations for teachers	The half of credit will be done under supervision, while the rest is individual student's work. Prerequisite for participants: basic knowledge of biostatistics.
Assessment of students	Oral examinations or written essays are preferable to multiple choice types of exams.

COHORT STUDIES

Slavenka Janković

"There are only a handful of ways to do a study properly but a thousand ways to do it wrong."

Saccket, 1986 (1)

THEORETICAL BACKGROUND

Cohort study definitions and overview

The cohort study is one of the two principal types of observational studies (the other one is the case-control study). A classical cohort study examines one or more health effects of a single exposure. Investigator defined the study subjects according to their exposure status (exposed group and comparison, non-exposed group) and followed them over time to determine the incidence of the health outcome (e.g. incidence or mortality rates of disease) (Figure 1).

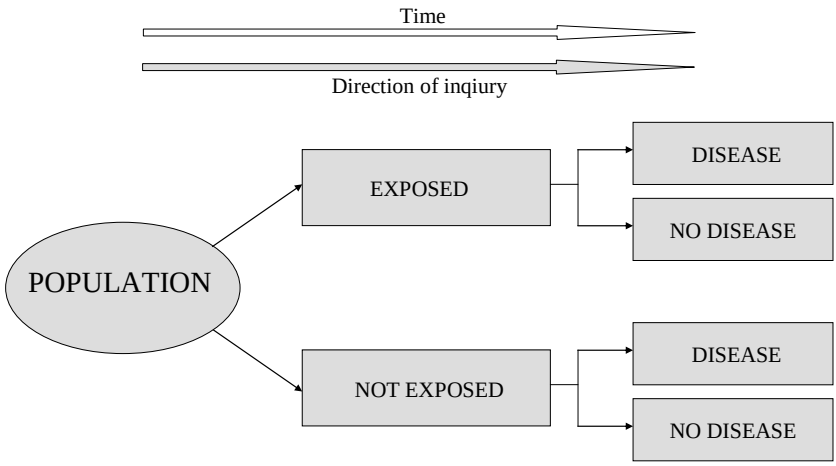


Figure 1. Design of a cohort study. Modified from Beaglehole et al (2).

Because cohort studies measure events in chronological order they can be used to distinguish between cause and effect.

A term *cohort* comes from the Latin *cohors*, plural *cohortes*, and meant a large military unit. Today we use the word *cohort* for "any designated group of persons who are followed or traced over a period of time" (3). The term is also used for a group of people who share a common characteristic or experience within a defined time period (e.g., are born, leave school, etc.). Some investigators use the next synonymous for cohort study: *follow-up*, *incidence*, and *longitudinal* study. The longitudinal study is a research design in

which subjects are followed over time (for months or years) with continuous or repeated monitoring of risk factors or health outcomes, or both.

Selection of cohort study populations

Types of population studied

In cohort studies, the investigators follow people over time. They obtain information about people and their exposures at baseline, let time pass, and then assess the occurrence of outcomes.

In an *open* or *dynamic population* the membership is defined by the changeable characteristics, such as age, marital status, place of residence, smoking and drinking alcohol. In open cohorts the study population is dynamic: people enter and leave the population at different points in time (for example inhabitants of a town).

A *fixed cohort* is defined by an irrevocable event (e. g., undergoing a medical procedure, eating contaminated food at party, or being present at a disaster). The exposure does not change over time. Fixed cohort enrols a defined number of participants at study onset. These participants are then followed-up from that time forward, often at set intervals, up to a fixed end date. This type of cohort does not gain members, but losses to follow-up may occur (4). One of the best-known fixed cohort studies is the study of biological effects of acute radiation exposure in Japanese atomic bomb survivors (5).

Selection of the exposed population

The sampling strategies used to define the cohort differ between themselves. The two main types of cohorts – special cohorts and general cohorts – are distinguished by the exposure frequency. *General cohorts* are assembled for common exposure such as dietary factors, smoking or use of oral contraceptives. General cohort includes an entire population or a representative sample of the population. Population-based cohort approach was used in the Framingham study, one of the best-known cohort studies of cardiovascular disease, initiated in 1948 in Framingham, Massachusetts. The original cohort consisted of 5,209 respondents of a random sample of 2/3 of the adult population of Framingham, who were between 30 and 62 years of age and free of cardiovascular disease at that time (6). The ability to generalize from population-based cohort studies makes them highly desirable. However, such strategic approach is not efficient for rare exposures (e.g. uncommon occupational chemical, natural or man-made disaster), when is more appropriate to use an *exposure-based cohort study* approach, it means to assemble *special cohorts.* One commonly used special cohort is people working in a particular industry or occupation who often have exposure of particular interest (7).

Selection of comparison (non-exposed) group

In a population-based cohort studies the natural comparison group is an *internal comparison group* consisted of people from the same sample who do not have the exposure. If exposed individuals are selected on the basis of a particular exposure, an

external comparison group must be sought. Often the comparison group is the general population of the area from which the exposed group is obtained.

Types of cohort studies

Although the basic characteristic of cohort studies is measurement of exposure and follow-up for outcome, there are several types in cohort studies based on temporal differences in cohort design (Figure 2).

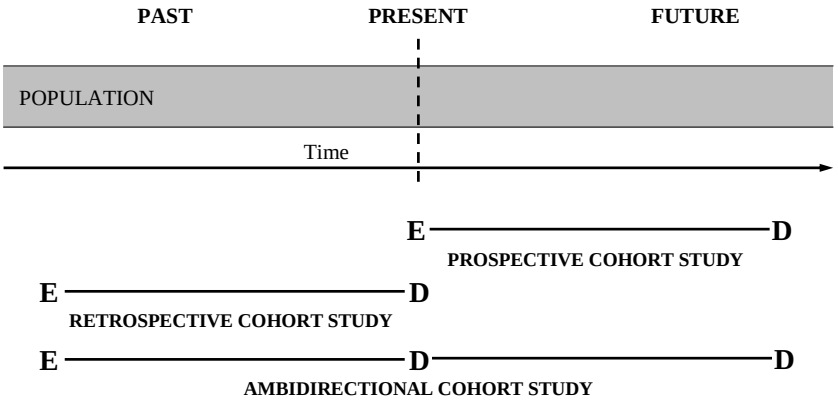


Figure 2. Timing of cohort studies. Modified from Friis and Sellers (8). LEGEND: E=exposure, D=disease.

Prospective cohort study

In a prospective cohort study (*concurrent cohort study*), the investigator collects information on the exposure status of the cohort members at the time the study begins (or at the time the exposure occurs during the study), and identifies new cases of disease from that time forward (the cohort is "followed up" prospectively) (9).

Example: One of the best known and very important studies is the prospective cohort study on mortality in relation to smoking, which was begun in 1951 in United Kingdom (10). The cohort was a group of 34,439 British male doctors listed in the British Medical Register. Information about their smoking habits was obtained in 1951, and periodically thereafter; cause specific mortality was monitored for 50 years. Study revealed that among the men born around 1920, prolonged cigarette smoking from early adult life tripled age specific mortality rates, but cessation at age 50 halved the hazard, and cessation at age 30 avoided almost all of it.

A problem when the cohort method is applied to the study of chronic diseases such as coronary heart disease, or cancer is that large numbers of people must be followed up for long periods before sufficient cases accrue to give statistically meaningful results. The difficulty is further increased when, as for example with most carcinogens, there is a long induction period between first exposure to a hazard and the eventual manifestation of

disease. An approach that can help to counter this problem is to carry out the follow up retrospectively. Obviously, such a study is only feasible when the health outcome of interest can be measured retrospectively.

Retrospective cohort study

In a retrospective cohort study (*historical, non-concurrent cohort study*) the exposure status is established from information recorded at some time in the past, and disease incidence is determined from then until the present (the cohort is "followed up" retrospectively). By this type of investigation only prior outcomes are studied and not future ones (Figure 2).

Example: To determine whether the frequency and pattern of use of the accident and emergency department (A&E) by individuals with diabetes is different from that of the general population a historical cohort of 696 individuals with diabetes and a non-diabetic comparison cohort was performed in 1997. A cohort of individuals with diagnosed diabetes was identified from the computerized repeat prescribing data of six general practices in the catchment area of Leicester Royal Infirmary A&E department, the only A&E department in the city. Each individual was matched with the non-diabetic patient closest in age and of the same sex from the same practice. Records of all A&E visits from November 1984 to June 1996 were extracted by manual searches on the A&E computerised database which has been in use for recording all new registrations since 1984. The use of A&E department by the two cohorts was compared for number of visits and pattern of use between 1984 and 1996 (11).

Historical prospective cohort study (ambidirectional cohort study) is a combination of both retrospective cohort study and prospective cohort study (Figure 2).

Example: The purpose of the Air Force Health Study of the men who were involved in the aerial spraying of herbicides during the Vietnam War is to determine if these men have an increased risk of adverse health and reproductive outcomes. The retrospective component of the study conducted analyses of mortality that occurred from the men's exposure in Vietnam through the 1980s. The prospective component will monitor the health of these men in the future (4).

Follow-up in cohort studies

The follow-up is very important issue in cohort studies. The follow-up period depends on the natural history of the outcome disease and the frequency of disease occurrence.

Active follow-up denotes the situation in which the investigator must obtain data on subsequent incidence of the outcome (e.g. disease), through direct contact with the cohort members (e.g. through mailings, phone calls, or written invitation to return to study sites for medical evaluation).

Passive follow-up does not require direct contact with cohort members. It is possible when databases containing the outcomes of interest are collected and maintained (e.g. cancer registries, national death databases) (8).

Analysis of cohort studies

The primary objective of the analysis of cohort study data is to compare disease occurrence in the exposed and non-exposed groups.

In the simplest case of two level of exposure (Yes/No) two incidence rates are calculated: incidence in the exposed group and incidence in the non-exposed group (Figure 3, Equations 1 and 2).

	DISEASE STATUS		
EXPOSURE STATUS	YES	NO	TOTAL
YES	***a***	***b***	***a + b***
NO	***c***	***d***	***c + d***

Figure 3. Design of a cohort study

$$Incidence_{EXPOSED} = \frac{a}{a+b} \qquad \textbf{Equation 1.}$$

$$Incidence_{NON-EXPOSED} = \frac{c}{c+d} \qquad \textbf{Equation 2.}$$

Incidence in the exposed cohort is then compared with the incidence in the unexposed cohort. This ratio is called *Relative Risk*. It is considered the best measure of effect (9).

Relative Risk

The *Relative Risk* (RR) of an event, such as the occurrence of a specified disease or a death from a specified cause, is the ratio of the risk of a disease or death among those exposed to a specified risk to those not exposed to this risk. It is calculated from the incidence or death rate of the specified disease (Equation 3). The synonym for relative risk is risk ratio (also being abbreviated as RR).

$$RR = \frac{Icidence_{EXPOSED}}{Incidence_{NON-EXPOSED}} \qquad \textbf{Equation 3.}$$

If the level of risk in both, exposed and unexposed group is the same, the RR will equal 1. If an exposure is harmful (e.g., cigarette smoking), the RR is

expected to be greater than 1. If an exposure is protective (e.g., vaccine), the RR will be less than 1.

In addition to the *Relative Risk* the *Attributable Risk* (AR) and *Population Attributable Risk* (PAR) can be calculated.

Attributable risk

The attributable risk (AR) is the portion of the incidence of a disease in the exposed that is due to the exposure. It is the incidence of a disease in the exposed that would be eliminated if exposure were eliminated. It can be calculated as *rate difference* (the rate in the exposed group minus the rate in the unexposed group) or *risk difference* (the difference between the risks in exposed and unexposed groups). When the level of risk in both, exposed and unexposed group, is the same, the risk difference is 0. If an exposure is harmful (e.g., cigarette smoking), the risk difference is expected to be greater than 0. If an exposure is protective (e.g., vaccine), the risk difference will be less than 0 (9) (Equation 4).

$$AR = Icidence_{EXPOSED} - Incidence_{NON-EXPOSED} \qquad \textbf{Equation 4.}$$

The AR is sometimes referred to as attributable risk in the exposed because it is used to quantify risk in the exposed group that is attributable to the exposure. It is the measure of association that is most relevant when making decisions for individuals.

The *attributable risk percent* (AR%) is the percent of the incidence of a disease in the exposed that is due to the exposure. It is the percent of the incidence of a disease *in the exposed* that would be eliminated if exposure were eliminated (Equation 5).

$$AR\% = \frac{Icidence_{EXPOSED} - Incidence_{NON-EXPOSED}}{Incidence_{EXPOSED}} \times 100 \qquad \textbf{Equation 5.}$$

Population Attributable Risk

The *population attributable risk* (PAR) is the risk of a disease of interest in a defined population (exposed and non-exposed) that can be attributed to an exposure of interest (12) (Equation 6).

$$PAR = Incidence_{EXPOSED+NON-EXPOSED} - Incidence_{NON-EXPOSED} \qquad \textbf{Equation 6.}$$

The *population attributable risk percent* (PAR%) is the portion of the incidence of a disease in the entire study population that is due to exposure. It is the incidence of a disease in the population that would be eliminated if exposure were eliminated. The PAR% is calculated by subtracting the incidence in the

unexposed from the incidence in total population (exposed and unexposed) (13) (Equation 7).

$$PAR\% = \frac{Icidence_{EXPOSED+NON-EXPOSED} - Incidence_{NON-EXPOSED}}{Incidence_{EXPOSED+NON-EXPOSED}} \times 100$$ **Equation 7.**

Strengths and weaknesses of cohort studies

Each type of design of cohort studies has its strengths and weaknesses (Box 1), which tend to be complementary (4).

Box1. Strengths and weaknesses of prospective and retrospective cohort studies. Modified from Aschengrau and Seage (4).

Strengths

- *efficient for rare exposures*
- *good information on exposures (prospective)*
- *can evaluate multiple effects of an exposure*
- *efficient for diseases with long induction and latent periods (retrospective)*
- *can directly measure disease incidence and risk (RR, AR)*
- *clear temporal relationship between exposure and outcome (prospective)*
- *less vulnerable to bias (prospective)*

Weaknesses

- *inefficient for rare outcomes*
- *poor information on exposure (retrospective)*
- *expensive and time consuming (prospective)*
- *inefficient for diseases with long induction and latent periods (prospective)*
- *more vulnerable to bias (retrospective)*
- *loss to follow-up*
- *changes over time in criteria and methods (prospective)*

CASE STUDY: CIGARETTE SMOKING AND LUNG CANCER

This case study is modified from: A Disease Detectives Exercise adapted from a CDC case study: Cigarette Smoking and Lung Cancer. Teacher's Guide and Answer Key, by Dr. Natale A Carasali (14).

A causal relationship between cigarette smoking and lung cancer was first suspected in the 1920s on the basis of clinical observations. To test this apparent association, numerous epidemiologic studies were undertaken between 1930 and 1960. Two studies were conducted by Richard Doll and Austin Bradford Hill in Great Britain. The first was a case-control study begun in 1948 comparing the smoking habits of lung cancer patients with the smoking habits of other patients, while the second was a cohort study begun in 1951 recording causes of death among British physicians in relation to smoking habits (15,16).

Data for the cohort study were obtained from the population of all physicians listed in the *British Medical Register* who resided in England and Wales as of October 1951. Information about present and past smoking habits was obtained by questionnaires which were mailed to 59,000 physicians. They were then categorized according to their exposure to cigarette smoking. Non-smokers were defined as persons who had never consistently smoked as much as one cigarette a day for as long as one year.

Usable responses to the questionnaire were received from 40,637 (68%) physicians, of whom 34,445 were males and 6,192 were females. The next section of this prospective cohort study is limited to the analysis of male physician respondents, 35 years of age or older.

The occurrence of lung cancer in physicians responding to the questionnaire was documented over a 10-year period (November 1951 through October 1961) from death certificates filed with the Registrar General of the United Kingdom and from lists of physician deaths provided by the British Medical Association. All certificates indicating that the decedent was a physician were abstracted. For each death attributed to lung cancer, medical records were reviewed to confirm the diagnosis.

Of 4,597 deaths in the cohort over the 10-year period, 157 were reported to have been caused by lung cancer; in 4 of the 157 cases this diagnosis could not be documented, leaving 153 confirmed deaths from lung cancer.

The following table (Table 1) shows numbers of lung cancer deaths by daily number of cigarettes smoked at the time of the 1951 questionnaire (for male physicians who were non-smokers and current smokers only). Person-years of observation ("person-years at risk") are given for each smoking category. The number of cigarettes smoked was available for 136 of the persons who died from lung cancer.

Table 1. Number and mortality rate (per 1,000 person-years) of lung cancer by number of cigarettes smoked per day

Daily number of cigarettes smoked	Number of deaths from lung cancer	Person-years at risk	Mortality rate per 1000 person years
0	3	42,800	0.07
1-14	22	38,600	0.57
15-24	54	38,900	1.39
25+	57	25,100	2.27
All smokers	133	102,600	0.94
Total	136	145,400	1.30

From Table 1 it is possible to calculate Relative Risk (RR) for each smoking category (Equations 8-11):

- RR in the group in which daily number of cigarettes smoked was 1-14 cigarettes (Equation 8):

$$RR = \frac{0.57}{0.07} = 8.1 \qquad \textbf{Equation 8.}$$

- RR in the group in which daily number of cigarettes smoked was 15-24 cigarettes (Equation 9):

$$RR = \frac{1.39}{0.07} = 19.7 \qquad \textbf{Equation 9.}$$

- RR in the group in which daily number of cigarettes smoked was 25+ cigarettes (Equation 10):

$$RR = \frac{2.27}{0.07} = 32.4 \qquad \textbf{Equation 10.}$$

- RR in the group smokers irrespective the daily number of cigarettes smoked (all smokers) (Equation 11):

$$RR = \frac{1.30}{0.07} = 18.6 \qquad \textbf{Equation 11.}$$

Relative risk for all smokers (RR=18.6) means that people who smoke are 18.6 time more likely than non-smokers to develop lung cancer. Similar explanation is for each category of daily number of cigarettes smoked.

From the results we can see that lung cancer mortality rates increase with an increase in the amount of cigarettes smoked per day ("dose-response relationship").

From Table 1 we can also calculate the Attributable Risks (AR) for each smoking category (Equations 12-15):

- AR in the group in which daily number of cigarettes smoked was 1-14 cigarettes (Equation 12):

$$AR = 0.57 - 0.07 = 0.50 \qquad \textbf{Equation 12.}$$

In this group AR is 0.50 per 1,000 person-years.

- AR in the group in which daily number of cigarettes smoked was 15-24 cigarettes (Equation 13):

$$AR = 1.39 - 0.07 = 1.32 \qquad \textbf{Equation 13.}$$

In this group AR is 1.32 per 1,000 person-years.

- AR in the group in which daily number of cigarettes smoked was 25+ cigarettes (Equation 14):

$$AR = 2.27 - 0.07 = 2.20 \quad \textbf{Equation 14.}$$

In this group AR is 2.20 per 1,000 person-years.

- AR in the group smokers irrespective the daily number of cigarettes smoked (all smokers) (Equation 15):

$$AR = 1.30 - 0.07 = 1.23 \quad \textbf{Equation 15.}$$

In the group group of all smokers irrespective the daily number of cigarettes smoked AR is 1.23 per 1,000 person-years.

The excess deaths attributable to smoking increases from 0.50 to 2.20 per 1,000 person-years as the quantity of cigarettes smoked per day increases.

From Table 1 we can also calculate the Attributable Risk percent (AR%). In Equation 16 calculation of AR% for the total group of smokers is presented:

$$AR\% = \frac{1.30 - 0.07}{1.30} \times 100 = 0.946 \times 100 = 94.6\% = 95\% \quad \textbf{Equation 16.}$$

Overall 95% of lung cancers deaths among smokers are attributable to smoking. Therefore, if no one had smoked, 126 deaths (95% of 133 deaths) due to lung cancer would have been avoided.

The cohort study also provided mortality rates for cardiovascular disease among smokers and non-smokers. Table 2 presents lung cancer mortality data and comparable cardiovascular disease mortality data.

Table 2. Mortality rates (per 1,000 person-years from lung cancer and cardiovascular disease by smoking status

Disease	**Mortality rate per 1,000 person-years**		
	Smokers	**Non-smokers**	**All**
Lung cancer	1.30	0.07	0.94
Cardiovascular disease	9.51	7.32	8.87

From these data, calculating the RR we can estimate with which cause of death smoking is more strongly associated (Equations 17 and 18):

- RR for lung cancer (Equation 17):

$$RR = \frac{1.30}{0.07} \times 1{,}000 = 18.57 \quad \textbf{Equation 17.}$$

- RR for cardiovascular disease (Equation 18):

$$RR = \frac{9.51}{7.32} \times 1{,}000 = 1.30 \qquad \textbf{Equation 18.}$$

The RR indicates a much stronger association between cigarette smoking and lung cancer mortality than between cigarette smoking and cardiovascular disease mortality (18.6 vs. 1.3). Smokers are 14-times more likely to die from lung cancer than from cardiovascular disease.

From Table 2, we can calculate the PAR% for both diseases as well (Equations 19 and 20):

- PAR% for lung cancer (Equation 19):

$$PAR\% = \frac{0.94 - 0.07}{0.94} \times 100 = 92.6\% \qquad \textbf{Equation 19.}$$

- PAR% for cardiovascular disease (Equation 20):

$$PAR\% = \frac{8.87 - 7.32}{8.87} \times 100 = 17.5\% \qquad \textbf{Equation 20.}$$

92.6% of all deaths due to lung cancer and 17.4% of all deaths due to cardiovascular disease in the study population are attributable to smoking. PAR% for a given exposure can be interpreted as the proportion of lung cancer or cardiovascular disease in the entire population that would have been prevented if exposure had not occurred.

EXERCISE

This exercise and corresponding tasks are consisted first of individual work, and then of a group discussion.

Task 1

In a prospective cohort study of the relationship between oral contraceptive use and the subsequent risk of developing endometrial cancer, a cohort of 1,000 women were followed for 5 years. The results are shown in table below (Table 3).

Table 3. Disease status (endometrial cancer) in a cohort of 1,000 women in relation to oral contraceptives after 5-year follow-up.

Exposure status	**Disease status**		**Total**
	Yes	**No**	
Yes	(a) 245	(b) 75	(a + b) = 320
No	(c) 50	(d) 630	(c + d) = 680
Total	(a + c) = 250	(b + d) = 705	(n) = 1,000

Please, answer to the following questions:

1. What was the incidence rate of endometrial cancer among women who used oral contraceptives?
2. What was the incidence rate of endometrial cancer among women who did not use oral contraceptives?
3. What was the relative risk in this study?
4. What was the attributable risk in this study?
5. What was the incidence rate of endometrial cancer among women who used oral contraceptives in person-years?[11]

Task 2

1. Design a prospective cohort study to test one of the following hypotheses:
   - alcohol consumption during pregnancy affects the risk of preterm delivery;
   - use of oral contraceptives affects the risk of coronary hart disease;
   - teenagers who excessively use their cell phones are more prone to disrupted sleep, restlessness, stress and fatigue. There seem to be a connection between intensive use of cell phones and health compromising behaviour such as smoking, snuffing and use of alcohol.

2. Discuss the strengths and weaknesses of your proposed study from Task 2 with other students.

REFERENCES

1. Sackett DL. Rational therapy in the neurosciences: the role of the randomized trial. Stroke. 1986;17:1323-9.
2. Beaglehole R, Bonita R, Kjellström T. Basic Epidemiology. Geneva: WHO, 1993.
3. Last JM. editor. A Dictionary of Epidemiology. 4th.ed. New York: Oxford University Press; 2001.
4. Aschengrau A, Seage GR. Essentials of Epidemiology in Public Health. 2nd ed. Sudbury MA: Jones and Barlett Publishers; 2008.
5. Sauvaget C, Lagarde F, Nagano J, Soda M, Koyana K, Kodama K. Lifestyle factors, radiation and gastric cancer in atomic bomb survivors. Cancer Causes Control. 2005;16:773-80.
6. Higgins, Millicent W. The Framingham Heart Study: Review of Epidemiological Design and Data, Limitations and Prospects. In: Alan R, editor. Genetic Epidemiology of Coronary Heart Disease: Past, Present and Future. New York: Liss, Inc, 1984, p.51-64.

[11] **Answers:**

1. 245/(245 + 75);
2. 50/(50 + 630);
3. [245/(245 + 75)] / [50/(50 + 630)]
4. [245/(245 + 75)] – [50/(50 + 630)];
5. 245/(320 × 5)

7. Kelsey JL, Whittemore AS, Evans AS, Thompson DW. Methods in observational epidemiology. New York: Oxford University Press; 1996.
8. Friis RH, Sellers TA. Epidemiology for public health practice, 3rd ed. Sudbury MA: Jones and Bartllet Publishers; 2004.
9. Gordis L. Epidemiology, 3rd ed. Philadelphia: Elsevier-Saunders; 2004.
10. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. Br Med J. 2004;328:1519-28.
11. Goyder EC, Goodacre SW, Botha JL, Bodiwala GG. How do individuals with diabetes use the accident and emergency department? J Accid Emerg Med. 1997;14:371-4.
12. Last JM, editor. A dictionary of public health. Oxford New York: Oxford University Press; 2007.
13. Jekel JF, Elmore JG, Katz DL. Epidemiology Biostatistics and Preventive Medicine. USA: WB Saunders Company, 1996.
14. Centers for Disease Control and Prevention, Epidemic Intelligence Service. Cigarette Smoking and Lung Cancer. Teacher's Guide and Answer Key. Available from URL: http://www.cdc.gov/excite/classroom/smoking_a.pdf. Accessed: March 24, 2009.
15. Doll R, Hill AB. Mortality in relation to smoking: 10 years' observation of British doctors. Br Med J. 1964;1:1399-1410. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=14135164. Accessed: March 28, 2009.
16. Doll R, Hill AB. Mortality in relation to smoking: 10 years' observation of British doctors. Br Med J. 1964;1: 1460-1467. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1814697. Accessed: March 24, 2009.

RECOMMENDED READINGS

1. Aschengrau A, Seage GR. Essentials of Epidemiology in Public Health. 2nd ed. Sudbury MA: Jones and Barlett Publishers; 2008.
2. Centers for Disease Control and Prevention, Epidemic Intelligence Service. List of Case Studies. Available from URL: http://www.cdc.gov/eis/casestudies/casestudy-list.htm. Accessed: August 12, 2008.
3. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. Br Med J. 2004;328:1519-28.
4. Friis RH, Sellers TA. Epidemiology for public health practice. 3rd ed. Sudbury MA: Jones and Bartllet Publishers; 2004.
5. Gordis L. Epidemiology, 3rd ed. Philadelphia: Elsevier-Saunders; 2004.
6. Janković S. Observational studies. In: Wilhelm K, editor. Encyclopedia of Public Health. Leipzig: Springer Verlag, 2008. p.1008-11.
7. Kelsey JL, Whittemore AS, Evans AS, Thompson DW. Methods in observational epidemiology. New York: Oxford University Press; 1996.
8. Last JM, editor. A dictionary of epidemiology. 4th.ed. New York: Oxford University Press; 2001.
9. Last JM, editor. A dictionary of public health. Oxford New York: Oxford University Press; 2007.
10. Olsen J, Saracci R, Trichopoulos D. Teaching Epidemiology. 2nd ed. New York: Oxford University Press,; 2001.
11. Sackett DL, Wennberg JE. Choosing the best research design for each question. Br Med J. 1997;315:1636.
12. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS Med. 2007;4(10):e297. doi:10.1371/journal.pmed. 0040297.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**INTRODUCTION TO INTERVENTION (EXPERIMENTAL) STUDIES**
Module: 1.4.7	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Tatjana Pekmezović**, MD, PhD, Professor Institute of Epidemiology, School of Medicine, University of Belgrade, Serbia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Tatjana Pekmezović** Institute of Epidemiology, School of Medicine, University of Belgrade, Višegradska 26A, Belgrade, Serbia E-mail: pekmezovic@sezampro.rs
Keywords	Intervention study, experimental study, field trial, community trial, clinical trial, randomization
Learning objectives	At the end of this topic student will be able to: • describe and explain basic concepts of intervention studies; • understand and explain different types of intervention studies; • participate in study design, derivation and data analysis; • read epidemiological literature that use and refer to the concepts outlined above.
Abstract	First a distinction between observational and experimental studies is made. Afterwards current ideas and trends in experimental epidemiology are explored, as well as basic characteristics of different types of intervention studies. During this topic key interrelated components in intervention studies, such as planning, organization, selection of study participants, calculation of sample size, follow-up, detection of the effects, etc, are introduced. The potential biases in study design and measurement of outcomes make an important part of this topic, too. In addition, an ethical consideration of intervention studies is discussed.
Teaching methods	The teaching method recommended: • the introductory lecture related to topics mentioned above; • exercises which include distribution of selected papers (different types of intervention studies) to each student and their critical appraisal (discussion in small groups), • development of a written protocol (project proposal) of intervention study according to problem assigned in small group (up to 5 students), • presentation of project proposals and overall discussion related to advantages and disadvantages of each design.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire and short written essay according to assigned study problem.

INTRODUCTION TO INTERVENTION (EXPERIMENTAL) STUDIES

Tatjana Pekmezović, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Introduction

Previous modules in this book were dealing with different types of observational studies while in this module the experimental designs are briefly introduced.

Basic definitions

Prior discussing the characteristics of this group of epidemiological studies it could be worthy to give some definitions of terms frequently used in relation to experimental studies.

1. An experiment.
   Among definitions of the term "experiment" we can find the following:
   - according to Rothman et al. (1), an experiment is a set of observations, conducted under controlled circumstances, in which the scientists manipulates the conditions to examine/verify what effect such a manipulation has on the observations. In epidemiology, the term "experiment" usually means that the investigator manipulates the exposure assigned to participants in the study. Usually, an experiment refers to any trial or test,

2. An experimental study.
   Among definitions of the term "experimental study" we can find the following:
   - according to A dictionary of epidemiology (2), experimental study is a study in which conditions are under the direct control of the investigator. In epidemiology it is a study in which a population is selected for a planned trial of a regimen whose effects are measured by comparing the outcome of the regimen in the experimental group with the outcome of another regimen in a control group. As examples of experimental studies, a randomized controlled trial and a community trial are mentioned,
   - according to the TheFreeDictionary's Medical dictionary (3), an experimental study is a study in which all of the risk factors are under the direct control of the investigator.

3. An intervention.
   Among definitions of this term we can find the following:
   - according to Beaglehole et al. (4), an intervention or experimentation is an act that involves attempting to change a variable (an outcome) in one or more groups of people. The effects of an intervention are measured by comparing the outcome in the experimental group with the outcome in a control group,

- according to Medicine.Net Online Dictionary (5), an intervention is the act of intervening, interfering or interceding with the intent of modifying the outcome. In medicine, an intervention is usually undertaken to help treat or cure a condition. It comes from the Latin term "intervenire", meaning "to come between",
- according to the TheFreeDictionary's Medical dictionary (3), an intervention is a) the act or fact of interfering so as to modify, and b) any measure whose purpose is to improve health or alter the course of disease,

4. An intervention study.
   Among definitions of this term we can find the following:
   - according to A dictionary of epidemiology (2), intervention study is an investigation involving intentional change in some aspects of the status of the subjects, e.g. introduction of a preventive or therapeutic regimen, or designed to test a hypothesized relationship; usually an experiment such as a randomized controlled trial,
   - according to the TheFreeDictionary's Medical dictionary (3), an intervention study is a testing of a hypothesized epidemiological cause-effect relationship by intervening in a population and modifying a supposed causal factor and measuring the effect of the change.

   In fact the term "intervention study" is an alternative for the term "an experimental study" (4).
5. To control.
   Among definitions of this term we can find the following:
   - according to A dictionary of epidemiology (2), the term "to control" among others means "to regulate" (in terms of experimental studies this means that a researcher gains the mastery over the situation in opposite to observational studies where a researcher only observe the situation/process and do not intervene).

6. A trial.
   Among definitions of this term we can find the following:
   - according to the TheFreeDictionary's Medical dictionary (3), the term "trial" is a synonym for the term "experiment"; it refers usually to the trying out of a substance or a material in order to determine its effect,
   - the same states Aschengrau and Seage (6).

Description of intervention studies

General characteristics of intervention studies

Intervention (experimental) studies, also known as intervention trials (7), are the epidemiological studies that are most similar to laboratory experiments. They are characterized by the property that in these studies an investigator directly controls on experimental circumstances, and more precisely, intercedes with new therapeutic agent,

vaccine, or preventive procedure (7,8). As we mentioned previously, the effects of interventions are measured by comparing the outcome in both experimental and control groups.

Major distinction between observational and intervention studies underlies in the fact that in observational studies investigator accepts the conditions as they are and makes observations with aim to answer questions related to studied problem, while in intervention studies he/she gains the mastery over the situation.

If treatments are allocated randomly in a sufficiently large sample, intervention studies have the potential to provide a degree of certainty about the validity of results that is absolutely not possible with any observational study (9). In other words, intervention studies provide the strongest evidence with which to test hypothesis. In addition, it is considered to be the ideal design for evaluating the effectiveness and effects of new treatment or intervention. Besides that, intervention studies can be used for many purposes: evaluating new drugs or other treatments, testing of new health and medical care technology, assessment of new programs for screening, or finding new ways of organization and delivery health services (10). However, due to mainly ethical and practical reasons, this approach is relatively rare study design in epidemiology and public health (11). Other limitations include difficulties in generalization of results, limited feasibility, response and attrition problems, extreme expenses because of large number of participants and engaged investigators, etc.

Randomisation - the essential characteristics of intervention studies

The essential characteristic of intervention studies is the randomisation although an experiment can also be non-randomised what will be discussed later in this section.

Randomisation (or random allocation) of individuals (or groups or populations) is a process of allocation (assignment) of participants in the study (individuals or communities) to the experimental and the control group (Figure 1) by chance (2). After the random allocation the intervention procedure is applied to the experimental but not to the control group. After the follow-up period is completed the effect is assessed according to previously defined outcome (11).

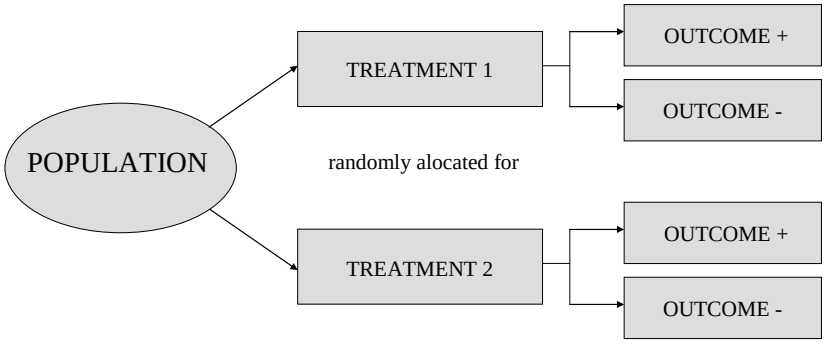


Figure 1. General design of intervention (experimental) studies. Adapted from Dos Santos Silva (7).

Within the limits of chance variation, randomisation is intended to make the control and the experimental groups similar at the start of the trial (2). In other words, it eliminates selection bias on the part of the participants and investigators. Because of this property it is one of the methods of controlling the potential confounding factors (7).

Various methods can be used to randomize the study subjects to different study groups such as:

- simple randomization - the most elementary method of randomization, equivalent of tossing a coin (7). Selection of the subjects occurs randomly. Randomization list could be obtained by using a table of random numbers, or it could be computer-generated by using random number generator,
- stratified randomization, including match-pair design (frequently used in a community trials) - used when the results of the trial are likely to vary between categories of certain characteristic (i.e. between sexes/genders or between different age groups). First strata are formed and randomization occurs separately for the subjects in each stratum (7),
- cluster randomization - this method is characterized by the property that randomization occurs at the group (cluster) level. All individuals within a given cluster are assigned to same study group (so-called arm). Cluster sampling is typically used when the researcher cannot get a complete list of the members of a population they wish to study but can get a complete list of groups or "clusters" of the population. It is also used when a random sample would produce a list of subjects so widely scattered that surveying them would prove to be far too expensive, for example, people who live in different districts (12).

  Cluster randomized trials are less efficient statistically than individually randomized trials because the responses of individuals in a cluster tend to be more similar (intracluster) than those individuals in different clusters (intercluster). The sample size required is accordingly larger and the analysis techniques have to be adjusted by the level of association among members of the cluster (intracluster correlation coefficient) (13).

Despite the randomization is the most optimal design for evaluating effectiveness, in practice (14), conducting a randomized controlled trial is not always feasible. Some alternatives to randomization include historical controls and non-randomized controls.

1. Historical controls can be used in case that we have therapy today that we believe will be quite effective, and would like to test it in a group of patients. So therefore, for comparison, we will go back to the records of patients with the same disease who were treated before new therapy became available. This type of design seems inherently simple and attractive, but we can not be sure that the differences between such groups are due to the therapy, because many things other than therapy change over time.
2. One of alternative approaches is also to use controls that are not selected in a randomized manner. In such a case we are talking of non-randomized trials**.**

   In the case of clinical trials, nonrandomized trial is a clinical trial in which the participants are not assigned by chance to different treatment

groups. Participants may choose which group they want to be in, or they may be assigned to the groups by the researchers (15).

The essential characteristic of non-randomized trials is that the participants are not assigned by chance to different treatment groups. Participants may choose which group they want to be in, or they may be assigned to the groups by the researchers.

Despite the fact that non-randomised trials do not yield the same sort of information than randomised, they have its own importance - the purpose is exploratory or hypotheses-generating. On the basis of their results something must be subsequently proved by randomised trials (14).

Non-randomized trials are often used to evaluate the effectiveness of surgical treatments (16). In this type of intervention study, the three aspects of design are the following:

- the relationship in time between groups being compared,
- whether or not participants were treated before the study was conceived, and
- the basis for allocating a treatment to a participant.

These characteristics are important because they are believed to influence the risk of bias. It is well known that in randomized controlled trials, treatment is allocated by chance, while in non-randomized studies, the decision to give a particular treatment to a particular participant is made by researchers. Both methods are likely to give rise to imbalances in prognostic factors between the groups being compared. Imbalances can be reduced by matching, or controlled by statistical methods, but their confounding effects can never be completely removed (7). Imbalances can bias estimates of treatment effects and almost always increase their uncertainty (16).

Types of intervention studies

Intervention studies (epidemiologic experiments) include (1,4): clinical trials, field trials, and community trials (community intervention studies) (Figure 2).

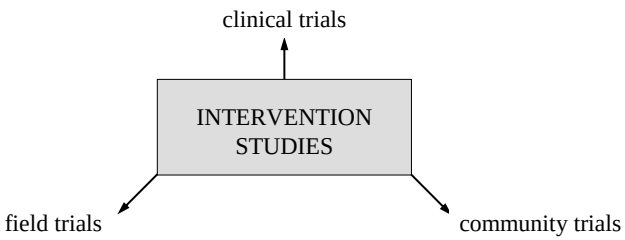


Figure 2. Three types of intervention studies.

This classification in fact comprises two classifications (6):

- clinical trial versus field trial, and

- individual trial versus community trial.

Clinical trials are trials performed on individual patients as subjects of investigation (mostly in hospitals/clinics), field trials are trials performed on individual (healthy) community members as subjects of investigation, while community trials are performed on whole communities as subjects of investigation.

This is only one of possible classifications of intervention studies. Here are some others (6):

- therapeutic trial versus prophylactic trial,
- randomized trial versus non-randomized trial,
- simple trial versus factorial trial, etc.

Clinical trials

Definition

Clinical trials are:

- according to the TheFreeDictionary's Medical dictionary (3), experiments performed on human beings in order to evaluate the comparative efficacy of two or more therapies,
- according to the U.S National Institutes of Health (17), clinical trials are studies to answer specific questions about vaccines or new therapies or new ways of using known treatments,
- according to Rothman et al. (1), experiments with patients as subjects, with a goal to evaluate a potential cure for disease, or to find a preventive of disease sequelae such as death, disability, or a decline in the quality of life,

Randomized controlled (clinical) trials

Clinical trials are mostly randomized. In this case are called randomized controlled (clinical) trials (RCTs):

- according to Lilienfeld and Stolly (8), randomized controlled (clinical) trials are epidemiologic experiments mainly conducted with the aim to study the efficacy of a drug or a medical procedure (e.g. surgical intervention) in the treatment of a disease. However, as Lilienfeld and Stolly state, they can also be used to evaluate a preventive agents (such as vaccines) or public health procedures (for example, screening) (8). While the different interventions are in question, the general methods and principles, for the most part, remain the same,
- according to Last et al. (2), RCTs are epidemiologic experiments in which subjects in a population are randomly allocated into a study and a control group to receive or not an experimental preventive or therapeutic procedure (or intervention).

To ensure that the groups being compared in RCTs are equivalent, patients are to them assigned randomly. RCTs are considered as the most scientifically rigorous method of hypothesis testing available in epidemiology (2).

Clinical trials designs

Clinical trials can be conducted according to different designs (6,7,14,18,19): parallel groups, cross-over, or factorial design.

1. Parallel group design (6,18,19).
   The most common design is the parallel group design in which patients are randomised to one of two or more so-called arms of a clinical trial. One arm is being allocated to an experimental group and the other to a control group. There exist several types of comparison of experimental group (an intervention under observation) to a control group (20). For example, it could be compared to:
   - placebo,
   - another intervention,
   - same intervention at a higher dose (or longer duration), or
   - no intervention.

   This design of clinical trials is schematically presented in Figure 3 (4).

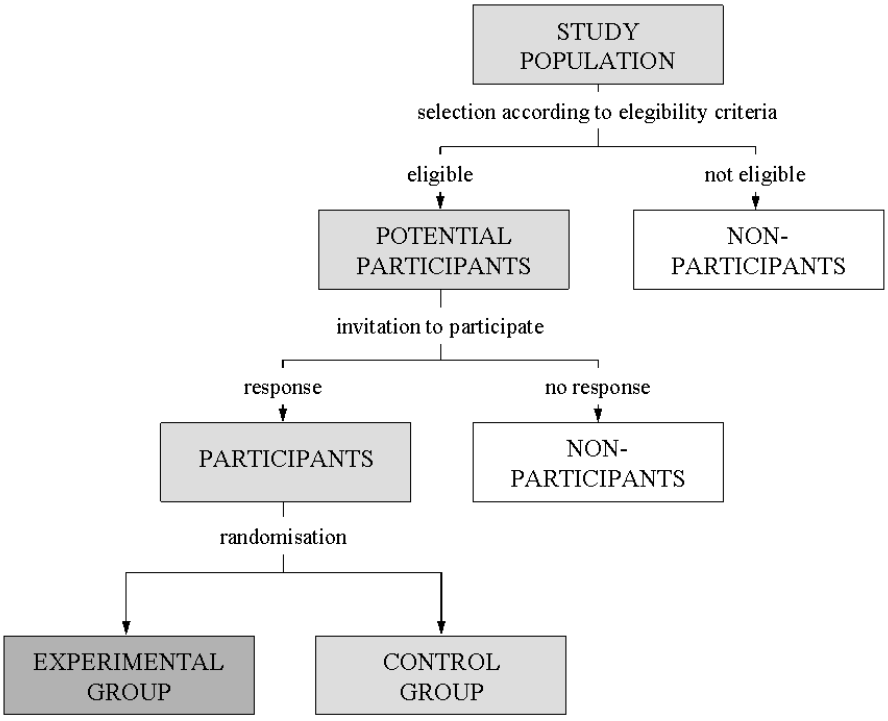


Figure 3. General design of randomized controlled (clinical) trials. Adapted from Beaglehole et al. (4).

2. Cross-over design (6,7,14,18,19).
   The cross-over design is one of within-patient comparison designs. In this design, each patient is randomised to a sequence of two or more treatments, and thus acts as his/her own control for treatment comparisons. This design is attractive primarily because it reduces the number of patients required to achieve a specific power, sometimes to a marked extent. In the simplest two-by-two cross-over design, each patient receives each of two treatments in randomised order in two successive treatment periods, often separated by a washout period.
3. Factorial design (6,7,14,18,19).
   In a factorial design two or more treatments are evaluated simultaneously in the same patient population through the use of varying combinations of the treatments. In other words, each group of patients gets two or more treatments.

   The simplest example is the two-by-two factorial design in which patients are randomly allocated to one of the four possible combinations of two treatments. If these two treatments are labelled as treatment A and treatment B, the combinations are: A alone; B alone; both A and B; neither A nor B.

   The usual intention of using factorial design is to make efficient use of clinical trial patients by evaluating the efficacy of the two treatments with the same number of patients as would be required to evaluate the efficacy of either one alone. In other words, factorial design allows answering to more questions in a single trial for minor increase in costs.

There exist also some other designs, being match pairs design, sequential design and other within-patient comparison designs like Latin and Greco-Latin square designs (14).

Detailed discussion on strengths and limitation of various designs are out of the scope of this module.

Clinical trials at different phases of experimental clinical research

Experimental clinical research usually progresses in an orderly series of steps, called phases. The trials at each phase have a different purpose and help scientists answer different questions (14,21,22):

1. Phase I trials.
   By these trials, researchers test an experimental drug or treatment for the first time. Main goals are to detect potentially harmful adverse effects of observed treatment, and to determine the metabolic and pharmacological actions, and safe dosage range.

   The observed group of participants is small (20-80). It is preferably to recruit healthy volunteers, because an unexpected and potentially dangerous reaction can occur, that is easily manageable in healthy participants.

   Duration of these studies is up to 1 month.
2. Phase II trials.

During this phase the potential treatment's therapeutic usefulness/effectiveness is evaluated. It is aimed also at determining the short-term side effects, and to identify common risks for a specific population and disease.

The experimental study drug or treatment is administered for a limited period (several months) to a smaller number of patients with target disease (100-300) that must be as homogenous as possible. Trial is usually offered to patients who have not improved with other available treatments.

3. Phase III trials.
   In phase III trials, the experimental study drug or treatment is extended to larger and less homogenous group of patients with target disease (1,000-3,000) to confirm its effectiveness, monitor side effects, compare it to commonly used treatments, and collect information that will allow the experimental drug or treatment to be used safely.

   Patients involved in this phase trials are patients with the same type of cancer who otherwise would receive best current treatment. The duration of administration increases (several years).
4. Phase IV trials.
   In phase IV trials, post marketing studies provide additional information including the drug's risks (long-term side effects), benefits (additional uses of the agent), and optimal use. The number of participants increases (thousands of individuals with target disease as well as new population groups), as well as the observation period (on-going process).

Types of clinical trials

There exist several types of clinical trials. According to classical classification by Lilienfeld and Stolly (8), there exist three types of clinical trials, being therapeutic, intervention and preventive clinical trials:

1. Therapeutic trials.
   Therapeutic trials are carried out with the aim to cure diseases, prevent recurrences and complications or increase survival. Subjects with a disease are involved in the study. An example of such a study is a study on a Zidovudine (AZT) treatment for acquired immunodeficiency syndrome (AIDS) (23) (Example 1).

Example 1.

AZT is a potent inhibitor of the replication of the human immunodeficiency virus type 1 (HIV), and it has been shown to improve survival in advanced HIV disease.

The objective of this study was to evaluate the efficacy and safety of AZT, early in the treatment of HIV infection.

A double-blind, randomized, placebo-controlled trial was performed, with subject stratification by pre-treatment CD4 T lymphocyte counts. It was a multicentre trial at AIDS Clinical Trial units in the USA. Included in the study were seven hundred eleven subjects with mildly symptomatic HIV infection. Three hundred fifty-one subjects were assigned to placebo and 360 to AZT, 200 mg orally every 4 hours. The median duration of follow-up was 11 months.

Example 1. Cont.

Fifty-one subjects developed the AIDS, advanced AIDS-related complex, or death as a first critical event. For the stratum of subjects with more than 200 but less than 500 CD4 T lymphocytes/mm3 before treatment, 34 events occurred in placebo recipients and 12 in AZT recipients (p=0.0002). For the stratum of subjects with 500-799 CD4 T lymphocytes/mm3 before treatment, 2 events occurred in placebo recipients and 3 in AZT recipients. Significant differences between the treatment groups in CD4 T-lymphocyte counts occurred in subjects with more than 200 but less than 500 CD4 T lymphocytes/mm3 after 4 weeks of therapy (p=0.002). Differences persisted through week 52. Less prominent changes occurred in subjects with 500 or more CD4 T lymphocytes/mm3. Serum levels of HIV antigen decreased significantly in AZT recipients. Serious anaemia and neutropenia occurred in 5% and 4% of AZT recipients, respectively, and in 0% and 1% of placebo recipients, respectively.

In conclusion, AZT delayed progression of HIV disease and produced little toxicity in subjects with mildly symptomatic HIV disease and less than 500 CD4 T lymphocytes/mm3.

2. Intervention trials.

In intervention trials investigator intercedes before a disease has developed in subjects at high risk of getting a disease. An example of such a study is s study on an AZT treatment of HIV-positive individuals without AIDS symptoms (24) (Example 2).

Example 2.

Since AZT has been shown to improve survival in advanced the human immunodeficiency virus (HIV) disease, the aim of present study was to estimate efficacy and safety of this drug in persons with asymptomatic HIV infection.

A randomized, double-blind trial was conducted in adults with asymptomatic HIV infection who had CD4+ cell counts of fewer than 500 per cubic millimetre on entry into the study. The subjects (92 percent male) were randomly assigned to one of three treatment groups: placebo (428 subjects); AZT, 500 mg per day (453); or AZT, 1500 mg per day (457). A mean follow-up was 55 weeks.

After a follow-up period, 33 of the subjects assigned to placebo had AIDS, as compared with 11 of those assigned to receive 500 mg of AZT (p=0.002) and 14 of those assigned to receive 1500 mg of AZT (p=0.05). In the three treatment groups, the rates of progression (per 100 person-years) to either AIDS or advanced AIDS-related complex were 7.6, 3.6, and 4.3, respectively. As compared with those assigned to placebo, the subjects in the AZT groups had significant increases in the number of CD4+ cells and significant declines in p24 antigen levels. In the 1500-mg AZT group, severe haematologic toxicity (anaemia or neutropenia) was more frequent than in the other groups (p less than 0.0001). In the 500-mg AZT group, nausea was the only toxicity that was significantly more frequent (in 3.3 percent) than in the placebo group (p=0.001).

The authors concluded that AZT is safe and effective in persons with asymptomatic HIV infection and fewer than 500 CD4+ cells per cubic millimetre.

3. Preventive (prophylactic) trials.
   Preventive trials are conducted with aim to estimate the efficacy of a preventive agent or procedure among subjects free of disease. An example of such a study is a study on an education in use of condoms in prevention of HIV transmission and infection (25) (Example 3).

Example 3.

In some parts of Africa, prostitutes and their clients represent the groups at greatest risk of human immunodeficiency virus (HIV) infection and the major disseminators of the virus. Condom use was assessed after a programme of education about the AIDS (acquired immunodeficiency syndrome) and a condom distribution programme in a well-characterised prostitute population in Nairobi. Women received their education at group meetings (barazas) and at individual counselling sessions during which they were given the results of serological tests for HIV (group 1) or at barazas only (group 2), or through very little of either (group 3). During the counselling sessions free condoms were distributed. Before either of the programmes started, 10%, 9%, and 7% of groups 1, 2, and 3 women, respectively, reported occasional use of condoms. During the first year of study, 80%, 70%, and 58% of groups 1, 2, and 3 women, respectively, reported at least some condom use. The mean frequency of condom use was 38.7%, 34.6%, and 25.6% of sexual encounters in groups 1, 2, and 3 women. 20 of 28 women who were non-condom-users seroconverted compared with 23 of 50 women who reported some use of condoms.

Beside the presented classification also other more recent classifications exist. For example, according to the U.S. National Institutes of Health (21) there are five types of clinical trials, being treatment, prevention, diagnostic, screening, and quality of life clinical trials:

- treatment trials test new drugs or new combinations of drugs, or new therapeutic approaches (i.e. new approaches to surgery or radiation therapy,
- prevention trials test new approaches, such as medicines, vitamins, minerals, or other supplements that are believed that may lower the risk of a certain types of diseases (i.e. cancer), or lifestyle changes. In other words they look for better ways to prevent disease in people who have never had the disease or to prevent a disease from returning,
- diagnostic trials are conducted to find better tests or procedures for diagnosing a particular disease or condition,
- screening trials test the best way to detect certain diseases or health conditions (i.e. cancer, especially in its early stages),
- quality of life trials (or supportive care trials) explore ways to improve comfort and the quality of life for individuals with a chronic illness.

Field trials

Field trials, in contrast to clinical trials, as a rule, deal with subjects who are free of disease but presumed to be at risk, and involve evaluation of agent or procedure with the aim to reduce the risk of developing disease in general population.

Data collection takes place “in the field”, usually among non-institutionalized people in the general population (4).

This design usually requires a larger number of subjects and longer follow-up period than clinical trials, since their purpose is to prevent the occurrence of diseases that typically occur with relatively low frequency (1,4,7). Additionally, since subjects are not under active health care in health care settings (e.g. under treatment in community health centre or even hospitalized in hospitals), they do not come to a central location for treatment. Consecutively, field trials often require visiting subject at home or on the work-place (school-place), or establishing study centres. All these characteristics mean that field trials are huge projects involving a lot of human and financial resources (1).

A random allocation of individuals to an experimental and a control group is again an ideal design. However, in practice there are a lot of difficulties for its implementation (1). Consecutively, other designs are frequently applied (e.g. cluster randomization). The problem is that these modifications can affect the informativeness and interpretation of experimental findings (1). Detailed description of these limitations is out of the scope of this module.

An example of such a study is a study on the effect of breast cancer screening on mortality from breast cancer (26) (Example 4).

Example 4.

With the aim to evaluate the effect of breast cancer screening on mortality from this disease, women, age 40-60 who were members of the Health Insurance Plan of town X, were randomly divided into two groups: intervention group (four mammography examinations were offered at annual intervals) and control group (receiving usual medical care). Each group comprised 31,000 women. The groups were very similar in the terms of demographic and other characteristics of interest (26).

With the aim to evaluate the effect of breast cancer screening on mortality from this disease, women, age 40-60 who were members of the Health Insurance Plan of town X, were randomly divided into two groups: intervention group (four mammography examinations were offered at annual intervals) and control group (receiving usual medical care).

Each group comprised 31,000 women. A critical factor in determining sample size was the interest in detecting at least a 20% reduction in mortality that might be attributed to screening. High levels of comparability between the study and control groups have been demonstrated for a wide range of demographic and other characteristics and for general mortality other than breast cancer.

Screening consisted of a clinical examination, usually by a surgeon; mammography, in which two views were taken of each breast (cephalo-caudal and lateral); and an interview to obtain relevant demographic information and a health history. Independence between the two examining modalities was strictly maintained so that it could be determined which was responsible for the chain of events that led to biopsy. Control women continued to receive their usual medical care.

By the end of 10 years after entry, the study group's mortality due to breast cancer was about 30% below the control group. However, there was no longer difference in mortality from causes other than breast cancer between intervention and control groups.

Community trials

Trying to find what community trials are we can find following definitions:

- according Rothman et al. (1), community intervention trials are an extension of the field trials that involves intervention on a community level,
- according to Dos Santos Silva (7), community trials are special form of field trials in which whole communities are the unit of allocation. Community trials thus involve population as a whole, i.e. the group as a whole studied collectively (7),
- according to Last et al. (2), a community trial is an experiment in which the unit of allocation to receive a preventive or therapeutic regimen is an entire community or political subdivision.

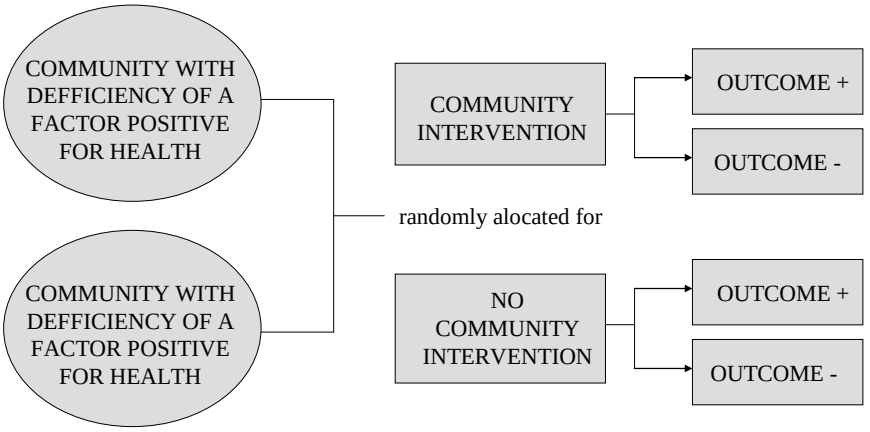


Figure 4. General design of community intervention studies.

General design of community intervention studies is similar to general design (parallel) in clinical trials (Figure 4). Again, a random allocation of study units (groups, communities) to an experimental and a control group is again an ideal design (in this case cluster randomization).

An example of such a study is the Newburgh-Kingston caries fluorine study (27) (Example 5).

Example 5.

The aim of Newburgh-Kingston dental caries study was to estimate whether the increased fluoride concentration in drinking water might reduce prevalence of decayed, missing or filled permanent teeth. One entry community (Newburgh) was allocated randomly to receive fluoride added to the water supply, while the other (Kingston) continued receiving water without supplementation (27).

The caries fluorine hypothesis which states that fluorine has a prophylactic effect on dental caries is supported by extensive epidemiological studies in the USA and in other parts of the world.

Example 5. Cont.

The aim of Newburgh-Kingston dental caries study was to estimate whether the increased fluoride concentration in drinking water might reduce prevalence of decayed, missing or filled permanent teeth.

In 1944 the study plan was made a reality when the, cities of Newburgh and Kingston in New York State agreed to participate in such a program, as study and control areas respectively. This study was started in June, 1944, when basic dental examinations were begun. One entry community (Newburgh) was allocated randomly to receive fluoride added to the water supply, while the other (Kingston) continued receiving water without supplementation. Both cities are situated on the Hudson River about 30 miles apart. Each has a population of approximately 30,000. The climate of both cities is also similar, and their water supplies at the outset of this study were comparable and have remained so, except for the addition of sodium fluoride to Newburgh's supply.

On May 2, 1945, sodium fluoride was added to Newburgh's water supply to bring its fluorine content up to 1.0-1.2 p.p.m., while Kingston's water supply remains fluorine-free.

All of the dental examinations in Newburgh and the first series in Kingston were made with mouth mirror and sharp explorer by the same examiner. The subsequent examinations in Kingston, using the same technique, were made by two dental hygienists trained in the method of examination and the charting of defects. In both areas the examiners called off the defects which were recorded by a staff clerk on a dental record card designed specifically for this study.

It is expected that the study will take 10-12 years to determine adequately the efficacy and safety of this caries prophylactic measure.

The proportion of erupted permanent teeth with evidence of caries experience (decayed, missing, or filled) decreased in each successive examination period in Newburgh, from 21 per 100 before water fluoridation to 14.8 per 100 at the time of the last survey. This rate in Kingston remained approximately 21 for the examination period. The difference between Newburgh and Kingston at the last examination suggests a 30% improvement in Newburgh.

Another well known example is a trial of heart disease prevention in North Karelia, Finland (2,28).

Course of intervention studies

Study protocol

Before the start of the study, protocol in written form must be developed. It contains rationale and specific objectives, precise description of methods for selecting and allocating study groups, number of participants, randomization schemes, criteria of including and excluding participants, type and duration of intervention, major and minor outcomes, as well as methods for their monitoring and registration. In addition, methods for data collection and result analyzing, procedures for obtaining the informed consent of

subjects must be described. The protocol should be accompanied by all forms which will be completed during the study (usually as appendix).

Selection of participants

Next step is selection of study population. Participants must be similar in regard to many characteristics that could significantly influence the outcome, such as sex, age, degree of previous exposure, stage of disease or its absence, etc. Investigator selects target population, which comprises individuals or groups with set of characteristics related to the problem investigated. When target (or reference) population is defined, one needs to select the actual population in which the study will be conducted. It is the experimental population. Inclusion and exclusion criteria must be clearly defined before the study begins. It is very important, especially as a method for elimination of all unsuitable subjects (mainly those with characteristics which can interfere with the outcome). Eligible subjects must be invited to participate in the study, after being fully informed about the purposes, procedures, possible risks and benefits of the study. After exclusion of refusals, the study population is defined. In community trials selected communities should be stable, with little migration and have self-contained medical care system (8).

Sample characteristics

The sample size, defined as a number of individuals, necessary to detect effect of intervention, is an essential part of experiment preparation. Inadequate samples can cause lack of improvement, which exists and is confirmed with large sample (8). The sample size is computed applying various statistical procedures. The following steps in sample size calculation are generally accepted:

- detection of difference in response rates,
- estimation of the response rate in one of the groups,
- detection of level of statistical significance (alpha),
- detection of the value of the power desired (1-beta), and
- detection whether the test should be one-sided or two-sided.

In circumstances when a rare type of exposure or outcome is in question, a sufficient number of participants may pose a big problem. In such situation, multicentric trial or meta-analysis can be useful. In multicentric trials, many hospitals (or groups) in the community, country, or throughout the world are included in the study using the same study protocol. Meta-analysis is a method in which data from similar studies are pooled in a statistically rigorous manner (9).

Randomization

Next step in the design of intervention study is allocation of participants into the test and comparison groups with the aim to ensure that those treated and those untreated are exactly similar in almost all aspects before intervention. It is the best method if assignment to study groups should be done at random.

Compliance

The important issue of experimental design is monitoring compliance and side effects. Noncompliance can decrease the statistical power of study to detect exact effect of

intervention, although a certain degree of noncompliance is acceptable, especially in estimation of effectiveness of intervention in real-life conditions.

Patients may be randomized, but following randomization they may not comply with the assigned treatment. Noncompliance may be overt or covert. In the first case, people may overtly articulate their refusal to comply or may stop participating in the study. These non-compliers are dropouts. On the other side, people may just stop taking the agent assigned without admitting this to the investigator or the study staff. Another problem in clinical trials is drop-ins. In this case, patients in one group may inadvertently take the agent assigned to the other group.

Outcome assessment

The outcome of interest must be clearly defined before the study beginning. In clinical trials, effects should be evaluated in each patient. In community trials, outcomes would be expressed as a reduction of incidence of disease or cost to health services (11). If investigator knows whether participants were in treatment or control group it can result with a biased assessment of effect. To eliminate this problem, three procedures have been developed: single-, double- or triple-masking. In a single-masking study, participants are not given any indication to whether they belong to treatment or control group. The aim is to prevent participants from introducing bias into observations; it can be achieved by use of a placebo. In a double-masking study, neither participants nor investigator have knowledge of the participant group allocation. In a triple-masking study, participants, investigator and reviewer of data are all masked with regard to the group individuals belong to (8).

Measurements of outcome include both improvement (the desired effect) and any side effects that may appear. Therefore, there is a need for explicitly stated criteria for all outcomes to be measured in a study.

Follow-up

Procedures during the follow-up period are the same for all study participants. In this part of intervention study important issues include equal and rigorous follow-up in both groups, simple but sufficient methods for detecting of all relevant events, and high quality cooperation (in that way loss from the study population should be minimized) (11).

Data analysis

The data analysis is performed with the aim to assess the efficacy of intervention. For example, in vaccine trials, the efficacy is the proportion (or percentage) of the expected incidence of disease which is prevented by intervention. In case when observed benefits are high or possible injury effects are serious, results must be analyzed sequentially. This means to continue data analysis and stop the study when a significant benefit or adverse effect has been demonstrated (11).

Ethical considerations

In intervention studies, since investigator deliberately intervenes, ethical considerations are more important in comparison to other types of epidemiological studies. Before study is carried out, many questions should be considered. Hill mentioned some of them such as: whether the proposal treatment is safe or possible harmful for study participants, whether it is

ethical to use placebo treatments, etc. (29). As mentioned above, each participant in study must be fully informed about the purposes and potential adverse effects of intervention. If subjects provided with this information decide to participate, their informed consent must be obtained. Personal privacy and confidentiality must be respected at all time. Nowadays, almost all research and health institutions have Ethical Committees, formed with the aim to control and survey ethical aspects of experimental studies. Each experiment including human subjects must be approved by Ethical Committee. The general guidelines for biomedical research are contained in the Declaration of Helsinki: Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects (30) prepared by World Medical Association (1996 version) and international guidelines published by the CIOMS (Council for International Organization of Medical Sciences (4,31).

EXERCISE

Teaching methods for this topic would include distribution of several published papers; in small groups, students will discuss on appropriateness of used design, the validity of study and the authors' conclusions in the light of the stated objectives.

Tasks 1-5 refer to Examples 1-5 presented previously in the module, while Tasks 6-8 are adapted from Biglan, Norell, and Omenn (32-34).

Task 1

This task refers to the Example 1. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What are characteristics of stratified randomization?
3. What does placebo mean?
4. What does mean double-blinded trial?
5. What was intervention?
6. How were effects of intervention assessed?
7. What do the results obtained suggest and has the objective been achieved?
8. Discuss the ethical considerations of the study.

Task 2

This task refers to the Example 2. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. How were effects of intervention assessed?
3. What do the results obtained suggest and has the objective been achieved?
4. Discuss the ethical considerations of the study.

Task 3

This task refers to the Example 3. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.

2. What was essential in prophylactic trials?
3. What was intervention?
4. How were effects of intervention assessed?

Task 4

This task refers to the Example 4. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What was essential in field trials?
3. Why women aged 40-60 were included in this study?
4. What was intervention?
5. How were effects of intervention assessed?

Task 5

This task refers to the Example 5. Carefully read it again and discuss the following questions:

1. Describe and discuss study objective and design.
2. What was essential in community trials?
3. Why women aged 40-60 were included in this study?
4. What was intervention?
5. How were effects of intervention assessed?
6. What could be another way of controlling an experiment, apart from measuring the dental health status in similar but untreated low-fluoride community?

Task 6

Carefully consider the paper entitled "A randomized controlled trial of a community intervention to prevent adolescent tobacco use" by Biglan et al. (32). Discuss the following questions:

1. What was the objective of this study?
2. Describe and discuss study design.
3. Which type of randomization was used?
4. What was intervention?
5. Were objective outcome criteria developed and used?
6. How were effects of intervention assessed?
7. What do the results obtained suggest and has the objective been achieved?
8. What were the limitations of this study?

Task 7

Mass screening with mammography for early detection and treatment of breast cancer in women can reduce the risk of advanced stages of the disease and death. To investigate this, a randomized trial was conducted among 162,981 women age 40 or more and living in two counties in Sweden at the time of randomization. Each county was divided into 19

blocks selected to give relative socioeconomic homogeneity within each block. In one of the two counties, each block was divided into two units of roughly equal size. One of these units was selected randomly to receive, and the other not to receive, the screening program (33). Discuss the following questions:

1. What was the unit of randomization?
2. What were the advantages and disadvantages of this approach?

Task 8

Read carefully the article "Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease" by Omenn et al. (34). Try to find answers to the following list of questions:

1. Specify the main hypothesis, main outcome and main exposure.
2. Discuss appropriateness of the study design.
3. Rephrase could any other study design be used for this research?
4. What was the target population?
5. What was the main result?
6. Was adjustment for potential confounders carried out?

ASSESSMENT OF STUDENTS

1. Major purpose of randomization in a intervention study is to:
   A. facilitate double-masking
   B. reduce selection bias
   C. reduce information bias
   D. facilitate measurement of outcome variables
   E. avoid sampling variation.

2. Problems pertaining to intervention studies include the following items, except one:
   A. ethical considerations
   B. response and attrition problems
   C. high likelihood of comparability of study groups
   D. limited feasibility
   E. high expenses.

REFERENCES

1. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
2. Last JM, editor. A dictionary of epidemiology. Fourth edition. New York, Oxford, Toronto: Oxford University Press; 2001.
3. TheFreeDictionary's Medical dictionary [homepage on the Internet]. *Available from URL:* http://medical-dictionary.thefreedictionary.com/. Accessed: September 5, 2009.
4. Beaglehole R, Bonita R, Kjellstrom T. Basic Epidemiology. Geneva: World Health Organization; 2007.
5. MedicineNet.com. MedTerms dictionary [homepage on the Internet]. Available from URL: http://www.medicinenet.com/script/main/hp.asp. Accessed: September 5, 2009.

6. Aschengrau A, Seage GR III. Essentials of epidemiology in public health. Sudbury, MA: Jones&Bartlett Publishers; 2003.
7. Dos Santos Silva I. Cancer epidemiology. principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
8. Lilienfeld DE, Stolley PD. Foundation of epidemiology. New York, Oxford: Oxford University Press; 1994.
9. Hennekens CH, Buring JE. Epidemiology in medicine. Boston/Toronto: Little, Brown and Company; 1987.
10. Gordis L. Epidemiology. Philadelphia: W.B. Sounders Company; 2000.
11. Farmer R, Miller D, Lawrenson R. Epidemiology and public health medicine. Oxford: Blackwell Science; 1996.
12. Singleton RA Jr, Straits BC. Approaches to social research. Third edition. New York: Oxford University Press; 1999.
13. Donner A, Klar N. Design and analysis of cluster randomization trials in health research. London: Arnold Publishers; 2000.
14. Spriet A, Dupin-Spriet T, Simon P. Methodology of clinical drug trials. Basel: Karger; 1994.
15. U.S. National Institutes of Health, National Cancer Institute. Dictionary of cancer terms [homepage on the Internet]. Available from URL: http://www.cancer.gov/dictionary/. Accessed: September 5, 2009.
16. Reeves BC. Principles of research: limitations of non-randomized studies. Surgery. 2008;26:120-4.
17. U.S. National Institutes of Health, Clinical Trials.gov. Glossary of clinical trials terms [homepage on the Internet]. Available from URL: http://clinicaltrials.gov/ct2/info/glossary#ttrials. Accessed: September 5, 2009.
18. European Union. Biostatistical methodology in clinical trials. Available from URL: http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/vol-3/pdfs-en/3cc7aen.pdf. Accessed: September 5, 2009.
19. Friedman LM, Schron EB. Methodology of intervention trials in individuals. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. Oxford: Oxford University Press; 2002. p.569-582.
20. Franceschi S, Plummer M. Intervention trials. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. Berlin: Springer; 2005. p.345-370.
21. U.S. National Institutes of Health, Clinical Trials.gov. Understanding clinical trials [homepage on the Internet]. Available from URL: http://clinicaltrials.gov/ct2/info/understand#Q18. Accessed: September 5, 2009.
22. University of Pittsburgh, Research Coordinator Orientation. Comparison of clinical trial phases. Pittsburgh: University of Pittsburgh, Research Coordinator Orientation; 2002. Available from URL: http://www.clinicalresearch.pitt.edu/docs/comparison_of_clinical_trial_phases.pdf. Accessed: September 5, 2009.
23. Fischl MA, Richman DD, Hansen N, Collier AC, Carey JT, Para MF, et al. The safety and efficacy of Zidovudine (AZT) in the treatment of subjects with mildly symptomatic human immunodeficiency virus type 1 (HIV) infection: A double-blind, placebo-controlled trial. Ann Inter Med. 1990;112:727-37.
24. Volberding PA, Lagakos SW, Koch MA, Pettinelli C, Myers MW, Booth DK, et al. Zidovudine in asymptomatic human immunodeficiency virus infection: a controlled trial in persons with fewer than 500 CD4-positive cells per cubic millimeter. N Engl J Med. 1990;322:941-9.
25. Ngugi EN, Plummer FA, Simonsen JN, Cameron DW, Bosire M, Waiyaki P, et al. Prevention of transmission of human immunodeficiency virus in Africa: effectiveness of condom promotion and health education among prostitutes. 1988;332:887-90.

26. Shapiro S. Evidence on screening for breast cancer from a randomized trial. Cancer. 1977;39:2772-82.
27. Ast DB, Finn SB, McCaffery I. The Newburgh-Kingston Caries Fluorine Study. Dental findings after three years of water fluoridation. Am J Publ Health. 1950;40:716-24. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1528792. Accessed: September 5, 2009.
28. Hoffmeister H, Mensik GBM. Community-based intervention trials in developed countries. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. Oxford: Oxford University Press; 2002. p.583-597.
29. Hill AB. Short Textbook of Medical Statistics. Philadelphia: JB Lippincott Company; 1977.
30. World Medical Association. World Medical Association Declaration of Helsinki: Recommendations Guiding Medical Doctors in Biomedical Research Involving Human Subjects. Available from URL: www.hku.hk/facmed/images/document/04research/institution/Declaration_of_Helsinki_1996_version.pdf. Accessed: August 24, 2008.
31. Council for International Organizations of Medical Sciences (CIOMS). International Ethical Guidelines for Biomedical Research Involving Human Subjects. GENEVA: CIOMS, 2002. Available from URL: http://www.cioms.ch/frame_guidelines_nov_2002.htm. Accessed: August 24, 2008.
32. Biglan A, Ary DV, Smolkowski K, Duncan T, Black C. A randomized controlled trial of a community intervention to prevent adolescent tobacco use. Tob Control. 2000; 9:24-32.
33. Norell SE. Workbook of Epidemiology. New York, Oxford: Oxford University Press; 1995.
34. Omenn GS, Goodman GE, Thornquist MD. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular disease. New Engl J Med. 1996;334:1150-1155.

RECOMMENDED READINGS

1. Dos Santos Silva I. Cancer epidemiology. principles and Methods. Lyon: IARC; 1999. Available from URL: http://www.iarc.fr/en/publications/pdfs-online/epi/cancerepi/CancerEpi.pdf. Accessed: June 24, 2009.
2. Olsen J, Saracci R, Trichopoulos D. Teaching epidemiology: a guide for teachers in epidemiology, public health and clinical medicine. New York, Oxford, Toronto: Oxford University Press; 2001.
3. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
4. Vetter N, Matthews I. Epidemiology and public health medicine. Edinburgh, London, New York: Churchill Livingstone; 1999.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers, and Health Professionals	
Title	**ACQUIRING QUALITATIVE SKILLS FOR PUBLIC HEALTH RESEARCH: USING INTERVIEWS TO GENERATE DATA**
Module: 1.5.1	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Danica Rotar Pavlič**, MD, PhD, Assistant Professor Department of Family Medicine, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Danica Rotar Pavlič** Department of Family Medicine, Faculty of Medicine, University of Ljubljana, Slovenia; Poljanski nasip 58, Ljubljana, Slovenia E-mail: danica.rotar@guest.arnes.si
Keywords	Qualitative study, fieldwork, interview
Learning objectives	After completing this module students should: • know the definition and characteristics of interviews; • be familiar with the design phase of a qualitative study; • be familiar with data reduction and concept construction; • be familiar with validation and triangulation.
Abstract	Qualitative research interviews are defined as attempts to understand the world from the subjects' point of view, to unfold the meaning of peoples' experiences, to uncover their lived world prior to scientific explanations. Interviewing is a well-established research technique. Much qualitative research is interview-based. There are three main types: structured, semi-structured, and in-depth interviews. In practice, open-ended, qualitative interview questions are often combined with more closed-ended, structured interview formats. Qualitative interviews may be used as an exploratory step before designing more quantitative, structured questionnaires to help determine the appropriate questions and categories. The module describes principles of interview-based qualitative research.
Teaching methods	An introductory lecture gives the students their first insight into the characteristics of qualitative research. Theory is illustrated through a case study. After the introductory lectures, the students carefully read the recommended readings. Afterwards, they check the findings in the fieldwork. In the next steps, they analyze the transcribed interview and present their findings to the other students.
Specific recommendations for teachers	• proportion of fieldwork under teacher supervision/individual students' work: 30/70%; • facilities: a lecture room, a computer room, rooms for group work; • equipment: computers (1 computer per 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: Bologna-type master's students.
Student assessment	Multiple-choice questionnaire.

ACQUIRING QUALITATIVE SKILLS FOR PUBLIC HEALTH RESEARCH: USING INTERVIEWS TO GENERATE DATA

Danica Rotar Pavlič

THEORETICAL BACKGROUND

About qualitative methods and interviews

Recent years have seen the development of interest in the usefulness of qualitative methods in primary care, public health research, and health services research (1). These qualitative methods have a long history in the social sciences and education, but a relatively short one in medicine. They are multi-method in focus, involving an interpretive, naturalistic approach to their subject matter (2). This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them. However, such research has utilized, often uncritically, a "cookbook" of methods for data collection, and common-sense principles for data analysis. It is therefore necessary to study qualitative methods in just as much depth as we do quantitative research methods (3). Qualitative studies are studies that go beyond numbers. Qualitative studies commonly employ interviewing and observation as their research methods, and the data is often text rather than numbers (4).

A major deciding factor when conducting research is the type of questions used. For example, a question such as "What percentage of patients would want antibiotics for their upper respiratory tract infections?" requires a quantitative approach. However, a question such as "What does the term 'antibiotics' mean to people who request it in a consultation for upper respiratory tract infections?" is far better answered by qualitative methods (5). Much qualitative research is interview-based. Interviews allow us to study areas that are difficult to analyze using quantitative methodology; for example, points of view regarding health or beliefs about illness (6). The qualitative research interview seeks to describe and discern the meanings of central themes in the subjects' lives and world. Respondents are encouraged to display their own understandings and opinions. The main task in interviewing is to understand the meaning of what the interviewees say (7).

Types of interviews

There are three main types: structured, semi-structured, and in-depth interviews.

1. Structured interviews.
   Structured interviews contain structured questionnaires. Interviewers are trained to ask questions in a standardized order. All interviewees are asked the same open-ended questions (8).
2. Semi-structured interviews.

Semi-structured interviews are conducted on the basis of a loose structure consisting of open-ended questions. The interviewer or interviewee may diverge in order to pursue an idea in more detail.

3. In-depth interviews.
   In-depth interviews are less structured than this, and may cover only one or two issues, but in much greater detail. These interviews might begin with interviewees freely narrating their stories while the researcher provides minimal guidance. An in-depth interview is useful for shaping basic theoretical models and ethnographic studies.

Aims

The qualitative paradigm:

- Aims to understand the social world from the respondents' viewpoints through detailed descriptions of their cognitive and symbolic actions, and through the richness of meaning associated with observable behaviour,
- Rejects both a cause-and-effect construct and universal laws devoid of any socio-historical context. (9)

Qualitative interviewers aim to go beneath the surface of the topic being discussed, explore what people say in as much detail as possible, and uncover new areas or ideas that were not anticipated at the outset of the research (7).

Methods

Sources of data

Recruitment process and types of sampling

Sampling for qualitative research is an area of considerable confusion for researchers experienced in the hypothetical-deductive model. Purposeful sampling selects cases that are rich in information for in-depth study. The size and the specificity of the cases are determined by the purpose of the study. There are about 16 different types of purposeful sampling. In many cases we are aiming to obtain interviews from subjects with a broad range of socio-demographic characteristics (10). Different types of purposeful sampling are (11):

- extreme and deviant case sampling
- intensity sampling
- maximum variation sampling
- typical case sampling
- homogeneous sampling
- stratified purposeful sampling (this illustrates characteristics of particular subgroups of interest).
- snowball or chain sampling
- criterion sampling
- theory-based or operational construct sampling
- opportunistic sampling
- individuals

Sample size

Adequacy of sample size in qualitative research is relative. A sample size of 10 may be judged adequate for certain kinds of homogeneous or critical case sampling. It may be assessed as too small to achieve maximum variation of a complex phenomenon or to develop theory, or too large for certain kinds of narrative analyses (12).

An appropriate sample size for a qualitative study is one that adequately answers the research question. For simple questions or very detailed studies, this might be in single figures; for complex questions, large samples and a variety of sampling techniques might be necessary. In practice, the number of required subjects usually becomes obvious as the study progresses, as new categories, themes, or explanations stop emerging from the data (data saturation) (13).

Interviewing process

Fieldwork

In ethnography and social anthropology, fieldwork is mainly associated with the technique of participant observation. Interviewing is either a complement of participant observation, or a major facet of it. Nowadays participant observation and interview techniques are paired as the "dynamic duo" of field research. Researchers that need to exert control over what they study design their own research strategies - both before going out and while in the field (14). Fieldwork is the central activity of qualitative data gathering. To be in the field means to have direct, personal contact with people in their own environments. It is the researcher's desire to conceptualize program or product implementation that allows him/her to capture important "results" (i.e., effects and/or outcomes) that standardized measures cannot (15).

The venue

The venue of interviews is defined in the protocol before the interviews are conducted. It might be at home, in the nursing home, in the hospital, or at the doctor's office.

Interview guide

All interviews need to have an interview guide (16)**.** The guide is developed to elicit detailed descriptions of perceptions of the interviewees' experiences. It should be piloted with other professionals external to the main study. Researchers deliver probes in an order that is based on how the interview unfolds. First, a brief demographic questionnaire is usually performed. Elements are presented in Table 1.

After this introduction, face-to-face interviews with the interviewees are conducted. At the start the interviewer can ask: "Could you tell me something about what you know about your illness?" When necessary, prompts or direct questions are used to try to cover the special sections. The guide includes a description of the conclusion of each interview.

Notes on the context (emotional content, etc.) should be written by the investigator. All interviews should be conducted in a sensitive, non-judgmental manner (18).

Table 1. Demographic questionnaireelements.

Element	Categories
Sex	M
	F
Age	
Marital status	Married/Common-law/Partner
	Widowed
	Separated/Divorced
	Single/Never Married
No. of children	
Occupation	
Highest level of education	Less than high school
	High school or equivalent (GED)
	Trade/Vocational/Comm. College
	College
	Beyond college
No. of years of education	
Religion	
Country of birth	
City/town of residence	
Language at home	
Language at work	

Time

The length of interviews is usually measured. They may last from a few minutes to more than an hour.

Interviewees and interviewers

Interviewee

A sample is expected to mirror the population from which it originates; however, there is no guarantee that any sample will be precisely representative of this population. In practice, it is rarely known when a sample is not representative and therefore should be discarded. Criteria for interviewees are defined in the study protocol. Examples include age, sex, diagnosis, registration with a certain health program or organization, specific ethnic or cultural background, and rural or urban setting.

Interviewer

Qualitative interviews require considerable skill on the part of the interviewer (7). The interviewers are chosen according to specific criteria linked to the nature of the task (experience with surveys in different conditions, or of contact with certain people; personal commitment, etc.). No two interviewers are alike and the same person may provide different answers to different interviewers. The novice research interviewer needs to notice how directive he or she is being, whether leading questions are being asked, whether cues are picked up or ignored, and whether interviewees are given enough time to explain themselves (7). Training materials should provide the

knowledge, skills, and attitudes for effective interviewing. The interviewer's training materials include a manual and/or a video. The trainer materials support the interviewer training and provide guidance and examples of activities that support the trainers in the training process. For example, it gives a tip, how to start an interview (Example 1) (17).

Example 1.

My name is _____, and I am a researcher at one of our neighbourhood health centres. I also analyze our hospital data, and have found that a lot of patients aren't receiving certain tests to prevent cancer. I'm trying to understand why this is the case. I'm very interested in hearing what you have to say about this issue.

Data processing

Recording

Recording is used to aid in the collection of information. Recording an interview is necessary to assure accurate reporting of the interview itself. Several research efforts have shown that if there is no adequate and full documentation of an interview, the interviewers' reports of the interview are inaccurate and misleading. Written accounts by the interviewer are not reliable.

Transcription

Verbatim transcriptionof interview data has become a common data management strategy in public health research and is widely considered to be integral to the analysis and interpretation of verbal data (19). The audiotapes should be submitted to an experienced qualified transcriber, who transcribes them verbatim. At least one investigator checks the transcripts for accuracy, listening to portions of the audiotape while reading the typed transcription.

Coding and data reduction

The interviews can be analyzed by using a qualitative method called grounded theory (20), whereby interviews are scrutinized to identify and categorize all ideas presented. Each idea is given a code, and the codes are linked to construct a conceptual framework explaining the collective experiences of those interviewed and the nature of the barriers to establishing code status. The interpretation of the data is thus inductive and "grounded in the words and experiences of the participants" (21). Integration of information is an ongoing cyclical process. Data can be managed using different types of software, such as ATLAS.ti, QSR N5, QSR Nvivo, and NUD*IST. Qualitative data analysis software facilitates extensive coding and text searches of documents.

After reading several interviews for comprehension of interview content, interviews are prepared for open coding. Open coding is a process of reading small segments of text at a time and making notations in the margins regarding content or analytic thought. This should be carried out with no constraints from existing theoretical explanations (22). The labels applied in the open coding process are then synthesized into a code list to remove redundancy. Similar labels are grouped together. The initial theme list closely follows the in-depth interview instrument, as it

is quite detailed in terms of lines of questioning and related probes. The code list is then used to code all of the interviews. The analysis is performed either using a computer program or manually; that is, independently for each interview. This is done using the following steps (23):

1. Theoretical categories are defined, with conceptual codes for organizing information;
2. Interviews are analyzed by assigning conceptual codes to data segments that share the same idea (open coding);
3. New categories with analytical value are identified from the data obtained;
4. Properties of each category are identified (axial coding);
5. Code families are integrated.

Validation

Validation consists of careful inspection of each interview to check whether there were features that would lead to assignment to another category. Investigators can use the consensus model - discussing and re-examining coding discrepancies - to ensure consistency in application of categories. Credibility is assessed by regular debriefing of the data with the team of researchers. A comparison of the results with previous qualitative studies (cumulative validation) might show that some factors have been identified in other settings, while others have not (24).

Triangulation

There are several types of triangulation. The four most commonly recognized include the following:

- data triangulation - the use of a variety of data sources in a study. For example, interviewing people in different status positions or with different points of view
- investigator triangulation - the use of several different evaluators or social scientists
- theory triangulation - the use of multiple perspectives to interpret a single set of data
- methodological triangulation - the use of multiple methods to study a single problem or program (e.g., interviews, observations, questionnaires, documents).

Biases and weaknesses

In interviews following biases can occur:

- generalizability is an issue when studies involve a limited number of respondents.
- patients may modify their responses.
- those patients who are interviewed by the doctor might be less overtly critical.
- one cannot exclude the possibility of a researcher-respondent interaction (i.e., the Hawthorne effect) (25) during interviews.

Advantages and disadvantages

In Table 2 some advantages and disadvantages of interviews are listed.

Table 2. Some advantages and disadvantages of interviews.

ADVANTAGES	DISADVANTAGES
1. High adaptability of the interviewer to the respondent's understanding of questions. 2. Allows the participant to describe what is meaningful or important to him or her using his or her own words rather than their being restricted to predetermined. categories; thus participants may feel more relaxed. 3. Respondent's non-verbal messages can be assessed. 4. Opportunity to monitor the environment in which the respondent is completing the questionnaire. 5. Allows evaluator to probe for more details and ensure that participants are interpreting questions the way they were intended. 6. Higher likelihood of spontaneous answers. 7. Opportunity to confirm the respondent's identity. 8. Possibility of posing more complicated questions. 9. Interviewers have the flexibility to use their knowledge, expertise, and interpersonal skills to explore interesting or unexpected ideas or topics raised by participants.	1. Training interviewers and conducting interviews can be expensive and time-consuming, because qualitative interviewing requires considerable skill and experience. 2. Conducting the interview may take a long time. 3. Possibility that there will not be enough time to fill in the questionnaire. 4. A non-unified method of posing questions. 5. Possibility of difficult access to respondents. 6. Possibility of interviewer bias. 7. More subjective than quantitative interviews because the evaluator/researcher decides which quotes or specific examples to report.

Presentation and interpretation of the results

There are a variety of ways to report the results of qualitative research/evaluation. Common among them is the sense of story, which includes attention to detail, descriptive vocabulary, direct quotes from those observed or interviewed, and thematic organization (15).

CASE STUDY: HOW OLDER PATIENTS AND THEIR GENERAL PRACTITIONERS ASSESS SHARED DECISION-MAKING IN HEALTHCARE.

Introduction

The traditional relationship and mode of communication between physician and patient has undergone radical changes over the last 40 years (26). Better access to information

about health, healthcare, and treatment options offer the opportunity for a new type of physician-patient partnership. As the user of health services, the patient has the primary role. Patients' involvement and participation in treatment decisions is increasingly important.

Current models and measures of patient involvement in treatment decision-making tend to focus on communication during consultations and/or on the patient's use of information to consider the selection of one treatment option from a well-defined set (27). Active participation leads to improved patient satisfaction, improved therapeutic compliance, better quality healthcare - including health status and satisfaction with care (28) - and a decrease in healthcare costs (29) in part due to reduced use of laboratory services and referrals (30). Many factors may influence patients' preferences for involvement (31), including some that relate to the patient, some to the physician, and some to the health system organization or cultural and historical background; all of these may change over time (32). Sometimes, attempts to achieve agreement between physicians and patients create conflict, as when patients' wishes are not in line with prevailing medical opinion. Patients may also interpret a physician's wish to involve them in decision-making as a sign of professional insecurity (34).

Design phase

Recognition and definition of the problem and justification of the survey objectives

The problem

This article examines the healthcare needs and expectations of the elderly. Life expectancy has increased. The proportion of the European population that is elderly (age 65 and older) increased from 13.9% in 1980 to 19.9% in 2000 (Eurostat population estimates). Older people's opinions about participation and involvement in medical treatment plans are poorly studied in Central and Eastern Europe. In Slovenia in 2003, 15% of the population was over 65, and by 2010 this share is expected to increase to 16.5% (9). Following independence in 1991, there were several changes in the Slovenian healthcare system, including the introduction of a physician of choice (patients may now select their family physician), changes in the health services financing system, and a different planning method (34). Primary care in Slovenia is mainly provided by general practitioners (GPs), paediatricians, gynaecologists, and dentists (35), all of whom control access to secondary care. The personnel delivering primary health care include: general practitioners or family physicians, dentists, nurses, pharmacists, physical therapists, speech therapists, occupational therapists, psychologists or psychiatrists, midwives and other health professionals necessary to carry out the work of the health centre. Following healthcare reforms in Slovenia, the number of consultations with GPs increased considerably while the average duration of consultations decreased (36,37).

Determination of aims and goals

This study is part of the international project IMPROVE, the purpose of which is to strengthen shared decision-making among patients over 70. Slovenia is also participating in this project. The first level of this three-year project consisted of

analyzing points of view and ideas about shared decision-making in healthcare among older patients and their GPs. The goal of this article is to analyze the role of shared decision-making in healthcare in Slovenia and thereby attain insight into the status of the active role of older patients in this Central European country.

Methods

This study was conducted in Slovenia as part of the international project IMPROVE, involving researchers from 11 countries: Austria, Belgium, Denmark, France, Germany, Israel, the Netherlands, Portugal, Slovenia, Switzerland, and the UK (37). The project protocol defined uniform procedures in all participating countries, including a description of the sample of older patients and a description of the sample of GPs. As a research tool**,** we used a problem-oriented interview (38), structured with an emphasis on patient involvement and conducted in a manner that did not allow deviation to unrelated matters. During the interviews the interviewer was required to check that all questions were properly understood. The interviews that we recorded with older patients started with the following introduction: "As you have been informed by your GP, we and other GPs are interested in knowing whether patients wish to participate in decisions concerning their healthcare planning and treatment. What do you think about this?" The follow-up questions are as follows:

- Have you ever had the feeling that you do not participate in your healthcare planning and treatment as much as you would like to?
- To what extent does your GP involve you in your healthcare planning and treatment, and give you the opportunity to be involved in the planning and decisions about your treatment?
- Can you describe an example of a particular situation or event?
- In your opinion, how could you be more involved?
- In general, would you like to participate more in your healthcare decisions?

The interview that we recorded with the GPs began with the following introduction: "In this project we wish to focus on how older patients participate in decision-making in general medicine. We would first of all like to determine which factors influence the use of methods to encourage older patients to participate in shared decision-making." The follow-up questions are as follows:

- What is your understanding of the idea of shared decision-making in healthcare?
- Can you provide one or two examples?
- Thinking about your patients, what are the advantages for you of patients' shared decision-making in healthcare?
- What obstacles prevent you from using methods to promote patient inclusion in healthcare at your office?
- What do you feel are the advantages and disadvantages of the tools that we use to promote shared decision-making?
- How do you view the future of patients' shared decision-making in healthcare?

Sampling and recruitment

The approximate number of participating older patients was defined at the international level. We interviewed 40 older patients stratified by age, sex, level of urbanization, and recent health status. The sample of older patients was chosen to achieve a balance by sex, age group (70 to 79 years old; 80 and older), place of residence (urban, suburban, rural), and health status (chronic disease, acute illness, life-threatening illness, and patients that had not consulted for a long time). Patients unable to speak Slovenian, in a terminal stage of a disease, or suffering from attention and concentration disorders were excluded. We had no difficulties recruiting older patients because their physicians of choice invited them to participate in the study. In Slovenia, the term "physician of choice" refers to a family physician that the patient has selected with his/her signature to ensure continuous and long-term care.

The sample of physicians that we interviewed in Slovenia was larger than in other countries. By including a larger number of physicians, we sought to include the points of view and evaluations of physicians that work in remote areas, those in private practice, those working in the public healthcare sector, school physicians, and those in partner or group practices. We included 26 physicians and recorded interviews with them.

Taped interviews were successfully conducted with 39 older patients. One patient interrupted the conversation during the interview, so his interview was excluded from the study. Most interviews with older patients were conducted in the patients' homes and a minority at health centres. All of the interviews with GPs were recorded at their offices.

First of all, DRP listened to the tapes to check their quality. After this initial quality check, the tapes were transcribed by AA and ML, who had already participated in similar studies. DRP compared the transcriptions with the recordings. This was followed by a text analysis using the ATLAS computer program according to the following principles:

- marking relevant parts (statements, quotes) and encoding;
- linking codes to relevant key content (categories, topics, subtopics);
- repeatedly checking quotes and codes (revision);
- presentation of codes, quotes, and categories at a researchers' meeting;
- final creation of text files with final quotes, subtopics, and topics.

Special attention was paid to data validation: five primary texts of both patients and GPs were analyzed and coded by two independent researchers (DRP, MK). Good congruity between them was demonstrated. The international consistency of the coding was evaluated based on two translated interviews (one patient interview and one GP interview) that were sent to Richard Baker. No major international differences were found.

Determination of ethical principles

Before the study began, the entire international research protocol and the protocol translated into Slovenian were examined and approved by the National Medical Ethics Committee (39).

Determination of database management and results dissemination channels

Thirty-nine older patients were interviewed between October 2000 and February 2001. Their ages ranged from 70 to 95. The interviews lasted 30 to 45 minutes. Interviews with 26 GPs were taped at their offices. These interviews lasted 20 to 40 minutes.

The recruited patients provided a detailed description of their perception of the relationship between GP and patient. They often interpreted involvement as a relationship in which the GP meets their expectations, whereas the GPs viewed these expectations as potential interference with their professional role. The patients evaluated involvement as a concept with four aspects:

- mutual activity and improved relationship,
- ethics,
- relationship,
- meeting expectations.

The synthesis of codes, categories, and themes depicted involvement as a concept, as illustrated in Figure 1.

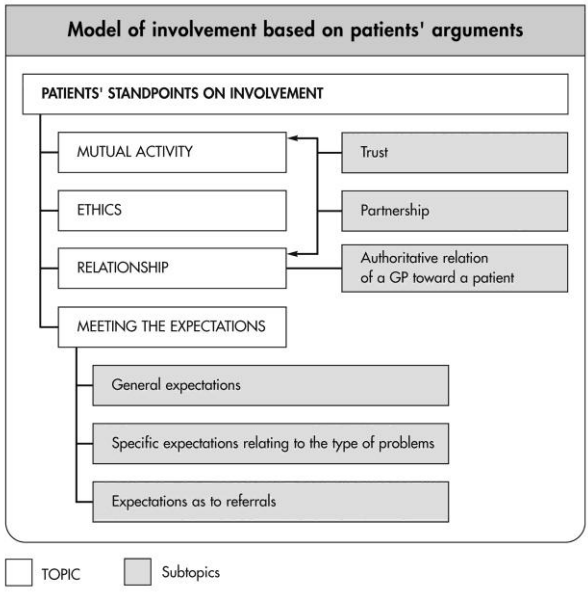


Figure 1. The synthesis of codes, categories, and themes depicted involvement as a concept.

The majority of GPs felt that patient involvement in healthcare was beneficial and should be encouraged. An understanding of the meaning of involvement embraces several areas:

- comprehension,
- expert knowledge,
- ethics,
- communication,
- relationship.

Most of the GPs defined involvement as a process of comprehension of dialogue during consultation. Involvement might be enhanced by the patient's relatives or the presence of a nurse. In contrast, the presence of a third party was seen as a threat by some GPs. Patients' proposals in making healthcare decisions can interfere in their professional area (Quote 1, Quote 2).

> *"Some elderly people actively interfere in matters, which means they're interfering not only with regard to themselves, but also in what the physician then says; then you need quite a lot of expert knowledge. They'd like to direct their treatment at a professional level." (GP, 202309)* **Quote 1.**

> *"We have problems with people, with patients that think they know a lot about healthcare and want to decide for themselves and don't accept any explanation. Just because a friend told him 'no,' he believes his friend, even though he's not a GP with expert knowledge. I find it difficult to convince him otherwise." (GP, 204509)* **Quote 2.**

The synthesis of codes, categories, and themes based on GPs' arguments is illustrated in Figure 2.

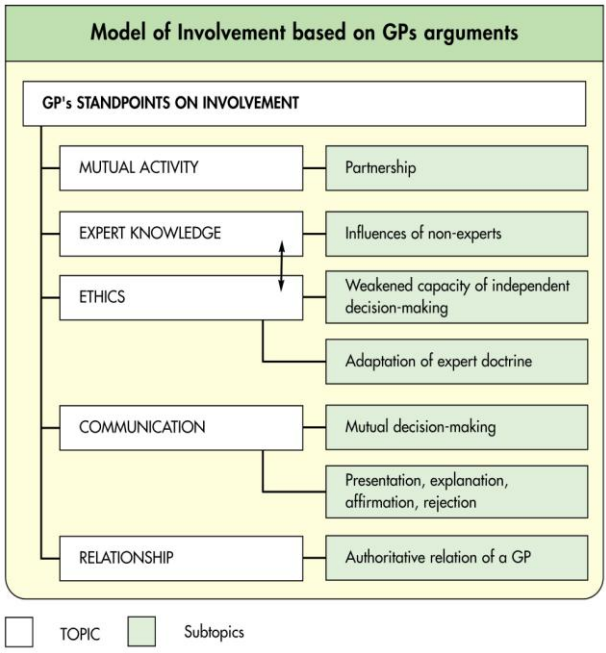


Figure 2. The synthesis of codes, categories, and themes based on GPs' arguments.

Involvement as mutual activity

The patients saw involvement as an improved relationship between the GP and patient that led to better treatment outcomes. Involvement in healthcare decisions is clearly dependent on the willingness and motivation to participate. Trust in the GP is essential. Even in the atmosphere of partnership, medical advice is not always accepted (Quote 3, Quote 4).

> *"Oh, well, a patient gives his opinion. Absolutely. Why not? A patient knows himself. It's good for a patient to give his opinion about the treatment." (patient, 76 years old, 101309)* **Quote 3.**

> *"By all means a person has to have general contact with the GP. A patient talks to him. And there's something else: the GP says, 'I'm giving you this medicine, and I advise this, but if YOU don't do anything yourself, the treatment will be no good either.' So a person must do a lot himself and also trust the GP because nothing can be done without mutual consent." (patient, 71 years old, 102309)* **Quote 4.**

The GPs consider involvement to be a process in which the GP advises the patient on diagnosis and treatment options. This process is based on consent. The following two quotes reflect examples of decision-making (Quote 5, Quote 6):

> *"Especially when I familiarize a patient with various possibilities, options he has in diagnostics, treatment, and of course my prognosis, and I obtain the patient's consent. Actually, I always inform him what it's all about, what disease he has or what disorder he's suffering from. I also explain the risk factors to him. I make suggestions and then ask him if he agrees. If he does, we cooperate. If he disagrees, I mark in his files that he doesn't agree." (GP, 201109)* **Quote 5.**

> *"By involvement, I understand that the GP informs a patient about his diagnosis, about possible treatment, and about the prognosis, and that he offers him some options. For instance, that he'll undergo treatment with medicine or physiotherapy. In short, the GP has to talk to him about the treatment mode. A patient can also disagree with the suggested examinations and treatment, of course." (GP, 202209)* **Quote 6.**

Comparatively few GPs questioned patients to find out what they had already done to treat their own problems, especially non-pharmacological measures (Quote 7).

> *"By involvement I understand that a patient tells me what he'd do for himself. Yes, I always ask him what he's done, and then I ask him what he'd do, and then we make a mutual decision." (GP, 204309)* **Quote 7.**

Ethics

The elderly understood involvement to include respect for their reluctance to receive treatment at a clinic or hospital and their preference for the familiar domestic environment. They expect GPs to adapt their professional advice to their particular circumstances (Quote 8).

> *"The GP examined me and he found out what was wrong. I gave blood, I went to have my urine checked. The GP gave me some medicine, but nothing improved. A few days later I came back for the same examinations, and he saw that the situation hadn't improved. He put his booklet on the table: 'You're going to the hospital.' Then I said: 'Doctor, I'm not going to the hospital. Just home, just home. I'm not going.' He wanted to know why. 'You know what, it's like this,' I said: 'I'm so old already.'" (patient, 96 years old, 103409)* **Quote 8.**

Several GPs treat elderly people in a stereotypical doctor-centred manner and do not allow them to become involved in healthcare. Others were concerned that patient involvement might lead them to deal with issues that were not relevant to optimum patient management (Quote 9).

> *"The option that he (the patient) decides on isn't good for him because he doesn't know himself." (GP, 203409)* **Quote 9.**

The physician-patient relationship

Some subjects felt that personal involvement in healthcare is not possible because the doctor is in charge and it is their job to follow instructions. Subordinate status in this authoritative relationship is shown by the following phrases used by patients: "obey", "permission to, do what the GP says". A similar interpretation of involvement was also found in the analysis of GPs' texts (Quote 10, Quote 11).

> *Interviewer: "How much do you participate at the doctor's?"* **Quote 10.**
>
> *Patient: "At the doctor's? Well, what the GP says I should do, I do. Let's take medicine, everything, I'd say I obey him." (patient, 81 years old, 107109)*

> *"The majority aren't prepared for this (involvement). The majority of elderly patients are used to past practice when the GP decided on the therapy and the patient had no say about it. No one even thought about making their own decision, about expressing their points of view." (GP, 209109)* **Quote 11.**

Trust in a GP is an important element of the relationship with a GP. The GPs did not highlight trust. They emphasized the tension and stress that patients experience when visiting a GP (Quotes 12-14).

"By involvement I understand that I trust a GP." (patient, 81 years old, 107109) **Quote 12.**

"You have to trust a GP, and I think this is a good contact." (patient, 71 years old, 102309) **Quote 13.**

"Patients are under some stress at the office. They're afraid and don't remember everything we say." (GP, 210509) **Quote 14.**

Meeting patients' expectations and requirements, and expert knowledge

Some elderly patients evaluated involvement according to how well their expectations were met during the consultation (Quote 15, Quote 16).

Interviewer: "Do you wish to be more involved in the decisions concerning your treatment?" **Quote 15.**

Patient: "So far, he [the GP] has done me a favour, he's really done me a favour." (patient, 76 years old, 112109)

"By involvement I understand that I also tell the GP what I want, that I get it from him." (patient, 79 years old, 106209) **Quote 16.**

Many patients see referral to the patient's chosen specialist as an example of good physician-patient cooperation. Participation also means that a GP facilitates specialist examinations for patients wishing to be referred (Quote 17).

"He [the GP] allows me everything, if I ask him for any kind of referral, anywhere. He's never refused to give one to me. I've always asked him for things that aren't stupid. He knew these doctors that I visited, and I had no problems with them because of that." (patient, 76 years old, 113109) **Quote 17.**

Some patients evaluated involvement passively, assuming the GP would address the patient's needs. From this perspective, the patient deliberately avoids taking any personal or even shared responsibility for his care (Quote 18).

Interviewer: "Would you like to make more decisions and to participate in your treatment?" **Quote 18.**

Patient: "No, because I don't have many problems. I wouldn't want to make it more difficult as far as the GP's work is concerned." (patient, 82 years old, 102109)

In patients' attempts to manifest specific needs and expectations, physicians see a danger that could impede healthcare. By expressing their nonprofessional expectations, a patient can "disturb" a professional treatment plan (Quote 19).

Quote 19.

"One disadvantage can be that a patient loses touch with reality. The GP is the one to tell him what reality is. The patient can go to extremes if he's overly involved. This means he deals with trivial issues that aren't vital, and trivial issues can become a major problem for a patient." (GP, 204309)

Discussion

A clear understanding between the GP and the patient and a confident relationship between them are two important factors in healthcare (31,38,40); this was confirmed in this study. As previously reported, elderly patients stressed the importance of the GP-patient relationship (41). The results of our study clearly express a need to build relations between doctors and patients. This can be explained by the fact that the structure of consultations has not significantly changed in the past decade, and that the building of partnership relations remains a minor part of the consultation (42). Our study also expresses the importance of developing a relationship before taking action (43,44). Older patients also wish to confidentially express trivial concerns and opinions that are important from a nonprofessional point of view, but are not part of narrowly-oriented professional medical treatment. A prominent feature of our research is its emphasis on an authoritative relationship. It is clear that older patients' healthcare is more oriented toward problem-solving phases of care; however, physicians feel uncomfortable if patients' subjective dimensions (i.e., opinions and expectations) are expressed in shared decision-making. This may be the result of insufficient instruction on patient participation, which is presented as a continuum between involvement and non-involvement at the undergraduate and graduate levels (45).

Our findings emphasize the diversity of perceptions and needs of the older population, and also confirm the model in which shared decision-making is defined as a continuum from the desire for the GP to make decisions for patients to patients' active participation (45). Patients differ in their needs for health-related information depending on their own priorities and their particular problems. Elderly people appreciate a relationship that embraces values such as trust, support, and discussion of feelings (46) and the individual, as opposed to an orientation toward disease. Patients' desire for shared decision-making must therefore be defined on an individual basis (47). The analysis of viewpoints showed mixed views on shared decision-making. Some patients want the GP to make decisions for them, whereas others would like to make decisions together, and still others stress one's personal responsibility for one's health. Shared decision-making is not possible (i.e., cannot be carried out) without co-determination by patients and GPs. It can occur only through the reciprocal relationships of dialogue and shared decision-making (48).

The communication method should be specifically adapted to the individual characteristics and preferences of the individual older patient.

Ethical issues were considered important by both older patients and GPs. Reliable information on the health issue and proper advice on treatment options is an ethical principle and a legal right (49,39). One ethical principle that is of special importance in this regard is patient autonomy. It is a GP's professional responsibility to permit this, and it is influenced by political decisions and the characteristics of a consumer society. An ethical dilemma is expressed not only with regard to patient autonomy, but also in adaptation and reduction of professionalism. "Forcing" GPs to change their professional doctrine and the perception of older patients as a special subgroup of people with limited cognitive abilities are current ethical issues. Tensions were discovered that arise due to GPs' focus on the best medical care for patients and expressions of patient autonomy. Studies (50) have shown that ethical aspects have often been treated in a theoretical manner, whereas empirical practice can be entirely different from theoretical points of departure.

Qualitative methodology as used here is particularly appropriate for research on perceptions, opinions, and personal experiences (40). The "one interviewer - one patient" interviews were evaluated to provide added value from the words and statements that the interviewed subjects chose. The more delicate, personal, and emotional contents of the interview can be identified (40). Direct observation of a group of older patients would not be possible without unanimous consent and would present unacceptable organizational difficulties. Nevertheless, the total sample met the stratification criteria and a broad range of views was achieved. Our findings cannot be generalized, but they do provide valuable indications and guides to further research.

Conclusion

Elderly people's understanding of involvement in healthcare is more focused on "building a relationship" than on making decisions. The perceptions and needs of elderly patients are characterized by considerable variation and diversity. Specific views and preferences regarding involvement in their healthcare should be explored during consultations. Dilemmas about patient autonomy and the adoption of a paternalistic approach come to the fore, and there is a need to educate GPs and their co-workers on these issues. Physicians ought to provide elderly patients with the opportunity to express very personal points of view, expectations, and fears, even though (due to occasional trivial points of view and concerns) these might not have a significant influence on professional medical treatment of health problems.

In their continuing medical education, GPs should dedicate more time to the GP-patient relationship, patient participation, and patient responsibility for health.

Funding for conducting the survey

The study was funded by the Slovenian Research Agency (ARRS, project code L3-3332).

EXERCISE

Task 1

Carefully read the part on the theoretical background of this module. Critically discuss the characteristics of "interweaving" as a research tool with your classmates.

Task 2

Find out if any other qualitative study has been performed in your country using international bibliographic databases (e.g., Medline, PubMed).

Task 3

Explore the advantages and disadvantages of the qualitative method used in the published qualitative study.

Task 4

Discuss the characteristics, strengths, and limitations of the selected qualitative study with your classmates.

REFERENCES

1. Murphy E, Mattson B. Qualitative research and family practice: a marriage made in heaven? Fam Pract. 1992;9:85-91.
2. Denzin NK, Lincoln YS, editors. Handbook of qualitative research. London: Sage Publications; 1994.
3. Harding G, Gantley M. Qualitative methods: beyond the cookbook. Fam Pract. 1998;15(1):76-9.
4. Greenhalgh T, Taylor R. Papers that go beyond numbers (qualitative research). Br Med J. 1997;315:740-3.
5. Pong LT. Qualitative research methods in family medicine. Fam Pract. 1992;9:85-91.
6. Pope C, Mays N. Qualitative research: reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. Br Med J. 1995;311:42-5.
7. Britten N. Qualitative Research: Qualitative interviews in medical research. Br Med J. 1995;311:251-3.
8. Britten N. Qualitative interviews. In: Pope C, Mays N. Qualitative research in health care. London: BMJ Books; 1999. p.12-20.
9. Myers M. Qualitative research and the generalizability question: standing firm with Proteus. The Qualitative Report [Internet]. 2000 March 4(3/4), Available from URL: http://www.nova.ed/ssss/QR/QR4-3/myers.html. Accessed: Apr 24, 2007.
10. Butler CC; Pill R, Stott NCH. Qualitative study of patients' perceptions of doctors' advice to quit smoking: implications for opportunistic health promotion. Br Med J. 1998;316:1878-81.
11. Patton MQ. Qualitative evaluation and research methods. 2nd ed. Newbury Park London New Delhi: Sage Publications; 1990.
12. Sandelowski M. Focus on qualitative methods. Sample size in qualitative research. Res Nurs Health. 1995;18:179-83.
13. Marshall, MN. Sampling for qualitative research. Fam Pract. 1996;13:522-5.

14. Sperschneider W, Bagger K. Ethnographic fieldwork under industrial constraints: toward design-in-context. Int J Hum Comput Interact. 2003;15(1):41-50.
15. Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks, CA: SAGE Publications; 2002.
16. Kirk P, Kirk I, Kristjanson LJ. What do patients receiving palliative care for cancer and their families want to be told? A Canadian and Australian qualitative study. Br Med J. 2004 Jun 5;328:1343-9.
17. Lasser KE, Ayanian JZ, Fletcher RH, DelVecchio Good MJ. Colorectal cancer screening barriers: diverging opinions. BMC Fam Pract. 2008;9:15.
18. Lawton J, Ahmad N, Hallowell N, Hanna L, Douglas M. Perceptions and experiences of taking oral hypoglycaemic agents among people of Pakistani and Indian origin: qualitative study. Br Med J. 2005;330:1247-9.
19. Halcomb EJ, Davidson PM. Is verbatim transcription of interview data always necessary? Appl Nurs Res. 2006;19:38-42.
20. Straus, A, Corbin J. Basics of qualitative research: grounded theory procedures and techniques. London: Sage Publications; 1990.
21. Grams GD, Herbert C, Heffernan C, Calam B, Wilson MA, Grzybowski S, et al. Haida perspectives on living with non-insulin-dependent diabetes. CMAJ 1996;155(11):1563-8.
22. Glaser B, Strauss A: The discovery of grounded theory**.** Chicago, IL: Aldine; 1967.
23. Pelcastre-Villafuerte BE, Tirado-Gómez LL, Mohar-Betancourt A, López-Cervantes M. Cervical cancer: a qualitative study on subjectivity, family, gender and health services. Reprod Health. 2007;4:2.
24. Spinewine A, Swine C, Dhillon S, Dean Franklin B, Tulkens PM, Wilmotte L, Lorant V. Appropriateness of use of medicines in elderly inpatients: qualitative study. Br Med J. 2005;331:935-9.
25. Macefield R. Usability studies and the Hawthorne effect. J Usability Stud. 2007;2:145-54.
26. Balint J, Shelton W. Regaining the initiative. Forging a new model of the patient-physician relationship. JAMA. 1996;275:887-91.
27. Entwistle VA, Watt IS. Patient involvement in treatment decision-making: the case for a broader conceptual framework. Patient Educ Couns. 2006;63:268-78.
28. Coulter A. What do patients and the public want from primary care? Br Med J. 1997;311:251-3.
29. Greenfield S, Kaplan SH, Ware J-EJ, Yano EM, Frank HJ. Patients' participation in medical care: effects on blood sugar control and quality of life in diabetes. J Gen Intern Med. 1988;3:448-57.
30. Say R, Murtagh M, Thomson R. Patients' preference for involvement in medical decision making: a narrative review. Patient Educ Couns. 2006;60:102-14.
31. Entwistle V, Prior M, Skea ZC, Francis JJ. Involvement in treatment decision-making: its meaning to people with diabetes and implications for conceptualisation. Soc Sci Med. 2007;66(2):362-375.
32. Logan RL, Scott PJ. Uncertainty in clinical practice: implications for quality and costs of health care. Lancet. 1996;347:595-8.
33. Dimovski V. Platform to develop a strategy of care of the elderly in the period from 2006 to 2010 (in Slovene). In: Festival za tretje življenjsko obdobje. Zbirka prispevkov okroglih miz in predavanj. Ljubljana. Ljubljana: Zveza društev upokojencev Slovenije and Infos d.o.o.; 2004. p.10-17.
34. Markota M, Svab I, Sarazin KK, Albreht T. Slovenian experience on health care reform. Croat Med J. 1999;40:190-4.
35. European Observatory of Health Care Systems. Health care system in Slovenia. 4th ed. Copenhagen: The World Health Organization Regional Office for Europe: 2002.

Available from URL: http://www.euro.who.int/document/e76966.pdf. Accessed: March 31, 2007.
36. Svab I. Progress in family medicine in Slovenia. Middle East J Fam Med. 2004;2:1-5.
37. Wetzels R, Geest TA, Wensing M, Ferreira PL, Grol R, Baker R. GPs' views on involvement of older patients: a European qualitative study. Patient Educ Couns. 2004;53:183-8.
38. Hall JA, Roter DL, Milburn MA, Daltroy LH. Patients' health as a predictor of physician and patient behavior in medical visits. A synthesis of four studies. Med Care. 1996;34:1205-18.
39. National Medical Ethics Committee. [Internet homepage]. Ljubljana: The National Medical Ethics Committee of the Republic of Slovenia; updated 2007 Nov 11;]. Available from URL: http://www.mf.uni-lj.si/kme-nmec/. Accessed: Dec 12, 2007.
40. Vermeire E. The study of research evidence synthesis, applied to adherence of treatment recommendations in people living with type 2 diabetes mellitus [dissertation]. [Antwerp]: Faculty of Medicine, Department of General Practice, University of Antwerp; 2005.
41. Ogden J, Ambrose L, Khadra A, Manthri S, Symons L, Vass A, et al. A questionnaire study of GPs' and patients' beliefs about the different components of patient centredness. Patient Educ Couns. 2002;47:223-7.
42. Deveugele M, Derese A, De Becquer D, van den Brink-Muinen A, Bensing J, De Maeseneer J. Consultation in general practice: a standard operating procedure? Patient Educ Couns. 2004;54:227-33.
43. Dewing J. Concerns relating to the application of frameworks to promote person-centeredness in nursing with older people. Int J Older People Nurs in association with J Clin Nurs. 2004;13:39-44.
44. Adams T. Communication and interaction within dementia care triads. Dementia. 2005;4:185-205.
45. Millard L, Hallett C, Luker K. Nurse-patient interaction and decision-making care: patient involvement in community nursing. J Adv Nurs. 2006;55:142-50.
46. Callahan EJ, Stange KC, Zyzanski SJ, Goodwin MA, Flocke SA, Bertakis KD. Physician-elder interaction in community family practice. J Am Board Fam Pract. 2004;17:19-25.
47. Florin J, Ehrenberg A, Ehlfors M. Patient participation in clinical decision-making in nursing: a comparative study of nurses' and patients' perceptions. J Clin Nurs. 2006;15(12):1498-508.
48. Thompson AGH. The meaning of patient involvement and participation in health care consultations: a taxonomy. Soc Sci Med. 2007;64:1297-310.
49. Edwards A, Elwyn G, Smith C, Williams S, Thornton H. Consumers' views of quality in the consultation and their relevance to 'shared decision-making' approaches. Health Expect. 2001;4:151-61.
50. Hoffmaster CB, Stewart MA, Christie RJ. A survey method for investigating ethical decision making in family practice. Fam Med. 1992;24:433-8.

RECOMMENDED READINGS

1. Kvale, S. InterViews: an introduction to qualitative research interviewing. Thousand Oaks, CA: Sage; 1996.
2. Rotar-Pavlič D, Švab I, Wetzels R. How do older patients and their GPs evaluate shared decision-making in healthcare? BMC Geriatrics 2008, 8:9. Available from URL URL: http://www.biomedcentral.com/1471-2318/8/9. Accessed: March 31, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**QUALITATIVE METHODS: FOCUS GROUPS**
Module: 1.5.2	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Rok Fink**, Bsc, MSc Candidate, Teaching Assistant Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia **Andreja Kukec**, Bsc, PhD Candidate, Teaching Assistant Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Mojca Jevšnik**, BSc, PhD, Senior lecturer Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia
Address for correspondence	**Mojca Jevšnik** Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana; Zdravstvena pot 5, Ljubljana, 1000 Slovenia. Tel:+386 1 3001149; Fax:+386 1 3001119 E-mail: mojca.jevsnik@vsz.uni-lj.si
Keywords	qualitative analysis, focus groups, public opinion
Learning objectives	After completing this module student should: • increase knowledge about qualitative analysis and focus groups as one of the method for public opinion studies, • understand why qualitative research is vital in public opinion analysis, • be capable to make own research with the method of focus groups, • get to know with sociological approaches.
Abstract	Qualitative research is one the basic research concepts in social sciences and in psychology. There is available a great variety of specific methods each of which starts from different premises, and pursues different aims. Each method is based on a specific understanding of its object. However, qualitative methods cannot be regarded independently of the research process. Special types of qualitative analysis are focus groups with purpose to listen and gather information. It is a way to better understand how people feel or think about an issue, product or service therefore participants are selected due to certain characteristics in common that relate to the topic of the focus groups.
Teaching methods	After introductory lecture students carefully read the recommended sources about the qualitative analysis technique and focus on the method of focus groups. Afterwards they discuss focus groups methodology. Students are separated in groups of six to eight persons. four to six represent public, meanwhile one is moderator and second one is keeper of the minutes. After transcription, each group should make (vertical analysis) and then compare with other groups (horizontal analysis). At last all results should be presented in thematic categories and discussed.
Specific recommendations for teachers	• work in small groups of 4-6 students; • proportion of fieldwork under teacher supervision/individual students' work: 50%/50%; • facilities: a lecture room, a computer room, rooms for small-groups work; • equipment: voice recorder, computer for transcript; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire and case study.

QUALITATIVE METHODS: FOCUS GROUPS

Rok Fink, Andreja Kukec, Mojca Jevšnik

THEORETICAL BACKGROUND

About qualitative research

Qualitative research can be seen on one hand as independent supplement, and on the other as contrast and special accentuation in relationship to experimental and theoretical models. Nevertheless, the qualitative method is also opposite to quantitative (Table 1), which is focused to paradigm of united science (1).

Getting information with focus groups is commonly used method in qualitative research. Howsoever, we are result of surroundings and peoples around us. Therefore this characteristic can be used in focus groups. Questionnaires, phone researches and e-mail methods are insufficient since they are presuming that individuals are realized what they feel. But a lot of human reactions are spontaneous and that is advantage of focus groups (2).

In history of social sciences two different approaches developed in exploring social phenomenon: Quantitative research is followed by natural science, meanwhile the qualitative is due to humanistic and cultural sciences (3,4).

Qualitative research is concerned with developing explanations of social phenomena. That is to say, it aims to help us to understand the social world in which we live and why things are the way they are. It is concerned with the social aspects of our world and seeks to answer questions about:

- why people behave the way they do;
- how opinions and attitudes are formed;
- how people are affected by the events that go on around them;
- how and why cultures and practices have developed in the way they have (5).

In a health or social care setting, qualitative research is particularly useful where the research question involves one of the situations below and people's experiences and views are sought:

- exploration or identification of concepts or views;
- exploration of "implement ability";
- the real-life context;
- sensitive topics where flexibility is needed to avoid causing distress (5).

Qualitative research is a highly rewarding activity because it engages us with things that matter, in ways that matter. Through qualitative research we can explore a wide array of dimensions of social world, including the texture and weave of everyday life, the understandings, experiences and imaginings of our research participants, the ways that social processes, institutions, discourses or relationship work, and the significance of the meaning that they generate. Qualitative research therefore has massive potential, and its practitioners face some major challenges. It deserves to be done well so that it can make fully justified claims for its own significance, effectiveness and meaning (6,3).

There have been many attempts to define qualitative research in the social sciences, and to determine whether or not it can, or should be, differentiated from something called quantitative research. However, there is no consensus on these questions, and we should not be surprised by this, because qualitative research, whatever it might be certainly is not unified set of techniques or philosophies, and ended has grown out of a wide range of intellectual and disciplinary traditions (7,8).

The most serious and central difficulty in the use of qualitative data is that methods of analysis are not well formulated. For quantitative data, there are clear conventions the researcher can use. But the analyst faced with a bank of qualitative data, has very few guidelines for protection against self-delusion, let alone the presentation of unreliable or invalid conclusions to scientific or policy-making audiences (4).

Table 1. Comparison of qualitative and quantitative methods (3).

QUALITATIVE RESEARCH	QUANTITATIVE RESEARCH
1. tends to focus on how people or groups of people can have different ways of looking at reality	1. tends to focus on ways of describing and understanding reality by the discovery of general "laws".
2. takes account of complexity by incorporating the real world context an take different perspectives on board	2. takes account of complexity by precise definition of the focus of interest and techniques that mean that external "noise" can be discounted.
3. studies behaviour in natural settings or uses people's accounts as data; usually no manipulation of variables	3. involves manipulation of some variables while other variables considered to be extraneous and confounding are held constant
4. focuses on reports of experience or on data which cannot be adequately expressed numerically	4. uses statistical techniques that allow us to talk about how likely it is that something is "true" for a given population in an objective or measurable sense
5. focuses on description and interpretation and might lead to development of new concepts or theory, or to an evaluation of an organisational process	5. focuses on cause & effect - e.g. uses experiment to test an hypothesis
6. employs a flexible, emergent but systematic research process	6. requires the research process to be defined in advance

Qualitative and quantitative approaches can complement each other. One simple way in which this can be achieved, is by using qualitative research as the preliminary to quantitative research. This model is likely to be the most familiar to those engaged in health services research (6,9). Typically, qualitative research will provide in-depth information into fewer cases whereas quantitative procedures will allow for more breadth of information across a larger number of cases (1,6,10,11).

Brief overview of methods of qualitative analysis

For the purposes of this module, we will use following classification:

1. Interviews.
   Interviewing can, at one extreme, be structured, with questions prepared and presented to each interviewee in an identical way using a strict predetermined order. At the other extreme, interviews can be completely unstructured, like a free-flowing conversation.

   Qualitative researchers usually employ "semi structured" interviews which involve a number of open ended questions based on the topic areas that the researcher wants to cover. The open ended nature of the questions posed, defines the topic under investigation but provides opportunities for both interviewer and interviewee to discuss some topics in more detail. If the interviewee has difficulty answering a question or provides only a brief response, the interviewer can use cues or prompts to encourage the interviewee to consider the question further. In a semi structured interview the interviewer also has the freedom to probe the interviewee to elaborate on an original response, or to follow a line of inquiry introduced by the interviewee (9,12).
2. Observation.
   Not all qualitative data collection approaches require direct interaction with people. Observation is a technique that can be used when data cannot be collected through other means, or those collected through other means are of limited value or are difficult to validate. For example, in interviews participants may be asked about how they behave in certain situations but there is no guarantee that they actually do what they say they do observing them in those situations is more valid: it is possible to see how they actually behave (12).

   Observation can also produce data for verifying or nullifying information provided in face to face encounters.

   In some research observation of people is not required but observation of the environment. This can provide valuable background information about the environment where a research project is being undertaken. In a health needs assessment, or in a locality survey, observations can (5,9):
   - provide broad contextual descriptions of the key features of the area, for example whether the area is inner city, urban or rural, the geographical location, the size of the population;
   - describe the key components of the area: the main industries, type of housing;
   - the availability of services can be identified: the number, type and location of health care facilities such as hospitals and health centres, care homes, leisure facilities, shopping centres.
3. Collection of documented material.
   Documented material such as letters, diaries, or photographs could be collected:
   - documentation – a wide range of written materials can produce qualitative information. These can be particularly useful in trying to understand the philosophy of an organisation as may be required in ethnography. They can include policy documents, mission statements,

annual reports, minutes of meetings, codes of conduct, web sites, series of letters or emails, case notes, health promotion materials, etc.;

- diaries - diary entries may be used retrospectively, or diaries may be given to research participants who are asked to keep an account of issues or their thoughts concerning diet, medication, interactions with health care services or whatever is the subject of the research. Audio diaries may be used if the written word presents problems. Notice boards can also be a valuable source of data (12);
- photographs - photographs are good way of collecting information which can be captured in a single shot or series of shots. For example, photographs of buildings, neighbourhoods, dress and appearance could be analysed in such a way as to develop theory about professional relationships over a given time period. Photographs may be produced for research purposes or existing photographs may be used for analysis. As with every method of data collection, any ethical implications of collecting documents should be considered (7,12).

4. Collection of narrative.
   A story, told by a research participant, or a conversation between two or more people, can be used as data for qualitative research. Data collected should be entirely naturally occurring, not shaped as in a semi structured interview or focus group (5,12).

   Narrative data can however be collected in the course of a form of interview. The "narrative interview" begins with a "generative narrative question" which invites the interviewee to relate his/her account of his/her life history or a part of it. This could be an account of living with a chronic illness or with a child with special needs or as a carer for an elderly relative. During the first part of the interview, the interviewee should listen actively but should not interject with further questioning.

   When the narrator indicates that the narrative is completed, a questioning phase where the interviewer elicits further information on fragments which have been introduced follows. This may be followed by a balancing phase where first "how" and then "why" questions are asked in order to gain further explanation of aspects of the narrative (5).
5. Open ended questions in questionnaires.
   Open ended questions, responses to which are to be analysed qualitatively, may be included in questionnaires, even though the majority of the questionnaire will generate quantitative data. The open ended questions usually require that responses, which reflect the opinions of the respondents, be written in blank spaces. This form of data may give useful guidance to a researcher planning an interview or focus group study. The outcome by itself may be a source of frustration as there is no opportunity to ask for clarification of any point made (9).

   Open ended questions are questions that can not be answered only with yes or no (e.g. "How do you feel today?").
6. Focus groups.
   A focus group is a special type of group in terms of purpose, size, composition, and procedures.

The purpose of a focus group is to listen and gather information, and gain relevant information about opinion within the focus group. It is a way to better understand how people feel or think about an issue, product or service.

Participants are selected because they have certain characteristic in common that relate to the topic of the focus group.

The researcher creates a permissive environment in the focus group that encourages participants to share perceptions and points of view, without pressuring participants to vote or reach consensus (10,13).

Focus groups are the most frequent research methods in qualitative analysis.

Focus groups

Background of focus groups

A deficiency of mails and telephone surveys, and even face to face interviews, is that those methods assume that individuals really do know how they feel. Evidence from focus group interviews suggests that people do influence each other with their comments and in the course of discussion the opinions of an individual might shift (10,14). This is due to the fact, that humans are a product of the environment and are influenced by people around.

Focus groups can be used in four different ways in relation to quantitative methods:

- can precede quantitative procedures. When used in this way, the focus group interview can help the researcher learn the vocabulary and discover the thinking pattern of the target audience;
- can be used at the same time as quantitative procedures. At times the researcher may wish to use triangulation; two or more different research methods to address the same issue to confirm findings and to obtain both breadth and depth of information;
- can follow quantitative procedures. Questionnaires typically yield a sizeable amount of data, and focused interviews can provide insights about the meaning and interpretation of the results;
- can be used alone, independent of other procedures. They can be helpful when insights, perceptions, and explanations are more important than actual numbers (2,15).

Advantages of focus groups

People are social creatures who interact with others. They are influenced by the comments of others and make decisions after listening to the advice and counsel of people around them. Focus groups use these facts, and others, as an advantage (2,8,15):

- they place people in natural, real life situations as opposed to the controlled experimental situations typical of quantitative studies;

- the format allows the moderator to probe. This flexibility to explore unanticipated issues is not possible within the more structured questioning sequences typical of mail-out surveys;
- focus group discussions have high face validity;
- the technique is easily understood and the results seem believable to those using the information;
- results are not presented in complicated statistic charts but rather in lay terminology;
- they have relatively low cost and results are gained fast.

Disadvantages of focus groups

Focus groups have also disadvantages (8,10,15):

- Moderator has less control in the group interview as compared to the individual interview.
- data are more difficult to analyze.
- group interaction provides a social environment, and comments must be interpreted within that context.
- the technique requires carefully trained interviewers. At times, an untrained moderator can achieve remarkable results, but it is better to influence the odds by using skilled interviewers.
- groups can vary considerably. Each focus group tends to have unique characteristics. One group can be lethargic, boring, and dull; the next selected in identical manner might be exciting, energetic and invigorating.
- groups are difficult to assemble.
- the focus groups require that people take time to come to a designated place at prescribed time to share their perceptions with others.

Realization of focus groups

The focus group is a special type of group in terms of purpose, size, composition and procedures. Typically is composed of 7 to 10 participants who are selected because they have certain characteristics in common that relate to the topic of the focus group. Furthermore, the focus group is repeated several times with different people (13).

Focus group study will consist of a minimum of three focus groups but could involve as many as several dozen groups.

The process of conducting a focus group study consists of three phases: planning the study, conducting the interviews, analyzing and reporting.

Planning begins with purpose of the study and is followed by organizing those thoughts in a local, sequential manner. Precise definition of the clientele may be essential to undertake the study. People who have the characteristics of your target audience have to be found. In next step location must be determinate; it can be restaurants, hotel, private homes, and public buildings. The place must be easy to find, free from outside distractions. Room should have chairs that can be arranged with participants facing each other. Eye contact among all participants equally spaced around a table is preferred. Plan of action should also take a place (10,14)

Questions are the heart of the focus group interview (2). They may appear spontaneous, but they have been carefully selected and phrased in advance to elicit maximum amount of information. The focus group goes through several different types of question like, opening, introductory, transition, key and ending question, each playing crucial role in analysis. Nevertheless all the questions must be open-ended, because they allow the respondent the opportunity to structure an answer in any of several dimension. In next step pilot testing of focus group interview is recommended, where all the rules already exist.

The interview is recorded on voice recorder and then type out with directly citation. Analysis can be transcript, tape, note and memory base. The participants' statements are then commented (vertical analysis) and compared with other statements and other focus groups (horizontal analysis). Comments are base for report and on the last stage for theory about analysed topic (16). At the end analysis of thematic categories can be done (Figure 1) where themes and subthemes that are common to all groups made in analysis are represent in scheme.

Steps in focus groups

Steps in focus groups method are as follows (17):

1. Planning:
    - participants that are helping to prepare focus groups,
    - intention and goals,
    - numbers of focus groups,
    - formation of questions,
    - who will participate,
    - time table,
    - location, date and time.
2. Realisation of focus groups:
    - arrangement,
    - realisation.
3. Analysis.
4. Focus group report.

CASE STUDY: FOCUS GROUPS ANALYSIS OF PUBLIC OPINION ABOUT GENETICALLY MODIFIED ORGANISMS IN FOOD IN SLOVENIA

Introduction

Focus groups method was used in Slovenia in analysis of public opinion about genetically modified organisms) in food among youth. The aim of research was to find out the view and perception upon GMO, ethical, moral and health concerns about the theme (18,19).

Methods

Analysis was based on qualitative analysis of 14 focus groups. The research took a place form December 2006 to March 2007. Groups were classified on level of education accurately:

- 4 groups from primary school,
- 3 secondary school groups,
- 3 student groups,
- 3 groups included dormitory students and also
- pilot focus group.

Each group included 6 respondents.

The structure of focus groups is based on principle from younger to older or from less educated to more educated. This structure was defined base on results of pilot focus group. Because of anonymity, citations in focus groups have abbreviation:

- FG – focus group,
- PS-primary school,
- SS-secondary school,
- ST-students,
- DS-dormitory students and
- P –pilot group.

The underline number determinate repetition of FG in every single age group (e.g. FG/PS_4:

"I have never heard about GMO." **Quote 1.**

In Table 2, an example of focus group analysis is presented.

Table 2. Example of focus group analysis.

PERSON	STATEMENT/QUESTION	COMMENT
Moderator	What is your opinion about genetically modified organism in food?	
Person "a"	I'm against GMO, I would never buy it and I'm for prohibition.	negative relation to GMO
Person "b"	And I'm pro, since the world population is growing, there wouldn't be enough food for all. The GMO could save the problem	Positive relation to GMO, because GMO can solve the problem of hunger.
Moderator	What is your opinion who is the most reliable person to give information about GMO?	
Person "c"	I thrust only scientists and professors, but I'm sceptic to media	Scientist and professors, but never media

Results

The results are as follows:

1. Pilot focus groups (19 years).
   Results of pilot focus groups are showing unfavourableness to GMO theme, respondents are even saying that they are force to buying GM food FG/P:

   "Yes, you have to, what can we eat today? Do we have any chance?" **Quote 2.**

   In discussion some uninformed situations can be seen FG/P:

   "All those people that are eating in Mc'Donald's, their food is not O.K." **Quote 3.**

   During the discussion different views on GMO information can be seen. On one hand, respondents would like to have more information, on other they are afraid that would lead to psychologically resistance for this kind of food FG/P*:*

   " If you know what are you eating, you are choosing that and that, ... on the end you don't know what to eat." **Quote 4.**

   Findings are showing lack of information among respondents

   "All those numbers are problem. I don't know what the point of this numbers is? Nobody knows what is E480. That's the problem." **Quote 5.**

2. Focus groups in primary school (14 years).
   The first thoughts about GMO in elementary school among pupils were mutations, changing food, tins, artificial food, capsules, food that can not rot, etc. In real situations, they would never buy a GMO product FG/PS_3:

   "*I would take the normal one."* **Quote 6.**

   Respondents are more or less in favour for immediate prohibition of GMO in food FG/PS_1:

   "I'm pro prohibition." **Quote 7.**

   Most of respondents do not know what the content of conversation is FG/PS_4:

   "I have never heard about GMO." **Quote 8.**

Therefore equating GMO with tins and pesticides, etc. is understandable. Pupils in primary school think more radically and are generally against GMO.

3. Focus groups in secondary school (18 years).
   Changes on genes, mutations, improvement of something, Russian apples, butterflies, Hiroshima, Chernobyl, tomatoes, carrots were the first thoughts about GMO in secondary school. The respondents believe that producing GMO food is morally and ethically unacceptable FG/SS_1:

   "Since we don't know GMO are harmful, everything is o.k." **Quote 9.**

   But there are also some cases, when manufacture is susceptible FG/SS_2:

   "On the one hand it is fine, that in countries with famine GMO are existing, although that kind of food is less qualitative." **Quote 10.**

   Pupils in secondary school do not think so radically, they are saying that production of GM food must be supervised and researched should be stimulated. FG/PS_1:

   "First some researches must be done, limits must be defined to control products on market." **Quote 11.**

   Respondents are saying that they are not for immediately prohibition because this would have impact on world market. FG/SS_2:

   "I'm not for prohibition, big changes would happen, economy would changed, every thing would be worse." **Quote 12.**

4. Focus groups among students (23 years).
   All the groups were talking about something negative and harmful to health. FG/ST_2:

   "I believe this is something bad." **Quote 13.**

   When asked for something specific they spoke of soya, tomato and maize. They were asked if they would ever buy GM food? At first there is rejection for the sake of health FG/ST_1:

   "I would never buy it, one hundred percentage." **Quote 14.**

   Asking if the price can influence on choice of purchase, they said yes FG/ST_1:

"*Yes, if GM food is much cheaper.*" **Quote 15.**

Psychological resistance can be seen in conversation FG/ST_2:

"This interferes with nature, everybody fears that." **Quote 16.**

Student's opinion is not radically, they are not for immediately prohibition FG/ST_2:

"No, I'm not for prohibition, because at first we have to know what the fears are, what can go wrong." **Quote 17.**

Obviously GMO is theme where objectiveness and relevance plays vital role in confidence, therefore professors and scientists are trustworthy FG/ST_2*:*

"I trust professors and scientists, but never media." **Quote 18.**

5. Focus groups among dormitory students (16 years).
   First thoughts about GMO among dormitory students were artificial food, additives, chemicals, something unnatural, Chernobyl. The price and publicity would have large influence on purchase FG/DS_1:

"I would take the cheapest one, beauty is selling today." **Quote 19.**

None of the group is strictly for prohibition FG/DS_2:

"I'm not for prohibition, if somebody wants to eat that ..." **Quote 20.**

They also expressed that the information has too many technical terms and that those that are presented in more simple terms are trustworthy FG/DS_3:

"I trust those that are talking simple." **Quote 21.**

Interpretation of the results

Focus groups were realised on described methodology. Public opinion toward GMO among younger's in Slovenia was determined during focus groups. The critical approach is reflected in categories and subcategories presented in Figure 1.

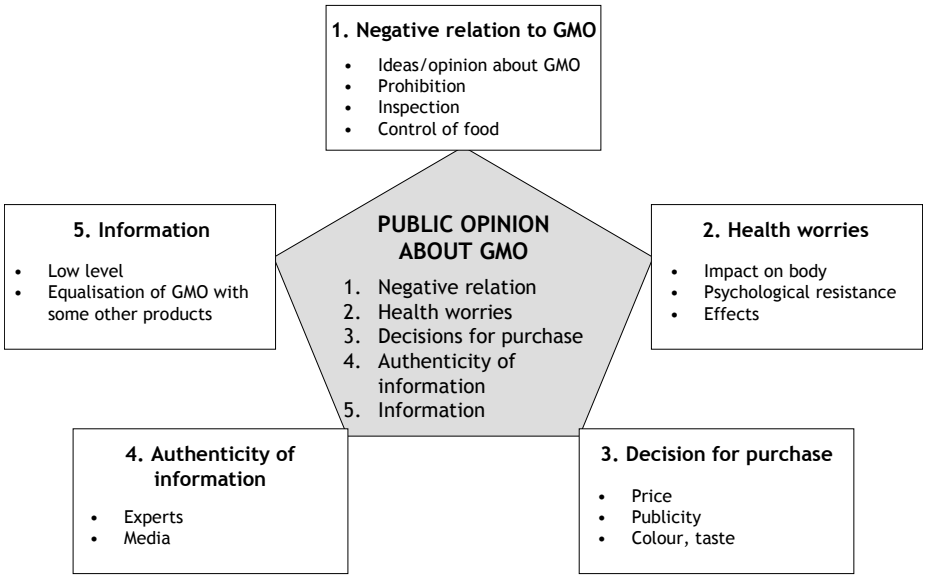


Figure 1. Five thematic categories and subcategories received during thematic analyses.

1. Negative Relation to GMO.
   Generally GMO opinion among respondents is negative. This can be seen as idea of possibility for immediately prohibition, negative relation to inspection workers and supervise of food. Resembling results can be found in research The societal aspect of food biotechnology: relative lack of trust and confidence in the effectiveness of EU and national regulation and institutions can be seen (20). Some negative thoughts appeared right at beginning of discussion. FG/PS_4:

   "I'm thinking about capsules." **Quote 22.**

   FG/SS_2:

   "Chernobyl" **Quote 23.**

   "Hiroshima" **Quote 24.**

   FG/SS_3*:*

   "Something that is add, colouring matter." **Quote 25.**

FG/DD_1:

"*All the E-s and chemicals, everything that's not O.K. for health"*. **Quote 26.**

Similar thoughts are expressed by South African respondents in research Consumer Perspectives on genetically modified food products containing genetically modified material in South Africa. They described GM food as (21):

"It is food in cans." **Quote 27.**

"Big vegetables and fruit. Beef that has been treated with hormones." **Quote 28.**

Psychologically resistance can also be seen in conversation (22):

"No, it is not right ...it is interfering with God's creation." **Quote 29.**

Citizen of Europe are saying that GM food should not be stimulated, that technology is useless, morally and ethically unacceptable and risky for society (23).

2. Health worries.
Respondents believe that GM food is harmful for health. GMO effects are positive and negative FG/SS_1:

"Everything has negative and positive effects, there is no different in case of GMO." **Quote 30.**

FG/ST_3:

" *I believe there are negative impacts on heath, but nobody knows*." **Quote 31.**

Some correlations form South Africa research can be seen:
FG/ST_3:

"It could have no short term disadvantages, but we are not aware of the long term implications on our health." **Quote 32.**

Uncertainty and lack of knowledge about this whole issue is a concern (21). 59% of Americans are concerned about impact of GMO on family's and relative's health (23).

During the conversation, some psychologically resistance can be seen FG/PS_3:

"*This is the food that should be better, bigger, but I don't know if this is true.*" **Quote 33.**

FG/SS_3:

"But you are eating that, they are adding things that last for ever, this is not normal." **Quote 34.**

FG/ST_2:

"*This is interfere in nature , everybody scars that."* **Quote 35.**

Respondents in South Africa express the same worries (21):

"No, it is not right ...it is interfering with God's creation" **Quote 36.**

Results of research Slovenia-a GMO free zone are showing that concerns for health are present in Slovenian media space (24).

3. Decision for purchase.

Factors like price, publicity, colour, taste and health are playing a crucial role in decision of buying GM food. On some individuals in research this factors have major influence, on other hand there is small group, that is not so under of influence FG/P_2:

"*I would buy the one, that is not GMO*" **Quote 37.**

FG/SS_3:

"At first I look up the price, and then make a comparison, If non GMO product is just a little bit expensive, I take non GMO, but if there is major difference, I buy the cheapest one." **Quote 38.**

FG/ST_1:

"*Yes, if GMO product is cheaper, I would take the one, if the difference is about 30%.*" **Quote 39.**

Consumers in South Africa have the same opinion:

"At first I look up the price, and then make a comparison, If non GMO product is just a little bit expensive, I take non GMO, but if there is major difference, I buy the cheapest one." **Quote 40.**

FG/ST_1 (21):

"If they taste exactly the same , the one is bigger or cheaper, then naturally I will buy the new one. It will not do anything to you." **Quote 41.**

4. Authenticity of information.
   Results are showing that scientists and professors have the most reliable information about GMO theme and that they would newer thrust a media. GMO is obviously delicate discussion, were the objectiveness should be taken FG/PS_3:

   "I would thrust only the scientist." **Quote 42.**

   FG/DS_3:

   "I believe only to expert." **Quote 43.**

   Meanwhile American research Review Of public Opinion Research in showing that respondents trust at first to family and friends (57%), experts and professors (32%) at on third place, and media is on last place (9%) (23).
5. Information.
   Respondents have lack of information; they are equating GMO with additives, pesticides and fast food. Similar results can be seen in Telephone opinion poll research: 14,5% have never heard about GM plant (25). Some of them are asking if there is any regulation about GMO FG/ST_2:

   "Is there any regulation?" **Quote 44.**

   FG/PS:

   "Pesticides are forbidden near water source." **Quote 45.**

   FG/P:

   " I believe that in Mc'Donald is not the best food." **Quote 46.**

   FG/DS_1:

   "*All the E's, chemicals, everything that is not good for health*" **Quote 47.**

   Lack of information can be seen not only among Slovenian respondents, but also in South Africa (21):

" We don't have enough information" **Quote 48.**

Slovenian consumers are not "per se" opposed to the use of gene technology in agricultural and food production. Their willingness (or lack there of) to accept genetically modified foods is driven primarily by their perception of risk, benefit, and safety of this technology (26).

Conclusions

Some patterns of repeating what are the first toughs about GMO can be seen. Similar samples can also be found in South African research Consumer Perspectives on genetically modified food products containing genetically modified material in South Africa. Pupils in primary school are talking about GM food with full of imagination, they are quoting exotics examples like capsules, cans, artificial food, food that can not root. Students in secondary school are making correlations with some catastrophes in history like Hiroshima, Chernobyl. Student would by GM product if it would be cheaper than conventional one. Also in this context GMO opinion in Slovenia and South Africa can be compared.

When people think and act like consumers GM is relatively insignificant consideration and negative attitudes can often be passed over in favour of lower prices or other consumer benefits (26).

In study can be seen that level of information is increasing as regards to education. Pupils in elementary school have understandable less information than students. Nevertheless equating GMO with some other products is decreasing regards to level of education. It seem that more information have individual, easier is decided pro or contra GMO. Pupils in primary school are more or less against GMO, but they do not argue the reasons meanwhile students and students in secondary school does. Opinion in South Africa and Slovenia is similar in both countries generally negative relation to GMO theme can be seen. Reasons can be found in low level of information. Consumers that have insufficient information about certain item are in principle confused and they express this via opposition like showed in the study with health worries and negative relation to GMO.

EXERCISE

Task 1

Carefully read the theoretical background of this module and discuss with other students about qualitative analysis and focus groups.

Task 2

The lecturer selects one topic that is close to all students. Afterwards he/she gives the students detailed instructions regarding the method of focus group about discussed topic. Students are divided in groups of six to eight persons. Four to six students

represent public, meanwhile one is moderator and second one is keeper of the minutes. When conversation is typed out at first make comparison with statements of all participants (vertical analysis) and then compare with results of focus groups of other students in class (horizontal analysis).

Task 3

Interpret your results with regard to analysed thematic categories following the example in Figure 1.

REFERENCES

1. Holloway W, Jefferson, T. Doing qualitative research differently: free association, narrative and the interview method. London: Sage Publication; 2003.
2. Richard AK, Casey MA. Focus groups. A practical guide for applied research. Second edition. London: Sage Publication; 2000.
3. Straus AL. Qualitative analysis for social scientist. Cambridge: University Press; 2003.
4. Keith FP. Introduction to social research. Quantitative and qualitative approaches. Second edition. London: Sage Publication; 2005.
5. Hancock B, Windridge K, Ockleford E. An Introduction to qualitative research. Nottingham: Trent Institute for Health Services Research; 2007.
6. Pope C, Mays N. Qualitative research in health care. Third edition. Malden, MA: Blackwell Publishing; 2006.
7. Mason J. Qualitative researching. Second edition. London: Sage Publication; 2002.
8. Pyzdek T. The six sigma handbook. Revised and expanded. New York: McGraw Hill Professional; 2003.
9. Mesec B. The introduction into qualitative reasurch in social work (in Slovene). Ljubljana: Faculty of Social Work; 1998.
10. Richard AK, Casey MA. Focus groups. A practical guide for applied research. Third edition. London: Sage Publication; 2000.
11. Flick U. An Introduction to Qualitative research. Second edition. London: Sage Publication; 2002.
12. Lacey A, Luff D. Qualitative data analysis. Nottingham: Trent Institute for Health Services Research; 2007.
13. Puchta C, Potter J. Focus group practice. London: Sage Publication; 2004
14. Fern EF. Advanced focus group research. London: Sage Publication; 2001.
15. Stewart DW, Shamdasani PN, Rook DW. Focus groups theory and practice. Second edition. London: Sage Publication; 2007.
16. Babbie ER. The basis of social research. Belmont, CA: Tomson Wadsworth; 2005.
17. Klemenčič S, Hlebec V. Focus groups as a method of judgement and development of quality of educational process (in Slovene). Ljubljana: Slovenian Institute for Adult Education; 2007.
18. Fink R, Jevšnik M, Raspor P. Public opinion about genetically modified organisms in food (in Slovene). Ljubljana: College of Health Studies, Department of Sanitary Engineering; 2007.
19. Fink R, Jevšnik M, Raspor P. Opinion poll about genetically modified organisms in food. London: Nutrition Society; 2007.
20. Barling D, Vriend H, Cornelese JA, Ekstrand B, Hecker EF, Howlett J, Jensen JJ, Lang T, Mayer S, Staer KB, Top R. The societal aspect of biotechnology: a European view. Environ Toxicol Farmalcol. 1999;7:93-85.

21. Kempen EL, Scholtz SC, Jerling JC. Consumer perspectives on genetically modified foods and food products containing genetically modified material in South Africa. Potchefstroom: University for Christian Higher Education; 2003.
22. Gaskell G, Allandsdottir A, Allum N, et al. Europeans and biotechnology in 2005: Patterns and trends. Eurobarometer 64.3. A report to the European Commission's Directorate – General for research. Luxembourg: European Commission, Public Opinion Analysis Sector; 2006.
23. The Mellman group. Memorandum on review of public opinion research. Washington, DC: The Mellman group; 2006. Available from URL: http://www.pewtrusts.org/uploadedFiles/wwwpewtrustsorg/Public_Opinion/Food_and_Biotechnology/2006summary.pdf. Accessed: April 28, 2008.
24. Kruszewska I. Slovenia GMO free zone? The only way to protect biodiversity and ensure organic agriculture can thrive. Ljubljana: Umanotera; 2001. Available from URL: http://www.umanotera.si/index.php?node=8&p=2&id=17. Accessed: April 28, 2008.
25. Umanotera. Public opinion about genetically modified organisms (in Slovene). Ljubljana: Umanotera; 2002.
26. Plahuta P, Tivadar B, Raspor P. Slovenian public opinion regarding genetically modified organisms in winemaking. Acta Alimentaria. 2007;36:61-73.

RECOMMENDED READINGS

1. Flick U. An Introduction to qualitative research. Second edition. London: Sage Publication; 2002.
2. Richard K, Casey MA. Focus groups. A practical guide for applied research. Third edition. London: Sage Publication; 2000.
3. Klemenčič S, Hlebec V. Focus groups as a method of judgement and development of quality of educational process (in Slovene). Ljubljana: Slovenian Institute for Adult Education; 2007.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**DELPHI ANALYSIS**
Module: 1.5.3	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Neda Milevska-Kostova**, MSc, MCPPM Centre for Regional Policy Research and Cooperation "Studiorum", Skopje, R. Macedonia **William N. Dunn**, PhD, Professor Graduate School of Public and International Affairs (GSPIA), University of Pittsburgh, PA, USA
Address for correspondence	**Neda Milevska-Kostova** Centre for Regional Policy Research and Cooperation "Studiorum" Nikola Parapunov BB, kompleks Makoteks 1 kat Skopje, R. Macedonia e-mail: nmilevska@studiorum.org.mk
Keywords	Delphi technique, Delphi study
Learning objectives	After completing this module students and public health professionals should be able to: • distinguish projections and predictions; • contrast potential, plausible, and normative futures; • evaluate judgmental forecasting methods; • use policy Delphi analysis to make point and interval forecasts • create a Delphi forecast for an issue.
Abstract	Policy Delphi analysis (PDA) is one of the most important and widely used methods of using expert judgment to make forecasts. The module gives overview to the three forms of forecasting used for Delphi analysis, and explores the techniques used in these types of forecasts. Based on the described assessment of the strengths and limitations of PDA, the module also describes the outline of the use of a 2-round Delphi exercise in class. The purpose is to forecast the probable outcomes in the pre-defined health policy case.
Teaching methods	An introductory lecture gives the students insight into characteristics of Delphi technique. The theoretical knowledge is illustrated with a pre-developed case study. Students are required to read some recommended readings before the introductory lectures. After the introductory lectures, they are given the assignment to prepare a Delphi forecast design, based on the theoretical background, which they discuss with other students, especially the designing and planning phase and expected outcomes of the PDA. In continuation, they need to create a virtual PDA, using published materials on the selected health issue and from own knowledge of the problem and present their findings to other students in the class.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 40%/60% ; • facilities: a lecture room, a computer room; • equipment: LCD projection, whiteboard (flipchart), access to the Internet and bibliographic data-bases; • training materials: recommended readings; • target audience: master degree students according to the Bologna process.
Assessment of students	Multiple choice questionnaire (MCQ) and case problem presentations.

DELPHI ANALYSIS

Neda Milevska-Kostova, William N. Dunn

THEORETICAL BACKGROUND

History and objectives of Delphi Method

A relatively old but still fairly often used method for objective and reliable exploration of new ideas through a consultation of independent experts in a creative way of forecasting to furnish arguments for decision making - is a short compiled definition of the Delphi method, which was developed in the 1940s by the RAND Corporation, as a tool that evolved in the process of experimentation in the technology forecasting studies. In 1944, General Arnold asked Theodor von Karman to prepare a forecast of future technological capabilities that might be of interest to the military (1). Later, in 1959 Helmer and fellow RAND researcher Rescher published a paper on "The Epistemology of the Inexact Sciences," which provided the philosophical base for forecasting (2). The paper argued that in fields in which the instruments are not yet developed to the point of scientific laws, the testimony of experts is permissible and should be acceptable. Thus the Delphi method was developed, recognizing human judgement as legitimate and useful input in generating forecasts - solving the problem of how to use human judgement and, specifically, how to combine the testimony of a number of experts into a single useful statement (3).

Objectives of the Delphi method

The Delphi method is a judgemental forecasting technique for obtaining, exchanging and developing informed opinion through a consensus about the most probable future by iteration.

Characteristics of the Delphi method

The Delphi method is an exercise in-group communication among a panel of geographically dispersed experts (4). The technique allows experts to systematically deal with a complex problem or task. In its essence, the Delphi technique is rather straightforward, with set rules and principles of its performance; it comprises of sets/series of questionnaires sent to the individual and independent experts, which have been pre-selected for the task. According to Fowles (2) anonymity, controlled feedback, and statistical response characterize Delphi. The group interaction in Delphi is anonymous, in the sense that comments, forecasts are not linked to their generator but are presented to the group in such a way as to suppress any identification (3).

Principles of the Delphi Method

Anonymity

Anonymity is the first and by far the most important principle of the Delphi. This approach allows for equality among the participating experts, avoiding the positions of authority or dominance of one's opinion or judgement. The panel of experts is selected based on their knowledge or opinion on the issue.

The questionnaires are designed is such a way to provoke and develop individual responses to the problems posed and to enable the experts to refine their views within the progress of the group work towards the goal of the task. The main point behind the Delphi method is to overcome the disadvantages of conventional committee action (3).

Iteration

Another important principle of the Delphi is the possibility of iteration - repetitiveness of the process as long as the experts, and especially the group coordinator feels that there is a need for refinement of the views and statements. Of course, this does not imply that all experts necessarily have to agree on certain point (see the principle of *stakeholders disagreement* below), but reaching a point at which all participants feel comfortable with the outcome is necessary precondition for considering the Delphi analysis to have undergone is a successful manner.

The rounds of exchanging judgements can be repeated as required, but most often the process is completed after 2 or 3 cycles; with either the *expert consensus* (see below), which is common for the *traditional* Delphi approach, or with *stakeholder disagreement* (see below), which is a newer principle often related to the *policy* Delphi approach.

Controlled feedback

The process is coordinated by a person called *facilitator*, who needs to have both analytical and managerial skills, as he/she is responsible for sending out questionnaires with instructions, collecting them back and summarizing views, as well as preparing a new set of questions that would serve as further distiller of the ideas and opinions, if and when consensus is not reached during the first or subsequent rounds. His/her analytical skills are needed for the process of synthesis of results and working towards building a group consensus.

Statistical group response

In order to represent the full range of opinions and not only the ones reached by consensual agreement of the panel of experts, the summary of individual responses are presented in the form of measures of central tendency (usually the median), dispersion (the interquartile range) and frequency distributions (histograms and frequency polygons) (4,5).

Expert consensus

Traditionally the Delphi method has aimed at a consensus of the most probable future by iteration; the number of cycles to reach the consensus was irrelevant, as long as there is obvious progress in moving towards the anticipated consensus. However, the weakness of this aspect is that not always the consensus can be reached, regardless of the number of repetitions, leading to necessary change in the selected panel of experts, which in turn shows certain imperfection of the method, and weakens the interest of the remaining expert members, as a result of the appearing sense of time being wasted. Thus, in late 1960s alternative Delphi approaches were introduced, explained further below.

Stakeholders' disagreement

The *Policy Delphi* (6) launched by Murray Turoff instead is a decision support method aiming at structuring and discussing the diverse views of the preferred future; the *Policy Delphi*, seeks to generate the strongest possible opposing views on the potential resolutions of a major policy issue. In the author's view, a policy issue is one for which there are no experts, only informed advocates and referees (6).

Policy Delphi begins by using snowball sampling to maximize (rather than minimize) differences among multiple perspectives, using these differences to inform the development of consensus and predictive accuracy in policy forecasting (5, 7, 8, 9). In the face of the policy issue, the expert becomes an advocate for effectiveness or efficiency and must compete with the advocates for concerned interest groups within the society or organization involved with the issue. The Policy Delphi also rests on the premise that the decision maker is not interested in having a group generate his decision; but rather, have an informed group present all the options and supporting evidence for his consideration (9).

Other types of Delphi

Besides the traditional and the policy Delphi, several other modified types have been proposed.

The *Argument Delphi* (10) developed by Osmo Kuusi focuses rather on the ongoing discussion and finding relevant arguments than on the output of the debate itself. The process is based on four level classification of statements, which are usually at least in part mutually exclusive; experts at first make a very simple evaluation on whether they approve or dissaprove the statement/topic, after what they develop arguments pro et contra for acceptance or rejection of given statement.

The *Disaggregative Policy Delphi* (11) developed by Petri Tapio uses cluster analysis as a systematic tool to construct various scenarios of the future in the latest Delphi round. The respondent's view on the probable and the preferable future are dealt with as separate cases.

Forecasting used in health policy

There are several types of forecasting techniques used in the health policy development; they often employ judgemental or statistical methods of forecasting. The statistical methods (extrapolation, multivariate forecasting and econometric forecasting) are used when there is sufficient statistical data; however, when there are insufficient data sources or there is low emphasis on the accuracy of the forecast, experts use judgemental methods, such as the *unstructured* (*unaided*) or *structured*, of which Delphi is one example.

The Delphi technique is used often in the health sector when there is insufficient or unreliable data to conduct a statistical forecast. Projections developed by Delphi panels are believed to be more accurate than forecasts based on unaided judgment. There is limited direct evidence of the accuracy of forecasts using the Delphi method (12).

Steps of Delphi

Conventional (traditional) Delphi

A Delphi method, as defined by its creators (2) has the following 10 steps:

1. Formation of a Delphi team to undertake and to monitor the project.
2. Selection of one or more panels to participate in the exercise. Customarily, the participants are experts in the investigation area.
3. Development of the first round Delphi questionnaire.
4. Testing the questionnaire for proper wording (e.g., ambiguities, vagueness).
5. Transmission of the first questionnaires to the panelists.
6. Analysis of the first round responses.
7. Preparation of the second round questionnaires (and possible testing).
8. Transmission of the second round questionnaires to the panelists.
9. Analysis of the second round responses. (Steps 7 to 9 are reiterated as long as desired or necessary to achieve stability in the results.)
10. Preparation of a report by the analysis team to present the conclusions of the exercise.

However, both the creators and other authors (13) argue that the most important issue in this process is *the understanding of the aim of the Delphi exercise* by all participants. Otherwise the panelists may answer inappropriately or become frustrated and lose interest (3).

On the other hand, the *Policy Delphi* can be conducted in a number of different ways, depending on the context and the skill and ingenuity of the persons using the technique. Since Policy Delphi is a major research undertaking, it involves a large number of technical questions, sampling, questionnaire design, reliability and vailidity, and data analysis and interpretation (5). The steps of the Policy Delphi, which are somewhat modified from the conventional approach, are:

1. issue specification,
2. selection of advocates,
3. questionnaire design,
4. analysis of first-round results,
5. development of subsequent questionnaires,
6. organization of group meetings,
7. preparation of final report.

As the Policy Delphi is more often used for health policy development, its steps are elaborated in the Table 1.

Table 1. Steps of Policy Delphi (6).

Step	Description	Example*
Issue specification	Analyst decides upon the specific issues for which the Delphi will be conducted.	To develop a range of possible national drug-abuse policy options

Table 1. Cont.

Step	Description	Example*
Selection of advocates	Selection of key stakeholders in the issue area, including experts (preferably with opposing or conflicting positions on the issue)	A list of about 100 experts was set, with invitations initially sent to 45 persons, for an expected positive response from 25 persons (in fact, positive responses obtained from 38 experts).
Questionnaire design	Development of questionnaire based on the selected issue; depending on the level of knowledge of the analyst and selected experts, the questionnaire is either done completely by the analyst (using scales for measurement, see example below) or the analyst develops an open-end questions set, that are further distilled by the experts through providing their opinions.	The first questionnaire was pretested, and the entire forecasting section deleted when it was determined that the time to complete the questionnaire was decreased considerably by deleting this section. The questionnaire consisted of 4 sections: development of objectives, transition matrix, policy issue statements and additional items.
Analysis of first-round results	Analyst attempts to determine the initial positions on the forecasts, issues, goals and opinions. Use of summary measures due to expected conflicting assessments of various advocates (avoiding presentation of central tendency only)	24 of 35 respondents (69%) returned the filled-in questionnaire; respondents were asked substantive-issue questions, self-rating of own expertise in drug abuse, expectations from the study, etc. Experts listed 78 policy issues that were distilled down to 55; the 187 key indicators were culled to 153.
Development of sub-sequent questionnaires	The results of prior rounds are used for development of subsequent ones; include summary of arguments for the most conflicting judgements. Thye analyst should count on decreased rate of response in the subsequent rounds, expecially those involved on voluntary basis.	The 2nd questionnaire developed after round One was completed; it included only two sections: National drug-abuse policy objectives and Policy issue statements. There were four issues that exhibited marked differences between policy experts and nonexperts in the importance of issues.
Organization of group meetings	Brinings advocates (stakeholders) in face-to-face discussion of reasons, assumptions and arguments of various positions; useful for immediate feedback.	-

Table 1. Cont.

Step	Description	Example*
Preparation of final report	Analyst is responsible for drafting the final report, based on the questionnaire results and face-to-face discussions; the final report reveals a review of various issues and options available, including a complete description of all conflicting positions.	Final report covering the main conclusions of the questionnaire analysis, as the group meting was not held in this example.

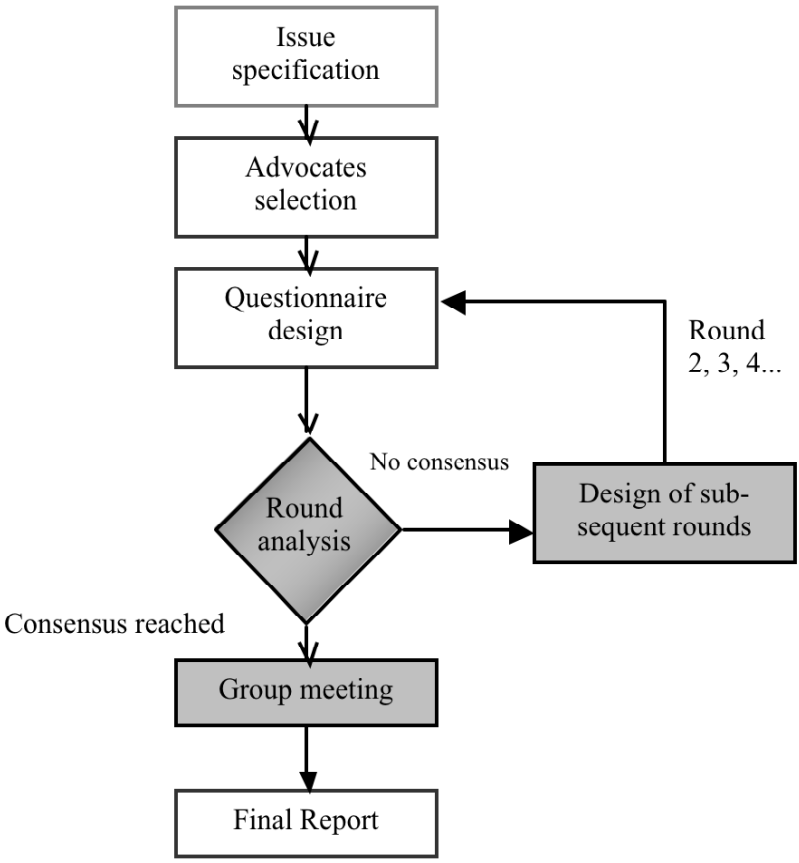


Figure 1. Policy Delphi steps and process

Design of questionnaire

First questionnaire

As explained earlier, the Delphi usually consists of several rounds; of all, only the first questionnaire can be prepared in advance and all other are derived based on the synthesized results of previous rounds.

In the first round, the Delphi process traditionally begins with an open-ended questionnaire. The open-ended questionnaire serves as the cornerstone of soliciting specific information about a content area from the Delphi subjects (14). For example, in the first questionnaire, participants might be asked to provide their judgment (15) on a most probable period by which a drug abuse policy will give visible effects. After receiving subjects' responses, investigators convert the collected information into a well-structured questionnaire. This first-round questionnaire is used as the survey instrument for the second round of data collection. It should be noted that it is both an acceptable and a common modification of the Delphi process format to use a structured questionnaire in Round 1 that is based upon an extensive review of the literature (16). Kerlinger (17) noted that the use of a modified Delphi process is appropriate if basic information concerning the target issue is available and usable.

Table 2. Types of items and scales used in Policy Delphi Questionnaire (5)

Type of item	Item	Scale
Forecast	According to a projection of researchers at the National Public Health Institute, over 20% of young people age 15-24 are smoking marijuana, and this percentage will be doubled in the coming 10 years. How certain are you that this projection is ***reliable***?	[1] Certainly reliable [2] Reliable [3] Risky [4] Unreliable [0] No judgement
Issue	Personal use of marijuana should/should not be legalized. How ***important*** is this issue relative to others?	[1] Very important [2] Important [3] Slightly important [4] Unimportant [0] No judgement
Goal	One goal of National drug abuse policy is to increase the awareness of difference between drug use (responsible) and drug abuse (irresponsible). How ***desirable*** is this objective?	[1] Very desirable [2] Desirable [3] Undesirable [4] Very undesirable [0] No judgement
Options	Drug abuse education is reported to contribute towards reduction of potential users. How ***feasible*** ins this policy option?	[1] Definitely feasible [2] Feasible [3] Possibly feasible [4] Definitely unfeasible [0] No judgement

Results analysis and providing feedback

Analysis of the results

When questionnaires are collected, the analyst or the facilitator is the one that summarizes the responses, synthetizes various positions and puts forward the arguments already pointed out by the advocates. Each sub-sequent round is structured and prepared based on the results of the previous one; however, the results are not presented to the panel of advocates/experts at each round in aform of report, but rather in a form of a new questionnaire (see above *Design of questionnaire*).

Data analysis can involve both qualitative and quantitative data; usually qualitative data is dealt with if conventional Delphi studies, which use open-ended questions to solicit subjects' opinions, are conducted in the initial iteration. Subsequent iterations are to identify and hopefully achieve the desired level of consensus among panelists.

The statistics used in Delphi studies are most commonly measures of central tendency (means, median, and mode) and level of dispersion (standard deviation and inter-quartile range) in order to present information concerning the collective judgments of respondents (18). Generally, the uses of median and mode are favored (16).

Providing feedback

As mentioned above, the results analysis from each round serves as feed into the next round of challenging experts'/advocates' positions and opinions on the issue; once the issue is exhausted, on there is seem no further argumentative development, the facilitator/analyst approaches to writing the final report. The final report is aimed primarily at informing of all opposing positions and alternatives presented by the advocates/experts, supported by arguments as presented by the panel. Yet, as this may become a long list of different standpoints that might confuse the policy- or decision-maker, other presentations are used, taking into consideration the possibility to present all opinions, including the outlayers. Some are mentioned below:

- difference graph,
- histogram or bar chart,
- box-and-whisker plot,
- lists of measures of central tendency and dispersion, etc.

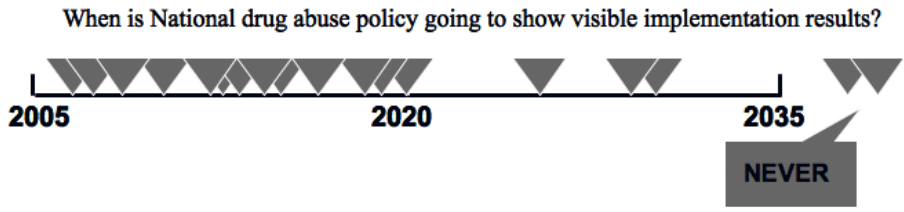


Figure 2. Providing feedback - Difference graph (5).

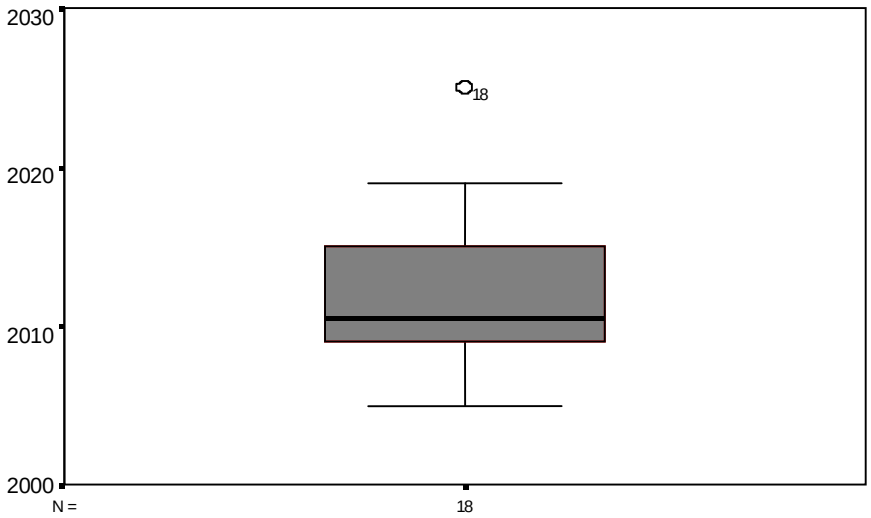


Figure 3. Providing feedback – box-whisker plot.

On-line tools for Delphi analysis

Available (free) software

The website of Forecasting Principles is offering on-line usable application free of charge for trying up the basic features of the Delphi method. Following the link http://armstrong.wharton.upenn.edu/delphi2/ you can register as administrator, start your own survey, in which the application will help you:

- select experts,
- develop questions and scales,
- obtain responses from the experts,
- summarize a report after each round.

This tool, developed by J. Scott Armstrong (19) is a very useful starter's kit for application of Delphi method into your own research.

Risks and disadvantages of Delphi analysis

Risks associated with Delphi

While offering a large possibilities for forecasting by use of experts opinion (Delphi) or advocates positions (Policy Delphi), this method has its own risks, associated mainly with the inseparable personal subjectivism and the partial or complete ignorance of the advocates in the issue. As the outcome of a Delphi

sequence is nothing but opinion, and the results of the sequence are only as valid as the opinions of the experts who made up the panel (20), if not carefully selected, the panel of experts, or advocates for that matter, can lead the discussion into predicition of a highly unlikely future outcome; in some cases this can further strengthen the confidence of the part of the experts or advocates which do not possess great knowledge or informed opinion in the issue.

Disadvantages of Delphi

As with other research methods, Delphi has been reported to have its own disadvantages; most of them associated with improper selection or invalid use of the method for a given type of outcome desired. It is up to the researcher/analyst to determine if this method will produce valid and usable results.

The challenges that the researcher should think of prior to selection of this method as their prefered tool, is related to the internal consistency and reliability of judgements among experts, which - if low or lacking - may lead to low reproduceability of forecasts based on the results elicited; sensitivity of results to ambiguity and respondent reactivity in the questionnaires used for data collection; difficulty in assessing the degree of expertise held by participating experts (21).

Another problem identified by research into the implementation and application of Delphi surveys has been the tendency for experts to over-simplify particular issues, and treat them as isolated events. This is particularly the case in forecasting, where experts tend to think in terms of linear sequential events, rather than applying a holistic view that involves complex chains and associations. Again, it is up to the researcher/analyst to extract this aspect from the answers of experts/advocates through the sub-sequent rounds of questionning. Other techniques for multi-dimensional analysis are also available, such as ‘cross impact matrix forecasting’ intended to compare a range of ‘possible futures’ against each other; but those are discussed elsewhere (4,22,23).

Advantages of Delphi

Although the approach was originally developed to capture expertise in uncertain and emergent domains, it tends to be used in evaluation when significant expertise exists on the subject, for example in the case of programmes that are not innovative. The method is recommended when the questions posed are simple (a programme with few objectives, of a technical nature) and for the purpose of establishing a quantitative estimation of the potential impacts of an isolated intervention (e.g. increase in taxes or in the price of health services). It is also recommended in an ex ante evaluation context if the evaluation concerns public intervention of a technical nature. But, it may also be used to specify relations of causes and potential effects in the case of innovative interventions. It is particularly useful when a very large territory is being dealt with since there are no experts' travel expenses, only communication costs.

It has been found to be particularly useful in programmes related to public health issues (such as, policies for drug use reduction and prevention of HIV/AIDS) and education (1,4). In general, the Delphi method is useful to explore and unpack specific, single-dimension issues. There is less support for its use in complex, multi-dimensional modelling. In these cases, the evidence does suggest that data

gathered by Delphi surveys is a useful input, when supported by data gathered from other sources, to complex scenario-building (24).

Nevertheless, according to some authors, still the context plays great role in deciding whether and when to use the Delphi method.

CASE STUDY: AIDS VACCINATION POLICY: A SCENARIO ANALYSIS USING THE DELPHI METHOD

Background

The Delphi method was used to explore and identify the potential implications associated with the introduction of a first AIDS vaccination in Switzerland (25,26). Thirty participants with an interest in the field contributed anonymously to the study. The study focused on an existing scenario which modelled the characteristics of a first preventive, partially effective, vaccination against AIDS.

The process

The Delphi consultation was carried out in three stages. In the first round, the participants were asked to:

- list the objectives to be achieved in the first five years;
- evaluate the acceptability and feasibility of proposals concerning the development of a public health strategy and the AIDS vaccination;
- estimate the potential use of the vaccination by different groups of users.

The outcome

The used Delphi method in the consultation process produced two main outcomes: firstly, a set of strategies and recommendations for the development of a framework of AIDS prevention campaigns and, secondly, an institutional framework for the setting up of a future AIDS vaccination strategy.

Follow-up

In parallel, in 2003, clinical trials of a new vaccine against HIV, started in Switzerland and the United Kingdom. EuroVacc, the foundation organising the trials, has tested two vaccines: DNA-C, developed by Professor Hans Wolf of the University of Regensburg, Germany, and its booster, NYVAC, developed by the French pharmaceutical company Aventis. About 160 healthy volunteers - half in London and half in the Swiss city of Lausanne, where EuroVacc is based – have been subjected to the test of the vaccine for safety.

In 2005 the combined vaccine consisting of DNA vaccine and NYVAC booster was tested in Switzerland, the Netherlands, the United Kingdom, Spain, Italy, Germany, and Sweden in hundreds of people seen as being at high risk of HIV infection, including gay men, drug users, and commercial sex workers. Volunteers' rate of infection was monitored and compared with the rate of infection in similar groups of people who have not been given the vaccine.

EXERCISE

Task 1

Before class, based on the Table 2 above, write as many as you can "forecast," "issue," "goal," and "options" questions related to the Case study above. These questions will be collected and used for the in-class Delphi exercise.

Task 2

After the in-class discussion of the collected questions, assume the role given by the teacher and fill out the sample questionnaire that will be distributed by the teacher. One of the students will be assigned the role of analyst who will chair the group meeting in the next class. Role-play exercise in the class.

Task 3

Based on the role-play held in the class, prepare a final report (assuming the role of analyst). The teacher will consider this paper as an assessment for the module.

REFERENCES

1. Cornish E. The study of the future: an introduction to understanding the art and science for understanding and shaping tomorrow's world. Washington, DC: World Future Society; 1977.
2. Fowles J. Handbook of futures research. Greenwood Press: Connecticut; 1978.
3. Illinois Institute of Technology. The Delphi method [homepage on the Internet]. Available from URL: http://www.iit.edu/~it/delphi.html. Accessed: April 6, 2009.
4. Adler M, Ziglio E. Gazing into the oracle. Bristol, PA: Jessica Kingsley Publishers; 1996.
5. Dunn WN. Public policy analysis. New Jersey: Pearson Prentice Hall; 2004.
6. Turoff M. The design of a policy Delphi. Technol Forecast Soc Change. 1970;2:2.
7. Mahmoud E. Accuracy in forecasting: a survey. J Forecast. 1984;3:139-59.
8. Rescher N. Predicting the future. Albany, NY: State University of New York Press; 1998.
9. Turoff M. The policy Delphi. In: Linstone H, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. New York: Adisson Wesley; 2002. p.80-98.
10. Kuusi O. Expertise in the future use of generic technologies. Research Reports 59. Helsinki: Government Institute for Economic Research Finland; 1999.
11. Tapio P. Disaggregative policy Delphi: using cluster analysis as a tool for systematic scenario formation. Technol Forecast Soc Change. 2002;70:83-101.
12. Forcastingprinciples.com [homepage on the Internet] Methods for health forecasting, forecasting principles; c1997-2007. Available from URL: http://www.forecastingprinciples.com/Health/methods.html. Accessed: April 6, 2009.
13. Delbecq AL, Van de Ven AH, Gustafson DH. Group techniques for program planning: a guide to nominal group and Delphi processes. Glenview, IL: Scott&Foresman; 1975.
14. Custer RL, Scarcella JA, Stewart, BR. The modified Delphi technique: a rotational modification. J Vocat Techn Educat. 1999; 15:1-10.

15. Gordon TJ. The Delphi method. Future research methodology. Washington, D.C.: AC/UNU Millenium Project; 1994. Available from URL: www.gerenciamento.ufba.br/Downloads/delphi%20(1).pdf. Accessed: April 6, 2009.
16. Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. Pract Assess Res Evaluat. 2007;12:1-8. Available from URL: http://www.pareonline.net/getvn.asp?v=12&n=10. Accessed: April 6, 2009.
17. Kerlinger FN. Foundations of behavioral research. New York: Holt, Rinehart, and Winston, Inc.; 1973.
18. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Advanc Nurs. 2000;32:1008-15.
19. Imerial College London [homepage on the Internet].The development of the Delphi software was funded in part by International Institute of Forecasters. Available from URL: http://www3.imperial.ac.uk. Accessed: April 6, 2009.
20. Martino J. The Precision of Delphi estimates. Technol Forecast Soc Change. 1970;1:293-9.
21. Makridakis S, Wheelwright SC. Forecasting methods for management. 5th ed. Chichester: Wiley; 1989.
22. Gordon TJ, Hayward H. Initial experiments with the cross-impact matrix method of forecasting. Futures. 1968;1:100-16.
23. Gatewood RD, Gatewood EJ. The use of expert data in human resource planning: guidelines from strategic forecasting. Human Resource Planning. 1983;5:83-94.
24. European Commission. Evaluating socio economic development. Sourcebook 2: methods & techniques, Delphi method. Luxembourg: European Commission; 2005. p.110-9.
25. Zuber P. Introducing a first AIDS vaccination in Switzerland: a Delphi policy analysis. Lausanne: Institut de médecine sociale et preventive; 1994.
26. Fleck F. Trials of AIDS vaccine to start in Switzerland and United Kingdom. Br Med J 2003; 326:952.

RECOMMENDED READINGS

1. Dunn WN. Public policy analysis. New Jersey, NY Pearson Prentice Hall; 2004 p.180-98.
2. Jillson IA. The National drug abuste policy Delphi: Progress report and findings to date. In: Linstone HA, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. New York: Addison-Wesley; 2002. p.124-59
3. Turoff M. The Policy Delphi. In: Linstone H, Turoff M, Helmer O. editors. The Delphi method: techniques and applications. New York: Addison-Wesley; 2002 p.80-98.
4. Zuber P. Introducing a first AIDS vaccination in Switzerland: a Delphi policy analysis. Lausanne: Institut de médecine sociale et preventive; 1994.

Chapter 2

SPECIAL EPIDEMIOLOGICAL AND OTHER METHODS

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**PRINCIPLES AND METHODS OF ENVIRONMENTAL EPIDEMIOLOGY: AN OVERVIEW**
Module: 2.1.1	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Ivan Eržen** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Zaloška 4, Ljubljana, Slovenia E-mail: ivan.erzen@gmail.com
Keywords	Environmental health, environmental epidemiology, study design
Learning objectives	After completing this module students should: • know the definition and basic characteristics of environmental epidemiology; • are familiar with aims of environmental epidemiology; • are familiar with basic study designs in environmental epidemiology.
Abstract	Environmental epidemiology represents a very important part of epidemiology branch which is developing very fast. At present a shift from disease to exposure centred environmental epidemiology has being noticed. The main reasons for this shift are identification of new environmental hazards and the need for estimation of individual exposure to environmental hazards. Further on the field of work in environmental epidemiology and the epidemiological principles and methods used are presented.
Teaching methods	An introductory lecture gives the students insight in characteristics of environmental epidemiology and especially different types of study design, used in environmental epidemiology. After introductory lectures students try to find from bibliographic data bases as many studies concerning environmental health problems as possible, but not less than five each. They analyze study designs used and present the results to other students.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	An essay on study design used in a set of articles publishing results of environmental health studies.

PRINCIPLES AND METHODS OF ENVIRONMENTAL EPIDEMIOLOGY: AN OVERVIEW

Ivan Eržen, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Epidemiological methods were developed initially to investigate the distribution and determinants of communicable diseases, but their scope has now been widened to include all aspects of health and wellbeing in relation to biological or non-biological agents.

There is widespread concern that contaminants in our environment may be making us ill. Many well known serious diseases arises as a consequence of exposure to harmful environmental factors. New environmental health risks are reported almost daily. However, the information we receive is often confusing and contradictory due to uncertainties and gaps in scientific knowledge. That is why the answers to the following questions often remain absent. The frequently asked questions are such as:

- are the risks real or imaginary;
- at what level is a contaminant considered unsafe;
- which people are most at risk;
- what are the most cost-efficient ways of protecting the environment and improving human health.

A branch of epidemiology which aims to answer these questions is so called environmental epidemiology.

Definitions of basic terms

Environment

According to A dictionary of epidemiology (1), environment is all that which is external to the individual human host. It can be divided into physical, biological, social, cultural, etc. environment. Any of them can influence health status of the populations

Environmental medicine and environmental health

In the mainstream scientific literature, environmental medicine is the work of clinicians and has been generally defined as the evaluation, management, and study of detectable human disease or adverse health outcomes from exposure to external physical, chemical, and biologic factors in the general environment. Although environmental medicine includes discussion of epidemiology, population health, and nonmedical interventions, the overwhelming focus is still on the clinical assessment and medical management of patients with individual illness due to hazardous exposures Institute of Medicine (2).

Environmental health is defined by the World Health Organisation (WHO) as those aspects of human health and disease that are determined by factors in the

environment. It also refers to the theory and practice of assessing and controlling factors in the environment that can potentially affect health. Environmental health includes both the direct pathological effects of chemicals, radiation and some biological agents, and the effects (often indirect) on health and wellbeing of the broad physical, psychological, social and aesthetic environment which includes housing, urban development, land use and transport (3).

Environmental epidemiology

According to A dictionary of epidemiology (1), environmental epidemiology is the study of health effects on populations of exposure to physical, chemical, and biological agents external to the human body, and of immediate and remote social, economic, and cultural factors (e.g. urbanization, agricultural development, energy production/combustion) related to these physical, chemical, and biological agents).

Agent

According to A Dictionary of epidemiology (1), agent is defined as a factor, such as a microorganism, chemical substance, or form of radiation, whose presence, excessive presence, or (for example in deficiency diseases) relative absence is essential for the occurrence of a disease.

In more broad sense the term "agent" is a neutral one with no intrinsic implication of "beneficial" or "adverse" characteristics. Most agents have the potential for one or other or both of these effects, varying with the precise nature of the agent, the level and duration of exposure, and the state of nutrition and other acquired or inherited characteristics of the subject. Thus, for example, the chemical agents constituting vitamins and their analogues, that may serve as essential food factors, are claimed to offer protection against certain diseases (e.g., vitamin A against carcinogenesis), or to have severe toxic effects, according to dose, to the state of nutrition and acquired characteristics of the subject, and to other agents operating coincidentally (e.g. the same vitamin A could cause toxic effects). When studying these biological, chemical and physical agents, it is necessary to characterize them and to determine their absorption, concentration in air, water, etc. with careful attention to precision. Special attention has to be given to demographic and socio-cultural factors that may affect the degree of exposure or uptake as well as to special host characteristics, including immunological status.

Exposure

According to A Dictionary of epidemiology (1), exposure is defined as among others as:

- proximity and/or contact with a source of a disease agent in such a manner that effective transmission of the agent or harmful effects of the agent may occur,
- the amount of a factor to which a group or individual was exposed; sometimes contrasted with dose, the amount that enters or interacts with the organism, and
- the process by which an agent comes into contact with a person in such a way that person may develop the relevant outcome (e.g. disease).

Exposure is usually meant in the terms of harmful effects, but it could be beneficial as well. An example of such an exposure is exposure to immunization (1).

Dose

According to A Dictionary of epidemiology (1), dose is the amount of a substance available for interaction with metabolic processes or biologically significant receptors after crossing the relevant boundary (epidermis, gut, respiratory tract). The absorbed dose is amount crossing a specific absorption barrier.

Aim of environmental epidemiology

Environmental epidemiology is aiming, by studying populations in different exposure circumstances, at clarifying relationships between exogenous agents and/or factors and health by (1).

Environmental agents/factors

Health of each individual is, in broad sense, determined by two main groups of factors (4), being genetic and environmental:

- genetic factors play an important role in human health and disease. From the parents of an individual come genetic factors (genome), consisting of the DNA in each body cell. Genome sets the main features and boundaries within which life is to be experienced. It also provides the blueprint for how the human body interacts with the environment. The genome does not generally change during the course of one's life. If it does change (in case of mutation), it may lead to cancer or cell death. Some studies have suggested that genome provide a built-in "clock of self-destruction", as the body can only function properly for a limited time. The limit for most individuals is within the range of 70 to 100 years. An individual's genetic material is one of the mayor factors that determine how an individual is affected by environmental exposure. While everybody will have problems if subjected to high enough exposures to an environmental hazard, some people are affected at lover exposure due to existence of pre-existing or concomitant risk factors or conditions, and some people are affected at quite low exposure due to an inherited susceptibility;
- the second group is the group of factors of external environment of human beings. The human environment consist of very basic elements: the air we breath, the water we drink, the food we eat, the climate surrounding our bodies, and the space available for our movements. In additional, we exist in a social and spiritual environment, which is of great importance for our mental and psychical health (5). Unwanted effects (in humans, as well as in animals and plants) may derive from:
  - chemical, physical, and biological agents artificially (and often involuntarily) introduced into air, water, and/or soil in the course of, or as a consequence of, productive and other human activities, i.e. man-made pollution;
  - exposure deriving from personal behaviour, including diet, tobacco smoke, alcohol intake, sexual habits, etc.;

– naturally occurring hazards, like excesses or deficiencies of macro- or microelement in diet, inadequate concentration of elements in air or drinking water, background radioactivity, aflatoxins, etc.

Biological agents

Biological agents include bacteria, viruses, fungi, yeasts, protozoa, and higher animal agents or vectors recognized as contributing to human disease. They have the ability to adversely affect human health in a variety of ways, ranging from relatively mild, allergic reactions to serious medical conditions, even death. These organisms are widespread in the natural environment; they are found in water, soil, plants, and animals. Because many microbes reproduce rapidly and require minimal resources for survival, they are a potential danger in a wide variety of human settings.

Chemical agents

Chemical agents involved in environmental considerations have been characterized as natural and manufactured organic (but not living) and inorganic substances occurring in food, air, water, soil, and other media. While living materials are excluded from this category, their products are widely distributed in the environment, in the form of metabolites, cell bodies, or bio- chemical extracts. Thus, many foodstuffs are infested by, or require for their synthesis, micro-organisms that are also found in the wild and may contaminate the general environment.

Physical agents

Physical agents that impinge on man may occur naturally or be man-made or man-intensified. They include ionizing and non- ionizing radiation, the latter ranging from ultraviolet through visible light and infrared to microwave, radio frequency and extremely low frequency electromagnetic fields. Climatic conditions of temperature and humidity play important direct and indirect roles in environmental health. Noise and vibration at the intensities experienced occupationally are associated with objective evidence of damage; lesser intensities occurring outside occupational environments, apart from affecting amenity, are a source of concern in case they present health hazards.

The effect of environmental agents to human beings

Biological, physical and chemical agents may have various effects on human being. Some of them may not produce any adverse effects, while others, may be liable, if exposures are sufficient, to affect such basic phenomena as growth and development. Sometimes, environmental exposures may affect host susceptibility or resistance, or produce functional or pre-pathological changes. Behaviour may be modified by exposure, especially to physical agents such as noise, light, and heat. A wide range of pathological states in different organs may be induced by exposure to environmental agents, and even death may be caused or hastened by such exposures. There is immense variability in combination of different environmental factors that each person is exposed both in general environment and at workplace and therefore

assessment of exposure and the contribution of individual factor to health impairment, or improvement in some cases, is a very difficult task (6).

Historical perspectives of environmental epidemiology

The dangers of cigarette smoking, exposure to mineral dusts, ionizing radiation, lead and a host of other substances present in the environment have been discovered and quantified largely through epidemiological investigations. The broad working field must be the reason that increasingly, the scientific and policy-making communities are turning to the field of environmental epidemiology for answers.

Environmental epidemiology has developed the tools for identifying and measuring the influence of environmental factors (physical, chemical, and biologic) on human disease in a community. It provides the scientific evidence for sound environmental and health policies. Because epidemiology considers real exposures in real populations under real life conditions, it can be especially valuable in uncovering the causes of human disease. Most diseases are either caused or influenced by environmental factors. Knowing and understanding of the ways in which specific environmental factors can interfere with health is therefore of crucial importance for developing of preventive programmes and measures.

In the past, the term environmental epidemiology has been used to describe the use of conventional epidemiological techniques in order to hypothesize about, study, and interpret associations between disease and environmental agents (as opposed, for instance, to genetic epidemiology). A broader approach is currently envisaged, which is primarily focused on exposure circumstances and which considers as dependent variables all possible health effects of environmental agents to which populations are exposed (7-9).

There are several reasons for the shift from disease - to exposure-centred environmental epidemiology.

5. First, particularly in developed countries, degenerative, chronic diseases (such as cancer, lung emphysema, etc.) have become the prevailing pathology: the aetiology of many of these conditions is multifactorial, i.e. no specific hazard can be considered as a necessary cause. To further complicate the picture, many environmental hazards (e.g. excess dietary fat, asbestos, etc.) are causally associated with more than one disease.
6. Secondly, most environment-induced ill-effects are dose-related. For a given hazard, there may well be exposures either low enough, or of short enough duration, as to be negligible in terms of risk. It has also become obvious that ill-effects are frequently the result of interaction (addition, synergism, antagonism, etc.) between different hazards. For the same exposure to a given hazard, the risk may differ according to which other hazards are present or not.
7. Thirdly, analytical techniques for measuring pollutants in the environment have been used more and more, and their sensitivity has increased by several orders of magnitude. Consequently, there has been a dramatic increase in hazard-specific environmental data requiring risk evaluation.
8. Finally, health authorities, public opinion, and the scientific community have become increasingly concerned by the number of environmental contaminants for which potentially deleterious effects are unknown or poorly understood.

Environmental epidemiology objectives

Environmental epidemiology assesses the added risk (real or potential) to the population exposed to environmental pollutants with the purpose of identifying the sources responsible for the pollution. The identification of previously unrecognized hazards due to exogenous exposures is still central to environmental epidemiology, but it is apparent that other scientific activities have become equally important. Main objectives of environmental epidemiology include (10):

- the identification of previously unrecognized environmental exposures to agents known to be hazardous and, when required, the quantification a posteriori of the ensuing risks (either absolute, relative, or attributable);
- the estimation of individual exposures to environmental hazards and risk assessment;
- the introduction of control and preventive measures and evaluation of the effectiveness of those measures.

Environmental epidemiology scope

The main areas of environmental epidemiology are (10):

- identification of previously unknown environmental exposure to agents that have proved dangerous, and, if necessary, 'a posteriori' evaluation of risks to which they are followed (absolute, relative or attributable);
- assessment of individual exposure to environmental hazards and risk assessment;
- the introduction of preventive measures and assessment of their effectiveness.

An increasingly important area of environmental epidemiology is also proper dissemination of information about environmental risks to the people (risk communication). The new environmental and health risks that are increasingly perceived are reported almost daily. The problem is that information, usually provided by different media, is often confusing and even contradictory due to uncertainties and gaps in expertise.

Characteristics of research in environmental epidemiology

Through research in the field of environmental epidemiology we study the relationship between environmental risks and health. Our purpose is to assess the relationship between the frequency of the disease and the level of exposure to certain substances (11). We use different types of studies at different levels. In doing so, not all types of studies useful in all cases – they are not alternative options, a free choice in any situation. Choosing the type of the study depends primarily on the issues that you are staging (objective research), and the limitations, such as:

- frequency of occurrence of the phenomenon under study,
- the time period within which to obtain at least approximate answers,

- availability of the study population,
- availability of funds and
- ethical requirements.

The art of good research plan is to find a compromise between the ideal and the possible in a way that will lead to the most useful data, and then the results of studies within the given options. Research in environmental epidemiology is accompanied by major problems.

Problems of environmental research

Researchers in the field of environmental health in their work are faced with problems which are almost exclusively own this branch of epidemiology, among which are (12,13,14):

- people are not exposed to only one individual adverse factor of the environment at the time, but to many factors in complex mixtures of harmful substances,
- an additional obstacle is that these substances are very often present in low concentrations and,
- very often, these concentrations are not significantly different between individuals within the observed area, making it necessary to explore the phenomena at the population level,
- very often we do not have data on actual levels of pollutant concentrations in the environment, so they should be evaluated by means of mathematical modeling,
- the values of rates to estimate the power of relationship between two (or more) phenomena (e.g. relative risk) are often low (below 1.5),
- health problems set in relationship with environment pollution, are usually generated as a result of long-term action of harmful substances, often even not by direct action of these agents,
- in addition to local pollution also global pollution may be present (there exist evidence of the transfer of pollutants such as ozone, to the places, which may be far away from the source of contamination, even hundreds of miles away),
- exposure may have been a short term one, and may have occurred far in the past. Some diseases that occur as a result of adverse environmental factors, have a very long preclinical (latent) stage of disease (for example, different types of cancer),
- diseases that are observed as a result of harmful substances from the environment, have, in addition to environmental determinants, present other determinants, which are often even stronger. Controlling for confounders means a major problem.

Special features of environmental epidemiology

In environmental epidemiology one of the main objectives is to identify sources of pollution. Pollution of air, water and earth is connected with the sources of pollution, which have a particular geographic location and which could be active during

specified periods. The two most important features of studies in environmental epidemiology are consequently the geographical and temporal dimensions of studies. Additional special feature is mathematical modeling for the purpose of quantifying exposure (6).

1. Spatial/geographical dimension of studies.
   Mapping the level of environmental exposure is a very useful tool in environmental epidemiology (6). In addition to environmental exposure various aspects of health status of people living at the treated area is mapped as well. This procedure is also known as geographic analysis.

   In this analysis we use as an analytical tool so-called geographical information system (GIS) analysis. GIS is the merging of cartography, statistical analysis, and database technology. It is defined as any information system that integrates, stores, edits, analyzes, shares, and displays geographic information for informing decision making (15).

   For spatial presentation of the phenomena usually two types of maps are used:

- spot map - this type of mapping is usually used for displaying accumulations or outbreaks with a limited number of situations. Dot (point) or the symbol "X" is placed in a spot where the affected person had lived or worked. An example of such a map is a map, in which John Snow depicted a cholera outbreak in London in mid-19th century (16,17) (Figure 1). His study of the distribution of cholera led to the source of the disease, a contaminated water pump located in the Broad Street/Soho.

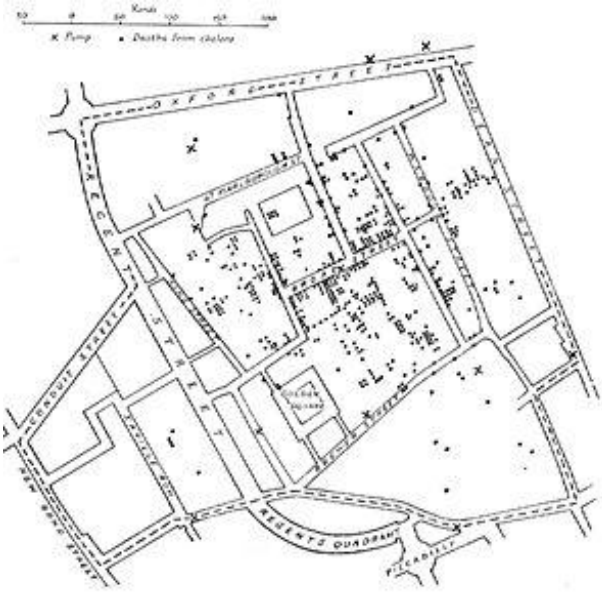


Figure 1. An example of a spot map - E. W. Gilbert's version (1958) of John Snow's 1855 map of the Soho cholera outbreak showing the clusters of cholera cases in the London epidemic of 1854 , published in Wikipedia, the free encyclopedia (17).

- choroplethic map – this surface diagram is a type of thematic maps to show the quantitative phenomena of territorial units in which areas are colored in different shades of one color or in multicolored scale or are differently shaded, and thus reflect the value of the observed phenomenon) (Figure 2). The word "choroplethic" comes from the Greek (choros = city; pleth = value) (18,19). Such maps, for example, are used by World Health Organization in their reports (20) (Figure 2).

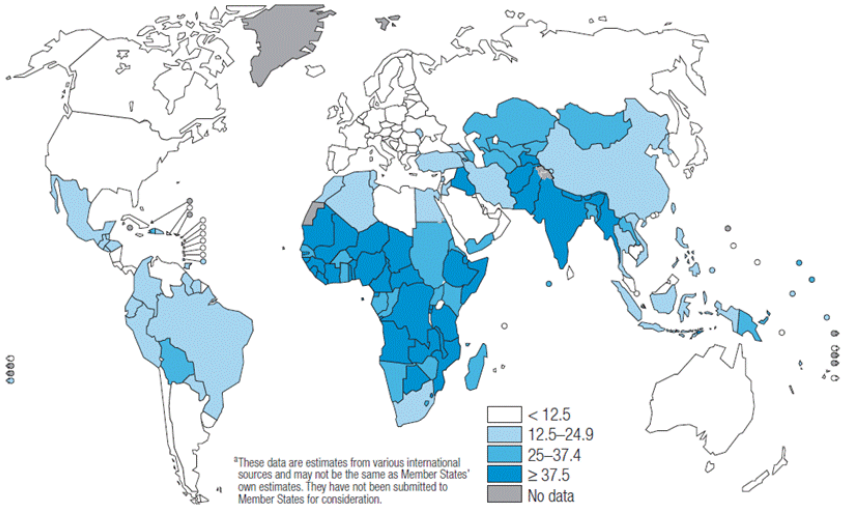


Figure 2. An example of a choroplethic map – Neonatal mortality rate per 1000 live births in 2000, published in The World Health Report 2005 (20)

2. Temporal dimension of studies.
   Description of events in time is often presented in the form of two-dimensional diagrams, which may be different. The most common are diagrams in which on the vertical axis number of epidemiological events or phenomena (eg incidence rate) is applied, while on the horizontal axis the time is applied. The occurrence of cases of the disease in relation to the time could be presented as a histogram or line graph.

   Time unit of observation is selected according to characteristics of the observed phenomena: the phenomena that evolve slowly in time, can be observed in years, while for phenomena that evolve rapidly in time, the unit of time could be even the day (21).
3. Mathematical modeling for quantifying exposure.
   Mapping of levels of environmental exposure is related to problems of availability of relevant data on exposures. Measurements are usually associated with high costs due to high cost of measuring devices. Therefore, these are not placed everywhere, where this would be necessary, or not

permanently installed (mobile measuring device), and very critical situations (eg loss of the filter device at the source of pollution) may therefore not be measured. In the last two decades therefore mathematical modelling of environmental pollution is increasingly used (22-24).

Overview of characteristics of studies in environmental epidemiology

Studies in environmental epidemiology can be, as in epidemiology in general, classified by the number of features that we have described in previous modules of this book. The most important in the context of environmental epidemiology is certainly a classification according to the level of research.

Environmental epidemiology is basically a science that deals with the analysis of the interaction between health events and environmental risk factors for their emergence and development at the population level. However, the problems we identified earlier, and in many others, has focused the research on other levels than population level (13,25,26). Levels of research in environmental epidemiology are 1) population level (macro level), 2) individual level (mezzo level), and molecular level (micro level) (13,25,26).

1. Population level.
   The level of studying the causes of health phenomena at the population level can be characterized as macro-level (26). The population is a dynamic construct, which consists of individual people. These formations behave as more or less single "organism" (depending on the variability of the phenomenon at the individual level). At this level of research, the unit of observation is the population, and the distribution of health among different populations is studied. Studies at this level are called with a single name the "ecological studies".

   Study of health at this level is important, even if accompanied by a number of criticisms relating to the transfer of results to population level to individual level. Studies on the population level arein some respects in medicine neglected, and are regarded as suitable only for making hypotheses (27). One of their biggest problems many years seemed to be the so-called "ecological fallacy". Ecological fallacy is the bias that occurs if the false assumption that the statistical relationship between two variables at the population level is the same as relationship between the relevant variables at the individual level (1,28-30). The fundamental problem of ecological fallacy is that no population group is not completely homogeneous with regard to exposure. If each observation area would be consisted only of people who were exposed or only people who were not exposed, then the ecological fallacy would not exist. However, this problem is no longer a problem, if we stay only at the population level (31).
2. Individual level.
   Notwithstanding the fact that populations are made up of individuals, the determinants of health at the population level differ from the determinants at the individual level. It is therefore necessary to study the phenomena at this level (26). Studies of this level are significantly more numerous than the

studies at the population level. This level is precisely the level at which the risks for exposures were traditionally most commonly assessed.

Focusing primarily on studies at the individual level can lead to various problems. Reliance solely on the results of studies on individual level ignores the population context in which exposures occur. It also ignores the uniform exposure within each population. Standardized comparisons of populations may reveal important factors which individual levels can not (13). Furthermore, it could lead to a point in the focus only on the individual level, where it is likely to ignore the major problems of environmental health. Consequently, even in interventions we focus on the individual level. This is perfect way how to ignore the major problems of environmental health.

3. Molecular level.
In mid-nineties we began to think that this so-called "epidemiology of risk factors" reached the limits of its range, and that it is necessary to introduce a more sensitive and specific method - the method of biomarkers of exposure, sensitivity and health events, in order to improve knowledge about the mechanisms of disease and risks (13). This level is the domain of molecular epidemiology (1,13). It is a discipline that seeks to combine a sophisticated, highly developed and highly sensitive laboratory method by analytical epidemiological methods (32). It is used for example in the study of occupational and environmental exposure to carcinogens mutagenom and (32).

At the beginning this level promised that could help at least partly alleviate the weaknesses of research in environmental epidemiology (33,34). Consequently, molecular markers were increasingly used to assess the exposure of an individual human being. It seemed that the research at this level, primarily through the knowledge of molecular mechanisms of association between exposure and health phenomenon, could increase the likelihood of the observed biological relationship. These methods were supposed also to reduce certain types of bias and interference, and hence potentially to increase the power of environmental research. However, over time, noted weaknesses of this approach. One of the major disadvantages is that the molecular markers are costly and research, in which they apply, and are therefore made on small samples. The power of these studies is therefore actually be lower than in studies of the other two levels (13).

Each level contributes something to the compilation of a comprehensive picture, which can be studied whenever we study relations between health events and environmental factors. In any case, we can not consider one level more important than the other. Unfortunately, in the past this was not the case and in different periods, epidemiologists gave one time greater importance to one, again tho other time to the other level. This problem has already raised two decades ago by Susser, who argued that all levels should be regarded as equivalent (35). This view is shared by an increasing number of researchers in the field of environmental epidemiology. Consequently, we are increasingly faced with multilevel studies (36). If we have data on population and individual level, we can reduce the interference (cross-level confounding) and the effect of changing the dimensions (cross-level effect modification) by using the multilevel modeling (36). This approach combines the best

features of the analysis at individual and population level (36). Of course we have to be extremely careful. We have already mentioned ecological fallacy, but there exist also so-called "atomistic fallacy" that we can commit, judging causality in the broader population level on the basis of the results results provided by studies at the individual level (35).

Examples of studies in environmental epidemiology at the population level

General remarks on studies at population level

Some general features

The common name for studies, in which the basic unit of analysis is population or group of people, and not the individual, is the "ecological studies" (1,28,37). Some of the main features of ecological studies are as follows (6,29,30,38):

- data are usually observed in well-defined spatial or administrative units (e.g. countries, regions, administrative units, municipalities, etc..), or in different institutions (eg schools, etc.);
- information can be obtained from various permanent sources, as well as from periodic sources (special surveys). Periodic surveys (such as cross-sectional surveys) may be an important source of information on the confounders at the population level;

Advantages and disadvantages

The advantages are that they are rapidly carried out and cheap that they can be the best approach for studying exposures that are measurable only on a population level. In fact, in such a situation they are the only choice available (39). They can also combine data from different databases and they are suitable to monitor the effectiveness of public health measures on the population level (6,29,30,38). In environmental epidemiology ecological studies are often the most appropriate design or the studying exposure such as air pollution, water quality and ultraviolet radiation (29).

Disadvantages are that they typically rely on data collected for other purposes, and information on various confounders are often not available. Moreover, since the unit of analysis is a group of people and not the individual, from the results of ecological studies, we can not draw conclusions about the relationships on an individual level (6,29,30,38).

Purpose

Very often we find that studies of this type are intended primarily to be descriptive and for the generation of hypotheses, which are then checked with other types of studies (29,30,38). This is true in the case that the results obtained in studies on the population level, we are trying to pass on an individual level and from them infer the biological effects and risks at this level. In this case there is a strong likelihood that committing ecological fallacy.

However, their results can also be used exclusively on the population level and from them the effects of the epidemiological phenomena (eg, incidence of disease observed in a population) could be inferred (37). In this case, they may also be used for analytical purposes. WHO has even issued a recommendation on when and how to use them for the purpose of improving the information support of environmental health policies and decision making in this area in general (22,23).

Types of variables in studies at the population level

In chapter on ecelogical studies in this book we discussed the nature of variables that enter the ecological studies. They can be basically measured on an individual (and aggregated if necessary) or on a population level. A set of variables in one ecological study may consist of variables of the same type, but may be also of mixed type. If we consider one otcome (dependent) and one exposure (independent) variable we have got four different possibilities that are presented in Figure 3 (40). The situation is more complex if we consider also confounding variables (41).

		Exposure variable: Population level	Exposure variable: Individual level
Outcome variable: Population level		PL/PL	PL/IL
Outcome variable: Individual level		IL/PL	IL/IL

Figure 3. Possible combinations of levels of outcome (dependent) and exposure (independent) variables in population level (ecological) studies.

Studies that involve variables of the same level (Figure 3: PL/PL and IL/IL cells) are called unmixed studies. They present few problems except of problems of scale and complexity of confounding. Some authors consider as true ecological studies only those in which all variables are measured at the population level. Mixed studies they consider as semi-individual or semi-ecological (41,42).

Types of studies in environmental epidemiology at the population level

Studies on the population level are classified in the observational and experimental, and in the descriptive and analytical (see module on the features of epidemiological studies from this book) (29,37). Additionally, further they are classified according to whether they examine (compare) populations in space or in time (12,37). Given all of this classification, we distinguish the following six types of studies: exploratory studies of geographical patterns, exploratory studies of temporal patterns, multiple group studies, time trend studies, mixed-type studies, and community trials (Figure 4):

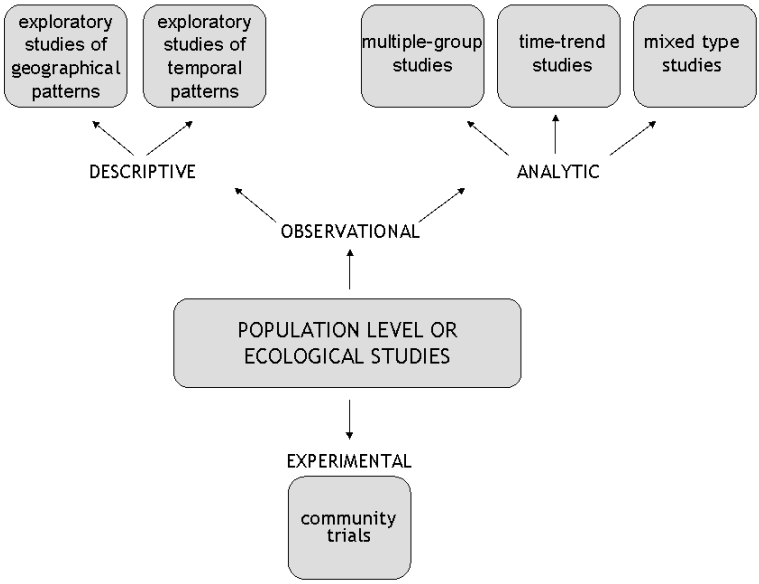


Figure 4. Types of studies in environmental epidemiology at the population level.

Exploratory studies of geographical patterns

These studies are studies by which we describe the frequency of health events in relation to space. They give us insight into the geographical dimension of the health problem and its geographic variation (21). Such studies are also known as geographical studies or mapping studies or disease mapping studies of medical events (12,23). They help us to determine the areas where there is an increased risk for a disease.

In these studies, theterm "space" has different meanings. It can mean the place of residence, place of birth, place of employment, place of education, place of treatment, depending on which health phenomena we observe. Different kinds of "space" can be also combined into categories, such as rural and urban areas (21).

The units of observation at which we collect data are different geographic or administrative units (e.g. country, region, municipality, local community or census district, street, etc.) (21).

Such studies have certain characteristics that need to be taken into consideration in analysis of the data. Two of these features are (12):

- areas with a small number of observed cases of disease tend to have greater variability in epidemiological phenomenon (rate) that we observe, so we can see the most extreme values of the rates is in these areas;
- adjacent areas tend to have more similar values of rates than more remote areas (positive autocorrelation).

To cope with these problems, there exist specific statistical methods (eg Bayesian methods) that go beyond the scope of the contents of this module.

An example of this type of study is the study of cancer incidence in the Brežice municipality in Slovenia on the basis of the Cancer Registry of Slovenia data (43,44).

Exploratory studies of temporal patterns

These studies are studies by which we describe frequency of health phenomena in time by following to the birth cohorts. These exploratory studies are usually performed by comparing the rates (e.g. incidence rates of disease, or mortality rates) for a geographically defined population over time (e.g. 20 years) (12).

From an analytical point of view the data in these studies are a special type of longitudinal data. The usual analytical approach for analyzing such data is the so-called "cohort analysis", which should not be confused with the analysis of data from cohort studies (45). The aim of this approach is to assess the separate effects of three variables that are all related to time: age of people, the period during which these people live (calendar time), and year of birth of people (which defines the birth cohort, the onset of the observed health phenomenon (12). In this analysis, a tabulation of epidemiological phenomena (e.g. the incidence rates of observed health phenomenon, or mortality rates) in relation to the age of certain groups of people (cohort), which was identified as an interesting for observation in a given time period (calendar time), and which is followed when going through different ages in part of their lives or even in their entire life. The aim of this type of studies is to detect the effects of age and calendar period as well as generational effects on observed health phenomenon (1).

An example of this type of study of Okamoto et al. on the effects of age and calendar period, and generational effects on the prevalence of asthma among school children in Japan (46).

Multiple group geographical studies

In multiple-group studies relationship between the average level of exposure and morbididy between a larger number of geographical areas (or administrative regions) on a population level is assesed (12). Thus, in this study design on the one hand information on health phenomenon (morbidity rate) and an on the other information on exposure (average exposure level) is collected (47,48). Such a design is the one the most frequently used in environmental epidemiology (12).

The usual analytical approach is to analyze whether between the health phenomenon (variable Y) and exposure (variable X) there is a statistically significant and substantively meaningful relationship, if possible with taking into account confounding variables.

The data in this type of ecological studies are often analyzed by statistical method called correlation analysis, Consequently, the term "correlation studies" became a synonym for all ecological studies (6,38). This type of analysis was discussed in more details also in the module on basic principles of ecological studies in this module. Morgenstern, however, prefers the regression analysis over the correlation analysis (47,48).

An example of this type of study is the study on the geographical variability of fatal injuries resulting from traffic accidents in Spain in 2002-2004, conducted by Rivaz-Ruiz et al (49).

Time-trend studies

Time-trend studies are studies in which we observe a single population, but in time. Specifically, we monitor changes in rates of observed health phenomenon (e.g. incidence) (variable Y) in respect to changes in the exposure (variable X) over the same period. The relationship between the phenomena exists, if changes in variables Y are parallel to similar changes in the variable X (47,48).

This study design is appropriate for use in cases where exposure to some environmental hazards varies significantly over short periods of time, having paralelly the impact on some acute biological phenomena in humans. Changes in the observed health phenomena can be observed in different time units (minutes, hours, days, weeks, months or years) (36,50).

When analyzing the changes observed in these studies, every population is a control for itself. Periods when pollution is high, are assigned as the exposure periods, while periods when pollution is low, are assigned as the non-exposure periods.

Studies of this type can be used in observational or experimental study designs. In the first case, for example, we are observing the dynamics of two events at a time in which we do not intervene. In another case, we intervene with the dynamics of the phenomenon of (negative) health phenomena, but they must be beneficial interventions for human health (eg vaccination).

In principle, measurements are carried out at large number of consecutive points (1,36). In analysis, for example, we compare the trend of health phenomenon in periods of exposure to periods of non-exposure (or before and after the intervention in the experimental design) (47). However, time series analysis can be very complex, as data for the individuals are not mutually independent. The values of phenomena that are measured on a continuous measurement scale on particular measurement day may correlate with the values of the previous day (36,51). Also, a delay in appearance of health phenomena could be present (12,47), which further complicates the analysis. Latency of disease in relation to exposure may be very long, and may also substantially vary between individuals that form the observed population (12).

In studies of this type we usually try to assess the health risks associated with short-term exposure, for example to high ambient temperatures, high concentrations of small airborne particles (particulate matter, PM), NOx, ozone, and certain other pollutants (52).

Monitoring health events in time is important because, for diseases that occur seasonally, health officials can anticipate their recurrence, and implement control and preventive measures (e.g. vaccinations against seasonal flu, spraying against mosquitoes, etc.).. For diseases that occur from time to time, researchers can conduct studies to identify the causes and modes of transmission, and then develop appropriate targeted measures to control and prevent a recurrence of the disease (21).

The first example of this type of study is the study on the effects of a heat wave in France during the 2 and 15 August 2003 on human health (53,54). Another example is a study of Burnett et al. (55), who studied the relationship between ozone concentrations and admissions of children younger than 2 years to the hospital treatment for acute respiratory disease. Similar to this study, is the study of Mouras et al. (56), who investigated the relationship between admissions of children aged 1 month to 12 years in hospital due to acute respiratory diseases and air pollution.

Mixed-type studies

Mixed type studies are studies that combine features of multiple-group studies and time-trend studies (12,47,48).

Community trials

The majority of studies in environmental epidemiology are observational. Experiments are used much less frequently (57). The reasons are primarily ethical in nature, especially because in this branch of epidemiology experiments often involve large groups of people.

Community trials are experiments in which the units of observation and randomization entire communities (eg total population of a given geographical area) (1,29). They are a special type of field trials (29,58) (common field trials are conducted on an individual level) which was in more details presented already in module on intervention studies in this book.

General design of community trials is similar to the overall design of other experiments (see module on intervention studies in this book). In principle, we have a group of units, in which intervention is applied and the control group, in which intervention is not applied. To assess the effectiveness of intervention we compare between these two groups values of an indicator that we have chosen in advance.

The units are assigned to the observed and the control group randomly or not. However, also in this type of study, randomization is ideal process for assignment of units to the observed and to the control group. The difference is that here whole communities are t to randomly assigned rather than the individuals within them. This process is called cluster randomization (1).

Notwithstanding the fact that the experiments in environmental epidemiology are not frequently used, especially community trials, there exists a number of important studies of this type. The most famous studies are studies of effectiveness of water fluoridation interventions. The first such trial was a trial of effectiveness of water fluoridation in the city of Grand Rapids in the U.S., which started in January 1945. As a control city served a nearby city Muskegon (59). In the next year, several similar experiments followed elsewhere in the U.S. and Canada (Brantford, Ontario, Evanston, Ill., Newburgh, New York) (60). These attempts were followed by many around the world. An example of recent study of this type is study of Mariño et al., who studied the effects of milk powder fluoridation in Chile (61).

Examples of studies in environmental epidemiology at the individual level

Epidemiological studies at the individual level are much better known and far more appreciated than just presented types of studies at the population level. In this group are very well known study types like cross-sectional, case-control or cohort studies.

These studies can be classified in a manner similar to studies on the population level (6,12,29). Figure 5 schematically presents the types of studies in environmental epidemiology at this level.

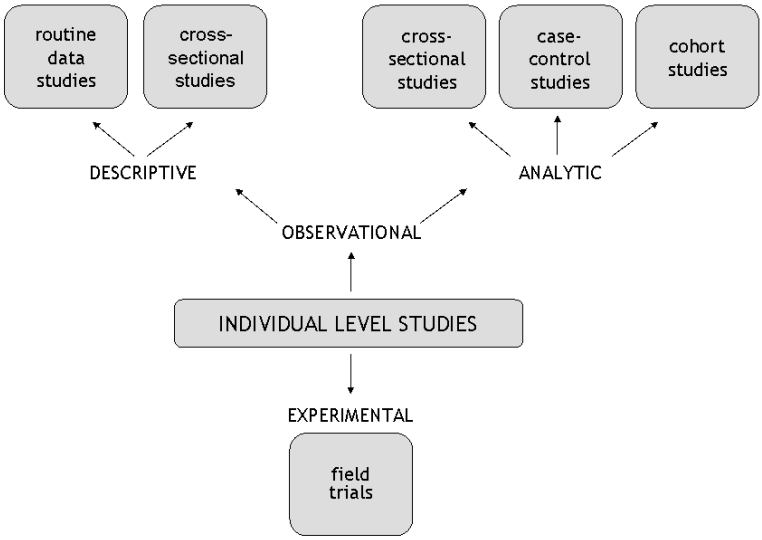


Figure 5. Types of studies in environmental epidemiology at the individual level.

Since the types of studies at the individual level are in details described in previous modules in this book, we present in this place only examples of their use in environmental epidemiology.

Cross-sectional studies

An example of a cross-sectional study in environmental epidemiology is a study of Langkulsen et al. about respiratory symptoms and lung function of primary school children in Bangkok in relation to PM_{10} concentration (62). In Bangkok inhalation of high concentrations of PM_{10} in ambient air is one of the major environmental problems.

Case-control studies

An example of a case-control study in environmental epidemiology is a study of Karakatsani et al. on the relationship between air pollution and incidence of chronic lung disease in the population of Athens (63).

Another example of this type of study is a study of Liou et al., who studied the relationship between environmental factors and Parkinson's disease (64).

Cohort studies

An example of a cohort study in environmental epidemiology is a study of Zanobetti and colleagues, who studied whether long-term exposure to particulate matter in the air is associated with the survival of people with chronic obstructive pulmonary disease (65).

Cohort studies are particularly useful in epidemiology of occupational diseases in the monitoring of groups of workers exposed to chemical or physical agents suspected of increased risk for cancer or other diseases. An example of such a study is the study Albers et al. who investigated the relationship between occupational exposure to klorpirifos (a widely used organophosphorous insecticid) and peripheral nervous system disorders (66). Cohort studies in the workplace are not only relevant for studying of exposure in the workplace, but are also relevant for studying the effects of general exposure in the environment. If in a group of workers who handle a substance, there is no increased risk, it is highly unlikely that exposure to this substance in the open space, where concentrations are usually significantly lower, increase the risk of disease.

Field trials

An example of a field trial in environmental epidemiology is a study of Gamble et al. who have studied if the supplementation of folic acid lowers levels of arsenic in the blood (67). The latter is the most common source of contamination with metals. Current estimates suggest that more than 100 million people in India, Bangladesh, Vietnam, Cambodia and Nepal drink water in which concentrations of arsenic are up to 100-fold higher than the directional concentration of 10 μg/L proposed by WHO.

Concluding remarks

Studies in environmental health are often first directed to the assessment ("proof") relationship between a specific source of pollution and the frequency of adverse health events associated with the pollutants from that source. Since the problems we have mentioned earlier in this chapter are not only present, but may be due to small populations under study, even strengthened, such studies are often doomed to failure. Worse yet, the results of such studies may be worn even in favor of polluters. They are abused as "proof" that the relationship between their pollution and disease in the population, does not exist.

The implementation of such studies is therefore very doubtful. It is not only questionable from a methodological point of view, but also of substance. Does it make sense to rediscover what is already discovered? If the connection between illness and substance from the environment has been demonstrated in many earlier (and stronger) study abroad, it makes sense in a domestic environment studies focused elsewhere. Surely it would be more appropriate for example, studied, if a harmful substance is present in our environment and in what concentrations. If it is present, it is necessary to act, regardless the relationship between disease and substance in the study environment is maybe not "proven". This task requires from researchers in environmental health to take an active role as advocates for the health of people in the area under study. This task is far from easy, as it is necessary to fight the powerful opposing forces.

EXERCISE

Task 1

Carefully read the theoretical background of this module, and recommended readings.

Task 2

Using snowball technique, discuss the characteristics of environmental epidemiology. Special attention, please, pay on problems in exposure assessment and determination of the dose of the specific substance. Discuss the differences in susceptibility to of different population groups and consideration of that fact in the environmental and occupational study.

Task 3

In the WHO web page, which is dealing with training in environmental epidemiology you will find following case study:

> Etzel RA. Problem-based training exercises for occupational and environmental epidemiology. Case 9. Parathion poisoning in Sierra Leon**.** http://www.who.int/peh/Occupational_health/OCHweb/OSHpages/OSHDocuments/EpidemiologyTraining/training_problembased/case9.htm. Accessed: August 24, 2009.

Please carefully go through the questions and try to find answers.

It is advisable to work in small groups where different aspects can be discussed.

Task 4

This taske refers to the understanding of different types of studies in environmental epidemiology and their application in practice.

Students should be divided into as many groups as we have types of studies to be considered. In our case it's nine teams.

Each team shall consider one type of study in practice (study one article, which refers to the type of study). Draw lots to decide which group will consider what type of study.

Types of studies and the corresponding articles are (number of references in the list of references):

1. Population level studies:
    - exploratory studies of geographical patterns: reference #44
    - exploratory studies of temporal patterns: reference #46
    - multiple group studies: reference #49
    - time trend studies: reference #53 and #54
    - community trials: reference #59

2. Individual level studies:
    - cross-sectional studies: reference #62
    - case-control studies: reference #63
    - cohort studies: reference #65
    - field trials: reference #67

Each team shall:

- read carefully the article,
- summarize the information on background, participants/material and methods, results, and discussion,
- prepare a critical view of the study,
- prepare a 15 minute presentation for other students,
- stimulate discussion on the contents of the study.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Institute of Medicine. Environmental Medicine: Integrating a Missing Element into Medical Education. Washington, DC: National Academy Press; 1995.
3. Novick R. Overview of the environment and health in Europe in the 1990s. Geneva: World Health Organisation; 1999. Available from: http://www.euro.who.int/document/e66792.pdf. Accessed: August 16, 2008.
4. Kaufman JS, Cooper RS. Seeking causal explanations in social epidemiology. Am J Epidemiol. 1999;150:113-20.
5. Buffler PA. Epidemiology needs and perspectives in environmental epidemiology. Arch Environ Health. 1988; 43:130-2.
6. Bonita R, Beaglehole R, Kjellström T. Basic epidemiology. 2nd edition. Geneva: World Health Organization; 2006.
7. Terracini B. Environmental epidemiology: a historical perspective. Am J Public Health. 1996;86:668-73.
8. Harvey T, Mahaffey KR, Velazquez S. Dourson M. Holistic risk assessment: an emerging process for environmental decisions. Regul Toxicol Pharmacol. 1995;22:110-7.
9. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S. editors. Modern epidemiology. Philadelphia: Lippincott-Raven; 1998. p.459-80.
10. Stenland K, Savitz DA. Future trends in environmental epidemiology. In: Stenland K, Savitz DA, editors. Topic in environmental epidemiology. New York: Oxford University Press; 1997. p.350-7.
11. Principles for the safety assessment of food additives and contaminants in food. IPCS-Environmental health criteria 70. Geneva: WHO; 1987.
12. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect 1993;101(Suppl 4):23-38.
13. Pekkanen J, Pearce N. Environmental epidemiology: chalenges and opportunities. Environ Health Perspect. 2001;109:1-5.
14. Walter SD. The ecologic method in the study of environmental health. II. Methodologic issues and feasibility. Environ Health Perspect. 1991;94:67-73.
15. Goodchild MF. Twenty years of progress: GIScience in 2010. JOSIS. 2010;1:3–20. Available from URL: http://www.josis.org/index.php/josis/article/viewDownloadInterstitial/32/33. Accessed: August 16, 2009.
16. Anonymous. John Snow's map 1 (1854). Available from URL: http://www.ph.ucla.edu/epi/snow/snowmap_1854.jpg. Accessed: August 16, 2009.

17. Wikipedia, the Free Encyclopedia. Geographic information system. http://en.wikipedia.org/wiki/Geographic_information_system. Accessed: August 16, 2009.
18. Yang TC. GIS resource document 05-72. Choropleth map. University Park, PA: The Pennsylvania State University, Population Research Institute; 2005. na URL: http://www.pop.psu.edu/gia-core/pdfs/gis_rd_05-72.pdf. Accessed: August 16, 2009.
19. Sluga G. Anamorfoze v tematski kartografiji. Ljubljana: Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo; 2005. Available from URL: http://www.fgg.uni-lj.si/sugg/referati/2005/SZGG_05_Sluga.pdf. Accessed: August 16, 2009.
20. World Health Organization. The World Health Report 2005. Make every mother and child count. Geneva: World Health Organization; 2005
21. Centers for Disease Control and Prevention. Principles of epidemiology in public health practice. An Introduction to applied epidemiology and biostatistics. Third edition. Atlanta: Centers for Disease Control and Prevention; 2006. Available from URL: http://www.ihs.gov/medicalprograms/portlandinjury/pdfs/principlesofepidemiologyinpublichealthpractice.pdf. Accessed: August 16, 2009.
22. Briggs D, Corvalan C, Nurminen M (uredniki). Linkage methods for environment and health analysis. 1996. General guidelines. Geneva: World Health Organization, Office of Global and Integrated Environmental Health; 1996.
23. Corvalan C, Nurminen M, Pastides H. Linkage methods for environment and health analysis. 1997. Technical guidelines. Geneva: World Health Organization, Office of Global and Integrated Environmental Health; 1997. Available from URL: http://whqlibdoc.who.int/hq/1997/WHO_EHG_97.11.pdf. Accessed: August 16, 2009.
24. Brindley P, Maheswaran R, Pearson T, Wise S, Haining R. Using modeled outdoor air pollution data for health surveillance. In: Maheswaran R, Craglia M. GIS in public health practice. Boca Raton, FL: CRC Press; 2004.
25. Pearce N. Traditional epidemiology, modern epidemiology, and public health. Am J Pub Health. 1996;86:678-83.
26. Susser E. Eco-epidemiology: thinking outside the black box. Epidemiology. 2004;15:519-20.
27. Ben-Shlomo Y. Real epidemiologists don't do ecological studies? Int J Epidemiol. 2005;34:1181–2.
28. Bailey L, Vardulaki K, Langham J, Chandramohan D. Introduction to epidemiology. Maindenhead: Open University Press; 2005.
29. Dos Santos Silva I. Cancer Epidemiology: Principles and Methods. Lyon: IARC; 1999.
30. Jekel JF, Katz DL, Elmore JG. Epidemiology, biostatistics, and preventive medicine. Philadelphia: W.B. Saunders company; 2001.
31. Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. Am J Public Health. 1994;84:819-24.
32. Šram RJ, Binkova B. Molecular epidemiology studies on occupational and environmental exposure to mutagens and carcinogens, 1997-1999. Environ Health Perspect. 2000; 108(suppl 1):57-70.
33. Hunter DJ. The future of molecular epidemiology. Int J Epidemiol 1999; 25:1012-15.
34. Potter JD. Reconciling the epidemiology, physiology, and molecular biology of cancer. JAMA 1992;268:1573-77.
35. March D, Susser E. The eco- in eco-epidemiology. Int J Epidemiol. 2006;35:1379-83.
36. Pearce N. A short introduction to epidemiology. Second edition. Wellington, New Zealand: Centre for Public Health Research, Massey University Wellington Campus; 2005.
37. Morgenstern H. Ecologic Studies in Epidemiology: Concepts, Principles, and Methods. Annu Rev Public Health. 1982;72:1336-44.

38. Kirch W (ur.). Encyclopedia of Public Health. Leipzig: Springer Verlag; 2008.
39. Susser M. The logic in ecological: II. The logic of design. Am J Public Health. 1994;84:830-5.
40. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994;84:825-9.
41. Künzli N, Tager IB. The semi-individual study in air pollution epidemiology: a valid design as compared to ecologic studies. Environ Health Perspect. 1997;105:1078-83.
42. Ribeiro MC, Pereira MJ, Soares A, Branquinho C, Augusto S, Llop E, e tal. A study protocol to evaluate the relationship between outdoor air pollution and pregnancy outcomes. BMC Public Health. 2010;10:613.
43. Zadnik V, Žagar T, Primic-Žakelj M. Pojavljanje raka v občini Brežice v primerjavi z ostalo Slovenijo: geografska analiza incidence raka v občini Brežice na podlagi podatkov Registra raka za Slovenijo: zaključno poročilo. Ljubljana: Onkološki inštitut, 2006. Available from URL: http://www.partnerstvo.si/upload/File/Rak_Brezice_2006.pdf?phpMyAdmin=lPZgCejTM0T8vHJ364pjAO-POQ0. Accessed: August 16, 2009.
44. Zadnik V, Žagar T, Drobne S, Primic Žakelj M. Estimation of cancer burden in Brežice Municipality, a community neighboring Krško Nuclear Power Plant in Slovenia. Croat Med J. 2008;49:257-66. Available from URL: http://www.cmj.hr/2008/49/2/18461681.htm. Accessed: August 16, 2009.
45. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research: principles and quantitative methods. New York: John Wiley&Sonc, cop.; 1982.
46. Okamoto E, Hata E, Kobayashi M, Hayashi K. Age-period-cohort analysis of asthma prevalence among school children. Environ Health Prev Med. 2007;12:119–28. Available from URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2723228/pdf/12199_2008_Article_BF02898026.pdf. Accessed: August 16, 2009.
47. Morgenstern H. Uses of ecologic analysis in epidemiologic research. Am J Public Health. 1996;86:668-73.
48. Walter SD. The ecologic method in the study of environmental health. I. Overview of the method. Environ Health Perspect. 1991;*94:61-65.*
49. Rivas-Ruiz F, Perea-Milla E, Jimenez-Puente A. Geographic variability of fatal road traffic injuries in Spain during the period 2002–2004: an ecological study. BMC Public Health. 2007;7:266. Available from URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2045108/pdf/1471-2458-7-266.pdf/?tool=pmcentrez. Accessed: August 16, 2009.
50. Dockery DW, Brunekreef B. Longitudinal studies of air pollution effects on lung function. Am J Respir Crit Care Med. 1996;154:S250-S256.
51. Pope CA, Schwartz J. Time series for the analysis of pulmonary health data. Am J Respir Crit Care Med 154: S229-S233. 1996
52. Bell ML, Samet JM, Dominici F. Time-series studies of particulate matter. Annu Rev Pub Health. 2004;25:247-80.
53. Ledrans M, Isnard H, (urednika). Impact sanitaire de la vague de chaleur d'août 2003 en France. Bilan et perspectives. Octobre 2003. Paris: Institut de veille sanitaire; 2004. Available from URL: http://www.invs.sante.fr/publications/2003/bilan_chaleur_1103/vf_invs_canicule.pdf. Accessed: August 16, 2009.
54. Pirard P, Vandentorren S, Pascal M, Laaidi K, Le Tertre A, Cassadou S, Ledrans M. Summary of the mortality impact assessment of the 2003 heat wave in France. Euro Surveill. 2005;10(7):554. Available from URL: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=554. Accessed: August 16, 2009.

55. Burnett RT, Smith-Doiron M, Stieb D, Raizenne ME, Brook JR, Dales RE, Leech JA, Cakmak S, Krewski D. Association between ozone and hospitalization for acute respiratory diseases in children less than 2 years of age. Am Journal Epidemiol. 2001;153:444-452. Available from URL: http://aje.oxfordjournals.org/cgi/reprint/153/5/444. Accessed: August 16, 2009.
56. Moura M, Junger WL, Mendonça GA, De Leon AP. Air quality and acute respiratory disorders in children. Rev Saude Publica. 2008;42:503-11. Available from URL: http://www.scielosp.org/pdf/rsp/v42n3/en_6299.pdf. Accessed: August 16, 2009.
57. Rothman KJ. Methodological frontiers in environmental epidemiology. Environ Health Perspect 1993;101:19-21.
58. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Third edition. London: Lippincott, Williams&Wilkins; 2008.
59. Lennon MA. One in a million: the first community trial of water fluoridation. Bull World Health Organ. 2006;84: 759–60. Available from URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2627472/pdf/17128347.pdf. Accessed: August 16, 2009.
60. Center for Disease Control and Prevention CDC. Public Health Focus: Fluoridation of Community Water Systems. MMWR. 1992;41;372-375. Available from URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/00016840.htm. Accessed: August 16, 2009.
61. Mariño R, Villa A, Guerrero S. A community trial of fluoridated powdered milk in Chile. Community Dent Oral Epidemiol. 2001;29:435-42. Available from URL: http://www.ncbi.nlm.nih.gov/pubmed/11784286. Accessed: August 16, 2009.
62. Langkulsen U, Jinsart W, Karita K, Yano E. Respiratory symptoms and lung function in Bangkok school children. Eur J Public Health. 2006;16:676-81. Available from URL: http://eurpub.oxfordjournals.org/cgi/content/full/16/6/676?view=long&pmid=16684784. Accessed: August 16, 2009.
63. Karakatsani A, Andreadaki S, Katsouyanni K, Dimitroulis I, Trichopoulos D, Benetou V, Trichopoulou A. Air pollution in relation to manifestations of chronic pulmonary disease: A nested case–control study in Athens, Greece. Eur J Epidemiol. 2003;18:45-53. Available from URL: http://www.springerlink.com/content/t2t524vv24370607/. Accessed: August 16, 2009.
64. Liou HH, Tsai MC, Chen CJ, Jeng JS, Chang YC, Chen SY, Chen RC. Environmental risk factors and Parkinson's disease: a case-control study in Taiwan. Neurology. 1997;48:1583-8. Available from URL: http://www.ncbi.nlm.nih.gov/pubmed/9191770. Accessed: August 16, 2009.
65. Zanobetti A, Bind ABC, Schwartz J. Particulate air pollution and survival in a COPD cohort. Environ Health. 2008, 7:48. Available from URL: http://www.ehjournal.net/content/pdf/1476-069X-7-48.pdf. Accessed: August 16, 2009.
66. Albers JW, Garabrant DH, Schweitzer SJ, R Garrison, Richardson RP, Berent S. The effects of occupational exposure to chlorpyrifos on the peripheral nervous system: a prospective cohort study. Occup Environ Med. 2004;61:201-11. Available from URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1740727/pdf/v061p00201.pdf/?tool=pmcentrez. Accessed: August 16, 2009.
67. Gamble MV, Liu X, Slavkovich V, Pilsner JR, Ilievski V, Factor-Litvak P, Levy D, Alam S, Islam M, Parvez F, Ahsan H, Graziano JH. Folic acid supplementation lowers blood arsenic. Am J Clin Nutr. 2007;86:1202–9. Available from URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2042963/pdf/nihms32634.pdf/?tool=pmcentrez. Accessed: August 16, 2009.

RECOMMENDED READINGS

1. Morgenstern H, Thomas D. Principles of study design in environmental epidemiology. Environ Health Perspect. 1993;110:23-38.
2. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S. editors. Modern epidemiology. Philadelphia: Lippincott-Raven; 1998. p.459-80.
3. Rothman KJ. Methodological frontiers in environmental epidemiology. Environ Health Perspect. 1993;101:19-21.
4. Stenland K, Savitz DA. Future trends in environmental epidemiology. In: Stenland K, Savitz DA, editors. Topic in environmental epidemiology. New York: Oxford University Press; 1997. p.350-7.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**THE INTERGRATED EXPOSURE UPTAKE BIOKINETIC (IEUBK) MODEL FOR LEAD IN CHILDREN**
Module: 2.1.2	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Rok Fink**, Bsc, MSc Candidate, Teaching Assistant Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia **Andreja Kukec**, Bsc, PhD Candidate, Teaching Assistant Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Matej Ivartnik,** Bsc, MSc Regional Institute of Public Health Ravne na Koroškem, Slovenia **Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Ivan Eržen** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Zaloška cesta 4, 1000 Ljubljana, Slovenia e-mail: ivan.erzen@gmail.com
Keywords	IEUBK model, environmental risk, human exposure, lead, health effects
Learning objectives	After completing this module student should: • increase knowledge about metals and their impact on human health, • understand models for simulation human exposure to some environmental risks, • be capable to make correlation between exposure and health effects.
Abstract	Metals, especially lead have important impact on human health. The most affected are children due to their body characteristics specific and specifies hygiene habits. Integrated Exposure Uptake Biokinetic (IEUBK) model offers simulation of lead concentration in blood and calculation of risk that child living on specific area will have exceeded level of lead. Nevertheless, the model is also useful tool for risk assessment of children that live on polluted area from different sources with focus on water, food, dust and air lead concentration.
Teaching methods	After introductory lecture students should carefully read the recommended literature about lead exposure. Afterwards they discuss above lead toxicity and practice on IEUBK model. Finally they compare and discuss their results.
Specific recommendations for teachers	• work in small group; • work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2 students); access to the Internet and IEUBK model software available from URL: http://www.epa.gov/superfund/health/contaminants/lead/products.htm; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire and case study.

THE INTEGRATED EXPOSURE UPTAKE BIOKINETIC (IEUBK) MODEL FOR LEAD IN CHILDREN

Rok Fink, Andreja Kukec, Matej Ivartnik, Ivan Eržen

THEORETICAL BACKGROUND

The Intergrated Exposure Uptake Biokinetic (IEUBK) model for lead in children

About the Intergrated Exposure Uptake Biokinetic (IEUBK) model

The Intergrated Exposure Uptake Biokinetic Model (IEUBK) for Lead in Children is a model, used to predict intake of measured concentrations of lead in basic factors of environment (soil, air, water and dust) for each location of interest. These concentrations in correlation with background data form good base for estimation of blood lead level in case of children (1,2).

The IEUBK model was designed by U.S. Environmental Protection Agency (EPA) (1,3,4). It describes exposure scenario for children aged 0-84 months (0-7 years). Model software can be free downloaded from the EPA website: http://www.epa.gov/superfund/health/contaminants/lead/products.htm, where is also entire documentation, necessary for working with this software available.

The IEUBK model is a computer added tool for prediction of the elevated blood lead levels in children under the age of seven who are exposed to environmental lead from many sources. It consists of four modules (1).

IEUBK model modules

The IEUBK model consists of four modules or components (Figure 1) (1,5):

1. Exposure module or component.
   The first module is "Exposure module" which calculates lead (Pb) intake rate from environmental lead concentrations and media specific consumption rates.
2. Uptake module or component.
   The second is "Uptake module" where the amount of lead which is absorbed into child's bloodstream is calculated from Pb intake into the lungs and digestive tract.
3. Biokinetic module or component.
   The "Biokinetic module" follows. It estimates the transfer rates for Pb moving between compartments and through elimination pathways. A variety of complex equations are used to calculate compartmental Pb transfer times. Based on site-specific environmental exposures input by the user, a geometrical mean blood lead concentration is predicted.
4. Probability Distribution Module or component.
   The last one is "Probability Distribution Module" which estimates a plausible distribution of Pb blood concentration that is centred on the geometric mean Pb blood concentration calculated by the "Biokinetic Module". The module calculates

the probability or the risk that the level of concern (default = 10 μg/dl) will be exceeded.

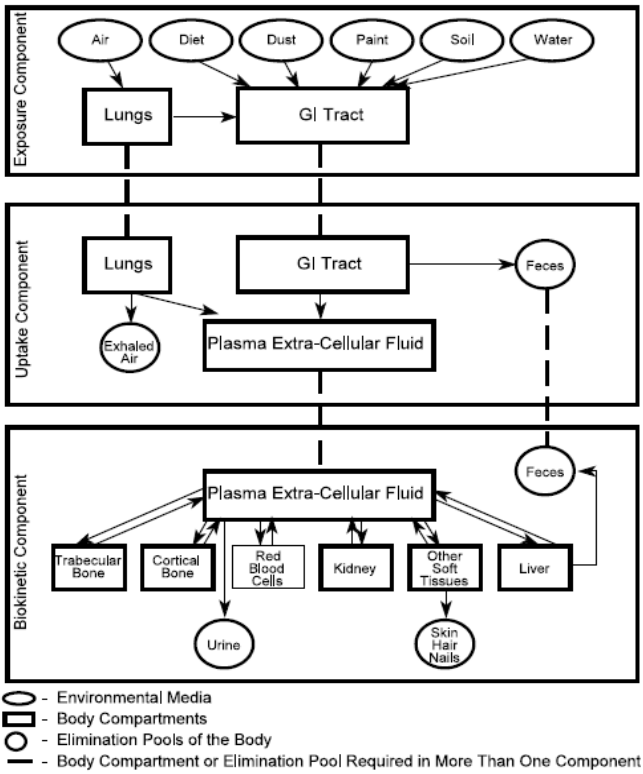


Figure 1. Biological Structure of the IEUBKwin Model (1,5).

Using IEUBK model

The user starts with default values for each parameter and then replaces default values with available specific exposure information for child or population of children in question. The default values (e.g., dietary lead concentrations and consumption values) are EPA-s best estimates for urban residents with no unusual lead exposures (6). The estimated blood lead levels with the default parameters represent best estimate of the blood lead "background" levels that cannot be avoided. These default values are not necessarily appropriate for every site.

The site-specific information usually consists of environmental media concentrations such as soil lead or house dust concentrations, however almost other parameter can be changed if proper data is available.

The model simulations represent chronic exposure and do not incorporate the variability in consumption patterns and media concentrations on a daily or seasonal basis.

For better results it is necessary to use site specific data and to concentrate on smaller area. In that way exposure of a typical child living in that area is better defined.

After site specific data is inputted by the user, calculations could be started. The model calculates the probability that a blood lead concentration derived from the model's specified parameters will exceed a level of concern specified by the user. The predicted blood lead concentration is the geometric mean of the distribution of blood lead concentrations that may occur for each child with the specified exposure scenario. Risk is calculated from this distribution as the probability that a hypothetical child living at this site, with the specified exposure scenario will have a blood lead concentration exceeding the blood lead level of concern. The upper tail of the probability distribution provides an estimate of the risk of exceeding blood lead level of concern.

If the same scenario describes exposure for more children risk is calculated by aggregating the calculated risk for each child as the percentage of hypothetical children living at this site or at these sites, with the specified exposure scenario, which will have a blood lead concentration exceeding the blood lead level of concern. The upper tail of the probability distribution represents the fraction of children exceeding the chosen blood lead level of concern in this case. The calculation is exactly the same as the single-child assessment, but there is an important shift in interpretation of the output (1,4,5,7,8).

When risk for population of children living in multiple sites with different exposure is being estimated, a single run of the IEUBK model is not enough. It is necessary to construct an exposure scenario for each distinct exposure subgroup in the population. For each exposure subgroup, risk is calculated in a single run of the IEUBK model with the specified exposure scenario. To estimate risk for population the risks for each exposure subgroup should be aggregated across all subgroups and weighted by the number of children with that exposure scenario or by the percentage or likelihood of the exposure scenario (7).

Theoretical background on lead toxicity, sources of lead exposures and health risks

Introduction

Metals have adverse effects on both human health and environment. The most concerned are lead, mercury, arsenic, etc. those that have no positive impact on human body. Concentrations in organisms are increasing as regards to time due to its tendency to bioaccumulation. Higher concentrations are impending normal functions of organisms and lead to illness (9,10).

In the past few years, human nail has been recognised as an invaluable tissue for monitoring human environmental exposure, as it provides a good indication of exposure to many toxic and essential trace metals over a period of time. Public concern over the effects of toxic metals on human health and the environment has led to many countries adopting the legislation that restrict or ban their uses. Lead and cadmium (Cd) have raised concern due to their relatively high toxicity and elevated quantity in the environment caused by their widespread use (11).

Heavy metals are ubiquitous in the environment, as a result of both natural and anthropogenic activities, therefore humans are exposed to them through various pathways. Excessive accumulation of heavy metals in agricultural soils through wastewater

irrigation, may not only result in soil contamination, but also lead to elevated heavy metal uptake by crops, and thus affect food quality and safety. Heavy metal accumulation in soils and plants is of increasing concern because of the potential human health risks. This food chain contamination is one of the important pathways for the entry of these toxic pollutants into the human body. In order to assess the health risks, it is necessary to identify the potential of a source to introduce risk agents into the environment, estimate the amount of risk agents that come into contact with the human-environment boundaries, and quantify the health consequence of the exposure (12).

Lead, with atomic number 82, atomic weight 207.19 g mol^{-1} and a specific gravity of 11.34, is a bluish or silvery-grey metal with a melting point of 327.5°C and a boiling point at atmospheric pressure of 1740 °C. It has four naturally occurring isotopes with atomic weights 208, 206, 207 and 204 g mol^{-1} in decreasing order of abundance (13,14).

Lead occurs naturally in the Earth's crust. However, it is rarely found naturally as a metal. It is usually found combined with two or more other elements to form lead compounds (15). In the environment lead is known to be toxic to plants, animals and microorganisms. Lead bio accumulates in the skeleton and wet tissue in mammals and in aquatic algae and invertebrates (16).

It has been used since prehistoric times, and has become widely distributed and mobilized in the environment. The amount of lead in the environment increased during the industrial revolution, and again significantly in the 1920s with the introduction of leaded gasoline (17). Therefore exposure and uptake of this non essential element have consequently increased. Both occupational and environmental exposures to lead remain a serious problem in many developing and industrializing countries, as well as in some developed countries. In most developed countries, however, introduction of lead into the human environment has decreased in recent years, largely due to public health campaigns and a decline in its commercial usage, particularly in petrol. Acute lead poisoning has become rare in such countries, but chronic exposure to low levels of the metal is still a public health issue, especially among some minorities and socioeconomically disadvantaged groups (18).

It seems that lead has no apparent biological function. Nevertheless, it is able to enter biological systems via food, water, air and soil. Drinking water can become contaminated, either at the source due to deposition from environmental sources or in the water distribution system. Lead emissions from the combustion of leaded fuels and metals smelting have contributed significantly to the accumulation of atmospheric and soil lead (19).

Historical background on lead toxicity

Lead's toxicity was recognized and recorded as early as 2000 BC and the widespread use of lead has been a cause of endemic chronic plumbism in several societies throughout history. The Greek philosopher Nikander of Colophon in 250 BC reported on the colic and anaemia resulting from lead poisoning. Hippocrates related gout to the food and wine, though the association between gout and lead poisoning was not recognized during this period (450-380 BC).

Later, during the Roman period, gout was prevalent among the upper classes of Roman society and is believed to be a result of the enormous lead intake using leaden cooking utensils and pots, leaden wine urns, lead plumbing, vessels used to concentrate grape juice, containers used to store wine, and lead-based makeup. There are many

distinguished historians who now believe that this high exposure to lead was a contributing force in the decline of the Roman Empire (20,21).

In the German city of Ulm, during the late 1690s, there was a severe outbreak of colic, an illness characterized by a variety of symptoms, including excruciating abdominal pain. The monks were using the culprit to sweeten the wine, litharge, a white oxide of lead In 1696, Duke Ludwig issued a decree forbidding the use of lead-based additives in any wine product. For anyone who violated this decree, the punishment was death.

In Hungary, in 1994, a major health problem occurred when red oxide lead was mixed into paprika to brighten the colour of the spice.

Today, occupational exposure to lead remains a big problem in developing countries. Occupational lead exposure is likely unregulated in these countries with little monitoring of poisoning being done. What has become a growing concern among health officials is the prevalence of home-based cottage industries in these countries (22).

Sources of lead in the urban environment

Lead occurs naturally in the environment. However, most of the high levels found throughout the environment come from human activities; lead alloys are commonly found in pipes, storage batteries, weights, shot and ammunition, cable covers, and sheets used to shield us from radiation. The largest use for lead is in storage batteries in cars and other vehicles. Lead compounds are used as a pigment in paints, dyes, and ceramic glazes and in caulk. Sources of lead in dust and soil include lead that falls to the ground from the air, and weathering and chipping of lead-based paint from buildings, bridges, and other structures (Figure 2) (15,23).

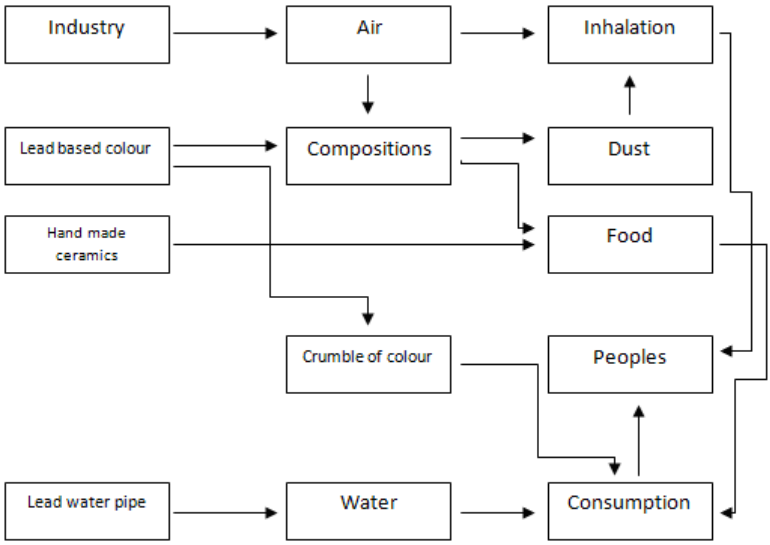


Figure 2. Usage and range of lead (7).

Howsoever, the major sources of lead pollution in urban environments are exhaust emissions from petrol vehicles, flaking of Pb-based paints, contamination of food and water from lead pipes and soldered joints, mining, smelting and foundry activities, former use of lead arsenate pesticides, and application of sewage sludge to agricultural soils (24).

Unlike overt lead toxicity, where there is usually one identifiable source, low level environmental exposure to lead is associated with multiple sources (e.g. petrol, industrial processes, paint, solder in canned foods, water pipes) and pathways like air, household dust, street dirt, soil, water, food (18). Because of low solubility and resistance to microbial degradation, lead accumulates in the surface horizons in soil (24).

Lead exposure is mainly due to concentrations of lead in air, food, water, soil and dust (Figure 2). Higher concentrations are often linked with emissions of lead in air (Table 1). In countries where usage of lead containing gasoline is already forbidden, air intake is characteristic for work exposure. In Slovenia lead containing gasoline is not in use since 2001. For common population oral intake is the most important one. This is even more important for children as they are more often in contact to soil and dust. Exposure in this population group is higher than in common population. Nevertheless food and water can be contaminated through environment or vessel (18).

Table 1. Representative relationship of blood lead median level to intake of lead for the general population (18).

	Median blood lead level among:	
Medium	**Children**	**Adults**
Air[b]	0.09 μmol Pb/litre (1.92 μg Pb/dl) per μg Pb/m^3 air	0.079 μmol Pb/litre (1.64 μg Pb/dl)[c] per μg Pb/m^3 air
Water	_	0.003 μmol Pb/litre (0.06 μg Pb/dl) per μg Pb/litre
Food	0.001 μmol Pb/litre (0.16 μg Pb/dl) per μg Pb/day	0.002-0.003 μmol Pb/litre (0.04-0.06 μg Pb/dl) per μg Pb/day
Dust[b]	0.09 μmol Pb/litre (0.18 μg Pb/dl) per 1000 μg Pb/g dust	_
Soil[b]	0.11 μmol Pb/litre (2.2 μg Pb/dl) per 1000 μg Pb/g soil	_

LEGEND: a = these data are provided for illustrative purposes only. The relationships are curvilinear and are broad guidelines that are not applicable at lower or higher levels of exposure; b = a value in the range 0.144-0.24 μmol Pb/litre or in the range 3-5 μg Pb/dl per μg/m^3 is obtained when one considers indirect contribution through deposition on soil/dust; c = the air to blood relationship in occupational settings is best described by a curvilinear relationship with slopes in the range 0.02-0.08 μg/m^3 air. The slope is variable but lower than that for humans in the general environment (1.6-1.9 μg/m^3).

Lead is removed from the atmosphere by dry or wet deposition. The residence time of lead containing particles in the atmosphere varies according to a number of factors, such as particle size, wind currents, rainfall and height of emission (13).

Point sources, such as primary or secondary lead smelters, may create local pollution problems. The level of contamination of the surrounding air and soil depends on the amount of lead emitted, the height of the stack, the presence of fugitive sources, topography and other local features. In addition, the refining and manufacture of lead-containing compounds and goods and refuse incineration also give rise to lead emissions. Most of the lead in the air is in the form of fine particles with a mass median equivalent diameter of less than 1 μm. The fraction of organic lead (predominantly lead alkyls that escaped combustion) is generally below 10% of the total atmospheric lead, the majority of lead from leaded petrol emissions being inorganic particles such as lead bromochloride (PbBrCl). In the immediate vicinity of smelters, the particle size distribution usually shows a predominance of larger particles. However, these particles settle at distances of a few hundred metres or 1–2 km, so that further away the particle size distribution is indistinguishable from that of other urban sites (13,24).

The presence of lead water pipes in old houses can be an important source of lead exposure for humans, particularly in areas with soft water. Houses can be an additional source of exposure, as are diverse uses such as lead solders, ceramic glazes, cosmetics and folk medicines (13).

Air

Most of the lead in ambient air is in the form of sub micron sized particles. Some 30–50% of these inhaled particles are retained in the respiratory system. Virtually all of this retained lead is absorbed into the body. Particles in the size range of 1–3 μm are also efficiently deposited in the lungs. Larger particles are deposited with variable efficiency, mainly in the upper respiratory tract with incomplete absorption. All lead particles that are cleared by the lung can be swallowed and result in further lead absorption from the gastrointestinal tract (12,13).

Drinking-water

Lead concentrations in drinking water and groundwater vary from 1-60 μg/l. In most European countries, the levels of lead in domestic tap water are relatively low, i.e. normally 20 μg/l. Consequently, exposure to lead through water is generally low compared with exposure through food. Nevertheless, in old houses with lead pipes used for the domestic drinking water supply, blood lead levels in six year old children were found to be elevated by about 30% relative to houses without lead pipes. In areas with soft water, where lead water pipes and lead plumbing are common, the contribution of lead in drinking water to the total lead intake may even be more pronounced (13,25).

Food

Many people receive the largest portion of their daily lead intake via food. Most lead enters food during storage and manufacture, e.g. in canned food and in alcoholic drinks. The most important pathway whereby atmospheric lead enters the food chain is thought to be direct foliar contamination of plants. This contamination depends on the rate of fallout of lead in the districts where food is grown; it tends to be higher in heavily industrialized areas (13,18).

Additionally, air deposits raise the level of lead in soil, which, in the course of decades and centuries, may result in an increased uptake of lead through the roots. The concentrations of lead in various food items are highly variable. Several studies have reported average lead intakes in the range of 100–500 μg/day for adults, with individual diets covering a much greater range. More recent data indicate total daily intakes of about 100 μg or less. For young children, estimates of total daily intake are about one half the figures for adults (13,15).

Dust and soil

Dust and soil can be significant lead exposure sources, especially for young children. Lead in soil can come from the air or from erosion of lead-bearing rocks, and may be carried indoors as dust. Lead dust can also be generated within the home, especially older homes that used lead based paints or lead solder. Lead dust is especially dangerous for babies and young children, because they tend to put things in their mouths and their breathing zone is closer to floor level (17,24).

Other routes

The former category includes high lead levels in dust and soil in residential areas near smelters or refineries, high-density traffic, and the consumption of vegetables and fruit grown on high lead soils or near sources of lead emissions (smelters or roads with high traffic density). The latter category includes occupational exposures; secondary occupational exposures of members of the families of lead workers; contamination of house dust in houses with interior lead paint; contamination of tap water in houses with lead water pipes or lead plumbing; the use of improperly glazed earthenware vessels; tobacco smoke and alcohol consumption (13,17).

Population groups at highest exposure risk

Children are at increased risk for lead exposure, as well as for adverse health effects, for the following reasons:

- children have behavioural characteristics (outdoor activity, less concern for hygienic conditions and hand-to-mouth activities), which increase the risk of lead exposure;
- children eat and drink more per unit of body weight than adults, so that their relative
- lead intake is increased;
- lead absorption in the gastrointestinal tract is substantially higher in children (about
- 50%, compared with about 10% in adults);
- there is a greater prevalence of nutritional deficiencies (e.g. iron and vitamin D);
- among children, which enhance absorption of lead from the gastrointestinal tract;
- the blood brain barrier is not yet fully developed in young children;
- haematological and neurological effects of lead occur at lower thresholds than in adults (13,18).

Children are more sensitive than adults to the effects of lead. Young children under the age of five are particularly vulnerable for a variety of reasons, including the fact that their body and brain are still developing. Two year-olds tend to have the highest blood level

concentration because they put many things into their mouth, including toys or other products that may become lead contaminated (26).

Environmental effects

Lead in the environment is mainly particulate bound with relatively low mobility and bioavailability. Lead bio accumulates in most organisms, in particular in biota feeding primarily on particles, e.g. mussels and worms. In general there is no increase in concentration of the metal in food chains (biomagnification). The distribution of lead within animals is closely associated with calcium metabolism. In shellfish, lead concentrations are higher in the calcium rich shell than in the soft tissue. In dolphins, lead is transferred from mothers to offspring during foetal development and breast feeding. Lead accumulates in the bones of animals like it does in humans (16,18).

Body intake

Absorption (bioavailability)

Absorption through the respiratory tract is influenced by the particle size distribution and the ventilation rate. For adults the retention rates of airborne particulates range from 20% to 60%. Although lead salts differ widely in terms of water solubility, the chemical form of lead is not considered an important factor for respiratory absorption. The proportion of lead absorbed from the gastrointestinal tract is about 10% in adults, whereas levels of 40–50% have been reported in infants (13,21).

Distribution

The non-excreted fraction of absorbed lead is distributed among three compartments: blood, soft tissues and the mineralizing tissues. About 95% of the lead body burden in adults is located in the bones, compared with about 70% in children. Some 99% of the lead in the bloodstream is bound to erythrocytes. The biological half time of lead in blood can be as short as 20–40 days, although longer half time values have been reported in lead workers and may depend on the lead body burden. Lead may be released from the bones in decalcification processes related to elderly people, pregnancy, acidosis, thyrotoxicosis or active remodelling processes in the bones of children. It has, however, been suggested that lead may be released from bone tissue after the menopause, and clearly higher blood lead levels have indeed been found in post-menopausal than in pre-menopausal women (13,21).

Human exposure and health effects

Exposure of human populations to environmental lead was relatively low before the industrial revolution but has increased with industrialization and large-scale mining. Lead contamination of the environment is high relative to that of other nonessential elements (18).

Lead, as ubiquitous and versatile metal, has been used since prehistoric times. It has become widely distributed and mobilized in the environment and human exposure to and uptake of this non essential element has consequently increased. At high levels of human exposure there is damage to almost all organs and organ systems, most importantly the central nervous system, kidneys and blood,

culminating in death at excessive levels. At low levels, haemal synthesis and other biochemical processes are affected, psychological and neurobehavioral functions are impaired, and there is a range of other effects (18,23).

Lead can increase blood pressure and cause fertility problems, nerve disorders, muscle and joint pain, irritability, and memory or concentration problems. It takes a significantly greater level of exposure to lead for adults than it does for kids to sustain adverse health effects. Most adults who are lead poisoned get exposed to lead at work. Occupations related to welding, renovation and remodelling activities, smelters, firing ranges, the manufacture and disposal of car batteries, and the maintenance and repair of bridges and water towers, are particularly at risk for lead exposure (26).

Elevated blood lead level (10 μg/dl or above) has been associated with toxicity in the developing brain and nervous system of young children, leading to lower intelligence quotient (Figure 3) (14).

Studies have found that Cd and Pb have a tendency to accumulate in vital organs and have a half life up to 30 years. Reports have indicated that absorption of toxic metals in the gastrointestinal tract depends on the species, dose, frequency, age, nutritional status and interaction with other dietary components such as magnesium, calcium, zinc, iron and phosphorous (11).

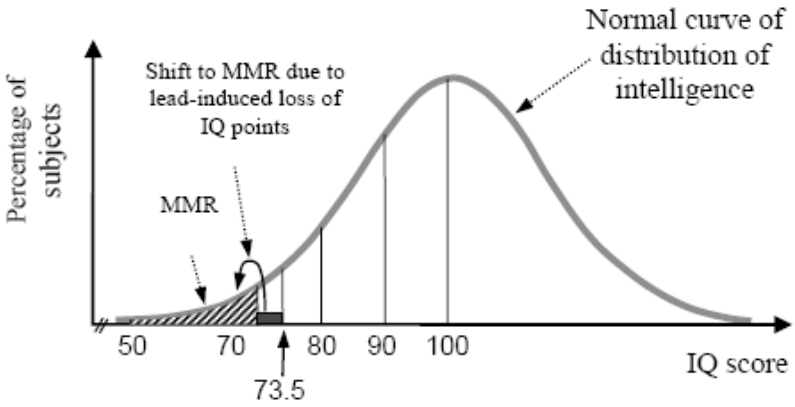


Figure 3. Shift to mild mental retardation (MMR) as a result of lead-induced IQ loss Health effects children (23).

The widespread environmental contamination, the propensity to cause a wide spectrum of toxic effects and the number of individuals affected worldwide makes this ubiquitous neurotoxicant a public health problem of global magnitude. In spite of the extensive documentation of the toxic effects of lead on human health, a complete and detailed explanation on the mechanisms by which lead exerts its effects on the central nervous system has yet to be defined (19).

The toxicity of lead may largely be explained by its interference with different enzyme systems: lead inactivates these enzymes by binding to Sulfhydryl (-SH)-groups of

its proteins or by displacing other essential metal ions. For this reason many organs or organ systems are potential targets for lead, and a wide range of biological effects of lead have been documented. These include effects on haem biosynthesis, the nervous system, the kidneys and reproduction, and also cardiovascular, hepatic, endocrinal and gastrointestinal effects (Table 2 and Table 3) (13).

Table 2. Summary of LOAELs (Lowest-Observed-Adverse-Effect Level) for lead – induced health effects in adults (13).

LOAEL at given blood lead level (µg/dl)	Haem synthesis haematological and other effects	Effects on the nervous system
100-120		Encephalopatic signs and symptoms
80	Frank anaemia	
50	Reduced haemoglobin production	Overt subencephalopathic neurological symptoms, cognition impairment
40	Increased urinary delta-aminolevulinic acid (ALA) and elevated coproporphyrin	
30		Peripherial nerve dysfunction
20–30	Free erythrocyte protoporphyrin (FEP) elevation in males	
15-20	FEP elevation in females	

Table 3. Summary of LOAELs for lead – induced health effects in children (13).

LOAEL at given blood lead level (µg/dl)	Haem synthesis haematological and other effects	Effects on the nervous system
80–100		Encephalopatic signs and symptoms
70	Frank anaemia	
40	Increased urinary ALA and elevated coproporphyrin.	
25–30	Reduced haemoglobin synthesis	
15–20	FEP elevation	
10–15	Vitamin D_3 reducion	Cognitive impairment
10	Delta-aminolevulinic acid dehydrase (ALAD) inhibition	Hearing impairment

According to IARC, evidence of the carcinogenicity of lead compounds in humans is inadequate (13).

In the general non smoking adult population, the major exposure pathway is from food and water. Airborne lead may contribute significantly to occupational exposure and exposure of smokers. Inhalation is the dominant pathway for lead exposure of workers in industries producing, refining, using or disposing of lead and lead compounds (16).

For infants and young children, lead in dust and soil often constitutes a major exposure pathway and this exposure has been one of the main concerns as to the exposure of the general population. The intake of lead will be influenced by the age and behavioural characteristics of the child and the bioavailability of lead in the source material. Baseline estimates of potential exposure of children to dusts, including intake due to normal hand to mouth activity, are 0.2 g dust/day for children 1–6 years old when both indoor and outdoor ingestion of soil and dust is considered, but for some children it may be up to 5 g/day. Depending on the lead content of dust/soil intake of lead with dust/soil can clearly exceed the PTWI (provisional tolerable weekly intake) value established by WHO (World health organization) (14,16).

In humans, lead can result in a wide range of biological effects depending upon the level and duration of exposure. Effects may range from inhibition of enzymes to the production of marked morphological changes and death (Table 4). Such changes occur over a broad range of doses (16).

Of particular concern for the general population is the effect of lead on the central nervous system. Epidemiological studies suggest that low level exposure of the foetus and developing child may lead to repro toxic effects, i.e. damage to the learning capacity and the neuropsychological development. Studies of children indicate a correlation between higher lead contents in the blood and a lower intelligence quotient (IQ). Slowing of nerve conduction velocity has been found at low lead blood levels. Impairment of psychological and neurobehavioral functions has also been found after long-term lead exposure of workers (16,19).

Lead has been shown to have effects on haemoglobin synthesis and anaemia has been observed in children at lead blood levels above 40 μg/dl. Lead exposure is associated with a small increase in blood pressure. There is no evidence to suggest that any association of lead blood levels with blood pressure is of major health importance (17,23).

Table 4. Consequences of lead impact on children and adults health (6).

Lead concentration (μg/dl blood)	**Children**	**Adults**
150	death	
100	comma, kidney damage	coma, anaemia, shorter life
50	anaemia, abdominal spasms, reduction in erythrocyte synthesis	reduction in erythrocyte synthesis
40	vitamin D_3 reduction	difficulties with nerves, perception problems, moving problems, infertility (men), kidney damage
30		hypertension, deafness, reduction in erythrocyte synthesis (men)
20	impact on central nervous system, perception problems, premature birth, low labour weight, disorders in brain cell development	reduction in erythrocyte synthesis (woman), hypertension
10	low IQ, deaf, stagnation in growth, lead can go through placenta	
0		

The reproductive effects of lead in the male are limited to sperm morphology and count. In the female, some adverse pregnancy outcomes have been attributed to lead. Lead does not appear to have deleterious effects on skin, muscle or the immune system (16,18).

The evidence for carcinogenicity of lead and several inorganic lead compounds in humans is inadequate. Classification of IARC (International Agency for Research on Cancer) is class 2B. The agent or mixture is possibly carcinogenic to humans. The exposure circumstance entails exposures that are possibly carcinogenic to humans (16).

As measurement techniques have increased in sophistication, however, it has become clear that lead can have effects on the human body at levels lower than initially suspected, and the nervous system appears to be the principal target although other organ systems are far from immune (23).

Mild mental retardation from lead related IQ loss

IQ loss is not considered to be a disease and therefore the analysis only addresses instances when IQ loss results in mental retardation. By convention, mild mental retardation (MMR) occurs when the IQ is below 70 points but above 50 points. Mental retardation (MR) includes cases in which the IQ is below 50 points. Intelligence in human populations approximates to a normal distribution (except for an excess below 50 IQ points, representing brain damage and disorder), generally with a mean of 100 and a standard deviation of 15 IQ points. Thus to attribute a disease burden to the loss of IQ points, it is necessary to calculate the number of people just above the threshold who would enter the MMR range through loss of IQ points due to lead exposure (Figure 3) (23).

It has been established that toxicity and absorption of these metals is elevated in children under the age of six years compared to adults due to their not having fully developed nervous system and other organs, more hand to mouth activities, untimely outdoor activities, not fully developed hygienic habits and active metabolism. Lead and cadmium exposure has also been associated with neurotoxicity, osteoporosis, and osteomalacia. Early brain growth is a critical period in the development of the central nervous system of the child and the negative environmental factors (lead in the environment, water and/or food) present at this developmental stage have a long term effect on the health and life of the child (11).

Preschool children may suffer long lasting adverse neurobehavioral effects when blood lead concentration exceeds 10 µg per dl of whole blood. Soil lead is therefore of particular concern with respect to potential chronic risks to human health, e.g. (24).

The effects of lead on human health have been recognized since antiquity. However, it was not until the 1970s that seminal epidemiological studies provided evidence on the effects of lead intoxication on cognitive function in children. During the last two decades, advances in behavioural, cellular and molecular neuroscience have provided the necessary experimental tools to begin deciphering the many and complex effects of lead on neuronal processes and cell types that are essential for synaptic plasticity and learning and memory in the mammalian brain.

Age at exposure is also considered a significant risk factor for lead intoxication and its effects. Children absorb significantly more lead in the intestine than adults. In addition, pre- and perinatal exposure results in higher brain lead levels than postnatal exposure due to an under developed blood brain barrier in early life (16,20). Elevated lead body burden has been associated with hypertension in adults. In children, the health effects of concern are neurobehavioral and endocrine alterations, such as an increase in hyperactivity and distractibility, delayed puberty and cognitive deficits in the form of IQ changes. Studies on cognitive ability have demonstrated a deficit between 0 and 5 points on the IQ scale for every 10 µg /dl increase in blood lead level (19).

Figure 4 shows average levels of lead in children's blood in 11 countries (Bulgaria, the Czech Republic, France, Germany, Hungary, Israel, Poland, Romania, the Russian Federation, Sweden and The former Yugoslav Republic of Macedonia) at different times between 1990 and 2006 (14).

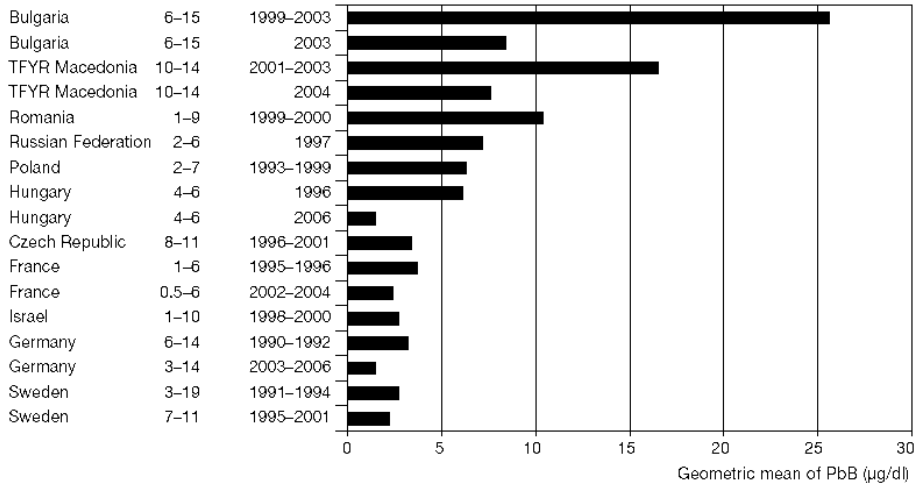


Figure 4. Mean blood lead levels in children measured in selected Europeans countries 1991–2006 (14).

Blood lead levels as a biomarker of exposure

However, because of rapid industrialization and the persistence of lead in the environment, exposure is likely to remain a significant public health problem in most developing countries for many years. Much work needs to be done to identify and treat children with elevated blood lead levels and reduce lead exposure in the community. Screening, monitoring, intervention and evaluation are critical for the development of rational, cost effective and science based public health policies aimed at achieving these goals (18).

Currently, the Centre for Disease Control (CDC) limit of concern for childhood lead intoxication is 10 μg/dl. This level is thought to be the threshold for childhood cognitive deficits. However, recent data suggest that effects on childhood cognition may be present at blood lead levels <10 μg/dl (19).

Lead concentrations in a variety of biological media, such as blood, urine, bone, tooth and hair, have been measured to serve as biomarkers of exposure. These different markers have different validity as surrogate measures of dose. Lead in urine and hair are of only limited value. Whole blood lead concentration is the most widely used and most generally accepted measure of absorbed dose. Blood lead is distributed among plasma and the erythrocytes, with less than 5% being in the plasma; most of the lead is bound to haemoglobin. Although it is likely that plasma lead concentrations may better reflect the "active" fraction of lead in blood and characterize the relationship between blood lead and tissue accumulation and effect, there is little experimental support because of the analytical problems (13,16).

Guidelines

Guidelines for lead in air will be based on the concentration of lead in blood. A critical level of lead in blood of 10 μg/dl is proposed by WHO (13). The following public health measures may be used to reduce the exposure of children to lead in the environment and thus to lower the level of lead in their blood (14).

In September 2006, the Intergovernmental Forum on Chemical Safety was the setting for the Budapest Statement on mercury, lead and cadmium, which recognizes that the risks from these three substances need to be addressed by further global, regional, national and local action, as appropriate. In the same context, the Declaration of Brescia on Prevention of the neurotoxicity of metals supported the revision of lead exposure standards and promoted an immediate reduction of the level of lead in children's blood to a concentration of 5 μg/dl worldwide (14,16).

In 2004, the Fourth Ministerial Conference on Environment and Health adopted the Children's Health and Environment Action Plan for Europe (CEHAPE), which includes four regional priority goals to reduce the burden of environment related diseases in children. In CEHAPE RPG IV specific action is set out to reduce the exposure of children to lead, such as the enactment of legislation on the content of lead in petrol and building materials, to develop and enforce regulations to minimize the risks from hazardous building materials and to carry out bio monitoring of lead in infants and mothers at risk (14).

Regularly conducted harmonized assessments of the levels of lead in children's blood are needed to identify and eliminate existing sources of environmental exposure to lead and to monitor the effectiveness of preventive action (14).

CDC considers children to have an elevated level of lead if the amount of lead in the blood is at least 10 μg/dl (15).

Practical work with IEUBK model

The process of working with the IEUBK model could be split in following steps:

1. Step one – starting the programme.
   The work with the programme starts with the basic communication window presented in Figures 5 and 6.

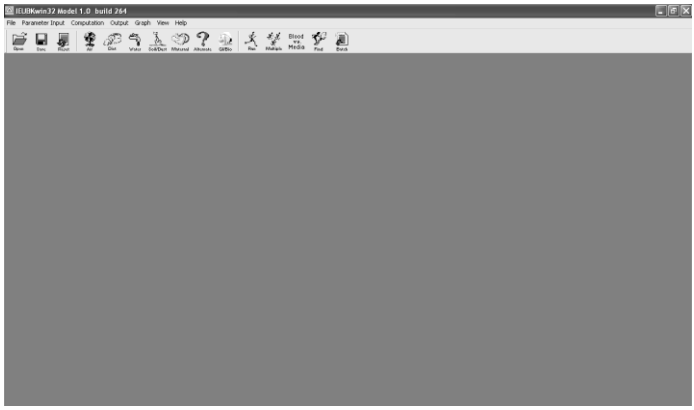


Figure 5. IEUBKwin32 Model 1.0 main communication window.

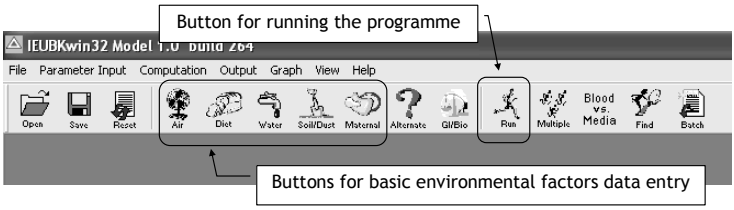


Figure 6. IEUBKwin32 Model 1.0 main communication window buttons.

2. Step two – basic environmental factors data entry.
   Data for five environmental factors should be entered via five different communication windows:
   - Air data – the values of lead concentration in indoor and outdoor air should be entered. It could be assumed that value is constant or values for time spent outdoors, ventilation rate and lung absorption could be entered (Figure 7).

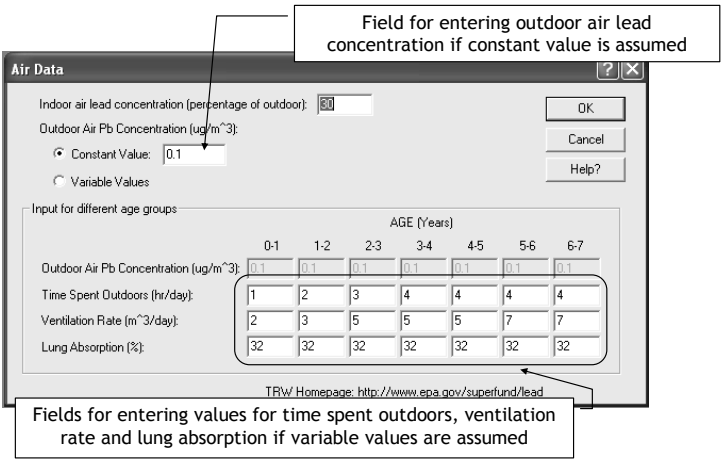


Figure 7. IEUBKwin32 Model 1.0 – air data communication window.

- Dietary data – the values of lead intake with food should be entered. Data on average dietary lead intake could be entered or alternate dietary values (e.g. home grown fruits and vegetables) (Figure 8).

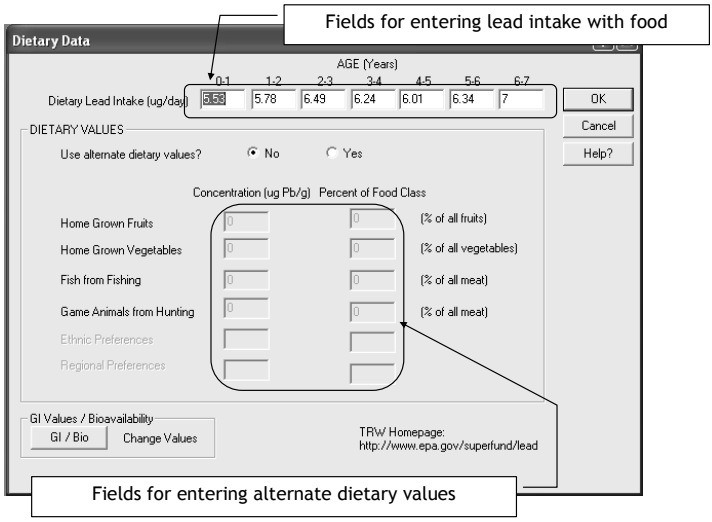


Figure 8. IEUBKwin32 Model 1.0 – dietary data communication window.

- Drinking water data – amount of water consumption should be entered, and lead concentrations, in which average value could be assumed or variable in different drinking water sources (Figure 9).

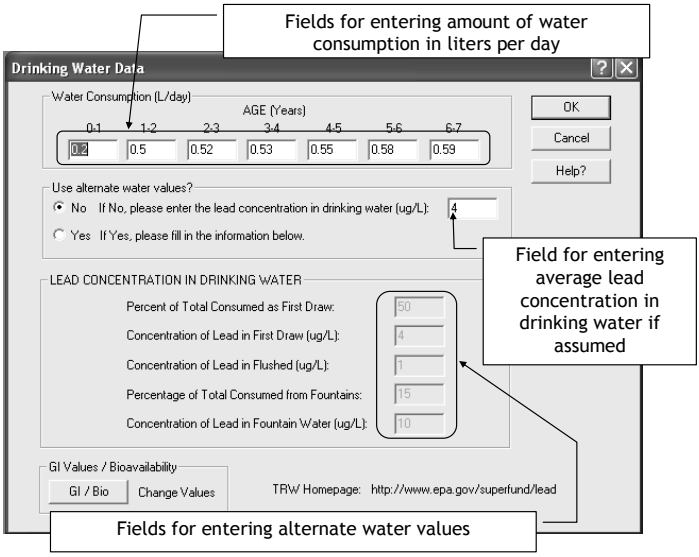


Figure 9. IEUBKwin32 Model 1.0 – drinking water data communication window.

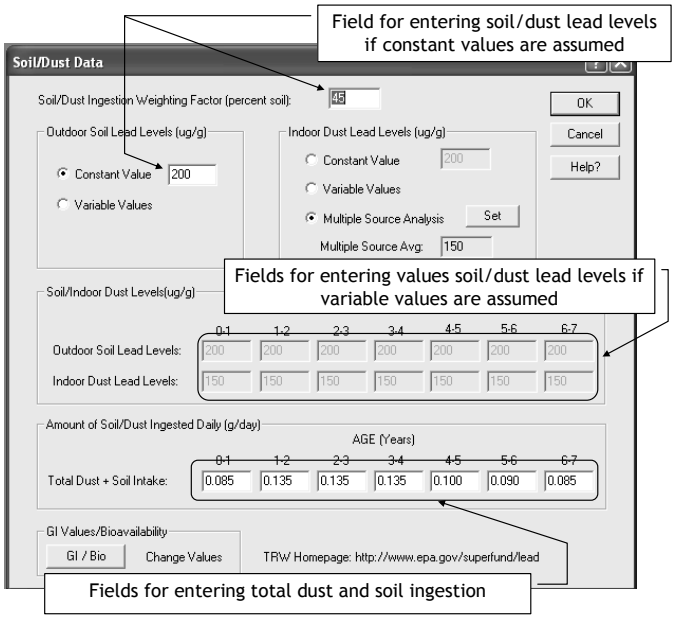


Figure 10. IEUBKwin32 Model 1.0 – soil/dust data communication window.

- Soill/dust data – data on outdoor soil and indoor dust should be entered. Constant or variable values could be assumed (Figure 10).
- Maternal data – data on mother's blood level concentration at childbirth should be entered (Figure 11).

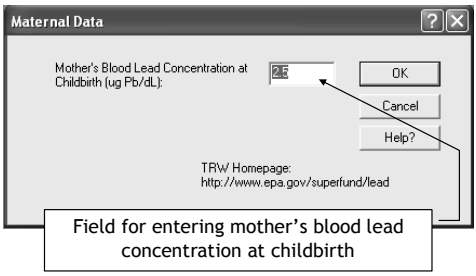


Figure 11. IEUBKwin32 Model 1.0 – maternal data communication window.

3. Step three – running the programme.
   The programme is run using the "Run" button (Figure 6):
4. Step four – the output.
   In IEUBK model results can be shown as text or figure (Figures 12 and 13). This function can be set up in each use of program. We can choose between curve line of distribution or density (Figure 12). Both relations can be used for graphical calculation when several curves are shown on the same figure.

CASE STUDY: CONCENTRATION OF LEAD IN BLOOD OF 3 – YEAR OLD CHILDREN: THE CASE OF UPPER MEŽICA VALLEY, SLOVENIA

Background

Data of lead concentrations in blood represent great advantages in health exposure evaluation. Lead amount in blood is showing exposure to lead in air, food, water and dust (27). Heavy metals, especially lead is the major pollutant in Upper Mežica valley. Half of millennium industrial mining and metal foundry of lead and zinc ore left special offprint in local environment. Through whole mining history lead dust was depositing in soil, water, springs, animals and plants as well as in bones, hairs and teeth of human (28).

Several studies were made since 1952 about impact of lead on population that live in the area of Upper Mežica valley. Results showed that environmental pollution and its health impact in this region are high above Slovenian average and that concentrations of blood lead are still exceeding levels that represent risk for human health (28).

Due to expert recommendations some arrangements must be taken when concentration of blood lead is exceeding 10 μg/dl of blood. Blood sampling, environmental and medical anamnesis, therapy etc. should be carried out (29). Thus, in the year 2002 extensive study took place. Exceeded concentrations in soil, forest ground,

spring and river were confirmed. Especially concerned amounts were found in house and road dust (29).

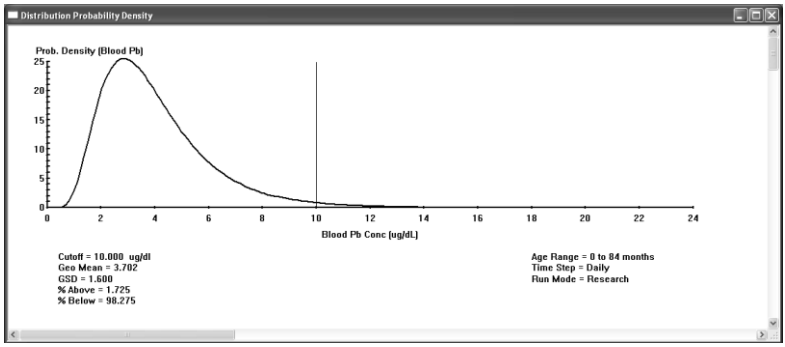


Figure 12. IEUBKwin32 Model 1.0 output – probability density curve.

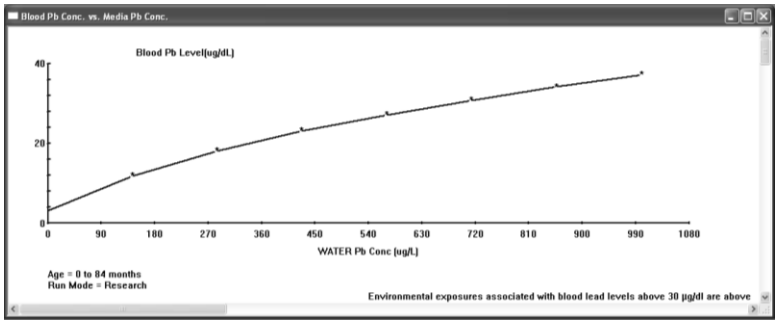


Figure 13. IEUBKwin32 Model 1.0 output – correlation between lead concentration in water and in blood.

Results of another study among three years old children where blood samples were taken during medical examination in period from December 2001 to June 2002 are showing high concentrations of lead in blood (29).

Due to results of different studies the Regional Public Health Institute Ravne started in the year 2003 and 2004 with information and educating campaign for the laity and expert public about health risk of lead, about exposure, intake and how an individual can control this. Regional Public Health Institute started in the year 2004 with monitoring of blood lead concentration in children age of three. Meanwhile polls were taken among parents of three years old children. With such approach important information was gained, which was supposed help to find better solution to solve the problem (29).

The use of the IEUBK model in the case of Upper Mežica valley

The last comprehensive study of lead and other heavy metals burden in Upper Mežica valley which addressed different areas from environment to food, animals and people is The Comparative study of the environment pollution in the Upper Mežica valley between 1989 and 2001 (29). The study results showed that concentrations of heavy metals and especially lead are still high in different media and also that lead burden in human is still high. In 2004 the Regional Institute of Public Health Ravne started the project "Living with Lead" (29). Part of the project was also the testing of blood samples of 3-years old children. In the period 2004 to 2007 179 children from the Upper Mežica valley area were included in the study. Results show that the lead blood levels were elevated (≥10 μg/dl) in 53.1% of children. Elevated blood lead levels were more common in boys (60.4%) than in girls (44.6%).

The aim was to found out if children with highest blood lead burden live in the most polluted area. Data for air, soil and dust lead concentrations from the Comparative study of the environment pollution in the Upper Mežica valley between 1989 and 2001 were put into the IEUBK model and values predicted by the model were compared with observed values from the project "Living with Lead" and from study among three years old children in period from December 2001 to June 2002 (29).

Since the IEUBK model requires specific data groups based on the geographic vicinity of sampling points for environmental media and permanent residences of children were established. For this purpose data were mapped on map of Koroška region. Map Info software package was used. All together 12 groups including 126 tested children were formed. Values predicted by the IEUBK model for blood lead geometrical mean and proportion of children with elevated blood lead levels were compared with the values calculated from observed data in children from same group. Values for each group were compared separately first and for whole area under study afterwards. Only results for groups where at least 5 children were included were discuss separately.

In the following table (Table 5) and figures (Figures 14-26), the calculations carried out are presented. Group No. 10 was chosen as an example. Site specific data for soil, indoor dust, air and drinking water lead concentration was available for this group. Group 10 included 21 children. The same calculations were carried out for all groups.

Table 5. Input data for one of group of Upper Mežica valley children (group 10).

Media	Mark	Concentration	
Dust	CS4		1087 mg/kg
Soil	C5		925 mg/kg
Air	AirR		0.083 μg/m^3
Water	United States		8.4 μg/l
Children	n=21	Blood lead concentration	
	(15 wit BLL over 10 μg/l)	Range: 42-303 μg/l	
		Geometrical mean:130.3 μg/l	
		Proportion with elevated BLL: 71.4%	

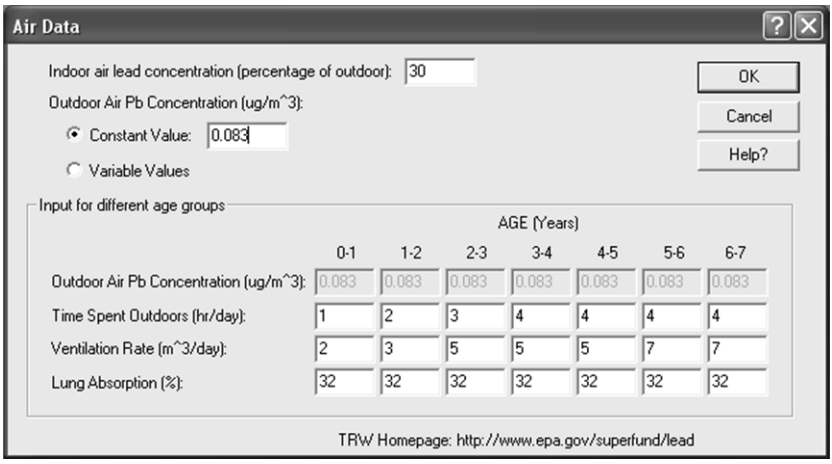


Figure 14. Air Pb concentration (μg/m^3) input for Group 10

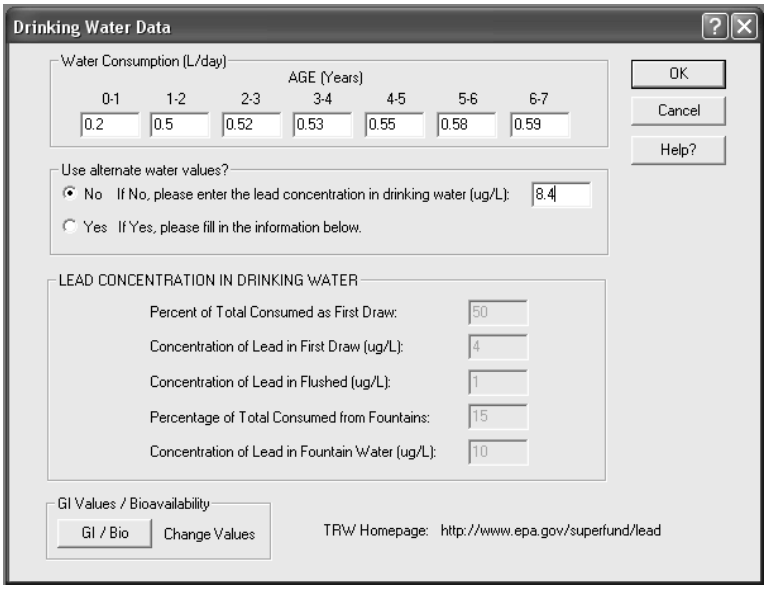


Figure 15. Drinking water Pb concentration (μg/l) input for Group 10

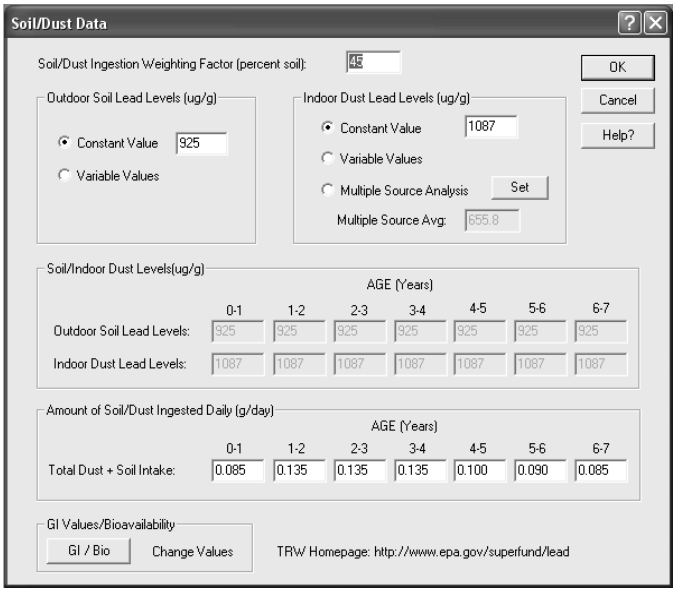


Figure 16. Outdoor Soil and Indoor Dust Pb concentration (μg/g) input for Group 10

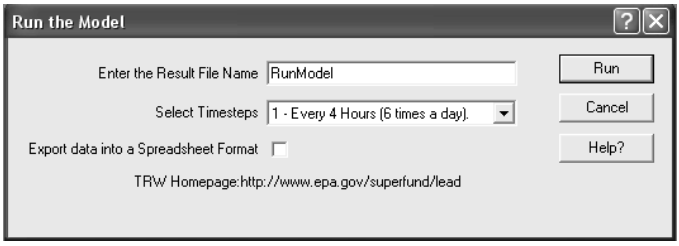


Figure 17. Run the model (for Group 10)

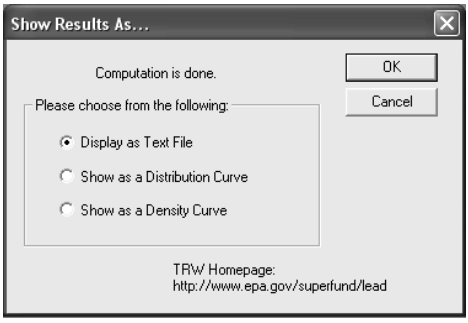


Figure 18. Results can be presented as text file.

```
****** Maternal Contribution: Infant Model ******

Maternal Blood Concentration: 2.500 ug Pb/dL

****************************************
CALCULATED BLOOD LEAD AND LEAD UPTAKES:
****************************************

Year          Air            Diet          Alternate       Water
            (ug/day)       (ug/day)        (ug/day)      (ug/day)
-------------------------------------------------------------------------
.5-1         0.017          2.119           0.000         0.644
1-2          0.029          2.123           0.000         1.542
2-3          0.051          2.469           0.000         1.662
3-4          0.055          2.454           0.000         1.751
4-5          0.055          2.537           0.000         1.950
5-6          0.077          2.750           0.000         2.114
6-7          0.077          3.082           0.000         2.182

 Year      Soil+Dust        Total          Blood
            (ug/day)       (ug/day)        (ug/dL)
-----------------------------------------------------------
.5-1         19.820         22.600          11.8
1-2          30.166         33.860          13.7
2-3          31.248         35.430          12.9
3-4          32.306         36.566          12.5
4-5          25.682         30.224          10.5
5-6          23.757         28.698           9.0
6-7          22.770         28.111           8.0
```

Figure 19. Results for Group 10. Only 3 years old children were included in study so our interest was on blood lead prediction for them. Group is marked with bold text.

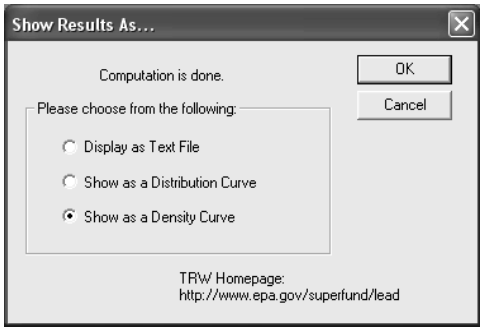


Figure 20. To predict proportion of children with elevated blood lead levels the option ”Show results as a density curve”.

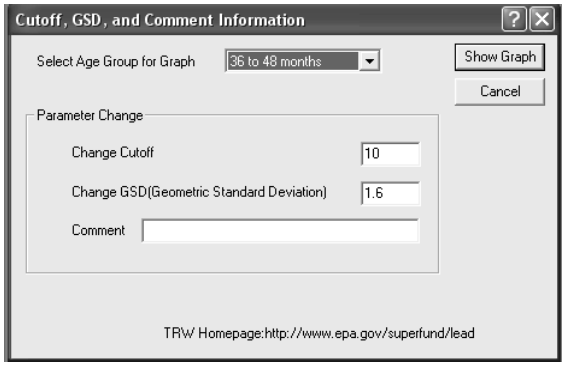


Figure 21. Age group of three years old children was chosen and default values for geometric standard deviation (GSD) and cut-off were used.

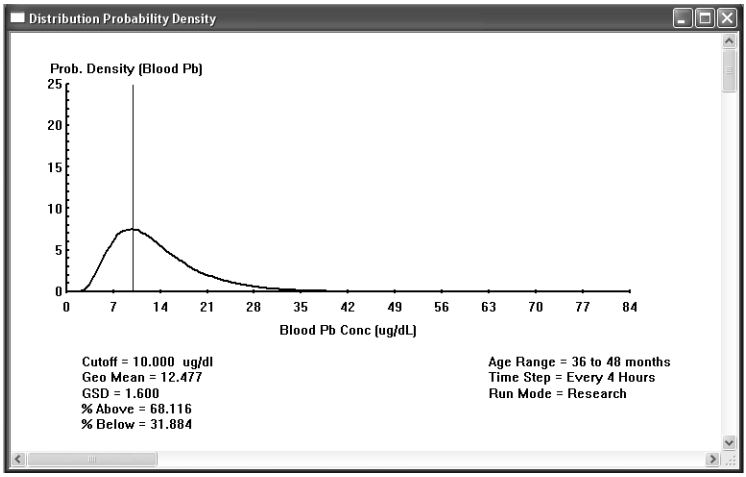


Figure 22. The same calculations were carried out for other groups.

Calculations were also carried out using Batch file which included data for all groups. Batch file was created in text (.TXT) format (Figure 23). Each line represents one group and each column contains data for this group: child's age in months and Pb concentrations for different environmental factors (soil, dust, water, air, paint, etc.). The Batch mode Output file is presented in Figure 26.

```
LEAD6 input data file.
MEZDOL 2007 - NOTE: actual data begins on line 4!!
  ID      FAM  BLOCK  AGE   SOIL  DUST  WATER  AIR  PAINT  PBB
MID001     1    301   42    SSC    SSC     4    0.1      0   2.5
MID002     2    301   42    SSC    SSC     4    0.1      0   2.5
MID003     3    303   42    SSC    SSC     4    SSC      0   2.5
MID004     4    303   42    SSC    SSC     4    SSC      0   2.5
MID005     5    304   42    SSC    SSC     4    0.1      0   2.5
MID006     6    306   42    SSC    SSC     4    0.1      0   2.5
MID007     7    306   42    SSC    SSC     4    0.1      0   2.5
MID008     8    306   42    SSC    SSC   SSC    0.1      0   2.5
MID009     9    307   42    SSC    SSC     4    SSC      0   2.5
MID010    10    307   42    925   1087    8.4  0.083     0   2.5
MID011    11    308   42    SSC    SSC   SSC    0.1      0   2.5
MID012    12    308   42    SSC    SSC     4    SSC      0   2.5
```

*(SSC – site specific concentration)

Figure 23. Batch file for running the procedure created in text (.TXT) format.

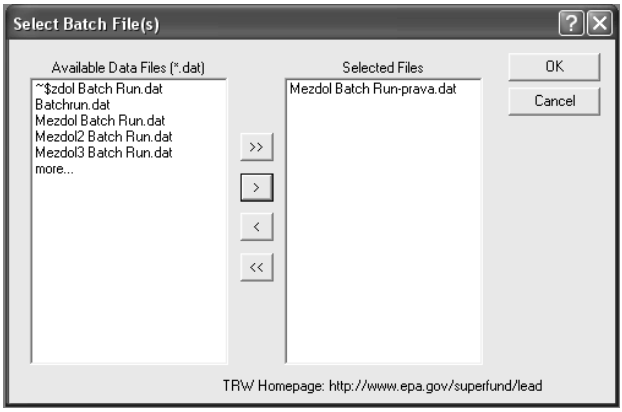


Figure 24. Option "Execute Batch mode run" chosen from menu and calculations carried out using default cut-off and GSD

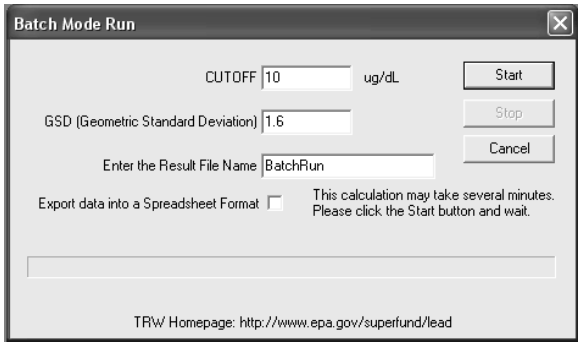


Figure 25. Results are shown in the form extended batch file used as input. Two columns – predicted geometrical mean blood lead level and predicted proportion of children with elevated blood lead levels are added.

Conclusions

To predict proportion of children with elevated blood lead levels for whole area under study predicted proportions for each group were multiplied with number of children in that same group. Results were added up and divided by number of children included in all groups. The goal was to assess the usefulness of the model as a resource of the information needed for the determination for directing the further research regarding the lead contamination linked with the health outcomes in the Upper Mežica valley.

LEAD BATCH MODE OUTPUT FILE

==========

Model Version: 1.0 Build 263
User Name:
Date:
Site Name:
Operable Unit:
Run Mode: Research

---

* : signify default values used in place of missing input data.
\# : signify surrogate values entered (determined) by user.
---: signify missing input data.
PBB & PRED are the observed and predicted blood Pb levels in ug/dL.

Percent exceedance was calculated using values of GSD and PbB Cutoff as follows:
GSD = 1.600
PbB Cutoff (C) = 10.000 ug/dL

Input File: Mezdol Batch Run-prava.dat

ID	FAM	BLK	AGE (mon)	SOIL (ug/g)	DUST (ug/g)	WATER (ug/L)	AIR (ug/m^3)	OTHER (ug/day)	PBB (ug/dL)	PRED (ug/dL)	P(PbB>C) (%)
MID001	001	301	42	SSC	SSC	4.0	0.10	0.000	2.5	PV	PV
MID002	002	301	42	SSC	SSC	4.0	0.10	0.000	2.5	PV	PV
MID003	003	303	42	SSC	SSC	4.0	SSC	0.000	2.5	PV	PV
MID004	004	303	42	SSC	SSC	4.0	SSC	0.000	2.5	PV	PV
MID005	005	304	42	SSC	SSC	4.0	0.10	0.000	2.5	PV	PV
MID006	006	306	42	SSC	SSC	4.0	0.10	0.000	2.5	PV	PV
MID007	007	306	42	SSC	SSC	4.0	0.10	0.000	2.5	PV	PV
MID008	008	306	42	SSC	SSC	SSC	0.10	0.000	2.5	PV	PV
MID009	009	307	42	SSC	SSC	4.0	SSC	0.000	2.5	PV	PV
MID010	010	307	42	925.0	1087.0	8.4	0.08	0.000	2.5	12.49	68.179
MID011	011	308	42	SSC	SSC	SSC	0.10	0.000	2.5	PV	PV
MID012	012	308	42	SSC	SSC	4.0	SSC	0.000	2.5	PV	PV

*(SSC – site specific concentration, PV– predicted value)

Figure 26. The Batch mode Output file.

Results showed that areas with the highest lead concentrations in yard soil and house dust do also have the largest proportions of children with elevated blood lead levels.

Values predicted by the IEUBK model were in satisfactory accordance with the values calculated from empirical data for 6 of 8 groups with at least 5 children. Predicted proportions of children with elevated blood lead level were a bit higher than observed ones for most groups and also for the whole area under study. According to prediction of the model 67.9% of children included in groups should had elevated blood lead levels, what would mean 86 of 126 children, in fact there were only 67 (53.2%) such children. Perhaps even better accordance of predicted values and observed data could be achieved by using more site specific data, which, to our regret, we did not have at the moment. Beside more sampling points for soil, dust, air and drinking water also data for the bioavailability of lead and abut behaviour patterns of local inhabitants would be welcome.

The results of calculations showed that house dust and backyard soil contribute large proportions in the overall exposure of children from the Upper Mežica valley to lead and that environmental remediation strategies should focus on them.

EXERCISE

Task 1

Carefully read the theoretical background of this module and discuss with other students:

- Toxic effects of metals on human body.
- IEUBK model as a tool for prediction of blood lead level.
- Cases related to metals.

Task 2

Make several cases with different concentrations of lead in environmental factors and compare results.

Task 3

Find similar models for exposure simulation in case of other environmental pollution on website.

REFERENCES

1. Health Canada. Effects of lead on human health, 2004:2-1. Available from URL: http://www.hc-sc.gc.ca/hl-vs/iyh-vsv/environ/lead-plomb-eng.php. Accessed: April 3, 2008.
2. Tong S, Yasmin SE, Prapamontol T. Environmental lead exposure: a public health problem of global dimensions. Environ Health. 2007;8:1068-77.
3. Toscano DC, Guilarte TR. Lead neurotoxicity: From exposure to molecular effects. Brain Research Rev. 2005;49:529-54.
4. U.S. Environmental Protection Agency. Lead poisoning: A historical perspective [homepage on the Internet]. Available from URL: http://www.epa.gov/history/topics/perspect/lead.htm. Accessed: May 12, 2008.
5. Grout J. Encyclopaedia Romana. Rome, the home of empire and of all perfection. Available from URL: http://penelope.uchicago.edu/~grout/encyclopaedia_romana/wine/leadpoisoning.html. Accessed: April 3, 2008.
6. Devine J. A History of Lead Poisoning in the Western World [homepage on the Internet]. Available at URL: http://ezinearticles.com/?A-History-of-Lead-Poisoning-in-the-Western-World&id=1647778. Accessed: February 27, 2008.
7. Fewtrell L, Kaufmann R, Prüss-Üstün A. Lead: Assessing the environmental burden of disease at national and local levels. Geneva: World Health Organization; 2003.
8. Hooker PJ, Nathanail CP. Risk based characterisation of lead in urban soils. Chemical Geology. 2006;226:340-51.
9. U.S. Environmental Protection Agency. Lead in drinking water in schools and non-residential buildings. Available from URL: http://www.epa.gov/dclead/leadinschools.pdf. Accessed: April 5, 2008.
10. San Francisco Department of Public Health. Lead in the body [homepage on the Internet]. Available from URL: http://www.sfdph.org/cehp/lead/framecontents.htm. Accessed: February 27, 2008.
11. National Institute of Environmental health Sciences. Lead and your health [homepage on the Internet]. Available from URL: http://www.niehs.nih.gov/health/docs/lead-fs.pdf. Accessed: February 15, 2008.

12. U.S. Environmental Protection Agency. Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for lead in children. Available from URL: http://www.epa.gov/superfund/health/contaminants/lead/products.htm. Accessed: June 21, 2009.
13. Gulson BL, Mahaffey KR, Jameson W, et al. Impact of diet on lead in blood and urine in female adults and relevance to mobilization of lead from bone stores. Environ Health Perspect. 1999;107:257-63.
14. Horvat M, Janet E, Ivartnik M. The level of lead in blood of three years children from Mežiška dolina (in Slovene). In: Cvahtetovi dnevi 2007. Zbornik prispevkov. Ljubljana: 28.september 2007. Medicinska fakulteta, Katedra za javno zdravje; 2007. p.16-24.
15. Eržen I, Janet E. The level of lead in blood of three years children that are living in area of Zgornje and Spodnje Mežiške doline (In Slovene). Zdrav Var. 2005;44: 18-25.
16. U.S. Environmental Protection Agency. Lead in paint, dust and soil. Available from URL: http://www.nachi.org/lead.htm. Accessed: January 13, 2008.
17. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry. Public health statement: Lead. Atlanta: Agency for Toxic Substances and Disease Registry; 2007. www.atsdr.cdc.gov/toxprofiles/tp13-c1-b.pdf. Accessed: January 13, 2008.
18. Kofi AD. Public health risk assessment for human exposure to chemicals [homepage on the Internet]. Available from URL: http://www.cplbookshop.com/contents/C875.htm. Accessed: December 23, 2007.
19. Powell MR. Science at EPA: Information in regulatory process. Washington D.C: Resources for the Future; 1999.
20. World Health Organization,European Environment and Health Information System. Blood lead levels in children. Fact sheet 4.5; May 2007 [homepage on the Internet]. Available at URL: http://www.euro.who.int/Document/EHI/ENHIS_Factsheet_4_5.pdf. Accessed: January, 15 2008.
21. U.S. Environmental Protection Agency. User's guide for the Integrated Exposure Uptake Biokinetic Model for lead in children (IEUBK) Windows®. Available from URL: http://www.epa.gov/superfund/lead/products/ugieubk32.pdf. Accessed: March 14, 2008.
22. U.S. Environmental Protection Agency. Short sheet: Overview of the IEUBK model for lead in children. Washington DC: U.S. Environmental Protection Agency, Office of Emergancy and Remedial Response; 2002. Available from URL: http://www.epa.gov/superfund/health/contaminants/lead/products/factsht5.pdf. Accessed: March 20, 2008.
23. National Research Council, Committee on Superfund Site Assessment and Remediation in the Coeur d' Alene River Basin. Superfund and mining megasities. Lesson from the Coeur D' Alene river Basin. Washington D.C.: National Academies Press; 2005.
24. Were HF, Njue W, Murungi J, Wanjau R. Use of human nails as bio-indicators of heavy metals environmental exposure among school age children in Kenya. Science Tot Environ. 2008;393:376-84.
25. Khan S, Cao Q, Zheng Y M, Huang Y M, Zhu Y G. Health risks of heavy metals in contaminated soils and food crops irrigated with wastewater in Beijing, China. Environmental Pollution. 2007;152:692-86.
26. Eržen I. Degree of exposure to lead, cadmium and mercury with food of inhabitants in Slovenia (in Slovene). Doctoral thesis. Ljubljana: Faculty of Medicine; 2004.
27. World Health Organization. Lead. In: World Health Organization. Air quality guidelines for Europe. Copenhagen: World Health Organization, Regional Office for Europe; 2001. p.1-17. Available from URL: http://www.euro.who.int/document/aiq/6_7lead.pdf. Accessed: March 20, 2008.
28. Nordic Council of Ministers. Lead review. Copenhagen: Nordic Council of Ministers; 2003: Available from URL:

http://www.chem.unep.ch/Pb_and_CD/SR/Files/Submission%20GOV/Submis_GOV_DNK/NMR_lead.pdf. Accessed: March 20, 2008.
29. Mandl N. Health promotion in local community with the purpose to decrease lead exsposure (in Slovene). Ravne: Regional Public Health Institute Ravne; 2007.

RECOMMENDED READINGS

1. Tong S, Yasmin SE, Prapamontol T. WHO. Environmental lead exposure: a public health problem of global dimensions. Environment Health. 2007;8:1068-77.
2. U.S. Environmental Protection Agency. User's guide for the Integrated Exposure Uptake Biokinetic Model for lead in children (IEUBK) Windows®. Available from URL: http://www.epa.gov/superfund/lead/products/ugieubk32.pdf. Accessed: March 14, 2008.
3. U.S. Environmental Protection Agency. Short sheet: Overview of the IEUBK model for lead in children. Washington DC: U.S. Environmental Protection Agency, Office of Emergancy and Remedial Response; 2002. Available from URL: http://www.epa.gov/superfund/health/contaminants/lead/products/factsht5.pdf. Accessed: March 20, 2008.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**EPIDEMIOLOGICAL INDICATORS OF ENVIRONMENTAL HEALTH**
Module: 2.1.3	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Alexandra Cucu**, MD, Ph.D, Lecturer Department of Public Health and Management, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania **Maria Nitescu**, MD, Ph D, Lecturer Department of Hygiene and Ecology, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania
Address for correspondence	**Alexandra Cucu** Department of Public Health and Management, University of Medicine and Pharmacy "Carol Davila", Bucharest, Dr Leonte str. 1-3, Bucharest Romania phone: +40 21 3180713; fax: +40 21 3180714 mail: alexandracucu2003@yahoo.com
Keywords	Environment, health, indicators
Learning objectives	After completing this module students should: • understand environmental health concept; • be aware of existence and usefulness of the environmental health information system; • describe the features of environmental health in their country; • be familiar with the existing national information on environment and health; • to use available information and tools on environmental conditions in the evidence-based policy making process.
Abstract	The core set of 29 environmental health indicators are the main component of the Environmental Health Information System (ENHIS) that represents comprehensive information and knowledge system, which will generate and analyze environmental health information to support relevant policies in Europe. They provide comparable information on environment and health in the pan-European Region and includes an analysis of core issues across the Region.
Teaching methods	An introductory lecture gives the students an overview of environmental health concepts and associated indicators developed by the European Environment and Health Information System project. The theoretical knowledge is illustrated by a fact sheet presentation and a case study. After introductory lectures students first carefully read the recommended readings. Afterwards they discuss the characteristics of environmental health indicators and elaborate a work plan for a country indicator based report on environmental health topics.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme, public health residents
Assessment of students	Seminar paper - elaboration of an indicator based report for environmental health national profile.

EPIDEMIOLOGICAL INDICATORS OF ENVIRONMENTAL HEALTH

Alexandra Cucu, Maria Nitescu

THEORETICAL BACKGROUND

Environmental factors, even less visible than other health determinants as unhealthy behaviours or health services, are responsible for approximately a quarter of the global burden of disease. Recent scientific evidences are contributing to a better understanding of the importance of environmental factors in health. The main questions looking for answer in environmental health are:

- what are the linkages between environmental exposures and human health outcomes?
- how do we identify and prioritize environmental health issues?
- how and when do we know whether we have an environmental health problem?
- how do we compare with other regions and countries?
- do current actions have any impact on environmental health?
- what does and what does not work?

In this context, reliable information, synthesised in a standardised format, as an essential tool for prioritizing actions related to environmental exposures and their health effects, as well as for monitoring the effectiveness of the actions taken, is a must.

Environmental health indicators

An environmental health indicator (EHI), according to D. Briggs definition (1) is “a measure which indicates the health outcome due to exposure to an environmental hazard, thus consisting of an environmental indicator or a health indicator plus a known environmental-exposure health effect relationship”. Rather clear but in the same time complex, due to broad sense of the concepts included. That’s why, for understanding and measuring environmental health effects the next term to be clarified is: environment -from the health relationship perspective.

Environment

According to Last (2) definition the environment is: "All that which is external to the human host. Can be divided into physical, biological, social, cultural, etc., any or all of which can influence health status of populations”. This is the broad view on environment as including anything that is not genetic, although even a wider perspective could be argued if we consider the fact that even genes are influenced by the environment in the short or long-term. This approach might be represented by Smith, Corvalàn and Kjellström framework (3), adapted in Figure 1, best illustrates the potential definition of the environment, from the most inclusive to the most restrictive definition.

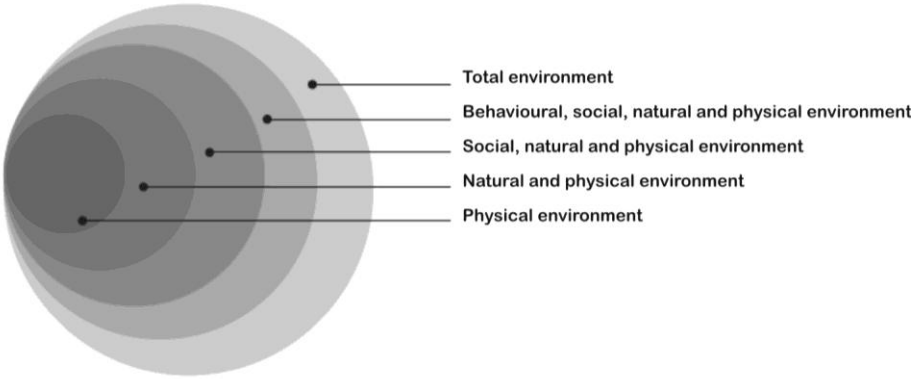


Figure 1. Definition of environment (Source: Smith, Corvalàn and Kjellström) (3).

The actual approach, used for measuring environmental health (EH) impact, focus on those parts of the environment that can be modified by short-term or longer-term interventions, excluding the social and cultural health determinants, which are usually independent of the environment (e.g. cultural pressures on lifestyle, unemployment). As result, the practical definition of the environment, provided by the World Health Organization (WHO) 2006 Report (4), defining environment as "all the physical, chemical and biological factors external to the human host, and all related behaviours, but excluding those natural environments that cannot reasonably be modified".

In this context the definition includes the main environmental modifiable factors as: pollution of air, water, or soil with chemical or biological agents; UV and ionizing radiation; noise, electromagnetic fields; occupational risks; built environments, including housing, land use patterns, roads; agricultural methods, irrigation schemes; man-made climate change, ecosystem change; behaviour related to the availability of safe water and sanitation facilities, such as washing hands, and contaminating food with unsafe water or unclean hands. Consequently the factors included by the current EH definition are: alcohol and tobacco consumption, drug abuse; diet (although it could be argued that food availability influences diet); the natural environments of vectors that cannot reasonably be modified (e.g. in rivers, lakes, wetlands); natural biological agents, such as pollen in the outdoor environment; person-to-person transmission that cannot reasonably be prevented through environmental interventions such as improving housing, introducing sanitary hygiene, or making improvements in the occupational environment.

Impact of environment on health

According to WHO 2006 Report (4) an estimated 24% of the global disease burden and 23% of all deaths can be attributed to environmental factors. More, from the 102 major diseases, disease groupings and injuries covered by the WHO Report, the

environmental risk factors contributed to disease burden in more than 80 categories. The diseases with the largest absolute burden attributable to modifiable environmental factors are: diarrhoea (94% of the diarrhoeal burden of disease); lower respiratory infections (20-42% in developing countries) injuries from road traffic accidents (40%), injuries, chronic obstructive pulmonary disease (42%), and malaria.

Additionally an uneven distribution was registered, concentrating in developed regions cardiovascular diseases (7-times higher than in developing regions), cancer (rates were 4-times higher) and diabetes mellitus, meanwhile developing countries, meanwhile, carry a heavier burden of disease from unintentional injuries and road traffic injuries attributable to environmental factors.

Globally, the per capita number of healthy life years lost to environmental risk factors was about 5-times greater in children under five years of age than in the total population. They represents the most vulnerable group, bearing over 40% of this burden (5-7), diarrhoea, malaria and respiratory infections, being among the biggest killers of children under five years old, especially in country in development. On average, children in developing countries lose 8-times more healthy life years, per capita, than their counterparts in developed countries from environmentally-caused diseases.

Main diseases contributing to the environmental

The list of the 24 diseases with the largest environmental contribution to overall burden of disease, according to the last WHO Report (8) on impact of environmental factors, is illustrated in Figure 2.

In this context, environmental health decision-making has become an increasingly complex and at times contested process and as such there is an increased need for tools to facilitate and enhance the outcome of this process.

As response, providing decision-makers with appropriate information regarding health effects attributable to environmental risks became a priority, confirmed and supported by a strong political commitment, most of the current global policies from Agenda 21 (9), the Millennium Development goals (10), the EU Public Health Programme, the European Environment and Health Strategy (11-13), and the specific Budapest 4th Ministerial Conference on Environment and Health Declaration (14) highlighting the need for good quality information for decision-making.

Consequently, through a series of projects aiming establishment of Environmental Health Information System (ENHIS) Supporting Policy Making projects (ENHIS1 and ENHIS2 projects) (15,16), coordinated by the World Health Organization, Regional Office for Europe, European Centre for Environment and Health, Bonn, Germany and co-funded by the Directorate-General for Health and Consumer Protection (DG SANCO) of the European Commission (EC) the ENHIS system was designed and developed aiming to address the current priority policies, including those related to children's environment and health, in accordance to Children's Environment and health Action Plan for Europe (17).

The actual Environmental Health Information System represents a comprehensive information and knowledge system, which generates and analyse environmental health information to support relevant policies in Europe. It provides

comparable information on environment and health in the pan-European Region and includes:

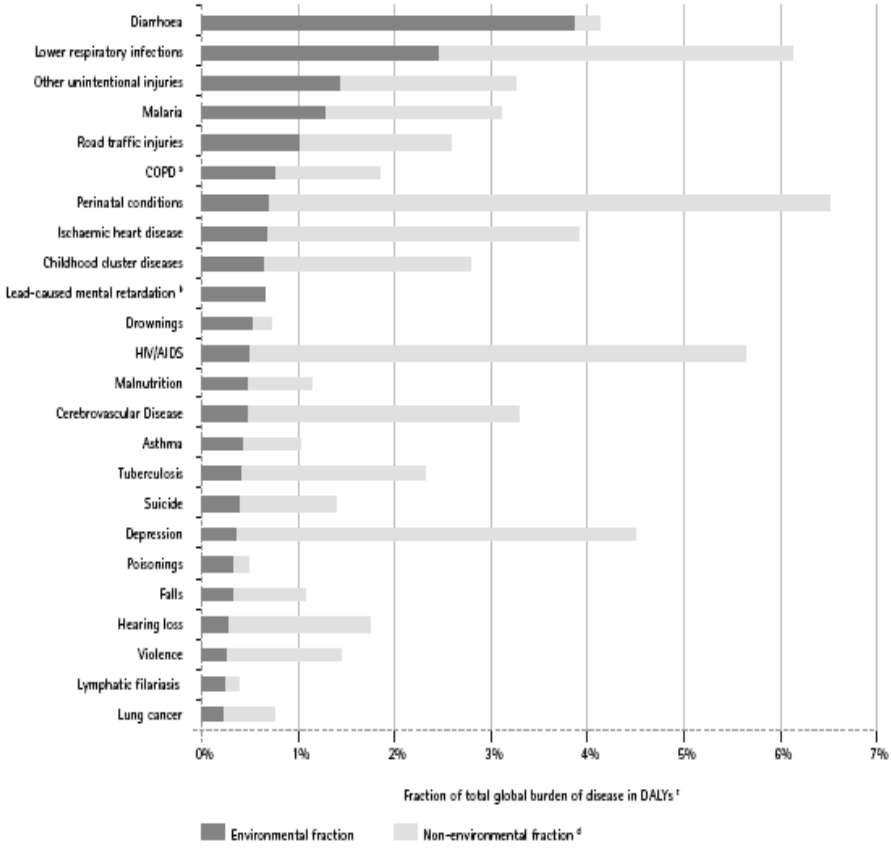


Figure 2. Disease with the largest environmental contribution (Source: Prüss-Ustün A and Corvalán C. 2006) (4). LEGEND: a=COPD- chronic obstructive pulmonary disease; b=lead-caused mental retardation is defined in the WHO list of diseases for 2002, accessed at: www.who.int/evidence; c=DALY represents a weighted measure of death, illness and disability; d=for each disease the fraction attributable to environmental risks is shown in dark colour, light colour plus dark colour represents the total burden.

- a core set of indicators, selected on the basis of relevance and availability of data, describing environmental exposures, health effects and policy measures for these issues.
- a series of indicator-based assessments ("fact sheets"), providing an analysis of core issues across the Region;
- country information for the 53 Member States of the WHO European Region;

- an overview of policies on core issues, at both national and international level, and topic-based comparative policy assessments across 18 countries;
- methodological guidance on the core set of indicators, to facilitate harmonization and dissemination of methods used within ENHIS and validated by international experts;
- guidance and examples of health impact assessments, showing the potential health benefits of policy actions and interventions aimed at reducing exposure to environmental risk factors.

Environmental health indicators

In general terms, health indicators are summary measures that capture relevant information on different health attributes and dimensions, and the performance of the health system. They represent the "building blocks for health situation analysis" (18) allowing: monitoring trends in the state of the environment, in order to identify potential risks to health; monitoring trends in health, resulting from exposures to environmental risk factors, in order to support policies; comparing different areas in terms of their environmental health status, identifying priority areas for action; evaluate the effects of policies or other interventions on environmental health; support awareness raising about environmental health issues; investigating potential links between environment and health

The, ENHIS indicators (8,19) were designed to:

- enable monitoring of children's environmental health risks, their determinants and effects of the intervention;
- provide appropriate information to countries to monitor the state of children's environmental health, allow trends to be established and to support national policies and action programmes;
- provide a sustainable basis for reporting and dissemination of evidence-based information (i.e. there is a policy need plus there is an established link between the exposure and health outcome) on children's environmental health avoiding duplication and ensuring continuity;
- provide a basis for improvement of existing monitoring and surveillance systems by pointing out priority data gaps in order to inform policy making decisions.

As any indicator, in order to be effective, they should satisfy a number of different criteria, as described in Box 1.

These criteria, defining the ideal indicator, are, some of them, at a certain extent, impossible to be met. For example, from the cost effectiveness point of view, ideally indicators must be developed on the basis of data which already exist, even if the appropriateness of those data, collected for other purposes, may affect the scientific validity of the indicator. By the other hand, the specific use perspective will influence the above mentioned qualities. For example an indicator intend to monitor trends over time, should be based on data which are spatially representative, but not necessarily spatially intensive or complete. The same indicator, used to examine geographic patterns and identify 'hotspots', will need to be based on data which are spatially detailed and comprehensive: temporal variations will be less important. Even

if some of the requirements could be reduced the scientific quality should be always ensured.

If to the broad purposes and quality criteria we add the complexity of the environmental factors and also the diversity of associated effects, we understand why elaborating environmental health indicators are one of the most challenging activity. public health monitoring tool. Based on these criteria the indicators were elaborated, in order to present the links between environment, health outcomes and actions in the DPSEEA framework developed by Corválan et al. (1996) (7), figure 3. This defines driving forces (D), that lead to pressures on the environment (P), which in turn change the state of the environment (S), resulting in human exposures (Ex) and then to health effects (E). Actions (A) can be taken at any point during the chain in order to prevent and mitigate health effects.

Box 1. Indicators criteria

- ***validity*** *(effectively measures what it attempts to measure)*
- ***reliability*** *(repeated measurements in similar conditions produce the same results).*
- ***specificity*** *(measures only the phenomena that it is meant to measure),*
- ***sensitivity*** *(has the capacity to measure changes in the phenomena that it is meant to measure),*
- ***measurability*** *(is based on available or easy to obtain data),*
- ***policy-relevance*** *(is capable of providing clear responses to key policy issues)*
- ***cost-effectiveness*** *(results justify the investment in time and other resources)*
- ***comprehensibility*** *- easy to use and interpret by analysts, as well as understandable by information users, such as managers and decision-makers. Important attributes to insure quality of the set of indicators are:*
- ***integrity*** *(good spatial and temporal coverage and no missing data)*
- ***internal consistency*** *(when seen alone or in a group, the values of the indicators are realistic and coherent and do not contradict themselves).*
- ***appropriate level of disaggregation*** *- availability of information at the specific required level*
- ***availability of data sources and regular operation of information systems****.*
- ***simplicity of the instruments and methods used in compiling***
- ***timeless*** *the time lag between data collection, analysis and dissemination of an indicator used in monitoring must be short for it to maintain its relevance*
- ***flexibility*** *-to be able to reflect changes in the observed phenomena*
- ***relevance****-to respond to the purpose it was designed*
- ***ethicality****-to ensure ethical use of information*
- ***transparency****-information to be accessible in a transparent manner*
- ***scientific sound-based on best scientific evidences***
- ***stabile and robust****- not influenced by minor differences in the source of - the data*

The ENHIS indicators responds to the commitment made thought the Declaration of the Fourth Ministerial Conference on Environment and Health, held in Budapest 2004 (‘The Budapest Conference’) (14), which focused on ‘The future for our children’, recognizing the need to address the rights of children, their health, and their particular vulnerability towards environmental risks, as well as to respond to

emerging environmental concerns. For implementation of the Declaration it was adopted the Children's Environment and Health Action Plan for Europe (CEHAPE), an international instrument negotiated with Member States to develop and manage environmental health indicators. CEHAPE sets four 'Regional Priority Goals (RPGs)' which encapsulate key themes for action on children's health in relation to environmental factors. The environmental health indicators were designed as tools not only for monitoring the status and trends in European countries but also for informing and communicating with a wide range of users. The methodology for indicators was developed, in order to ensure international comparability, focusing on environmental factors most relevant to health, health outcomes most influenced by the environment, and policy action deemed to reduce and prevent the risks.

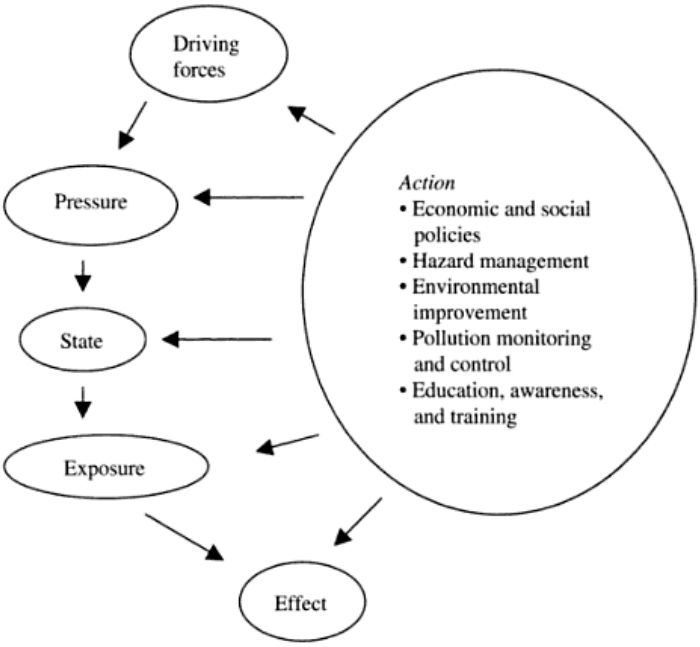


Figure 3. The DPSEEA framework (Adapted from Corvalan et al. 2000) (7).

A core set of 26 indicators (Table 1) was selected during a process involving multiple working groups and consultations (8), using the criteria of scientific credibility, a focus on children's EH and relevant policy action such as CEHAPE, and feasibility. These are available on the ENHIS web portal (20) includes:

An information base has been created for the 26 children's EH indicators. They provide evidence clearly and concisely to support the development of action which benefits public health and the environment, and to track the progress of its implementation.

Table 1. Overview of the core set of the environmental health indicators (Source: ENHIS web site) (21).

Group	Indicator
RPG I	**Gastrointestinal health related to safe water and adequate sanitation**
1.	Outbreaks of waterborne diseases
2.	Public water supply and access to improved water sources
3.	Wastewater treatment and access to improved sanitation
4.	Bathing water quality
RPG II	**Healthy and safe transport and mobility, home safety and physical activity**
5.	Mortality from road traffic injuries in children and young people
6.	Mortality in children and adolescents from unintentional injuries (falls, drowning, fires and poisoning)
7.	Prevalence of excess body weight and obesity in children and adolescents
8.	Percentage of physically active children and adolescents
9.	Policies to promote safe mobility and transport for children
10.	Policies to reduce and prevent unintentional injuries from falls, drowning, poisoning, fires and choking
11.	Policies to reduce and prevent excess body weight and obesity in children and adolescents
RPG III	**Respiratory health and clean air**
12.	Prevalence of asthma and allergies in children
13.	Infant mortality from respiratory diseases
14.	Exposure of children to outdoor air pollution (particulate matter)
15.	Exposure of children to environmental tobacco smoke
16.	Children living in homes with problems of damp
17.	Proportion of children living in homes using solid fuel
18.	Policies to reduce the exposure of children to environmental tobacco smoke
RPG IV	**Health through environment free of hazardous chemicals, physical and biological factors**
19.	Incidence of childhood leukaemia
20.	Incidence of melanoma in people aged under 55 years
21.	Persistent organic pollutants (POPs) in human milk
22.	Exposure of children to chemical hazards in food
23.	Levels of lead in children's blood
24.	Radon levels in dwellings
25.	Work injuries in children and young people
26.	Policies to reduce the excessive exposure of children to ultraviolet radiation

LEGEND: RPG- Regional Priority Goals group

Also Fact sheets have been prepared for the core set of indicators, providing information on the environment and health context, the policy relevance and context, as well as suggestions for further monitoring. When available, case studies of health impact assessment (HIA) are also presented.

Based on the information provided, an indicator based report for identifying country status and priorities can be elaborated, as in the following case study for Romania environmental health profile.

CASE STUDY: ROMANIA ENVIRONMENTAL HEALTH PROFILE

Environmental background:

Romania's past focus on heavy industry and unsustainable production and transport models has generated a general wide spread industrial pollution. According to environmental performance reviews, toxic air emissions, industrial waste pollution in waterways represented the most significant environmental hazard in Romania. However, beginning with 2004 impending membership in the European Union (EU) has stimulated substantial upgrading of environmental monitoring and legislation in order to comply with EU standards by 2007. In 2006, the European Commission listed Romania's two main remaining issues as increased transparency of environmental decisions and improved waste management (10). Romania has received probably the numerous and the longest transition period from all the newly integrated countries to reach full compliance on environmental matters.

RPG I indicators

This group includes indicators focusing on the quality, accessibility of drinking water, bathing water and their effects on health.

Despite the existence and tradition of surveillance environmental health surveillance systems for waterborne outbreaks, quality of drinking water and authorised natural bathing water the lack of data available in international databases where recorded as blank areas on the EH map.

The indicator on public water supply and access to improved water sources

The indicator on public water supply and access to improved water sources estimates the achievement of the minimum requirements for access to an adequate supply of piped and safe water in the home. It is a core indicator for risks related to water and hygiene. The EUROSTAT data, from Figure 4, covering 22 countries in Europe, reveals a clear East-West gap, Romania, being in on of the worst situation regarding the continuous access to adequate amounts of safe drinking-water at home.

The statistics of the UNICEF/WHO Joint Monitoring Programme database, gives a broader view of situation in the Region revealing that an important proportion of the population has poorer access to improved drinking-water sources, particularly in rural areas, where less than 20% of the population has adequate accessibility to improved water sources. Even the indicator estimates mostly the proportion of the general population with access to piped water in the home, it also provides an estimate of the number of people, including children who are potentially exposed to water-related health risks, which is one of the greatest of the EU, including Romania behind countries with as Tajikistan, Kazakhstan, Albania, with relative lower income and less economic potential.

If additionally we consider the quality criteria's compliance for those Romania received the longest transition period from the all integrated countries, we might have a clearer view of the associated health risks for the population health, accounting direct and mostly indirect to the significant values of the total burden of disease in the Eastern European region.

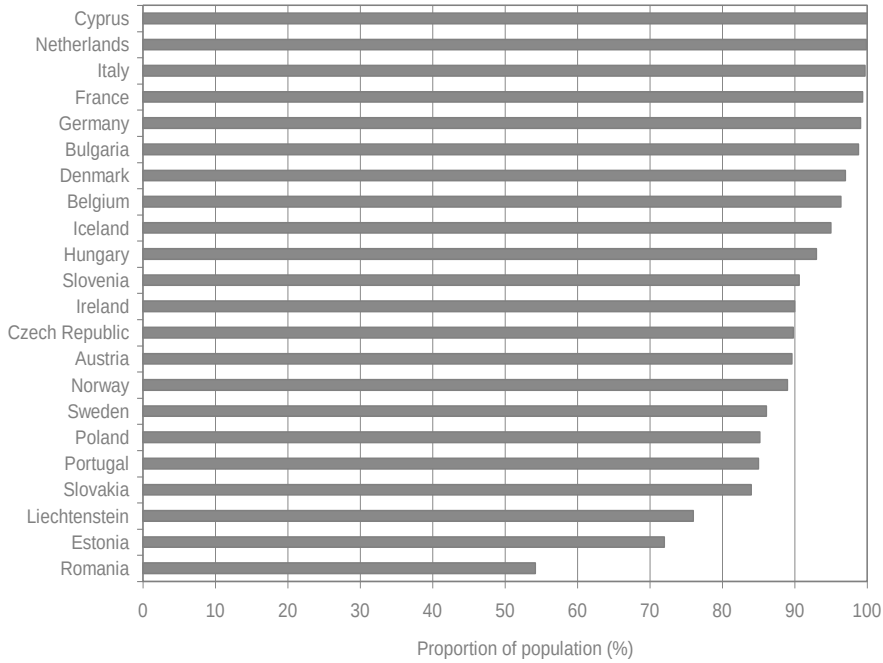


Figure 4. Proportion of population connected to public water supply in Europe, 2002, or latest available (Source: ENHIS web page) (21).

In this context, fostering advocacy and partnership for rural development, massive investments in most vulnerable regions must be a priority where action should be taken to increase the proportion of the (child) population with access to safe drinking-water.

Wastewater treatment and access to improved sanitation

Wastewater treatment and access to improved sanitation indicator estimates the potential level of pollution from domestic point sources, entering the aquatic environment and the associate percentage of the population at risk of infection via the faecal-oral route, especial diarrhoeal diseases, due to the absence of adequate sewage and sanitation disposal systems. The EUROSTAT data from Figure 5 situates Romania, again in one of the worst situation, especially for rural areas where the population connected to a sewerage system, (private septic tanks or dry sanitation where excluded) is less than 10%, on of the last places within the EU region.

In order to comply to the Directive concerning urban wastewater treatment requirements Romania will receive substantial support for improving the wastewater treatment for all agglomerations of more than 2,000 population equivalents (p.e.) with collecting systems. Secondary (biological) treatment must be provided for all agglomerations of more than 2,000 p.e. discharging into fresh waters and estuaries and for all agglomerations of more than 10,000 p.e. discharging into coastal waters. Also numerous rural zones, where many environmental risk factors are aggregated are not

included in those EU supported areas. That is why special emphasis should be paid for supporting their development, too.

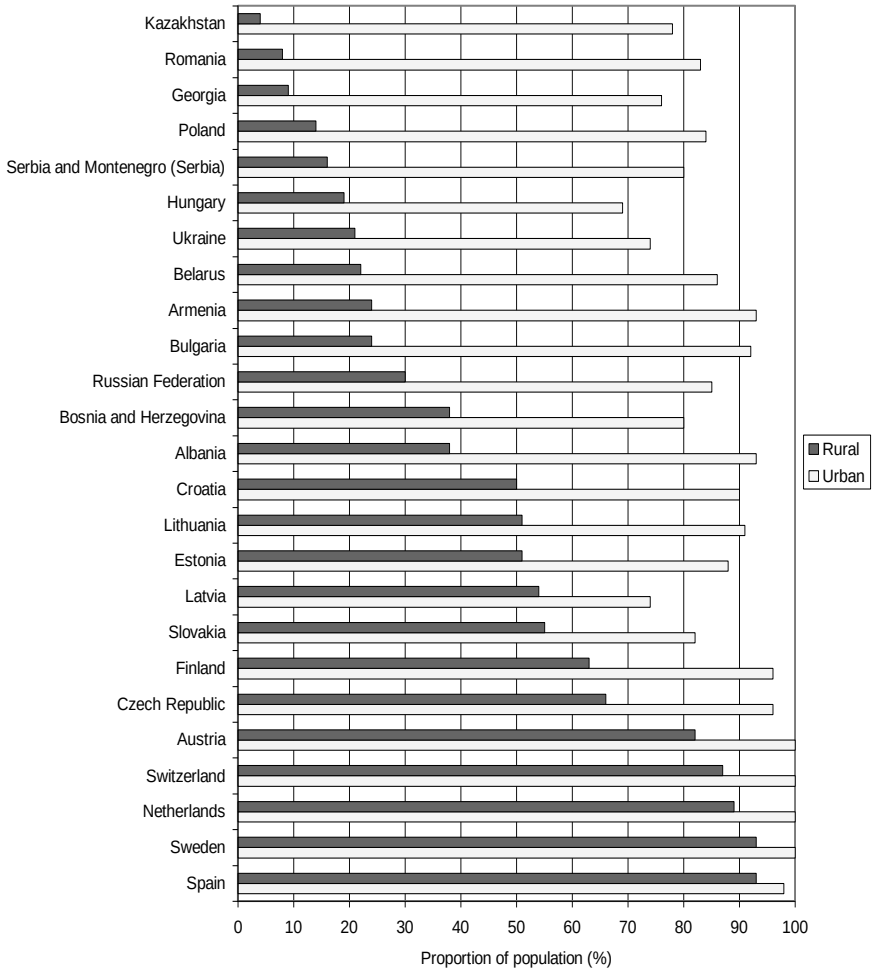


Figure 5. Percentage of the population connected to sanitation facilities in urban and rural areas, 2004 (Source: ENHIS web site) (21).

RPG II indicators

The second priority strain consist in indicators focusing on road traffic accidents, unintentional end external causes injuries, safe mobility and transport for children and policies to reduce and prevent excess body weight and obesity in children and adolescents.

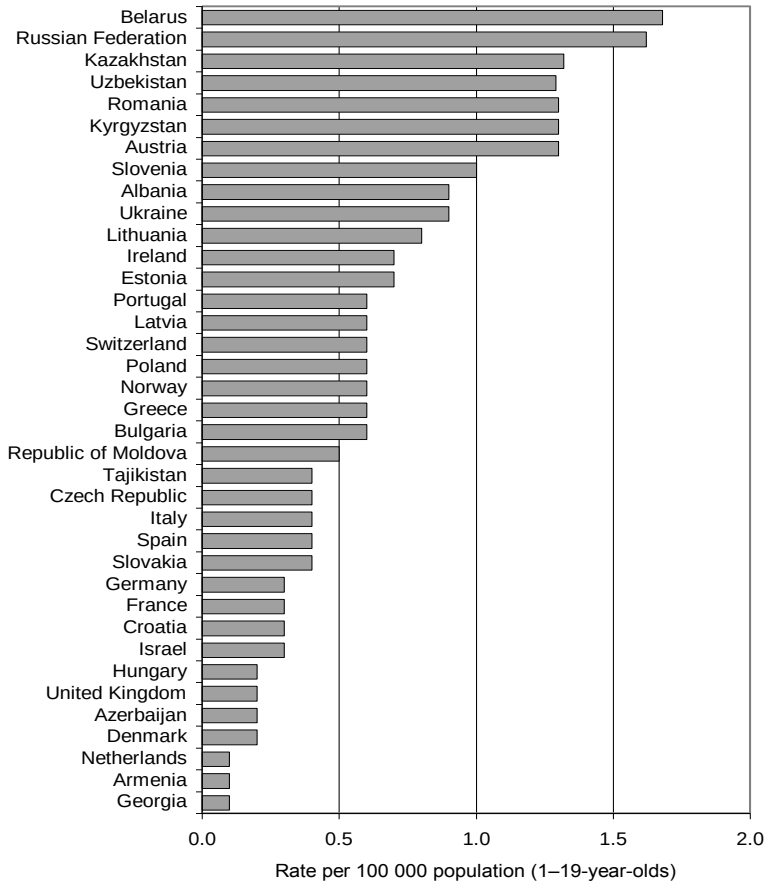


Figure 6. Standardized mortality rates (SMRs) for fallings (Source: ENHIS web site) (21).

Mortality from road traffic injuries in children

The indicator on mortality from road traffic injuries in children and young people focused on one of the preventable leading cause of death in children and young people (aged 5-24 years). The WHO health for all mortality database reveals in Figure 6 rather high standardized mortality rates (SMRs) occurring in both high- and low-income countries, Romania being apparently in a middle position. But if we consider the fact that in Romania the number of cars is about a quarter of the amount in developed countries, the intensity of phenomena is far bigger then is seems. As these are preventable through the concerted efforts for the implementation of effective measures addressing leading risk factors, further efforts should be made for saving innocent lives, through more concrete action for educating the youngsters.

Mortality in children and adolescents from unintentional injuries

Mortality in children and adolescents from unintentional injuries (falls, drowning, fires and poisoning) addresses another important contributor of the Global burden of disease in the Region (22). The data from the WHO European mortality database, Figure 6 even illustrating an East-West gap are constantly situating Romania the forth, fifth or sixth places, with higher mortalities due to fallings, as in Figure 6.

Policies to reduce and prevent excess body weight and obesity in children

The action indicator regarding policies to reduce and prevent excess body weight and obesity in children which are among the leading risk factors accounting for the burden of non-communicable diseases. The indicator reflects the degree of implementation of 12 specific policies in 5 broad policy areas aimed at preventing excess body weight and obesity in children. The most important policies assessed for implementation are presented in Table 2.

Table2. Key policies for preventing obesity.

Policy topic	Key policies to reduce and prevent obesity
Advertising	Legislation to practise responsible advertising and marketing of food, particularly with regard to promotion and marketing aimed at children of foods high in saturated fats, trans-fatty acids, free sugars and salt
Healthy diet and nutrition	National strategy to promote and increase the consumption of fruit, vegetables and legumes and to reduce the consumption of saturated fats, sugars and the elimination of trans-fatty acids Written policy document, adopted by a political body, explicitly concerned with nutrition Set of recommended nutrient reference values
Physical activity	Legislation requiring a minimum of 30 minutes of physical activity per day in schools
Education/ awareness /research	Health and nutrition education and awareness programmes in schools National health survey or participation in an international health survey that allows the monitoring of the prevalence of obesity, eating habits, physical activity and health in children
Implementation structures/ collaboration	Special administrative structure with responsibility for implementation of the policy Nutrition council or other advisory structure responsible for providing scientific advice to national policy-makers Any form of regular government-initiated collaboration between the parties responsible for food production, manufacture and sale, control and legislation and nutrition education Any form of regular consultation between the ministries of health and of agriculture on matters related to nutrition

Data computed on the basis of a specific questionnaire fill in by experts of the participating countries where the total score for degree of implementation is the sum of the scores for each policy: 0=no policy; 1=partly implemented or enforced; 2=substantially implemented or enforced, are presented in Figure 7

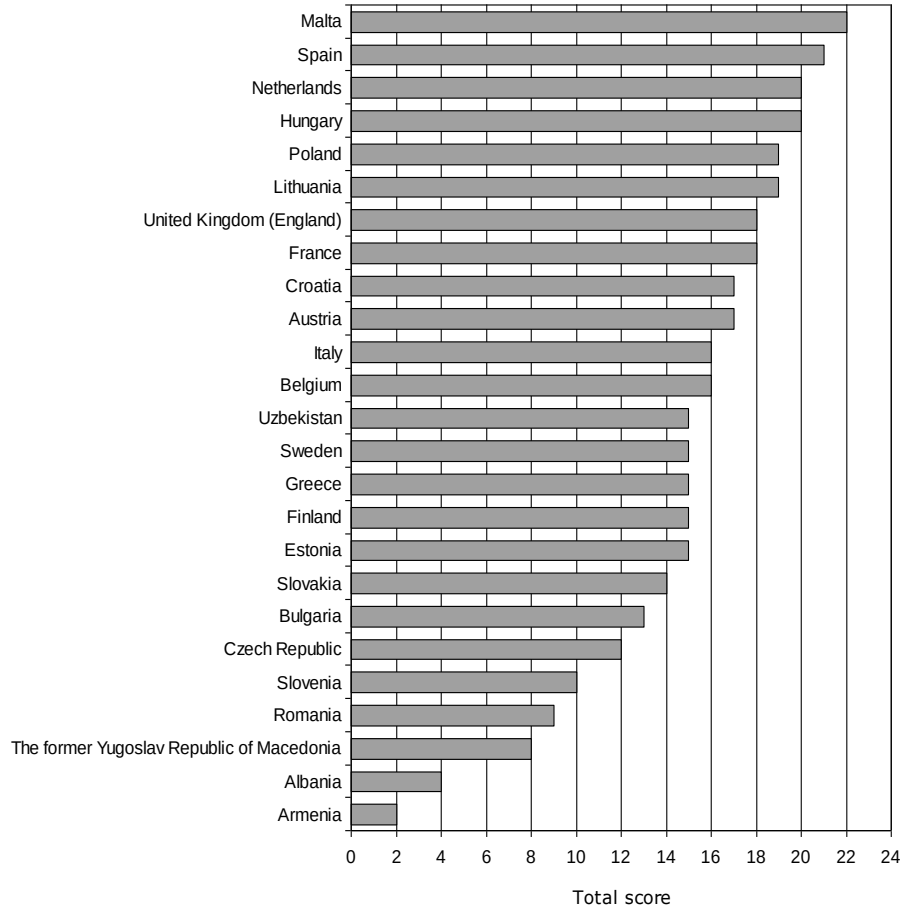


Figure 7. Degree of implementation of the 12 targeted policies 2005-2006 (Source: ENHIS web site) (21).

Romania is again in an unfavourable position, one of the last places, the majority of the score being associated with the requirements of the EU legislation, transposed legislation and partially enforced, than to concrete, public health interventions.

Facing new behavioural patterns, the limited available resources versus expensiveness of fruit and vegetables, the commercial success of the junk food

culture it is expected that the Romanian trends for overweight and even obesity, to raise sharply. In this context effective, coherent, multisectoral approach, including comprehensive long-term policy measures and active nutritional interventions in the school, workplace and community becomes a necessity.

RPG III indicators

The third priority goal indicators are concentrated on air quality and its effects on children health, including 7 indicators as illustrated in the table below (Table 3).

Table 3. Regional Priority Goals (RPG) III indicators.

Indicator
Prevalence of asthma and allergies in children
Infant mortality from respiratory diseases
Exposure of children to outdoor air pollution (particulate matter)
Exposure of children to environmental tobacco smoke
Children living in homes with problems of damp
Proportion of children living in homes using solid fuel
Policies to reduce the exposure of children to environmental tobacco smoke

Exposure of children to air pollution in outdoor air

Exposure of children to air pollution in outdoor air estimates the child population-weighted exposure based on the annual mean particulate matter (PM_{10}) concentration in cities and total (all age) city populations (23). The measure of exposure combines the PM10 concentration and the size of the population subject to the exposure. The health relevance of the indicator is one of the most valuable for the health and environment association. According to existing evidences children are particularly sensitive to air pollutants (21), effects from those to foetus reflected by the levels of postneonatal mortalities to well-known effects on lung function, aggravation of asthma or respiratory symptoms, increased prevalence and incidence of cough and bronchitis, to increased levels of mortality in adults, all have being scientifically proved. The indicator computed, based on internationally available data from AirBase and EUROSTAT, in Figure 8 reveals that most people in European cities (where PM_{10} is monitored) are exposed to PM_{10} levels exceeding the WHO air quality guideline (AQG) levels (of 20 $\mu g/m^3$).

The exposure feature for Romania situates us, again, in one of the worst position, with a PM_{10}-weighted mean of 53 $\mu g/m^3$, even above the European Union (EU) limit value (of 40 $\mu g/m^3$) for more than 80% of the reference population, generating a substantial risk to children's health. Even the PM_{10} data coverage reflects the situation only for 3 cities, and 21% of the urban population, it represents a issue requiring measures for integrate management of PM exposure, in order to reach the EU targets of a mean PM10 of 40 μg/m3 for 2005 and 20 μg/m3 in 2010, in line with the Clean Air for Europe programme and the proposal for the new EU air quality directive.

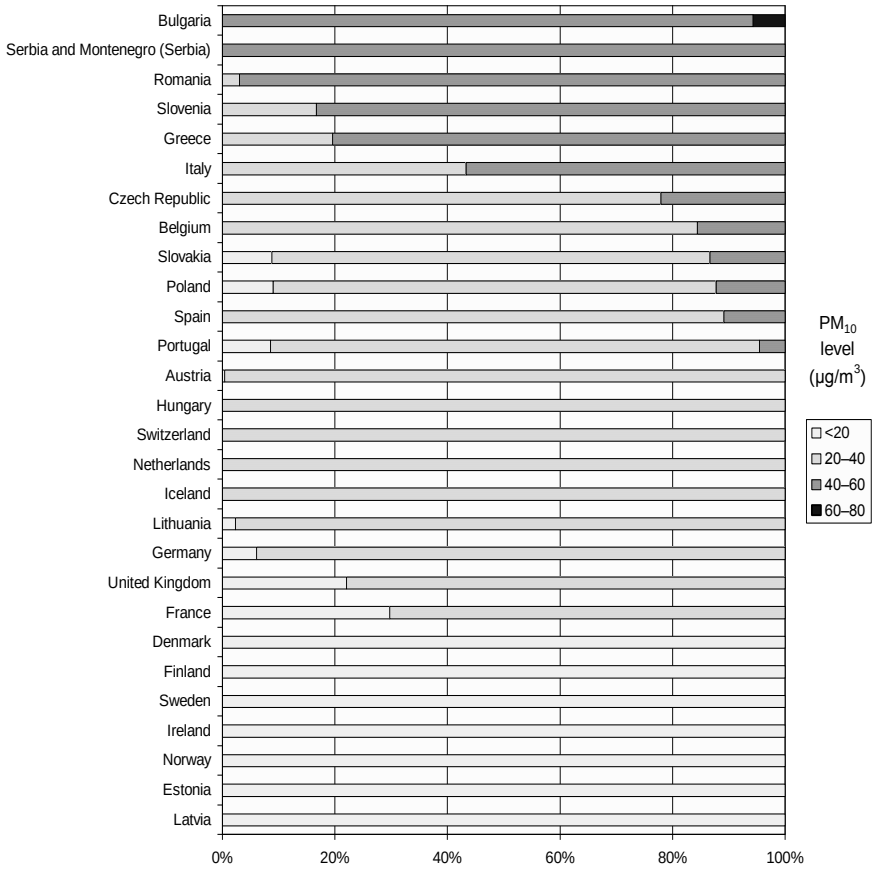


Figure 8. Percentage of children living in cities with various PM_{10} levels, 2004 (or last available year) (Source: ENHIS web site) (21).

Post neonatal mortality from respiratory diseases

The effect indicator on post neonatal mortality from respiratory diseases gives an indirect assessment of the adverse impact on health of environmental factors in a vulnerable age group. Studies have proved a positive association between the level of air pollutants and mortality in children due to respiratory causes. Several other social and environmental factors are contributing to those premature deaths, as quality of housing, energy type, heating system, tobacco smoke and seasonality.

The WHO mortality database figures presents the average of number of post-neonatal deaths due to respiratory diseases, showing for Romania one of the highest mortality rates for nearly all countries, Figure 9.

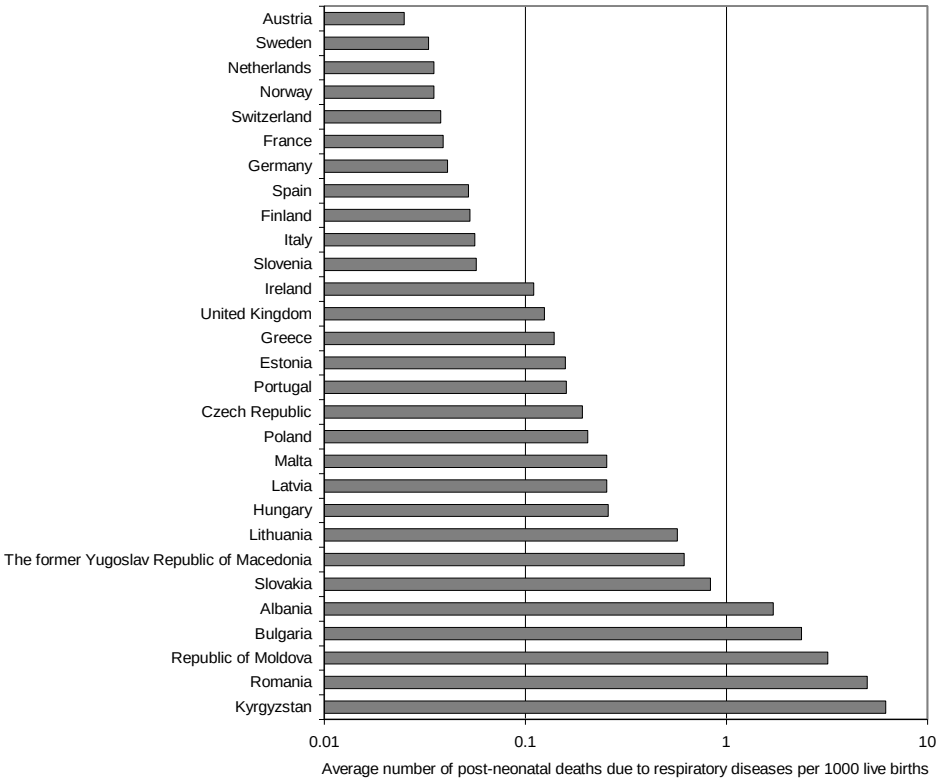


Figure 9. Post-neonatal mortality rate due to respiratory diseases, WHO European Region, 2001 (Source: ENHIS web site) (21).

If we consider the fact that respiratory mortality are one of the major contributors to the overall post-neonatal and also infant mortality rate, highlighting, in worse, Romania among the EU countries, a need for more effective intervention could be foreseen.

Prevalence of asthma and allergies in children

The prevalence of asthma and allergies in children, another effect indicator, illustrating the prevalence rates of symptoms of asthma and allergic rhinoconjunctivitis in children aged 6-7 years and 13-14 years, point out another priority issue of concern for the environmental health, closely associated with indoor and outdoor air quality. Early diagnosis and adequate treatment avoid long lasting effects, and associated costs and health impairments, that's why adequate management of disease should be completed with public health intervention both on health promotion and information but also on surveillance and control of

environmental conditions that contributes to asthma and allergies, as policies for indoor air quality or house-dust allergens as mites or those from pets. The data in Figure 10, results of the International Study of Asthma and Allergies in Childhood (ISAAC) (24), illustrates the prevalence of symptoms and diseases in selected centres (represented by cities/regions) participated in the study. Still the comparisons between the centres shows highest prevalence of asthma symptoms in children aged 13-14 years for asthma and rhinoconjunctivitis for Romania, not followed by positive trends of figures for the surveillance period of the study.

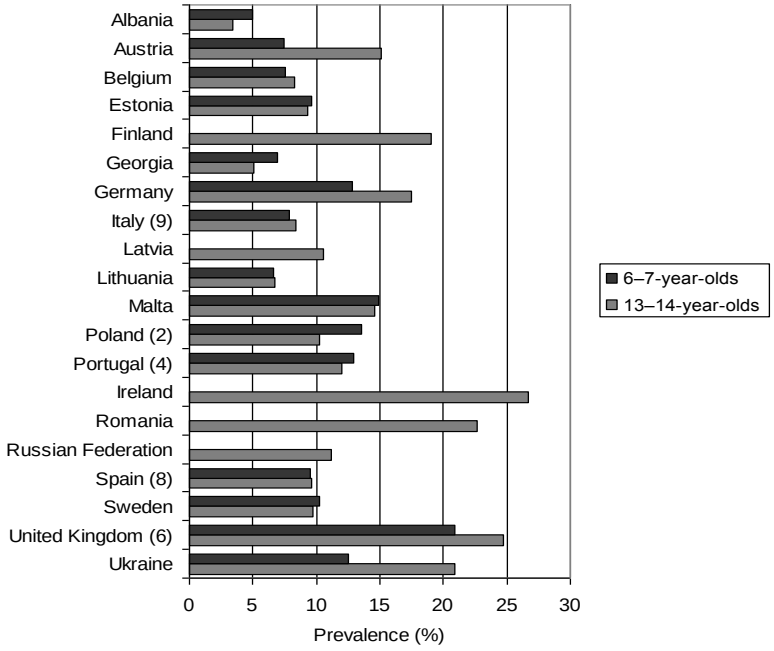


Figure 10. Prevalence of asthma symptoms in children aged 6-7 years and 13-14 years, ISAAC Phase Three, 1999-2004 (Source: ENHIS web site) (21).

Policies for reducing the exposure of children to environmental tobacco smoke

The action indicator on policies for reducing the exposure of children to environmental tobacco smoke is a composite index of commitment to reduce exposure to environmental tobacco smoke (ETS) and promote smoke-free areas for children. It is computed as a sum of the score of seven components, (0-for no restriction or prohibition, 1-for partial restriction, prohibition or voluntary agreement and 2-for complete ban or prohibition) as following:

1. Smoking prohibited in health care facilities
2. Smoking prohibited in education facilities
3. Smoking prohibited in bars and restaurants
4. Smoking prohibited in theatres and cinemas
5. Smoking prohibited in public transport vehicles
6. Advertisement of tobacco products in national mass media prohibited
7. Sale of tobacco to minors prohibited (combined component: sale of tobacco to persons aged under 16 years not allowed).

The indicator reflects the political commitment to transpose and enforce the WHO Tobacco Control Framework (FCTC), the first legal instrument designed to reduce tobacco-related deaths and disease around the world, that came into force in February 2005, ratified by Romania also in 2005, and in line with the European Union legislative measures on the prevention of smoking and tobacco control and the recently launched, in January 2007 Green Paper Towards a Europe free from tobacco smoke: policy options at EU level and opened a new strategy aiming "to launch a broad consultation process and an open public debate, on the best way forward to tackle passive smoking in the EU".

The indicator reflects the political commitment to transpose and enforce the WHO Tobacco Control Framework (FCTC), the first legal instrument designed to reduce tobacco-related deaths and disease around the world, that came into force in February 2005, ratified by Romania also in 2005, and in line with the European Union legislative measures on the prevention of smoking and tobacco control and the recently launched, in January 2007 Green Paper Towards a Europe free from tobacco smoke: policy options at EU level (25), and opened a new strategy aiming "to launch a broad consultation process and an open public debate, on the best way forward to tackle passive smoking in the EU".

The data computed in the WHO tobacco control database, for 2006, presented in Figure11, is situating this time Romania in a better position, in line with the EU legal requirements commitments. This score reflects a more extensive scope and comprehensive policies. Still the indicator assesses the extent to which regulations covering public spaces exists and are enforced, being more an indirect measure of the exposure of children to ETS, especially at home.

RPG IV indicators

The fourth priority goal indicators are concentrated on effects of physical and chemical hazards as: ultraviolet radiation, radon, lead, persistent organic pollutants (POPs). They are presented in Table 4.

Data on indicators assessing the complex exposure of children to potentially hazardous chemicals in food, natural or manmade as aflatoxins, antibiotics, additives, metals, cleaning agents, pesticide residues, packaging materials, arsenic, polychlorinated biphenyls (PCBs), otherwise strictly regulated by the EU food safety legislation due to their severe, chronic and irreversible effects, despite the long tradition public health existing monitoring system, in line with the WHO GEMS/Food system, they are not reported and internationally available.

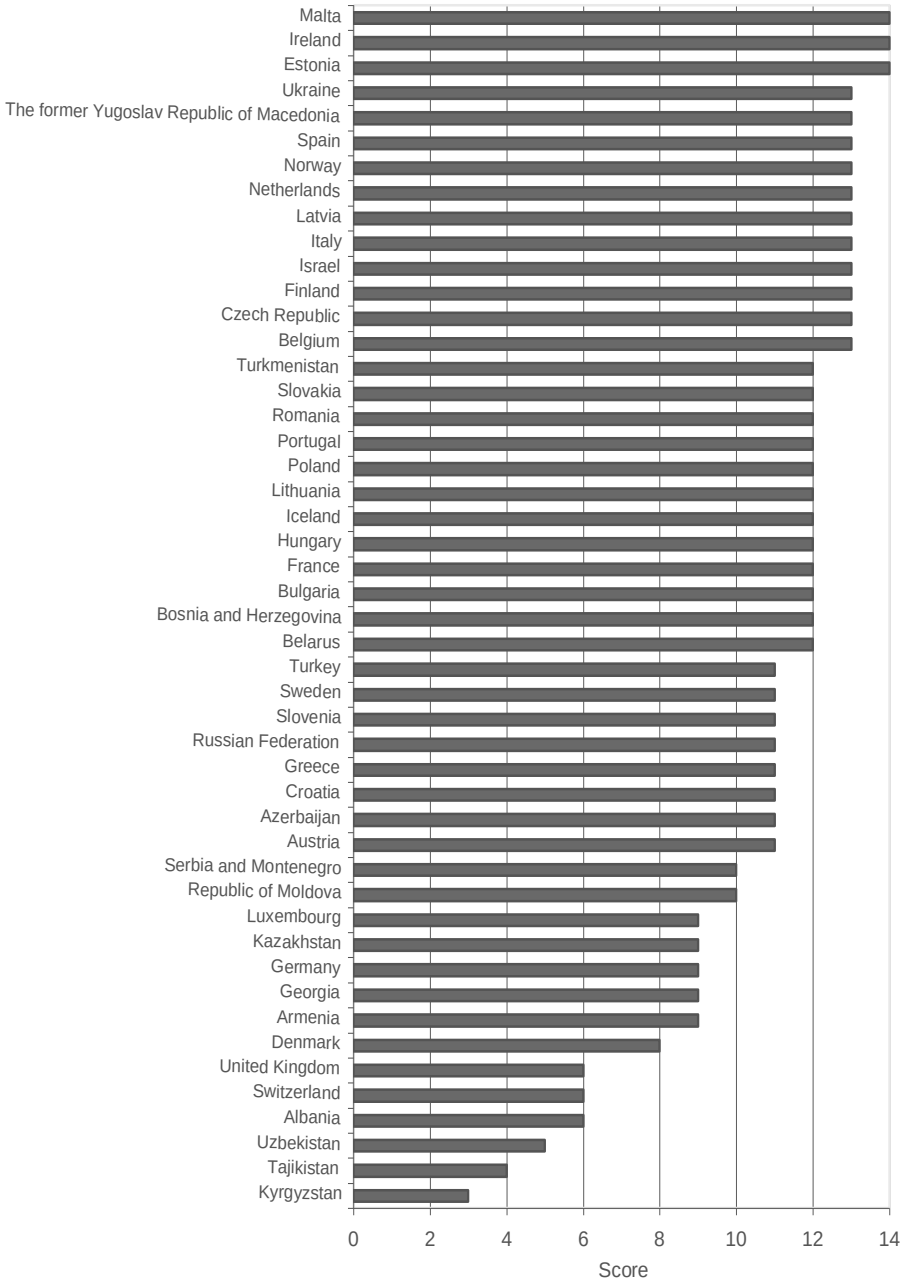


Figure 11. Implementation of policies to reduce exposure of children to environmental tobacco smoke in the WHO EURO Region (Source: ENHIS web site) (21).

Table 4. Regional Priority Goals (RPG) IV indicators.

Indicators
Incidence of childhood leukaemia
Incidence of melanoma in people aged under 55 years
Persistent organic pollutants (POPs) in human milk
Exposure of children to chemical hazards in food
Levels of lead in children's blood
Radon levels in dwellings
Work injuries in children and young people
Policies to reduce the excessive exposure of children to ultraviolet radiation

That's why, on the list of the public health system priorities regarding the development and strengthening of the national food safety and nutrition programme, in accordance to the EU White Paper on Food Safety (26), convergent to the Codex Alimentarius system, to update the chemical contaminants surveillance and control should be addressed as another priority.

Work injuries in children and young people

Data on work incidence rate of nonfatal injuries in children and young people, under 18 and between 18 and 24 years of age, recognized and legislative protected, (27) as a specific risk group, due to specific risk factors as lack of experience, limited awareness, immaturity, tools and equipment designed for adults, and by the other hand generated by the vulnerability of a immature body, where available for a limited number of countries, mostly form specific surveys.

Even if data where not available, and from the total number of 4,764 work accidents reported for 2006 by the Labour Inspection some cases by sure are affecting the vulnerable group of young, sometimes temporary employed, actions have been started for awareness and preventing those accidents by the series of activities and guidelines elaborated during the 2006 European week dedicated to a safe start in professional life.

Still, efforts should be continued for specific monitoring and enforcement of the useful guidelines elaborated and preventing youngster's workplace injuries and illness, in line with the WHO Global Strategy on Occupational Health for All and the EU Strategy on Occupational Safety and Health principles in a constructive health-labour -employers - employee partnership.

Level of lead in the blood of children

The exposure indicator on the level of lead in the blood of children is based on data from several WHO leaded surveys reflects the lead exposure occurring from multiple sources, especially lead-containing petrol as the most important source of atmospheric contamination in the countries where it is still used but also from industrial processes, paint, water pipes. The mean lead blood level shows for Romanian children a decreasing trend form levels of 18.2-18.9 μg/dl in the WHO 80s leaded studies to values about 10 μg/dl in 2000, mainly due to the beneficial effects of reducing the leaded petrol consumption. Unfortunately, the existing hot spots of industrial

emissions are still influencing the level of lead in children's blood in the vicinity of the plants, realities requiring an operational, flexible surveillance system for identification and elimination of the remaining sources of exposure to lead and monitoring of the effectiveness of preventive action, according to the 77/312/EEC Directive on Biological Screening of the Population for Lead (28) and the recently adopted strategic approach to international chemicals management for that minimize significant adverse impacts on the environment and on human health.

Incidence of childhood leukaemia

The effect indicator on incidence of childhood leukaemia the most common childhood malignancy, and a potential cluster for many environmental risk factors, as ionizing radiation exposure, electromagnetic fields, representing about 30% of all cancers diagnosed in children aged under 15 years, was included in the core set as a general environmental exposure marker. Data from the Automated Childhood Cancer Information System (ACCIS) (29), presented in Figure 12 shows one of the lowest rate of standardized incidence of leukaemia in children aged under 15 years- 30 cases per million per year in Romania, in line with values from the Eastern European countries. However data should be cautiously interpreted taking in account that the national cancer registry is not completely functional yet.

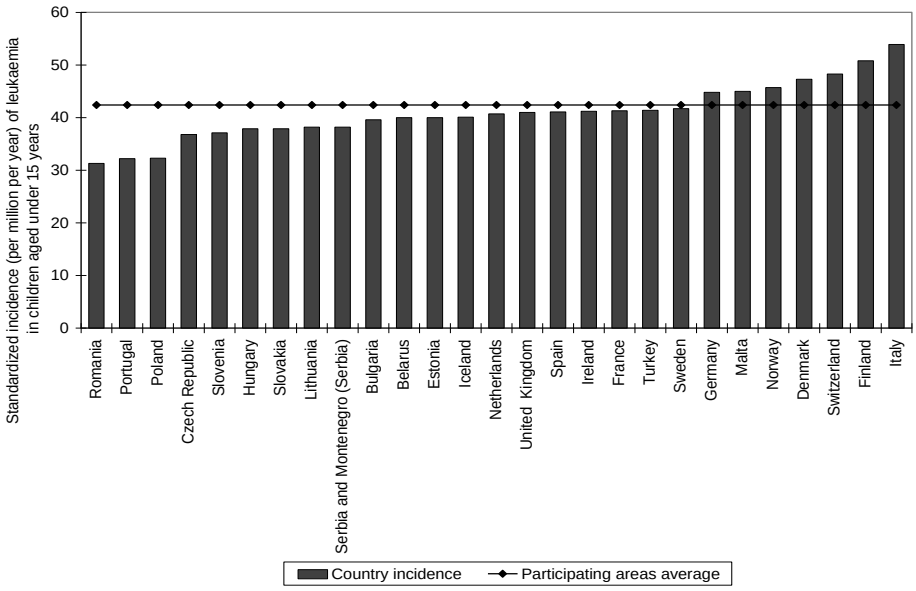


Figure 12. Standardized estimates of leukaemia in children aged under 15 years, selected countries, 1970-1999 (Source: ENHIS web site) (21).

As leukaemia is scientifically associated to various environmental exposures and the new Regulation of the European Parliament and the Council concerning the

Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) considers the carcinogenicity, mutagenicity and reproductive toxicity of chemical industrial substances as priority criteria for authorizing their use, the health system contribution in implementing the system could contribute to a better protection against chemical hazards.

Conclusions

Far from a holistic image, the indicators mapping is a very useful tool for identifying issues of interest for future interventions and their evidence base background reveals several priority areas of intervention for Romanian environmental health system.

EXERCISE

Task 1

Task 1 is about understanding the environmental health system components:

- list the core environmental health indicators;
- read: the indicator methodology, the specific fact sheets, the HIA examples;
- discuss: data availability for different issues; importance of legislative context on data availability; the health significance of different indicators based on FS and HIA information.

Task 2

Task 1 is about identifying environmental health features for specific countries. Based on the available environmental health indicators:

- describe the main characteristics for your country,
- identify and analyse trends and intervention priorities based on the indicators information.

REFERENCES

1. Briggs A. Making a difference: Indicators to improve children's environmental health. Geneva: World Health Organization; 2003. Available from URL: http://www.who.int/phe/children/en/cehindic.pdf. Accessed: June 6, 2009.
2. Last JM, Abramson JH, Friedman G1), Porta M, Spasoff RA, Thuriaux M, editors. A dictionary of epidemiology. New York: Oxford University Press; 1995.
3. Smith KR, Corvalán CF, Kjellstrom T. How much global ill health is attributable to environmental factors? Epidemiology. 1999;10:573-84.
4. Prüss-Ustün A and Corvalán C. Preventing disease through healthy environments. Towards an estimate of the environmental burden of disease. Geneva: World Health Organization; 2006. Available from URL: www.who.int/quantifying_ehimpacts/publications/preventingdisease.pdf. Accessed: March 20 2009.
5. Anonymous. Editorial: Children's health coming of age in Europe. Lancet. 2005;366:1052.
6. Bellinger DC. Lead. Pediatrics. 2004;113:1016-22.

7. Briggs D, Corvalán C, Nurminen M, editors. Linkage methods for environment and health analysis. General guidelines. Geneva: United Nations Environment Programme, U.S. Environmental Protection Agency, World Health Organization; 1996.
8. Pond K, Kim R, Carroquino MJ, Pirard P, Gore F, Cucu A, Nemer L, MacKay M, Smedje G, Georgellis A, Dalbokova D, Krzyzanowski M. Workgroup report: Developing environmental health indicators for european children: World Health Organization Working Group. Environ Health Perspect. 2007;115:1376-82.
9. United Nations. Agenda 21. Available from URL: http://www.un.org/esa/sustdev/documents/agenda21/english/agenda21toc.htm. Accessed: June 6, 2009.
10. United Nations. Millenium development Goals. New York: United Nations Statistics Division; 2006. Available from URL: http://mdgs.un.org/unsd/mdg/News.aspx?ArticleId=18. Accessed: June 6, 2009.
11. European Parliament. European Parliament resolution on a European Environment and Health Strategy (COM (2003) 338-C5-0551/2003-2003/2222(INI)). Strasbourg: European Parliament; 2004. Available from URL: (http://www2.europarl.eu.int/omk/sipade2?PUBREF=-//EP//TEXT+TA+P5-TA-2004-0246+0+DOC+XML+V0//EN&LEVEL=3&NAV=X. Accessed: June 6, 2009.
12. European Commission. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee on the "The European Environment & Health Action Plan 2004-2010" Brussels: Luxembourg Office for Offficial Publications of the European Communities; 2004. Available from URL: (http://ec.europa.eu/environment/health/pdf/com2004416.pdf. Accessed: June 6, 2009.
13. European Commission. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions on the sixth environment action programme of the European Community - Environment 2010: Our future, our choice. Luxembourg, Office for Official Publications of the European Communities; 2001. Available from URL: (http://europa.eu/scadplus/leg/en/lvb/l28027.htm. Accessed: June 6, 2009.
14. World Health Organization, Regional Office for Europe. Fourth Ministerial Conference on Environment and Health, Budapest, Hungary, 23-25 June 2004. Declaration. Copenhagen: World Health Organization, Regional Office for Europe; 2004. Available from URL: http://www.euro.who.int/document/e83335.pdf. Accessed: June 6, 2009.
15. World Health Organization. Environmental health indicators: Framework and methodologies. Geneva: World Health Organization; 1999. Available from URL: http://whqlibdoc.who.int/hq/1999/WHO_SDE_OEH_99.10.pdf. Accessed: June 6, 2009.
16. World Health Organization, Regional Office for Europe. 2004. Development of environment and health indicators for European Union countries: Results of a pilot study. Copenhagen: World Health Organization, Regional Office for Europe; 2004. Available from URL: http://www.euro.who.int/document/E85061.pdf. Accessed: June 6, 2009.
17. World Health Organization, Regional Office for Europe. Children's Health and Environment. Developing Action Plans. Copenhagen: World Health Organization, Regional Office for Europe; 2005. Available from URL: http://www.euro.who.int/document/e86888.pdf. Accessed: June 6, 2009.
18. Anonymous. Health Indicators: Building blocks for health analysis. Epidemiol Bullet. 2001;22:1-16. Available from URL: http://www.paho.org/english/dd/ais/eb_v22n4.pdf. Accessed: June 6, 2009.
19. Kim R, Dalbokova D, Krzyzanowski M. Development of environmental health indicators for European Union countries. Epi Marker. 2005;9(3):1-4.

20. European Commission, Health & Consumer protection General Directorate, Directorate C - Public Health and Risk Assessment. Operating a European Union public health information and knowledge system. Document for the meeting of the network of competent authorities for health information and knowledge. Luxembourg: European Commission, Health & Consumer protection General Directorate; 2004. Available from URL: http://ec.europa.eu/health/ph_information/implement/nca/docs/ev20040705_rd10_en.pdf. Accessed: June 6, 2009.
21. The European Environment and Health Information System (ENHIS) [homepage on the Internet]. Available from URL: http://www.enhis.org/object_class/enhis_gastrointestinaldiseases_.html. Accessed: May 30, 2009.
22. World Health Organization. 2007. WHO SIGHT. Summary information and global health trends [homepage on the Internet]. Geneva: World Health Organization; 2007. Available from URL: http://www.euro.who.int/foodsafety/Chemical/20040728_1. Accessed: June 6, 2009.
23. Schwartz J. Air pollution and children's health. Pediatrics. 2004;113:1037-1043.
24. International Study of Asthma and Allergies in Childhood - ISAAC [homepage on the Internet]. Available from URL: http://isaac.auckland.ac.nz/. Accessed: June 6, 2009.
25. European Commission, Health & Consumer protection General Directorate. Green paper. Towards a Europe free from tobacco smoke: policy options at EU level. Luxembourg: European Commission, Health & Consumer protection General Directorate; 2004. Available from URL: http://ec.europa.eu/health/ph_overview/health_forum/docs/ev_20071128_rd03_en.pdf. Accessed: June 6, 2009.
26. Commission of the European Communities. White paper on food safety. Brussels: Commission of the European Communities; 2000. Available from URL: http://ec.europa.eu/dgs/health_consumer/library/pub/pub06_en.pdf. Accessed: June 6, 2009.
27. World Health Organization, Regional Office for Europe. Work injuries in children and young people. Fact Sheet NO. 4.7, May 2007. Code: RPG4_Work_E1. Copenhagen: World Health Organization, Regional Office for Europe; 2007. Available from URL: http://www.euro.who.int/Document/EHI/ENHIS_Factsheet_4_7.pdf. Accessed: June 6, 2009.
28. Council of the European Communities. Directive on biological screening of the population for lead. 77/312/EEC. Luxembourg: Council of the European Communities; 1977. Available from URL: http://www.legaltext.ee/text/en/T50878.htm. Accessed: June 6, 2009.
29. Automated Childhood Cancer Information System - ACCIS [homepage on the Internet]. Available from URL: http://www-dep.iarc.fr/accis.htm. Accessed: June 6, 2009.

RECOMMENDED READINGS

1. Briggs D, Corvalán C, Nurminen M, editors. Linkage methods for environment and health analysis. General guidelines. Geneva: United Nations Environment Programme, U.S. Environmental Protection Agency, World Health Organization; 1996.
2. Prüss-Ustün A and Corvalán C. Preventing disease through healthy environments. Towards an estimate of the environmental burden of disease. Geneva: World Health Organization; 2006. Available from URL: www.who.int/quantifying_ehimpacts/publications/preventingdisease.pdf. Accessed: March 20 2009.

3. Smith KR, Corvalán CF, Kjellstrom T. How much global ill health is attributable to environmental factors? Epidemiology. 1999;10:573-84.
4. World Health Organization. Environmental health indicators: Framework and methodologies. Geneva: World Health Organization; 1999. Available from URL: http://whqlibdoc.who.int/hq/1999/WHO_SDE_OEH_99.10.pdf. Accessed: June 6, 2009.
5. World Health Organization, Regional Office for Europe. 2004. Development of environment and health indicators for European Union countries: Results of a pilot study. Copenhagen: World Health Organization, Regional Office for Europe; 2004. Available from URL: http://www.euro.who.int/document/E85061.pdf. Accessed: June 6, 2009.
6. World Health Organization, Regional Office for Europe. Children's Health and Environment. Developing Action Plans. Copenhagen: World Health Organization, Regional Office for Europe; 2005. Available from URL: http://www.euro.who.int/document/e86888.pdf. Accessed: June 6, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**ENVIRONMENTAL HEALTH RISK ASSESSMENT STUDIES**
Module: 2.1.4	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Dragan Gjorgjev**, MD, PhD, Professor Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius, Skopje, Republic of Macedonia **Vladimir Kendrovski**, MD, PhD, Assistant Professor Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius, Skopje, Republic of Macedonia **Fimka Tozija**, MD, PhD, Associated Professor Republic Institute for Health Protection and Faculty of Medicine, University Ss Cyril and Methodius-Skopje, Republic of Macedonia
Address for correspondence	**Vladimir Kendrovski** Republic Institute for Health Protection 50 Divizija 6, 1000 Skopje, Republic of Macedonia E-mail: kendro@t-home.mk
Keywords	Environmental health, environmental health risk assessment study
Learning objectives	After completing this module students should: • know the definition and characteristics of environmental health risk assessment (EHRA) studies; • be familiar with designing phase of EHRA studies; • be familiar with implementation phase of EHRA studies.
Abstract	EHRA is an essential element in environmental management and an important condition in precise priority setting to the necessary actions for environmental sanitation. The level of risk can be described qualitatively or quantitatively. The module is describing principles of EHRA studies, especially their designing and implementation phase.
Teaching methods	After introductory lectures students first carefully read the recommended readings. Afterwards they discuss the characteristics of EHRA studies with other students, especially the designing and implementation of this type of epidemiological studies. In continuation, they need to find published materials on EHRA studies and present their findings to other students, i.e. arsenic in drinking water or air pollution by $PM_{2,5}$ in urban area.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire.

ENVIRONMENTAL HEALTH RISK ASSESSMENT STUDIES

Dragan Gjorgjev, Vladimir Kendrovski, Fimka Tozija

THEORETICAL BACKGROUND

Introduction

Chemicals can be either beneficial or harmful for health of the people, depending on a number of factors, such as the amounts to which people are exposed. Low levels of some substances may be necessary for good health, but higher levels may be harmful. Health risk assessments are used to determine if a particular chemical poses a significant risk to human health and, if so, under what circumstances.

The environment in which people live, work and play is an important determinant of health and well being, but the extent of its importance is difficult to quantify, especially in developed economies. The non-communicable diseases present the biggest burden to public health analyzed by direct cost to the society as well as to the governance from aspect of disability adjusted life years (DALY). Due to fact that the proportion of elderly in total population is increasing in all countries and the prevalence of non-communicable diseases and disability has continuous upward trend, the needs for data which will reflect the quality of life, including the influence of environmental risk more precisely, is essential. The most common diseases - heart and circulatory diseases, cancer, respiratory diseases, injuries etc - have many risk factors which are often interconnected; including genetics, the condition people are in (via diet, exercise etc.), and the environmental circumstances to which they are exposed. Identifying cause-and-effect relationships is therefore very difficult, especially if the impact of the environment on health is delayed, or is the product of many, perhaps small, environmental factors acting together. There is a serious lack of data and information on exposures, effects and biological models that connect them. Therefore considerable uncertainty surrounds many issues of concern, such as air pollution, noise, water contamination, waste, climate change, chemicals (including endocrine disruptors and antibiotics), ionising and non-ionising radiation.

Risk and hazard are two distinct, but interrelated, concepts. A hazard represents a chemical, physical, or biological substance that has potential to produce harm to health if it is present in the environment and comes into contact with people. The hazardous properties of an environmental agent are defined according to the nature and severity of its harmful consequences. Fortunately, many hazards can be either contained or avoided, so not every potential environmental hazard poses an actual health risk. A risk, in turn, is defined as the likelihood of adverse health effects arising from exposure to a hazard in a human population, which is conceptually expressed as the product of two factors: the probability of exposure and the severity of the consequences (2).

Health Impact Assessment (HIA) is defined by different agencies in different ways, but there is a general consensus around a broad definition, published in 1999 as the "Gothenburg Consensus Paper" by the WHO Regional Office for Europe. That definition is: *"a combination of procedures or methods by which a policy,*

program or project may be judged as to the effects it may have on the health of a population." (3).

HIA may thus include assessment of high level policy and programs as well as individual developments, and encompass the vast array of assessment techniques used for each. In its broadest form, HIA seeks to predict the health impact of a policy, program or project (including a development) usually before implementation, and ideally early in the planning stage. It aims to facilitate the reduction or avoidance of negative impacts on human health and enhancement of the positive impacts, and in so doing promoting sustainable development - human health being central to the concept of sustainable development. Internationally, HIA has become a key component of informed decision making and is being undertaken by governments all around the world in a variety of circumstances and situations.

The term "health risk assessment" is often misinterpreted. People sometimes think that a risk assessment will tell them whether a current health problem or symptom was caused by exposure to a chemical. This is not the case. Scientists who are searching for links between chemical exposures and health problems in a community may conduct an epidemiologic study. These studies typically include a survey of health problems in a community and a comparison of health problems in that community with those in other cities, communities, or the population as a whole. Although they are both important, health risk assessments and epidemiologic studies have different objectives. Most epidemiologic studies evaluate whether *past* chemical exposures may be responsible for documented health problems in a specific group of people. In contrast, health risk assessments are used to estimate whether *current* or *future* chemical exposures will pose health risks to a broad population, such as a city or a community. Scientific methods used in health risk assessment cannot be used to link individual illnesses to past chemical exposures, nor can health risk assessments and epidemiologic studies prove that a specific toxic substance caused an individual's illness.

Environmental health risk assessment is an essential element in environmental management and an important condition in precise priority-setting to the necessary actions for environmental sanitation. At present there are not sufficient scientific data available for a large number of health-related environmental hazards representing risk on human health. In addition, even with the best possible information available on the nature and level of pollutants in the environment and about population exposure to different pollutants, environmental health risk assessment may not be complete because of difficulties in analyzing the complexity of possible interactions in the case of multiple exposures. Even more complex is the assessment and comparison of costs and benefits of health risk elimination. This is partly because environmental health risk assessment is still limited in its effectiveness by the inadequacy of the information available, especially on exposure. In addition, even with the best possible information, an environmental health risk assessment may not be complete because of difficulties in analysing the complexity of possible interactions in the case of multiple exposures (4).

Health risk assessment methodologies

In the context of environmental health, the risk management process can be organized into several distinct activities. The three core activities that constitute the essential decision-making steps in the risk management process are:

- risk assessment,
- risk evaluation and
- risk control.

Each of them is involved in examining various aspects of the risk problem (5).

Risk estimation

Risk estimation is the use of science-based risk information and analytical methods to characterize the nature and extent of environmental health risks in the human population.

Risk evaluation

Risk evaluation is consideration of the economic, social, political, and legal factors that influence a decision to adopt a particular course of action to reduce health risks - in some risk frameworks, the quantitative economic analysis of the benefits and costs of risk reduction is combined with results of the risk estimation process, so that risk evaluation may subsume parts or overall of risk assessment.

Risk control

Risk control is the selection of options and the commencing of actions intended to reduce risk to an acceptable or tolerable level; this activity is often referred to as risk management, but the term risk control is more specific and better reflects the objectives of the activities it denotes. Risk assessment is the process of estimating the potential impact of a chemical, physical, microbiological or psychosocial hazard on a specified human population or ecological system under a specific set of conditions and for a certain timeframe. Risk assessment is intended to provide complete information to risk managers, specifically policymakers and regulators, so that the best possible decisions are made. There are uncertainties related to risk assessment and it is important to make the best possible use of available information.

Methods

Retrospective epidemiological method was used in order the following four distinct and essential components of the risk assessment paradigm to be addressed:

1. hazard identification,
2. assessment of the dose-response relationships,
3. exposure assessment as the qualitative and/or quantitative assessment; and
4. risk characterization.

The logical consequence of the process of assessment of potential risk is the application of the information to the development of practical measures (risk

management) for the protection of human health. The migration pathways and sensitive receptors are as following:

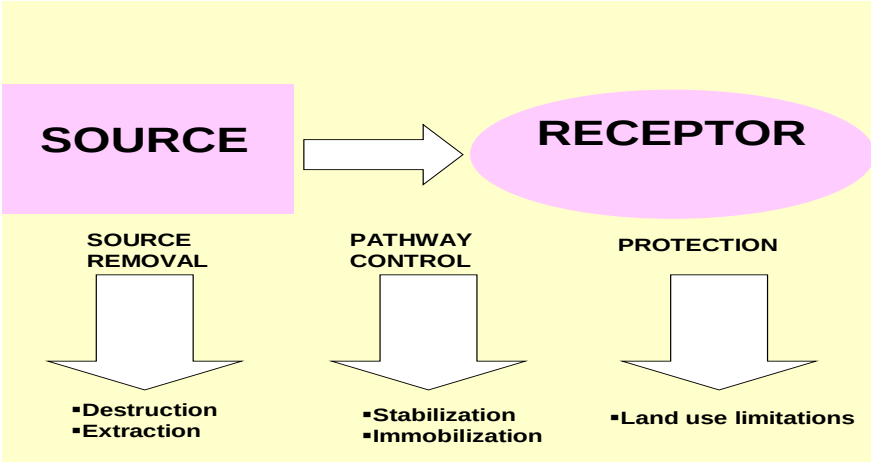


Figure 1. Contaminated land management (Adapted from: Nathanail & Bardos, 2004) (6).

Figure 1 shows the principles of contaminated land management as a general approach for identification of the sources of contamination and associated hazards (6).

Advantage of various methods

Many organisations are now actively involved in Environmental Risk Assessment, developing methodologies and techniques to improve this environmental management tool. Such list of organisations includes Organisation for Economic Co-operation and Development (OECD), World Health Organization (WHO), and European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC). One of the major difficulties concerning the use of risk assessment is the availability of relevant data. Furthermore, the data that are available are often loaded with uncertainty. The risk assessment may include an evaluation of what the risks mean in practice to those affected. This will depend heavily on how the risk is perceived. Risk perception involves people's beliefs, attitudes, judgements and feelings, as well as the wider social or cultural values that people adopt towards hazards and their benefits. The way in which people perceive risk is vital in the process of assessing and managing risk. Risk perception will be a major determinant in whether a risk is deemed to be "acceptable" and whether the risk management measures imposed are seen to resolve the problem. The procedures, methods and techniques for regulatory risk assessment of chemicals in the European Union (EU) is described in both legislation and supporting Technical Guidance Documents. Implementation is supported by the , part of the Joint Research Centre, in Ispra.

Most methodologies for human health risk assessment of chemicals are based on the U.S. National Academy of Sciences (NAS) model. A number of

methodologies exist due to differences in the toxic mechanisms exerted by different classes of chemical and the toxicological end-point being assessed. The end-point being assessed could be death, or a specific pathological condition relating to exposure to a chemical. When attempting to assess the risks from an immuno-suppressant toxin, specific end-points may be difficult to determine, as may be the role of other agents and stressors on the body. This will lead to risk assessment methodology for immuno-suppressants being different from assessments for irritants for instance. All human health risk assessments of chemicals include hazard identification, dose-response assessment, exposure assessment and risk estimation/ characterisation. If the assessment is site-specific, then a release assessment would be required in the absence of good data of environmental levels or to account for non-routine, accidental releases.

Risk and constraints of various methods

Risk assessments may assess individual or population risks. Individual risks may be for the average (i.e. typical) individual or the highly exposed or particularly susceptible individual and the risks may be estimated for various duration of exposure (e.g. per year or per lifetime) or for different locations. Individual risk can only be assessed for a hypothetical individual with assumed characteristics. Assessing the risk for any real individual will be frustrated by the fact that risk predictions for an individual can never be validated by experience. Any real individual will either experience the negative outcome or will not. Neither of these results can validate any risk prediction other than a probability of one or zero. Population risk may relate to the number of adverse health effects (e.g. fatalities, cancers, or illnesses) in a population over a specified period of time or the rate of adverse effects for a given location or sub-population.

The joint programme of the *United Nations Environment Programme* (UNEP), International labour organization (ILO) and WHO, the International Programme on Chemical Safety (IPCS), in collaboration with the U.S. Environmental Protection Agency (EPA), the European Commission (EC), the OECD, and other international and national organizations, developed a working partnership to foster the integration of assessment approaches to evaluate human health and ecological risks. The overall goal of this project was to promote international understanding and acceptance of the integrated risk assessment process. Three specific objectives were identified to meet this goal:

1. enhance understanding of the benefits of integration,
2. identify and understand obstacles to integration, and
3. engage key scientific organizations to promote discussion of an integrated approach to risk assessment (7).

A generic framework and associated documentation were developed to communicate how an integrated risk assessment could be conducted. Recognizing the similarities in risk assessment frameworks currently in use internationally, the integrated risk assessment framework is based on EPAs framework for ecological risk assessment and its associated terminology. Ecological risk assessment frameworks have greater general applicability than the human health frameworks

(or those environmental frameworks derived directly from human health frameworks). They have the following common characteristics:

- they were developed to deal with a range of environmental stressors beyond toxic chemicals,
- they have to describe the nature and role of the environment in the risk assessment process, and
- they must explicitly identify the endpoint to be assessed.

Further, a well-developed body of concepts and terminology exist in the literature treating the ecological risk assessment that supports integration. The integrated framework consists of three primary assessment phases:

1. During the first phase, *Problem Formulation*, the overall goals, objectives, scope, and activities of the assessment are delineated.
2. The *Analysis* step consists of data collection and modelling exercises to characterize exposure in time and space, and to define the effects on humans and ecological systems resulting from exposure. The methods appropriate for the Analysis step may be stressor-specific, but also depend upon the nature of the systems identified to be at risk.
3. Exposure and effect information are synthesized as estimates of risk in the *Risk Characterization* step. Ideally, these estimates are quantitative with respect to the level of risk expected under different exposure scenarios, although only qualitative estimates of risk may be possible in some circumstances.

The integrated risk assessment framework treats the relationships among risk assessment, risk management, stakeholder input, and data collection activities in a general parallel and concurrent manner. Essentially, risk characterisation is a summary of the data compiled in the risk assessment process including the uncertainties associated with each stage and the presentation of a risk estimate.

Parameters and indicators for choosing a methodology

Risks can be managed in many ways. They can be eliminated, transferred, retained or reduced. Risk reduction activities reduce the risk to an "acceptable" level, derived after taking into account a selection of factors such as government policy, industry norms, and economic, social and cultural factors. It is important to note that although risk assessment is used extensively in environmental policy and regulation it is not without controversy. This is also true for risk management. There are various criteria for assessing risk assessment including risk management (5).

The logical soundness of the method, e.g. its justification, is based on theoretical arguments or scientific knowledge, and the validity of the underlying methodological assumptions:

1. Completeness - (e.g. whether it can address all aspects of the problem and the degree to which it excludes issues because they are hard to accommodate);
2. Accuracy - (e.g. the precision reflected in the confidence level associated with the results; biases resulting from undue weight given to specific interests or considerations; and the sensitivity of results to untested assumptions or assumptions that cannot be tested);

3. Acceptability - (e.g. compatibility with existing processes; whether it is viewed as rational and fair; the level of understanding for all parties affected by it; and the confidence and familiarity of those who will use it);
4. Practicality - (e.g. the level of expertise, time and input data required);
5. Effectiveness - (e.g. usefulness of results; range of applicability across different risks and problem areas; the generalisability of the conclusion to other problem areas; and effectiveness and efficiency of linkage with other types of methods).

The level of risk can be described either qualitatively (i.e. by putting risks into categories such as 'high', 'medium' or 'low' as we used in our study as matrix showed in Figure 2) or quantitatively (with a numerical estimate). Current risk assessment methods do not enable accurate quantitative estimates of risk for low levels of exposure to environmental hazards. Numerical estimates of risk will rarely be feasible because of variability in the agent and population and limitations in toxicological and exposure data which will be reflected in the uncertainty assessment, but a degree of quantification may be possible for some components such as data collection and exposure assessment (8,9).

Contaminant Hazard
- Significant (H)
- Moderate (M)
- Minimal (L)

Migration Pathway
- Evident (H)
- Potential (M)
- Confined (L)

Receptors
- Identified (H)
- Potential (M)
- Limited (L)

Contaminant Hazard Factor	Receptor Factor	Migration Pathway		
		Evident	Potential	Confined
Significant	Identified	HHH	HHM	HHL
	Potential	HHM	HMM	HML
	Limited	HHL	HML	HLL
Moderate	Identified	HHM	HMM	HML
	Potential	HMM	MMM	MML
	Limited	HML	MML	MLL
Minimal	Identified	HHL	HML	HLL
	Potential	HML	MML	MLL
	Limited	HLL	MLL	LLL

Figure 2. Qualitative environmental health risk matrix (Adapted from: Kendrovski & Gjorgjev, 2008) (9). LEGEND: H-high, M-medium, L-limited (i.e. HHH (high, high. high); HHM (high, high, medium); HHL (high, high. limited); HMM (high, medium, medium); HML (high, medium, limited) HLL (high, medium, limited)…).

Regarding human health, the assessment was focused on exposure routes, both direct and indirect. For each of the two sites we established what exposure routes exist and what routes are significant. We also established the number and type of people that may be affected to a significant extent, depending upon location, age and profession.

CASE STUDY: ENVIRONMENTAL HEALTH RISK ASSESSMENT STUDY IN JEGUNOVCE, TETOVO

Introduction

HEK Jugohrom plant was established by the Government of the Republic of Macedonia in 1952 as the country's sole producer of chromium minerals and ferro-alloys, employing almost 2,000 workers and with an annual production capacity of around 69,000 tons. The uncontrolled disposal of waste material from the plant and the improper handling of material containing chromium salts have led to severe chromium contamination of groundwater and soil, including in the vicinity of the River Vardar. In 1982 the plant began monitoring soil and groundwater and the data confirmed contamination of the water by chromium. To address this problem the plant designed, installed and financed a groundwater abstraction system, which resulted in the concentrations of Cr^{+6} decreasing by 200-800 mg/l to total contamination levels of 5-15 mg/l. The plant's target is 1 mg/l (for comparison the target and intervention levels in the Netherlands are 0.001 mg/l and 0.03 mg/l) - if this is to be achieved remediation measures will need to be stepped up. Chromium production ceased in 1993 and the shed where chromium was produced and used has neither been cleared of chromium nor secured. Significant air pollution from the stacks, notably an estimated 9,000 to 17,000 tons of dust and fly ash a year, is due to the plant's electric furnaces operating without any form of gas cleaning. A project to reduce these emissions and recycle energy from a number of the furnaces has been proposed by the plant management, but requires funding.

According to the Macedonian National Environmental Action Plan (NEAP), total dust, black smoke and particle-borne chromium standards have been breached in the past years in the vicinity of the plant. Lack of data makes it impossible to assess any health impacts on workers. HEK Jugohrom was planed to undertake three remedial activities under the direction of the Ministry of Environment and Physical Planning with EU funding. The measures include keeping the bottom of the dumpsite drier, installing a drainage system to catch leachate and divert it to the waste treatment plant and further re-cultivating the dumpsite to abate rainwater infiltration. SILMAK is a company created with aim to restart the activities of ex Jugohrom, whose main activities were production of ferroalloys which contains certain chemical elements such as silicon, chromium and magnesium for the needs of steel industry. The production in Jugohrom was stopped in December 2001 (10).

In 2002, The Macedonian Government conducted privatization of Jugohrom in cooperation with the French investment group SCMM. Annual production now is 80.000 tons, Ferrosilicon (Si-content 75%, 65%, 90%) and Silicon metal (small tonnages), Four main types of granulation: 10-50 mm, 10-80 mm, 3-10 mm, 0-3 mm Silmak is a leading producer of Ferroalloys on Balkan peninsula Logistics: Annual

processed: Quartz/quartzite 140 000-160 000 tons Coal 78 000 tons, Lignite 82 000 tons. Total power consumption is 713 GW/h (11).

Hazard identification - Chromium

Chromium occurs in the environment primarily in two valence states, trivalent chromium (Cr^{+3}) and hexavalent chromium (Cr^{+6}). Exposure may occur from natural or industrial sources of chromium. Chromium-III is much less toxic than chromium-VI. The respiratory tract is also the major target organ for chromium-III toxicity, similar to chromium-VI. Chromium-III is an essential element in humans. The body can detoxify some amount of chromium-VI to chromium-III. The respiratory tract is the major target organ for chromium-VI toxicity, for acute (short-term) and chronic (long-term) inhalation exposures. Shortness of breath, coughing, and wheezing were reported from a case of acute exposure to chromium-VI, while perforations and ulcerations of the septum, bronchitis, decreased pulmonary function, pneumonia, and other respiratory effects have been noted from chronic exposure. Human studies have clearly established that inhaled chromium-VI is a human carcinogen, resulting in an increased risk of lung cancer. Animal studies have shown chromium-VI to cause lung tumours via inhalation exposure.

Chromium exists in several different oxidation states, the most stable and common of which are Cr^{+3} and Cr^{+6}. Due to the different chemical characteristics of Cr^{+3} and Cr^{+6} they behave differently in the environment. The redox-conditions in the aquatic environment are a very important factor in the precipitation of Cr^{+3} and the mobility of Cr^{+6}. An Eh-pH stability diagram of chromium in aqueous solutions shows that Cr^{+6} compounds are generally strongly oxidising agents that will only prevail in strongly oxidising conditions. In the pH-range from about 4 and above, Cr^{+6} will mainly occur as $HCrO_4^-$. Above pH 6 it occurs as CrO_4^{2-}. Cr^{+3} is believed to be the most stable form of chromium at pH-values above 4. The presence of dissolved oxygen will cause a very slow oxidation of Cr^{+3} to Cr^{+6}. This process will be very effective if MnO_2 is present (12).

Assessment of dose-response relationships

HEK Jugochrom's groundwater remediation programme has been operating since 1989. Throughout its operation, water samples from the abstraction wells, the drainage system and several monitoring wells have been analysed monthly during the period 1989-2000. Cr^{+6} concentrations have clearly decreased in the groundwater areas monitored.

Acute Effects

Chromium VI

Characteristics are as follows:

- chromium-VI is much more toxic than chromium-III, for both acute and chronic exposures;

- the respiratory tract is the major target organ for chromium-VI following inhalation exposure in humans. Shortness of breath, coughing, and wheezing were reported in cases where an individual inhaled very high concentrations of chromium trioxide;
- other effects noted from acute inhalation exposure to very high concentrations of chromium-VI include gastrointestinal and neurological effects, while dermal exposure causes skin burns in humans;
- ingestion of high amounts of chromium-VI causes gastrointestinal effects in humans and animals, including abdominal pain, vomiting, and hemorrhage. Acute animal tests have shown chromium-VI to have extreme toxicity from inhalation and oral exposure.

Chromium III

Characteristics are as follows:

- chromium-III is an essential element in humans, with a daily intake of 50 to 200 μg/d recommended for adults;
- acute animal tests have shown chromium-III to have moderate toxicity from oral exposure.

Chronic Effects (Non-cancer)

Chromium VI

Characteristics are as follows:

- chronic inhalation exposure to chromium-VI in humans results in effects on the respiratory tract, with perforations and ulcerations of the septum, bronchitis, decreased pulmonary function, pneumonia, asthma, and nasal itching and soreness reported. Chronic human exposure to high levels of chromium-VI by inhalation or oral exposure may produce effects on the liver, kidney, gastrointestinal and immune systems, and possibly the blood. Rat studies have shown that, following inhalation exposure, the lung and kidney have the highest tissue levels of chromium;
- dermal exposure to chromium-VI may cause contact dermatitis, sensitivity, and ulceration of the skin. The Reference Concentration (RfC) for chromium-VI (particulates) is 0.0001 mg/m^3 based on respiratory effects in rats. The RfC is an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without appreciable risk of deleterious non-cancer effects during a lifetime. It is not a direct estimator of risk but rather a reference point to gauge the potential effects. At exposures increasingly greater than the RfC, the potential for adverse health effects increases. Lifetime exposure above the RfC does not imply that an adverse health effect would necessarily occur. EPA has medium confidence in the RfC for chromium-VI (particulates) based on medium confidence in the study on which it was based because of uncertainties

regarding upper respiratory tract, reproductive, and renal effects resulting from the exposures;

- the Reference Concentration (RfC) for chromium-VI (chromic acid mists and dissolved chromium-VI aerosols) is 0.000008 mg/m^3 based on respiratory effects in humans. EPA has low confidence in the RfC based on low confidence in the study on which the RfC for chromium-VI (chromic acid mists and dissolved chromium-VI aerosols) is based. This is because of the uncertainties regarding the exposure characterization and the role of direct contact for the critical effect (1); and low confidence in the supporting studies which are equally uncertain regarding the exposure characterization (2). The Reference Dose (RfD) for chromium-VI is 0.003 mg/kg/d based on the exposure at which no effects were noted in rats exposed to chromium in the drinking water;
- EPA has low confidence in the RfD based on: low confidence in the study on which the RfD for chromium-VI was based because a small number of animals were tested, a small number of parameters were measured, and no toxic effects were noted at the highest dose tested; and low confidence in the database because the supporting studies are of equally low quality and developmental endpoints are not well studied.

Chromium III

Characteristics are as follows:

- although data from animal studies have identified the respiratory tract as the major target organ for chronic chromium exposure, these data do not demonstrate that the effects observed following inhalation of chromium-VI particulates are relevant to inhalation of chromium-III;
- EPA has not established an RfC for chromium-III.
- the RfD for chromium-III is 1.5 mg/kg/d based on the exposure level at which no effects were observed in rats exposed to chromium-III in the diet;
- EPA has low confidence in the RfD based on: low confidence in the study on which the RfD for chromium-III was based due to the lack of explicit detail on study protocol and results; and low confidence in the database due to the lack of high-dose supporting data.

Reproductive/Developmental Effects

Chromium VI

Characteristics are as follows:

- limited information on the reproductive effects of chromium-VI in humans exposed by inhalation suggest that exposure to chromium-VI may result in complications during pregnancy and childbirth;
- animal studies have not reported reproductive or developmental effects from inhalation exposure to chromium-VI. Oral studies have reported severe developmental effects in mice such as gross abnormalities and

reproductive effects including decreased litter size, reduced sperm count, and degeneration of the outer cellular layer of the seminiferous tubules.

Chromium III

Characteristics are as follows:

- no information is available on the reproductive or developmental effects of chromium-III in humans;
- a study of mice fed high levels of chromium-III in their drinking water has suggested a potential for reproductive effects, although various study characteristics preclude a definitive finding;
- no developmental effects were reported in the offspring of rats fed chromium-III during their developmental period (13-16).

Cancer Risk

Chromium VI

Characteristics are as follows:

- epidemiological studies of workers have clearly established that inhaled chromium is a human carcinogen, resulting in an increased risk of lung cancer. Although chromium-exposed workers were exposed to both chromium-III and chromium-VI compounds, only chromium-VI has been found to be carcinogenic in animal studies, so IARC has concluded that only chromium-VI should be classified as a human carcinogen;
- animal studies have shown chromium-VI to cause lung tumours via inhalation exposure;
- IARC has classified chromium-VI as a Group A, known human carcinogen by the inhalation route of exposure (17).

Chromium III

Characteristics are as follows:

- no data are available on the carcinogenic potential of chromium-III compounds alone;
- EPA has classified chromium-III as a Group D, not classifiable as to carcinogenicity in humans;
- EPA has stated that "the classification of chromium-VI as a known human carcinogen raises a concern for the carcinogenic potential of chromium-III" (18).

Exposure assessment

The production of the ferrosilicon is conducted in 7 electro furnaces, through the following processes: Preparation of the quartz (crushing and washing), melting, cooling and crashing of the ferrosilicon. Main raw materials used by Silmak for the production of FeSi come from Macedonia and nearby countries. 50% of

suppliers are situated in Macedonia. Slightly over half of word consumption of ferrosilicon is used in cast irons, which contain between 1% and 3% silicon. Silicon reduces the stability of iron carbide in cast iron and promotes the formation of graphitic carbon by a process of inoculation. Cast iron output has been stagnant for a number of years because of declining use in automobile engines, and growth of continuous casting making redundant the use of iron ingot moulds in steel making.

In the past, the plant's chromium processing activities produced large amounts of ferrochromium slag (446,000 tonnes), and other (chromate) slag (385,000 tonnes). Since 1996, this slag was dumped in an open dumpsite located close to the HEK Jugohrom plant. The dumpsite is 25 meters high and covers 7 hectares. Plant management estimates that its total mass is approximately 1,200,000 tons. The dump is covered with 2,000 m^3 of soil, which was allowed to naturally re-vegetate with grass and some small trees. A concrete pipe was installed to allow a small creek called Muzga Spring to run under the dumpsite. Muzga Spring has a mean flow of 3-5 l/s. The concrete pipe, however, is broken (probably due to the waste load), allowing the creek water to absorb Cr^{+6} contamination, which it discharges into the Vardar River. HEK Jugohrom plans to implement three remedial measures under the direction of the Ministry of Environment and Physical Planning of the Republic of Macedonia. According to plant management, European Union funds will support the implementation of these measures.

UNEP team has assumption that the groundwater outside the contoured area is not polluted, therefore, cannot be considered reliable. The precise size of the plume is unknown. For conceptual purposes, however, the plume's order of magnitude can be assumed to be (400 × 100 m) 40,000 m^2. Plant management reports that the deeper aquifers are not contaminated. Therefore, it can be assumed that the contaminated aquifer layer is less than approximately 8 meters thick. If the contamination were to reach the surface water, it could infiltrate the karst mountains (taking into consideration that Vardar River is both a discharging and recharging river). According to plant management, private homes outside of the site boundaries have suffered from chromium-contaminated groundwater creeping up the building walls through their capillaries (19,20).

Many walls and floors of the old chromium processing facilities are potentially contaminated with chromium compounds. However, the roofs of these buildings are in relative good condition. Therefore, the leaching of contamination into the soil and groundwater can be assumed to be limited.

The average monthly rate registered patient with respiratory diseases (J00-J99) without (J10-J18) among school children in Tetovo Region in 2006, shows that in rural area the rate is higher than in rural area, especially in winter due to air pollution in Jegunovce area which is unique for the country.

The distribution of Mortality from malignant neoplasms of liver in the Republic of Macedonia and Tetovo Region for the period 2000-2005 (rate/10.000) has showed lower rate in Tetovo compared to Macedonian rate. Also, the distribution of Mortality from malignant neoplasms of lung and bronchial tubes as well as kidneys in the Republic of Macedonia and Tetovo Region for the period

2000-2005 (rate/10.000) showed the lower rate in Tetovo compared to Macedonia rate (21-23).

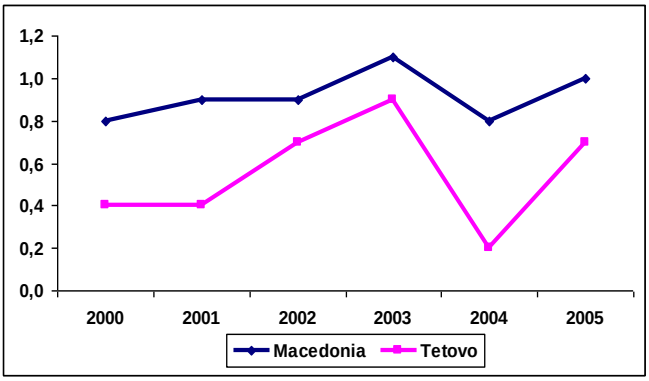
Table 1. Monthly rates ($^{o}/_{oo}$) of registered patient with respiratory diseases (J00-J99) without (J10-J18) among preschool children in Tetovo in 2006. Source: Republic Institute for Health Protection-Skopje, 2007.

Month	Tetovo	
	Urban	Rural
I	165.9	177.2
II	198.3	130.6
III	173.9	173.4
IV	157.7	106.6
V	151.9	64.9
VI	134.7	72.0
VII	84.2	69.9
VIII	152.7	60.0
IX	78.4	123.5
X	37.4	111.5
XI	156.0	117.1
XII	188.0	137.0
Average	139.92	119.40

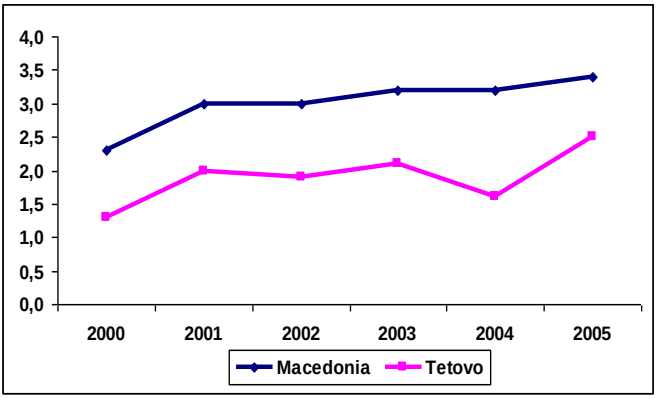
The average monthly rate of registered patients with respiratory diseases (J00-J99) without (J10-J18) among preschool children in Tetovo Region in 2006 shows that in urban area the rate is almost the same as in the rural area, especially in winter due to air pollution in Jegunovce area.

Table 2. Monthly rates ($^{o}/_{oo}$) of registered patient with respiratory diseases (J00-J99) without (J10-J18) among school children in Tetovo in 2006. Source: Republic Institute for Health Protection-Skopje, 2007.

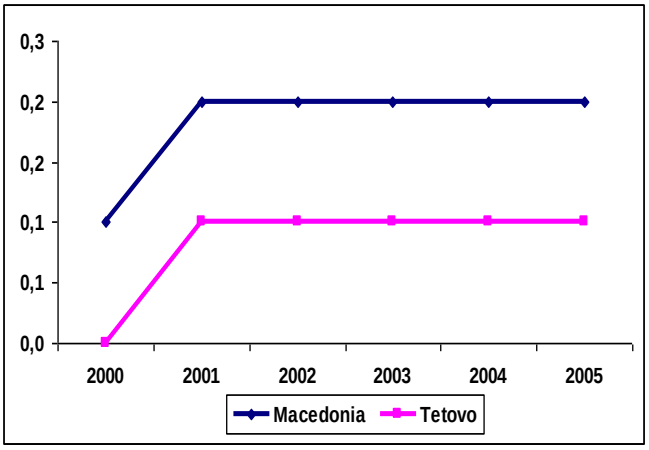
Month	Tetovo	
	Urban	Rural
I	65.0	97.7
II	47.7	64.3
III	37.1	70.7
IV	26.3	31.6
V	52.6	41.4
VI	24.1	20.9
VII	18.4	16.0
VIII	19.2	19.9
IX	16.1	24.8
X	17.9	33.6
XI	31.5	38.5
XII	56.8	77.0
Average	34.63	44.75



a) liver



b) lung and bronchial tubes

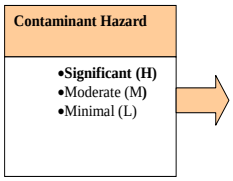


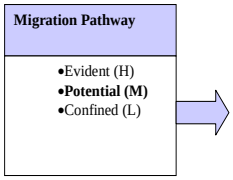
c) kidneys

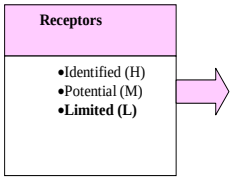
Figure 3. Distribution of mortality from specific malignant neoplasms in the Republic of Macedonia and Tetovo Region for the period 2000-2005 (rate/10.000) Source: Republic Institute for Health Protection-Skopje, 2007.

Risk characterization

The risk characterization is the synthesis of critically evaluated information and data from exposure assessment, hazard identification and dose-response considerations into a summary that identifies clearly the strengths and weaknesses of the database, the criteria applied to evaluation and the validation of all aspects of methodology, and the conclusions reached from the review of scientific information.







Contaminant Hazard Factor	Receptor Factor	Migration Pathway		
		Evident	**Potential**	Confined
Significant	Identified	HHH	HHM	HHL
	Potential	HHM	HMM	HML
	Limited	HHL	**HML**	HLL
Moderate	Identified	HHM	HMM	HML
	Potential	HMM	MMM	MML
	Limited	HML	MML	MLL
Minimal	Identified	HHL	HML	HLL
	Potential	HML	MML	MLL
	Limited	HLL	MLL	LLL

Figure 4. Qualitative environmental health risk at Jegunovce, Tetovo. (adapted from Kendrovski & Gjorgjev, 2008) (9) LEGEND: H-high, M-medium, L-limited (i.e. HHH (high, high. high); HHM (high, high, medium); HHL (high, high. limited); HMM (high, medium, medium); HML (high, medium, limited) HLL (high, medium, limited)…).

Regarding the Jegunovce Study the risk characterization is summarised in Box 1.

Systematic literature review

Thorough review of the relevant literature is required to provide a solid basis for Environmental health risk assessment studies.

Such a review identifies existing knowledge and key gaps. One approach is to convene an expert panel to conduct the review. It is important that the most appropriate

experts be identified and that they represent a range of skills and subject areas that are required for the assessment. With respect to Environmental health risk assessment studies, having academic experts in the various fields would be important. Clearly defining a search strategy is important. This would include specifying the search terms (such as exposure routes and health outcomes) and the databases that will be searched.

Box 1. Jegunovce Study risk characterization.

Migration Pathway:
air (dust and smelter emissions), water/groundwater, total ferrochrome dust to atmosphere.

Particulates:
Cr, Fe, Cu and Zn and total ferrochrome dust to atmosphere. Soil contamination by the dust deposited around the smelter. Cr and Fe occurred in the fine particulates of sizes less than 70 μm, Cu and Zn in the coarse particulates of size range 70-100 μm.

Other issues include waste/slag dumps and process chemical pollution. Obtained results from RIHP for Cr distribution in groundwater showed that the concentration is minimal. The results for Cr in new drinking water supply system in Jegunovce have showed very low concentration. The results from old drinking water supply system have showed higher values compared with new one but under the MPL. The groundwater analysis for Cr in water showed the higher value. Also the several years trend for Cr in Rasce drinking water showed lower Cr values under MPL.

Detected risk for respiratory diseases among children in Jegunovce is related to air pollution. The dust from the plant's exhaust appears not to be toxic by itself. However, dust generally causes coughing and creates a feeling of fatigue.

People who have weakened physical conditions or dust allergies, such as asthma and bronchitis, will suffer from exposure to the dust. Dust also creates visual contamination, which is usually opposed.

Depositional soil contamination, waste/slag dumps and process chemical pollution.

Receptors:
Limited.

Defined environmental health risk:
MODERATE to HIGH

The types of literature to be included should be decided at the beginning of the review. The assessment may include unpublished data from official sources (such as health statistics). An experienced literature searcher familiar with the relevant public health subject area should ideally be hired to perform these activities. Comprehensive literature review requires time and money. Gaining access to literature in countries with less well developed library and Internet systems or few literature or journal subscriptions may be difficult. A current WHO initiative is promoting access to international journals for developing countries. The Health Inter Network was created to bridge the digital divide in health, ensuring that relevant information - and the technologies to deliver it - is

widely available and effectively used by health personnel: professionals, researchers and scientists and policy-makers.

EXCERSISE 1: SYSTEMATIC LITERATURE REVIEW

The purpose of the exercise is to provide students with basic information about relevant literature as a solid basis for Environmental health risk assessment studies.

Students will be divided in two groups and will prepare essays in accordance to Task 1-3. Each of the group will oppose or accept the findings from others.

Task 1

Students are asked to determine the scope and type of the literature review.
Scope:

- inclusion criteria,
- exclusion criteria.

Types of literature:

- inclusion criteria,
- exclusion criteria (such as excluding newspaper articles or non-peer reviewed material).

Task 2

Determine the sources of relevant literature:

- primary sources (such as original peer-reviewed articles),
- secondary and tertiary sources (also called grey literature) such as review articles, reports, citations in journal articles, books, literature directories, Internet databases, newspapers, personal communications and unpublished data.

Task 3

Review and evaluate literature:

- develop evaluation criteria,
- evaluate each paper in relation to:
  - methods used,
  - relevance to local area,
  - validity of findings.

EXCERSISE 2: METHODS

The purpose of the exercise is to provide students with basic information about the environmental risk assessment studies through step by step approach.

Task 1

Hazard identification - identification of the inherent capability of a substance to cause adverse effects by sides;

Task 2

Assessment of dose-response relationships involves characterization of the relationship between the dose of an agent administered or received and the incidence of an adverse effect;

Task 3

Exposure assessment is the qualitative and/or quantitative assessment of the chemical nature, form and concentration of a chemical to which an identified population is exposed from all sources (air, water, soil and diet);

Task 4

Risk characterization is the synthesis of critically evaluated information and data from exposure assessment, hazard identification and dose-response considerations into a summary that identifies clearly the strengths and weaknesses of the database, the criteria applied to evaluation and the validation of all aspects of methodology, and the conclusions reached from the review of scientific information.

Students read the two files containing the different environmental risk assessment studies based by above mentioned literature review. After that, they should explain the:

- estimation of qualitative risk assessment of each studies,
- estimation of quantitative risk assessment of each studies,
- defining environmental health risk of each study.

REFERENCES

1. World Health Organization. Elemental speciation in human health risk assessment. Environmental health criteria No. 234. International Program on Chemical safety, Geneva: World Health Organization; 2006.
2. Environmental Protection Agency. Guideline for exposure assessment. Washington DC; Environmental Protection Agency; 1992.
3. World Health Organization. Health impact assessment: Main concepts and suggested approach. Brussels: European Centre for Health Policy; 1999
4. National Environmental Protectoral Council. Guideline of human risk assessment methodology. Canberra: National Environmental Protectoral Council; 1999.
5. Covello VT, Merkhofer MW. Risk assessment methods - approaches for assessing health and environmental risks. New York, NY: Plenum Press; 1993.
6. Nathanail PC, Bardos PR. Reclamation of contaminated land. Chichester: John Wiley & Sons Ltd; 2004.

7. World Health Organization. International programme on chemical safety. The WHO recommended classification of pesticides by hazard and guidelines to classification 1994-1995:UNEP/ILO/WHO. Geneva: World Health Organization; 1994.
8. Department for Health and Ageing. Environmental health risk assessment-guidelines for assessing human health risks from environmental hazards: Canberra: Department for Health and Ageing; 2002.
9. Kendrovski V, Gjorgjev D. Qualitative environmental health impact assessment in Veles, Republic of Macedonia. In: Burcu OG, editor. Environmental technologies: New Developments. Wien: I-Tech Education and Publishing; 2008. p.67-84.
10. Government of the Republic of Macedonia. The second national environmental action plan. Skopje: Government of the Republic of Macedonia; 2006.
11. Government of the Republic of Macedonia. National waste management plan 2006-2012. Skopje: Government of the Republic of Macedonia; 2005.
12. Agency for Toxic Substances and disease registry. Toxicological profile for chromium. U.S. Public Health Service: Atlanta, GA; U.S. Department of Health and Human Services; 1998.
13. California Environmental Protection Agency. Technical support document for the determination of noncancer chronic reference exposure levels. Draft for public comment. Berkeley, CA: Office of Environmental Health Hazard Assessment; 1997.
14. American Conference of Governmental Industrial Hygienists. TLVs and BEIs. Threshold limit values for chemical substances and physical agents, biological exposure indices. Cincinnati, OH: American Conference of Governmental Industrial Hygienists; 1999.
15. Rai D, Eary LE, Zachara JM. Environmental chemistry of chromium. Sci Total Environ. 1989;86: 15-23.
16. National Institute for Occupational Safety and Health. Pocket guide to chemical hazards. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention; 1997.
17. International Agency for Research on Cancer. Some halogenated hydrocarbons. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. Lyon: IARC; 1979.
18. Environmental Protection Agency and the Office of Air Quality Planning & Standards. Chromium and compounds. Washington DC: USEPA; 1998.
19. United Nation Environmental Programme. Feasibility study for urgent risk reduction measures at hot spots In FYR of Macedonia. Geneva: United Nation Environmental Programme; 2001.
20. United Nations. Environmental performed review for FYR of Macedonia. Skopje: United Nations; 2002.
21. State Statistical Office. Statistical yearbook of the Republic of Macedonia. Skopje: State Statistical Office of the Republic of Macedonia; 2005.
22. State Statistical Office. Statistical yearbook of the Republic of Macedonia. Skopje: State Statistical Office of the Republic of Macedonia; 2006.
23. Republic Institute for Health Protection. Yearbook of preventive programs in the Republic of Macedonia. Skopje: Republic Institute for Health Protection; 2007.

RECOMMENDED READINGS

1. World Health Organization. Elemental speciation in human health risk assessment. Environmental health criteria No. 234. International Program on Chemical safety, Geneva: World Health Organization; 2006.
2. World Health Organization. Health impact assessment: Main concepts and suggested approach. Brussels: European Centre for Health Policy; 1999.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**THE GEOGRAPHIC INFORMATION SYSTEM (GIS) USE IN ANALYSIS OF TRAFFIC AIR POLLUTION**
Module: 2.1.5	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Andreja Kukec**, Bsc, PhD Candidate, Teaching Assistant Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Rok Fink**, Bsc, MSc Candidate, Teaching Assistant Department of Sanitary Engineering, Faculty of Health Studies, University of Ljubljana, Slovenia **Saša Erlih,** BSc, MSc Institute for Comprehensive Development Solutions, Slovenia **Ivan Eržen**, MD, PhD, Assistant Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Ivan Eržen** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Zaloška cesta 4, 1000 Ljubljana, Slovenia e-mail: ivan.erzen@gmail.com
Keywords	Geographic information system, air pollution, health effects
Learning objectives	After completing this module students should: • understand why GIS is important tool in analysis of air pollution in environmental epidemiology, • be able to use AirGIS in practice with help of an instructor.
Abstract	Geographic Information System (GIS) is software that encompasses storage, retrieval, analysis and display of spatial-geographical data. In environmental epidemiology represents powerfull tool for presentation of results of epidemiological studies. AirGIS is a human exposure modelling system for traffic air pollution. As a case study, AirGIS model is used in case of air pollution in coastal community Koper in Slovenia.
Teaching methods	After introductory lecture students carefully read the recommended sources about traffic air pollution. In continuation they discuss air pollutants and practice on GIS software. At the end they compare and discuss their results.
Specific recommendations for teachers	• work in small groups (max. 3 students); • work under teacher supervision/individual students' work proportion 50%/50%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students); access to the Internet and GIS software; reliable data for GIS study; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire and case study.

THE GEOGRAPHIC INFORMATION SYSTEM (GIS) USE IN ANALYSIS OF TRAFFIC AIR POLLUTION

Andreja Kukec, Rok Fink, Saša Erlih, Ivan Eržen

THEORETICAL BACKGROUND

Geographic Information System

About the Geographic Information System

Geographic Information System (GIS) is software that encompasses storage, retrieval, analysis and display of spatial-geographical data.

GIS is a promising tool for exposure modelling due to the increase in coverage and quality of digital maps, developments in administrative databases managed by the authorities and developments of more user-friendly desktop GIS with increasing number of analytic features (1,2).

There exist different modules of this software, according what the observed outcome is. For analyzing human exposure modelling system for traffic air pollution AirGIS module is used.

About the AirGIS System

AirGIS is a human exposure modelling system for traffic air pollution. It has been developed by the National Environmental Research Institute in Denmark for application in Danish air pollution epidemiological studies, human exposure studies, as well as urban air quality assessment and management.

The model system allows for exposure estimates at a high spatial resolution (address level) and a high time resolution (hour). The system integrates air pollution dispersion models, digital maps, national and local administrative databases, concentrations of air pollutants at regional, urban background and street level, meteorological data, and GIS (3-5).

About using the Geographic Information System in analysis of traffic air pollution in practice

Various exposure models have been developed that combine microenvironment concentrations with individual time-activity patterns and extrapolation to the entire population to give population exposure distributions (Figure 1). Traffic emissions are estimated based on emission factors together with average daily traffic (ADT), the percentage of heavy vehicles and the travel speed for each street section applying a default seasonal, weekly and diurnal variation in traffic loads to obtain hourly traffic inputs as well as default values for cold starts.

An example of use of GIS in analysis of traffic air pollution in practice is a project of Centre for Transport Research on Environmental and Health Impacts and Policy entitled Traffic Air Pollution Exposure Modelling (1). This project takes a micro-environment approach to population exposure modelling but adds a geographic dimension by taking advantage of GIS, digital maps and administrative databases. The objectives of the research project are: to develop an exposure model that

combines modelled air pollution data using the Danish Operational Street Pollution Model (OSPM), population data using existing administrative databases, digital maps and GIS. A simple population dynamics model will be established to model the number of people present in a given area during a given time using simple profiles for time spent in the various areas at home, at work, and in transit. Additionally, ratios between indoor and outdoor concentrations will be taken into account (1,3,6).

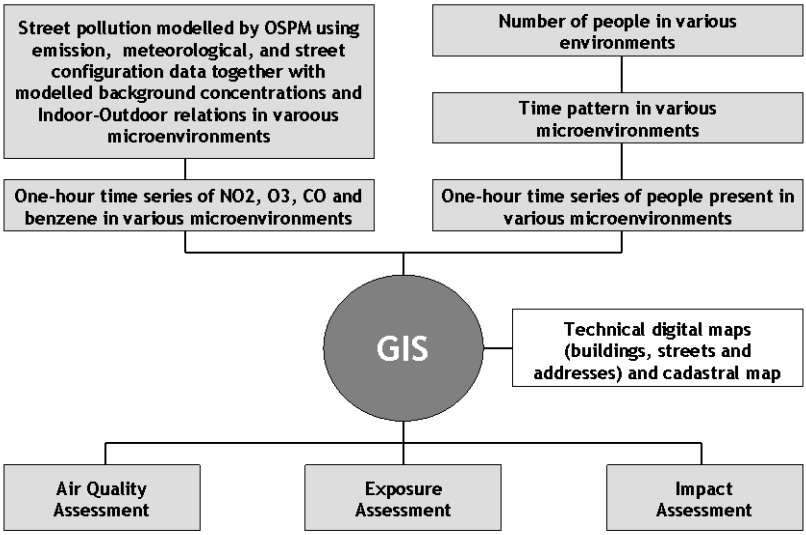


Figure 1. Outline of the methodology for the exposure modelling (Adapted from Jensen, 1998) (1). LEGEND: GIS- Geographic Information System; OSPM- Operational Street Pollution Model.

CASE STUDY: TRAFFIC AIR POLLUTION AND HEALTH CONCERNS IN KOPER COMMUNITY/SLOVENIA

Introduction

Each and every one of our actions demands energy. Already early in history, man attempted to gain control over more energy, primarily through the use of animals and slaves. Not long after we also learned through technical prowess to use nature's energy. It was only with Watts' invention of the steam engine in 1769 that it became possible for man to produce large amounts of energy on demand (7).

Transportation system has contributed significantly to the development of human civilization. On the other hand it has an enormous impact on the air quality in several ways (8). In Western Europe, the transport of people and freight has dominated road traffic for many decades. Emissions from road traffic, from both combustion and friction processes, result in a complex mixture of air pollution, which is known to have adverse the population's exposure and the effects on health (8,9,10).

Nowadays, in average 1.4 persons are travelling by one car, therefore for transportation 100 persons 70 cars are needed. From point of environmental consideration for 100 km 40 litres of gasoline are used up, with average consumption 8 litres/100 km is this 560 litres. For every single litre of gasoline harmful compounds are emitted (11,12). Fuel combustion is the primary source of a large number of health-damaging air pollutants, including fine and respirable particulate matter ($PM_{2.5}$ and PM_{10}), carbon monoxide (CO), sulphur dioxide (SO_2), nitrogen oxides (NO_x), volatile organic compounds (VOCs), ozone (O_3), and atmospheric lead (13). Some of these pollutants are direct by-products of fuel combustion, but others (such as O_3) are formed in the air through chemical reactions with other agents in the atmosphere. Despite these facts, world oil consumption is significantly rising over last decade. In year 2000, 76.6 million barrels per day were used, seven years later, in 2007, 85.7 million barrels per day. Trends are showing that in 2009 almost 90 million barrels will be used for transportation each day (14).

Of all the different types of pollution affecting human health, by far the most important is air pollution (7). Human exposure is believed to cause severe health effects, especially in urban areas where pollution levels often are high. The classic example is the severe London smog (smoke and fog) episode in 1952 where the mortality rate in the city increased dramatically (8). Although transport emissions are rising faster in low-income and middle-income countries than in those with high income, there remain massive global inequalities in transport energy use both between and within countries (15).

Environmental effects of traffic air pollution

Combustion of liquid fossil fuels causes emissions due to its composition or type of combustion. Resulting pollutants are:

- Sulphur oxides (SO_x): although SO_x is a symbol of all oxides of sulphur (e.g., SO_2 and SO_3), about 95% of all sulphur oxides are in the form of SO_2. In the atmosphere, however, SO_2 is a precursor of highly destructive sulphates (SO_4^{2-}), formed by the chemical addition of oxygen (O_2). SO_3 is not a stable compound and may react with water (H_2O) to form sulphuric acid (H_2SO_4), a component of acid rain (16). Although road transport is a minor source of SO_2 at the national level, in some urban areas it can be important. Raised concentrations of SO_2 have been detected alongside busy roads (17),
- Nitrogen oxides (NO_x): NO_x are powerful greenhouse gas with global warming potential 296 times that carbon dioxide (18). They are formed during high-temperature combustion, largely from the nitrogen and oxygen present in air, but also from the oxidation of nitrogen contained fuels. The main sources are internal combustion engines. Almost all NO_x are emitted as nitric oxide (NO) which is then rapidly oxidized to the more toxic nitrogen dioxide (NO_2) (17). During wintertime pollution episodes, NO_2 concentrations in urban areas may exceed internationally accepted air quality criteria set for the protection of human health. These cases were reported in some large industrial and urban population centres in northwest Europe, particularly in United Kingdom during stagnant wintertime weather conditions. During summertime pollution episodes, photochemical reactions driven by sunlight lead to the conversion of organic compounds and oxides of nitrogen into photochemical oxidants, in

particular ozone. This phenomenon was first reported in Los Angeles in 1940's and has subsequently been observed in almost urban population centres worldwide. These photochemical reactions also lead of the oxidation of SO_2 into fine haze of sulphuric acid (H_2SO_4) (19),

- Carbon monoxide (CO): most of anthropogenic carbon monoxide is generated in combustion processes. Internal combustion engines, both in on-road and in diverse off-road use, comprise principal source. The majority of the carbon in automotive fuels is oxidized to carbon dioxide, while a small fraction is incompletely oxidized to CO (17). Catalytic converters reduce CO emissions approximately by factor of 8 (7),
- Volatile organic compounds (VOCs): VOCs comprise a wide range of chemical compounds including hydrocarbons, oxygenates and halogen- containing species. However a few VOCs are considered toxic to human health, such as benzene and 1,3 butadiene. The main sources of benzene is production, distribution and use of automotive fuels. Petrol vehicle emit more benzene than diesel vehicle, even when catalytic converters are used. Petrol and diesel contains little 1,3 butadiene, but is formed during combustion in the engine. (17,20),
- Particulate matter ($PM_{2,5}$ and PM_{10}): PM is an air pollutant consisting of a mixture of solid and liquid particles suspended in the air. These suspended particles vary in size, composition and origin. The most commonly used size fractions are particles with an aerodynamic diameter between 2.5 μm and 10 μm. PM can either be directly emitted into the air (primary PM) or be formed in the atmosphere from gaseous precursors (mainly SO_2, oxides of nitrogen, ammonia and non-methane volatile organic compounds). The most important chemical constituents of PM are sulphates, nitrates, ammonium, other inorganic ions such as Na^+, K^+, Ca^{2+}, Mg^{2+} and Cl^-, organic and elemental carbon, crustal material, particle-bound water and heavy metals (9). A major contribution of particulate matter in urban areas is believed to be attributable to emissions from diesel-powered vehicles and traffic. The fine and ultra-fine particles are generated mainly by combustion from diesel exhausts (13).

  Since 1950 the world population has more than doubled, and the global number of cars has increased by factor of 10. In the same period the fraction of people living in urban areas has increased by factor 4 (21). The environmental impact of transport depend on transport volume (e.g. vehicle kilometres), technology, the way vehicles are used (speed, acceleration), the distribution of the use of vehicles over space and time, and the locations of exposed people, nature and buildings (22). But rapid urbanisation and increasing time spent in congested traffic means that exposure is increasing even where pollution levels are falling. The greatest burden is posed to the mega cities of developing countries (15).

  Due to Slovenian urbanisation, green house gasses emission increased, especially CO_2. Results are showing that share of traffic transport with individual cars is rising; meanwhile public railway and bus transport are decreasing (23).

  Climate change is occurring in the context of increased anthropogenic stress across a range of natural systems including stratospheric ozone depletion, loss of biodiversity, spread of invasive species, exhaustion of wild fisheries, and the depletion of freshwater supplies. In addition to its own resource use, transport facilitates the exploitation of other resources. From a

climate-change perspective the most important effect is opening up areas to deforestation. Through increasing the demands on our environment's carrying capacity a capacity not known in advance energy intensive transport reduces its human carrying capacity (15,24),

- Ozone (O_3): ozone is a gas composed of three oxygen atoms. It is not usually emitted directly into the air, but at ground-level is created by a chemical reaction between oxides of NO_x and volatile organic compounds in the presence of sunlight. Ozone has the same chemical structure whether it occurs kilometers above the earth or at ground-level and can be "stratospheric" or "ground -level," depending on its location in the atmosphere. In the lower atmosphere of the Earth, ground-level ozone has negative impact on environmental and human health. Motor vehicle exhaust and industrial emissions, gasoline vapours, and chemical solvents as well as natural sources emit NOx and VOC that help form ozone. Ground-level ozone is the primary constituent of a smog. Sunlight and hot weather cause ground-level ozone to form in harmful concentrations in the air. As a result, it is known as a summertime air pollutant. Many urban areas tend to have high levels of ground-level ozone, but even rural areas are also subject to increased ozone levels because wind carries ozone and pollutants that form it hundreds of kilometres away from their original sources. Nevertheless, stratospheric ozone occurs naturally in the stratosphere approximately 15-45 kilometres (10 to 30 miles) above the earth's surface and forms a layer that protects life on earth from the sun's harmful rays (25).

Health effects of traffic air pollution

Air pollution has both acute and chronic effects on human health. Health effects range anywhere from minor irritation of eyes and the upper respiratory system to chronic respiratory disease, heart disease, lung cancer, and death.

Air pollution has been shown to cause acute respiratory infections in children and chronic bronchitis in adults. It has also been shown to worsen the condition of people with pre-existing heart or lung disease. Among asthmatics, air pollution has been shown to aggravate the frequency and severity of attacks. Both, short-term and long-term exposures have also been linked with premature mortality and reduced life expectancy (13,26-28).

As we breathe, gases and particles of traffic exhaust are drawn into the lungs, where they contribute to a range of health problems. Pollutants can damage the lungs, as well as get into the bloodstream and travel to organs throughout the body (Figure 2) (29,30). When air pollution levels go up, there are more (29,30):

- emergency room visits,
- hospital admissions,
- asthma attacks,
- children absence from school,
- deaths from lung and heart diseases.

Air pollution can affect both the respiratory and cardiac systems. The health effects of air pollution can be seen as a pyramid (Figure 3), with the mildest but not common effects at the bottom of the pyramid, and the least common but more severe

at the top of the pyramid. The pyramid demonstrates that as severity decreases the number of people affected increases (29,30).

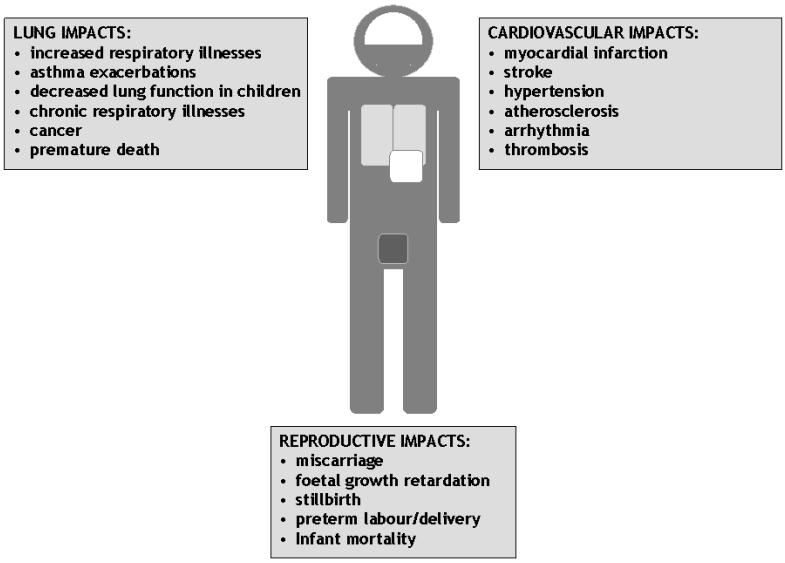


Figure 2. Air pollution effects in the body (Adapted from Mishra, 2003) (30).

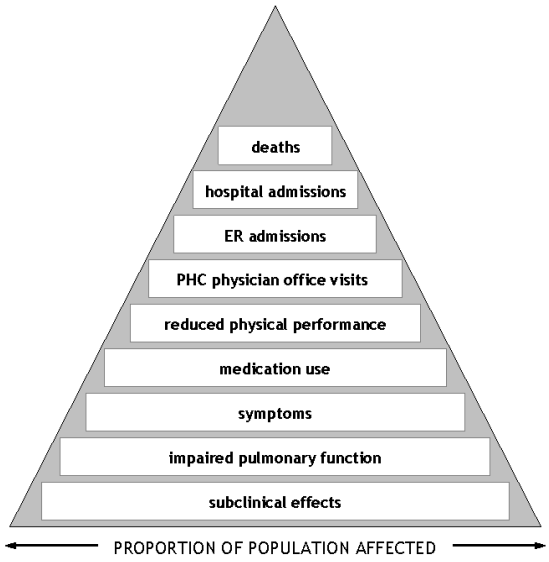


Figure 3. Pyramid of health effects of air pollution (Adapted from Mishra, 2003) (30).

Impact of pollutants on health are:

- Sulfur oxides (SO_x): sulphates and sulphur-containing acids are more toxic than SO_2 gas. They interfere with normal functioning of the mucous membrane in respiratory passages, increasing susceptibility to infection. At high concentration levels, both H_2S and organic sulphur compounds are toxic (16). SO_2 is a water soluble irritant gas (28). Sometimes causing discomfort and coughing in healthy people, and often causing severe respiratory symptoms in asthmatics. When people with asthma were exposed under controlled conditions to levels of SO_2 similar to those found near pollution sources such as ports, they developed an average decrease of 25-30% in their lung function (26). However, on occasional down-drafting of smokestack plumes or meteorological inversions near point-sources result in low parts per million (ppm) levels of SO_2 that may pose a hazard to some individuals. A two minute exposure to 0.4-1.0 ppm can elicit bronchoconstriction in exercising asthmatics within 5-10 minutes (16). Once deposited along the airway, SO_2 dissolves into surface lining fluid as sulphite or bisulphite and is readily distributed throughout the body. It is thought that the sulphite interacts with sensory receptors in the airways to initiate local and centrally mediated bronchoconstriction (28),
- Nitrogen oxides (NO_x): numerous studies have found that NO_x can cause toxic effects in the airways, leading to inflammation and to asthmatic reactions. In fact, people with allergies or asthma have far stronger reactions to common allergens such as pollen when they are also exposed to NO_x (26). In health effects, NO_2 can irritate the lungs and lower resistance to respiratory infection (31). Several risk assessment studies have shown that both short- and long-term exposure to NO_2 can induce effects to the human health and that, given the role of NO_2 as a precursor of other pollutants and as a marker of traffic related pollution, there should be benefits for the public health from keeping low NO_2 levels in the atmospheric air. Recent studies have shown that chronic exposure to the levels of air pollutants, such as NO_2, currently observed may have even higher impacts on mortality than acute exposure (32). Especially persons suffering from chronic respiratory diseases, such as asthma, are very sensitive to NO_2 at high concentrations.

  Several studies indicate that the combination of SO_x and NO_x in the air is particularly noxious - these compounds appear to act together to increase allergic responses to common allergens such as pollen and dust mites (26),
- Carbon monoxide (CO): CO is classified toxicologically as a chemical asphyxiant because its toxic action stems from its formation of carboxyhemoglobin, obstructing oxygenation of the blood for systemic transport (31). CO reduces the ability of blood to deliver oxygen to the cardiovascular and nervous system. Long-term exposure can cause brain damage due to the lack of oxygen reaching the brain. The symptoms of CO exposure are various and include dizziness, nausea, fatigue, and decreased muscular control (33),
- Volatile organic compounds (VOCs): benzene and butadiene are known to cause cancer in humans. Formaldehyde is very irritating to the airways, and is a probable carcinogen. Toluene at occupational exposure levels has been associated with birth defects and miscarriages. Other VOCs emitted by vehicles have also been linked to cancer, reproductive harm, asthma, or neurological disorders (26),

- Particulate matter ($PM_{2,5}$ and PM_{10}): fine particles are believed to be the most critical pollutant and main contributor to excess mortality, especially among people suffering from respiratory and cardiovascular diseases (1). Elevated PM_{10} concentrations were linked with increases in mortality and hospital admissions with respiratory symptoms, especially for sensitive population subgroups (9,34). Air pollution is a major environmental health problem causing approximately three million deaths per year in the world, as result of exposure to particulate matter (35). Correlation between increased exacerbation of respiratory diseases, cardiopulmonary morbidity, mortality and the levels of urban airborne particulate matter is now well established (36),
- Ozone (O_3): exposure to ozone as a secondary component formed in the atmosphere at some distance from traffic is widespread and not very different for people living in rural and urban areas, although it depends largely on the time spent outdoors (26). Thousands of scientific studies have been published on the health effects of ozone. Ozone can make people more susceptible to respiratory infections and can aggravate pre-existing respiratory diseases, such as asthma. Ozone can also cause irreversible changes in lung structure, which eventually lead to chronic respiratory illnesses, such as emphysema and chronic bronchitis (26,31).

Figure 4 shows where particles are deposited in the respiratory tract, depending on their size. Smaller particles (in particular $PM_{2.5}$) penetrate more deeply into the lung and may reach the alveolar region. Ultrafine particles contribute only slightly to PM_{10} mass but may be important from a health point of view because of the large numbers and high surface area. They are produced in large numbers by combustion especially internal combustion engines (29,30).

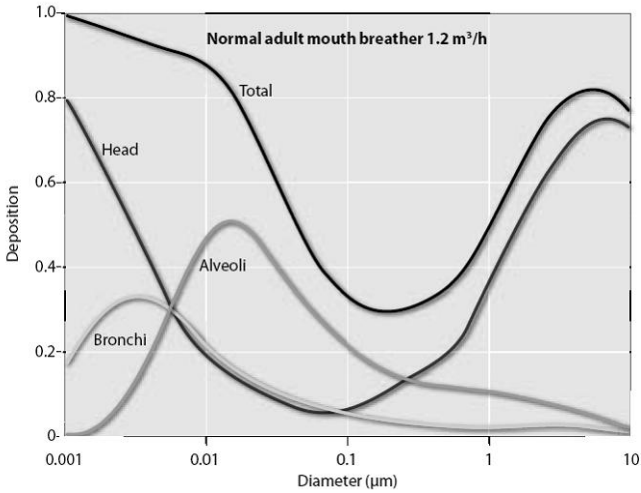


Figure 4. Deposition probability of inhaled particles in the respiratory tract according to particle size (Adapted from Mishra, 2003) (30).

Exposure and Health Effects

Exposure estimates to atmospheric pollutants can address individuals (personal exposure) or large population groups (population exposure), and can be based on direct (exposure monitoring) or indirect methods like (exposure modelling) (35). Exposure to ambient air pollution has been associated with a number of different health outcomes, starting from modest transient changes in the respiratory tract and impaired pulmonary function, continuing to restricted activity/reduced performance, emergency room visits and hospital admissions and to mortality. There is also increasing evidence for adverse effects of air pollution not only on the respiratory system, but also on the cardiovascular system. This evidence stems from studies on both acute and chronic exposure (Table 1) (9).

Table 1. Important health effects associated with exposure to different air pollutants (9).

Pollutant	**Effects related to short-term exposure**	**Effects related to long-term exposure**
1. Particulate matter	• lung inflammatory reactions • respiratory symptoms • adverse effects on the cardiovascular system • increase in hospital admissions • increase in mortality	• increase in lower respiratory symptoms • reduction in lung function in children • increase in chronic obstructive pulmonary disease • reduction in lung function in adults • reduction in life expectancy, owing mainly to cardiopulmonary mortality and probably to lung cancer
2. Ozone	• adverse effects on pulmonary function • lung inflammatory reactions • adverse effects on respiratory symptoms • increase in medication usage • increase in hospital admissions • increase in mortality	• reduction in lung function development
3. Nitrogen dioxide	• effects on pulmonary function, particularly in asthmatics • increase in airway allergic inflammatory reactions • increase in hospital admissions • increase in mortality	• reduction in lung function • increased probalility of respiratory symptoms

Some groups may receive much higher exposures than others, such as people who live and work near busy roads and those who travel or commute in heavy traffic. Also, the intake of pollutants by road users varies between drivers, bicyclists and pedestrians. Further, exposure to transport-related air pollution is difficult to separate from exposure to total air pollution (9,27,28).

People on road transport, in residencies and schools near the main roads are exposed to high levels of vehicle pollutant emissions (13).

Population exposure to traffic pollution can be classified into three general categories:

- contribution of traffic emissions to rural and urban background air pollution;
- locally elevated levels of traffic pollution near busy roadways and city zones; and
- still further elevated exposure concentrations in commuting/transport (29,31).

In Figure 5, three types of pollutants and their impact range is presented.

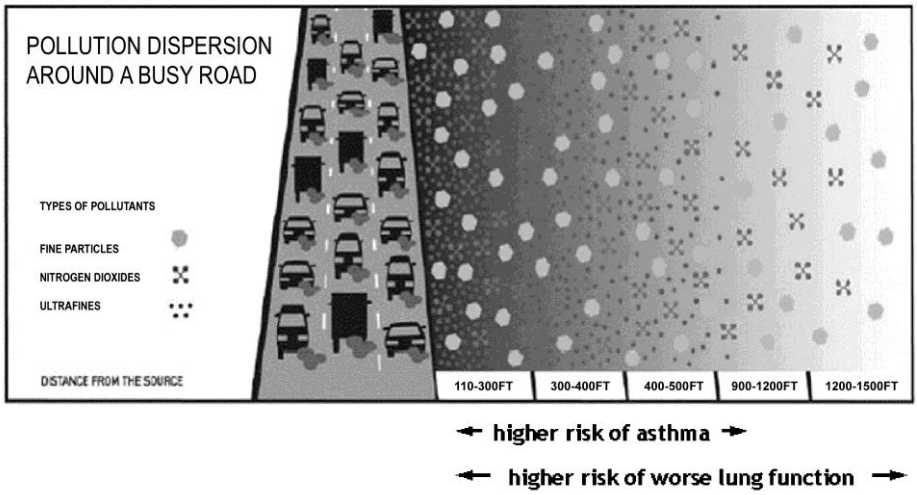


Figure 5. Traffic spreads pollution up to 1500 feet (about 450 meters) from the roadway (Adapted from Yeh, 2005) (31). NOTE: 1 foot equals to 0.3048 m. LEGEND: higher risk of asthma up to 500 feet (160 m); higher risk of worse lung function.

Geographical dispose of chronic lung diseases at children in Koper community

Introduction

Chronic lung diseases are becoming important public health concern mainly in developed countries. Asthma, chronic bronchitis and allergy are in many cases present in childhood. The most frequent lung illness among children in developed world and in Slovenia is asthma (37).

Inhabitants in west and north part of Koper community were complaining about lung diseases. How so ever analyses showed that mortality in Koper community in case of lung illnesses is higher than in others Slovenian communities. In year 2003 study, focused on health in different geographical party of Koper community was made. Aim of this study was to found out if health risk is higher in parts of Koper community where harmful factors exists in comparison to other parts where these environmental factors are less present (37).

Geographic review

Data

Different sources of data are used in GIS system:

1. For research purposes geographic information system was used, to show lung illness distribution.

   In the system mentioned above data were collected, saved, controlled, processed and analysed.
2. Two kinds of data can be joined in GIS - descriptive and geographic. From the data-base that contained address for each questionnaire data about location or geocode (X and Y coordinate; Gauss-Krueger coordinate system D48/GK) were made. Relation base was made from existing data, being number of questionnaire, geocode of location (X, Y), distribution regarding community area, distribution regarding postal codes, and distribution regarding health status.
3. Data were used to calculate prevalence of chronic diseases: asthma, chronic bronchitis and allergy. Nevertheless, another base of illnesses that is in GIS connected with geographic survey of different data groups was made (37).
4. Data base (vector data) was moved to GIS, so that questionnaires could be located. Cartographic pads of national topographic maps in measure 1:25,000 and 1:50,000 (DTK 25 and DTK 50) were used (raster presentation). Maps are in ownership of Surveying and Mapping Authority of the Republic of Slovenia. Also, artificial personal data that due to activity that can influence on environment, were incorporated in GIS. Data were collected in special study, which was made by OIKOS d.o.o. company in the year 2006.
5. In GIS system layers were entered as well as borders of Koper community, and borders of local communities of Koper community. For processing and reviewing of the data program software ESRI Arc View 3.2 1996 was used (37).

Air pollution

Industry, traffic and combustion devices are the most important sources of air pollution in Koper community. Permanent source of pollution during the year is industry. On the other side traffic, despite its permanent presence, is characterized by significant weekly as well as seasonal dynamic. The higher emissions are present in summer and at the week-ends, when it is the most dense.

Regulation and air pollution control

The most direct intervention for improving air quality is of course through regulation of emissions. The EU legislation on vehicle emission and fuel quality

standards has evolved greatly since the first directive in 1970. The early legislation had the dual purpose of reducing pollution and avoiding barriers to trade due to different standards in different member states.

The "Auto Oil I Programme" undertaken by the European Commission in conjunction with industry, has set up the targets for a series of traffic-related pollutants and assessed different technologies and fuel quality standards (38,39). The base for air quality regulation in Slovenia is Environment Protecting Act (RS) No. 41/2004, 17/2006, 20/2006, 28/2006, 39/2006, 49/2006, 66/2006, 112/2006, 33/2006 (40). Howsoever, in the year 2005 Slovenian government accepted Decree on nation emission ceilings for atmospheric pollutants (RS) No. 24/2005, 92/2005 where national maximum emission of SO_2, NO_x, VOC and NH_3 were established (41). Strategies and arrangements for enforcement of the decree are described in Operational programme on national emission ceilings for atmospheric pollutants in ambient air (RS) No. 24/2005 (42). Some specific claims for each individual pollutant is defined in Decree on benzene and carbon monoxide in ambient air (RS) No. 52/2002 and Decree on sulphur dioxide, nitrogen oxides, particulate matter and lead in ambient air (RS) No. 52/2002, 18/2003, 121/2006 (43,44).

Table 2. Maximum concentrations of some pollutants (WHO-Air quality guidelines) (45)

Pollutant	Time interval	Maximum concentration
PM 2,5	Annual	10 $\mu g/m^3$
	24-hour	25 $\mu g/m^3$
PM 10	Annual	20 mg/m^3
	24-hour	50 mg/m^3
O3	8-hour	100 $\mu g/m^3$
NO2	Annual	40 $\mu g/m^3$
	1-hour	200 $\mu g/m^3$
SO2	24-hour	20 $\mu g/m^3$
	10-minute	500 $\mu g/m^3$

Practical work with GIS

The process of working with the GIS model could be split in several steps. The addresses were converted into spatial location data, which is needed to present, analyse attribute data in spatial setting. We used point data attained from geocoding and regional data to give an overview o situation (diminish variability in population counts for regions).

Data insertion and processing

Geocoding

Geocoding is used as method for assigning a spatial location to an address record. Although many possibilities exist we used public available databases of Environmental Atlas of Slovenia (EAoS). Point shape files were created:

1. Location information (local community, street and number) was collected with poll in pilot research "Environmental impact on incidence of some illness and mortality in Koper community" in time period 2002-2003.
2. Since information about permanent residence is indirect data in each poll, data about location (geocoded address) are in two different data-bases :

- EAoS is available at website of Environmental Agency of the Republic of Slovenia (http://gis.arso.gov.si/atlasokolja/) where single person can be determined on the lever of permanent address. In program house numbers can be found with coordinates X and Y and resolution of 1: 1,500. Data are then copied in database (C/P),
- and for the better matching and indirect precision control of geolocation test with data base ENHIS (evidence of house numbers, dBase, GURS).

Geocodes can be checked up in case of every single address. In birth database matching local community can be found.

Relation database

Result of geocoding is relation database (contains following data; poll cipher, geocoded location (X,Y), local community, post number and health condition) in Microsoft Excel and then transformed in dBase for program ESRI ArcView 3.2 and new version of ESRI ArcMap 9.2.

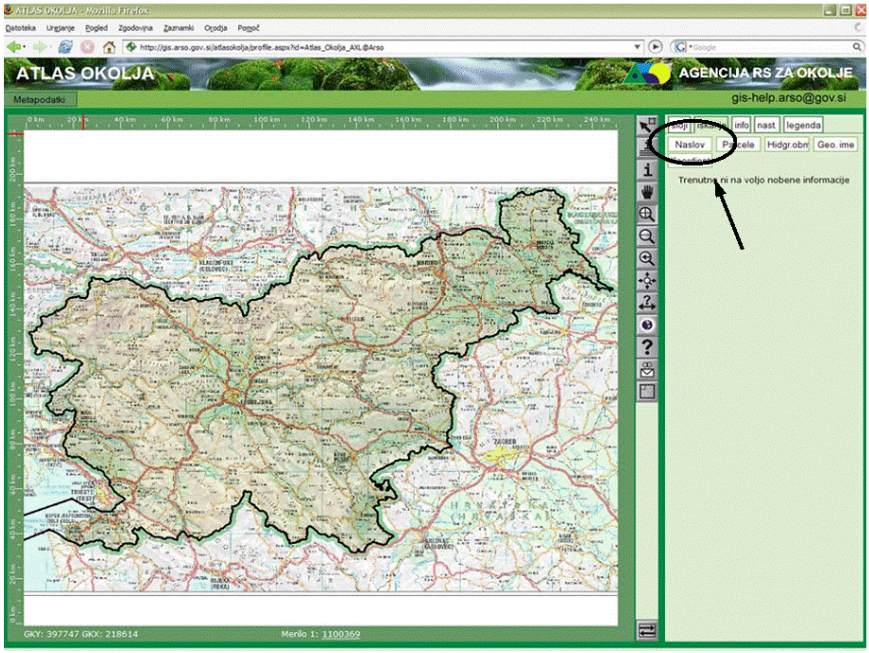
Data insertion in GIS

In GIS sphere geolocated raster map is inserted (.tiff) in adequate measure. Data that has to be in dBase are transformed in individual layer. Further data manipulation like community borders and addition of others data bases is dependent from program package.

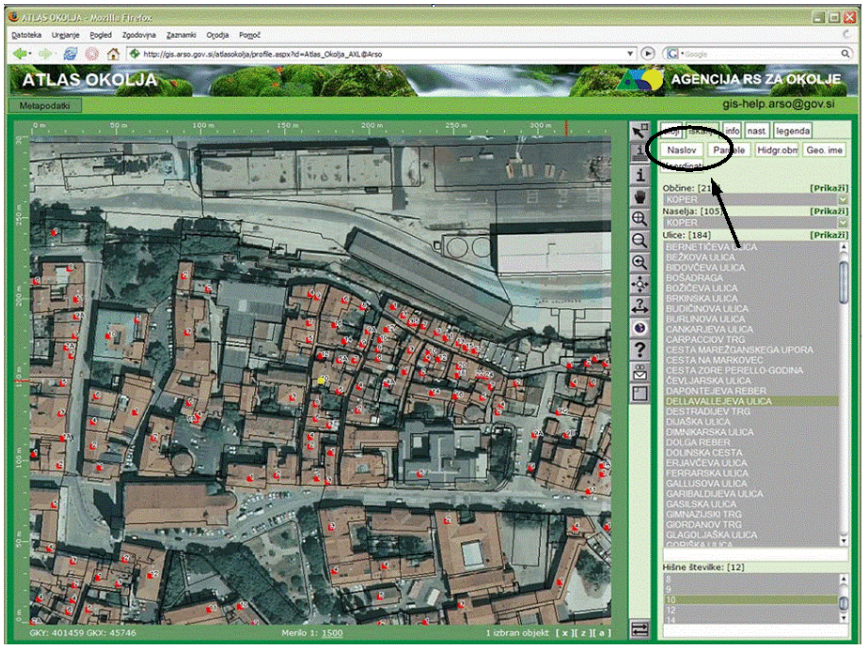
Algorithm

Algorithm is as follows:

1. Location control.
   In program GIS first choose option "Search" and then "Parcel address". Insert community and settlement. Figure 6 shows where in the communication window address in EAoS (Figure 6) is to be inserted (Note:geocoding address can be more easily carried out from the ENHIS but this data-base is more expensive than EAoS).
   In continuation, street and street number of the addresses of diseased children collected in a special survey (37) need to be found in the EAoS. The EAoS program will show exact location. Figure 7 shows where to choose a command to run procedure for information about coordinates (Figure7).
   Mark both coordinates and transform them in Microsoft Word or Excel. At the end, addresses are controlled against the survey data.
2. Composition of information for running the GIS programme.
   This information is prepared in a data-matrix form (it is simply formed in Microsoft Excel programme), and it is composed of: X,Y coordinates of location of address, local community code of the address, and zip code of the address. When prepared it is simply copied to GIS programme (Figures 8 and 9).
3. Insertion of raster map (1:50,000) in GIS (ArcMap) and posing of all previously prepared data into GIS (Figure 10).



a)



b)

Figure 6. Insertion of addresses of permanent residence addresses in Environmental Atlas of Slovenia.

Figure 7. Where to choose „i" for information about coordinates

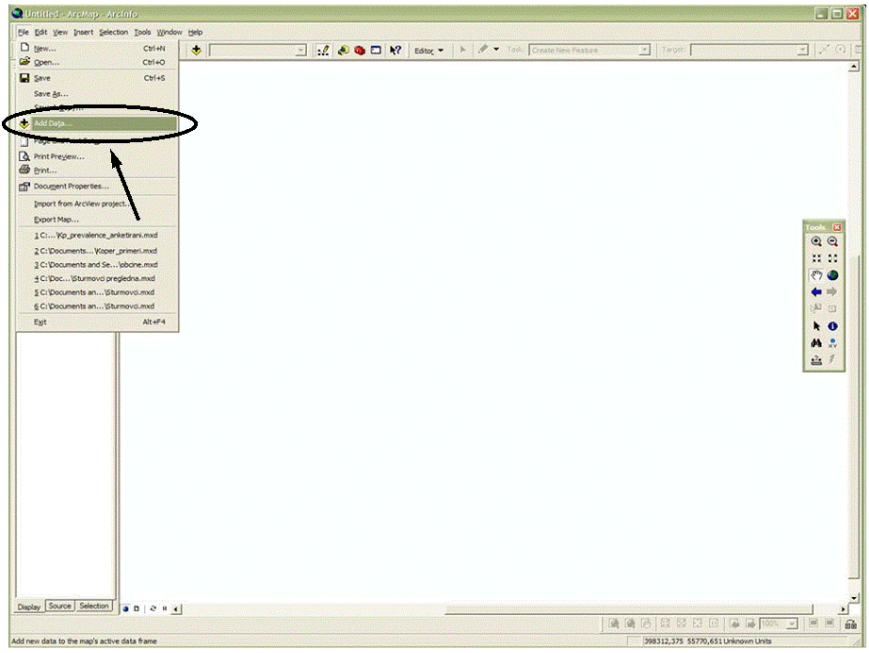


Figure 8. Function "File/Add data" recognise TIFF data for location of map in space. In left field title of map can be seen.

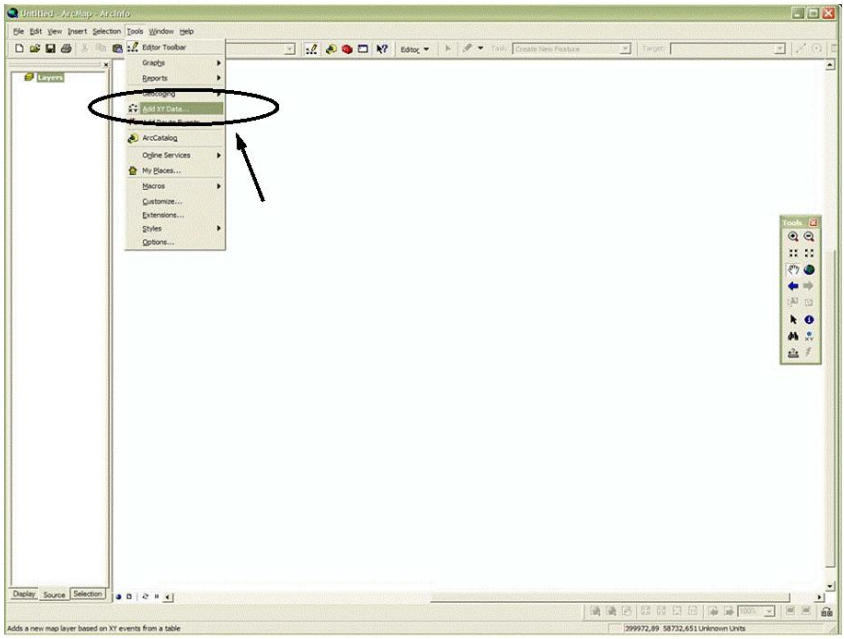


Figure 9. Coordinates X, Y are transformed in ArcMap 9.2 with simple function were only names of colons are necessary (Function Tools/Add X, Y data). Choose base /shape.

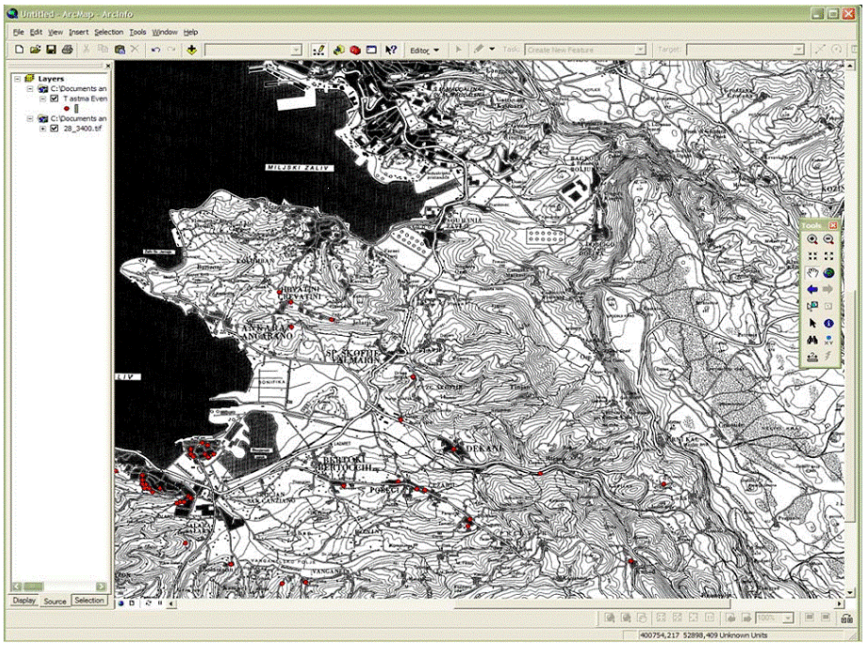


Figure 10. The basic view of coordinates can be seen. With right mouse button individual properties, correlations, changing properties can be done.

Algorithm of assigning area or regional data

Algorithm of assigning area or regional data is as follows:

1. In sphere GIS "GIS shape" communities (2 dimensional form of community borders) are transferred.
2. With function "Add data" (Figure 8):
    - in sphere GIS table need to be complement ("Attribute table" and function "Edit". Changeability, prevalence and constant risk hypothesis, etc.),
    - data manipulation (clusters of changeability, prevalence and different levels of review),
    - covering of different layers (Hypothesis about correlation of sea level can be check up. Sea level in insert in GIS and adjacency of single case with isohyps. Histogram of cases in isohyps is a result).

Appearance overview of chronic lung diseases

Overview of all types of chronic lung disease shows accumulation of cases in some local communities. For example, risk for some chronic lung diseases in childhood is 4.23-times higher in local community Pobegi Čežarji, and 3.67-times higher in local community Hrvatini, when compared to control community Šmarje pri Jelšah (37) that was used as a non-polluted reference local community (situated in the eastern part of Slovenia).

Geographic distribution of diseased children address

Overview of geographic distribution of permanent residence (children's addresses) that has chronic lung disease showed that some areas of Koper community are more risky for presence of chronic lung diseases than others (Figures 11-15). More detailed overview shows that children with chronic lung disease live mostly in higher locations (700-900 meters above the sea level) (37).

Conclusion

This case study is showing the possible use of GIS in presenting the geographic distribution of various health states, more specific the geographic distribution of various chronic lung diseases and related states in Koper community/Slovenia. This account is especially expressed in polluted areas. Results of analysis shows higher frequency of chronic lung diseases in more polluted than in less polluted local communities.

In this stage the geographical distribution of level of air pollution of local communities is missing.

In the Koper community further studies about pollution sources and their distribution as well as their impact on health should be taken, and joined with the data on prevalence of chronic lung diseases in children.

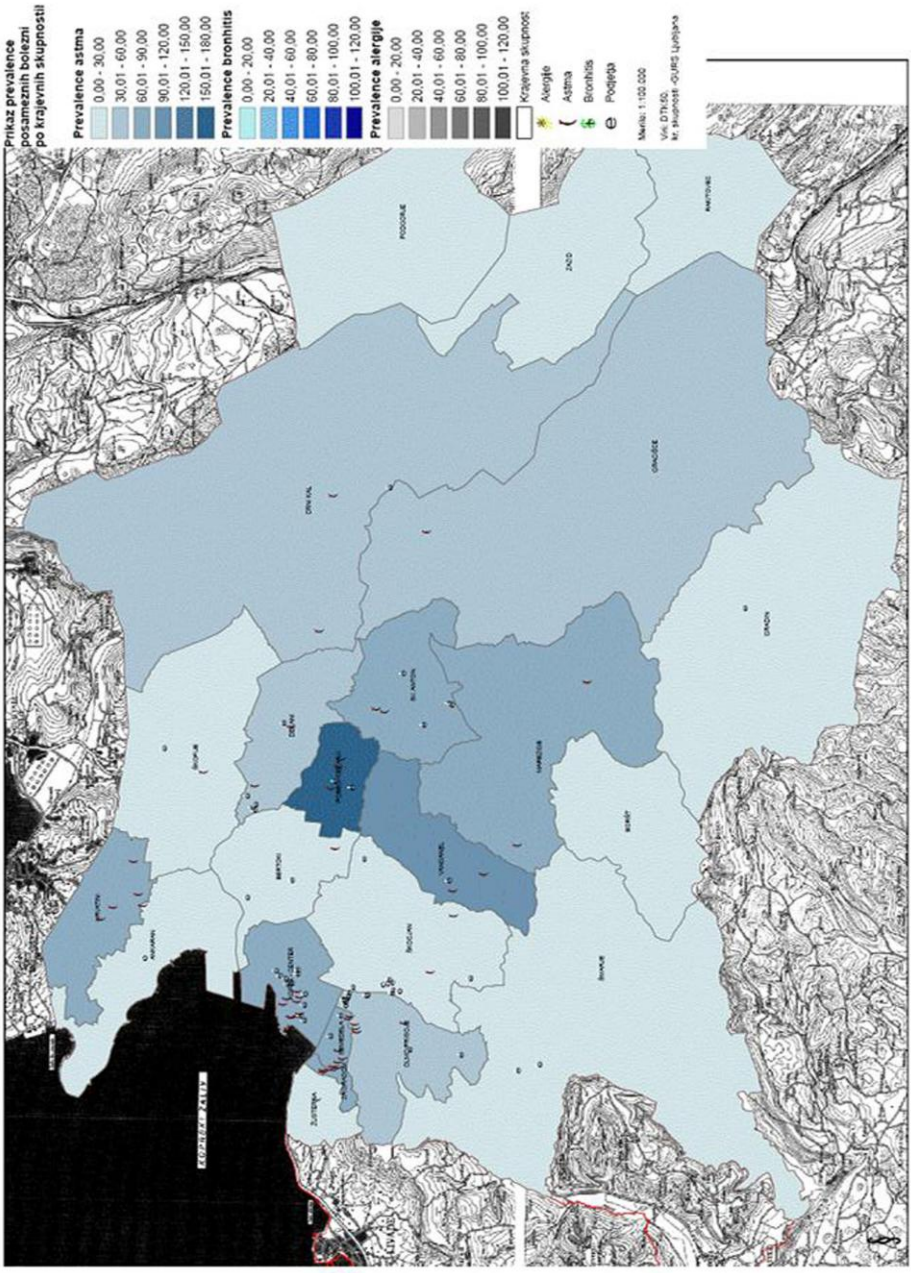


Figure 11. Geographic distribution of permanent residence addresses (corresponding local community) of children that were in year 2002/2003 in fourth class primary school and had asthma (37).

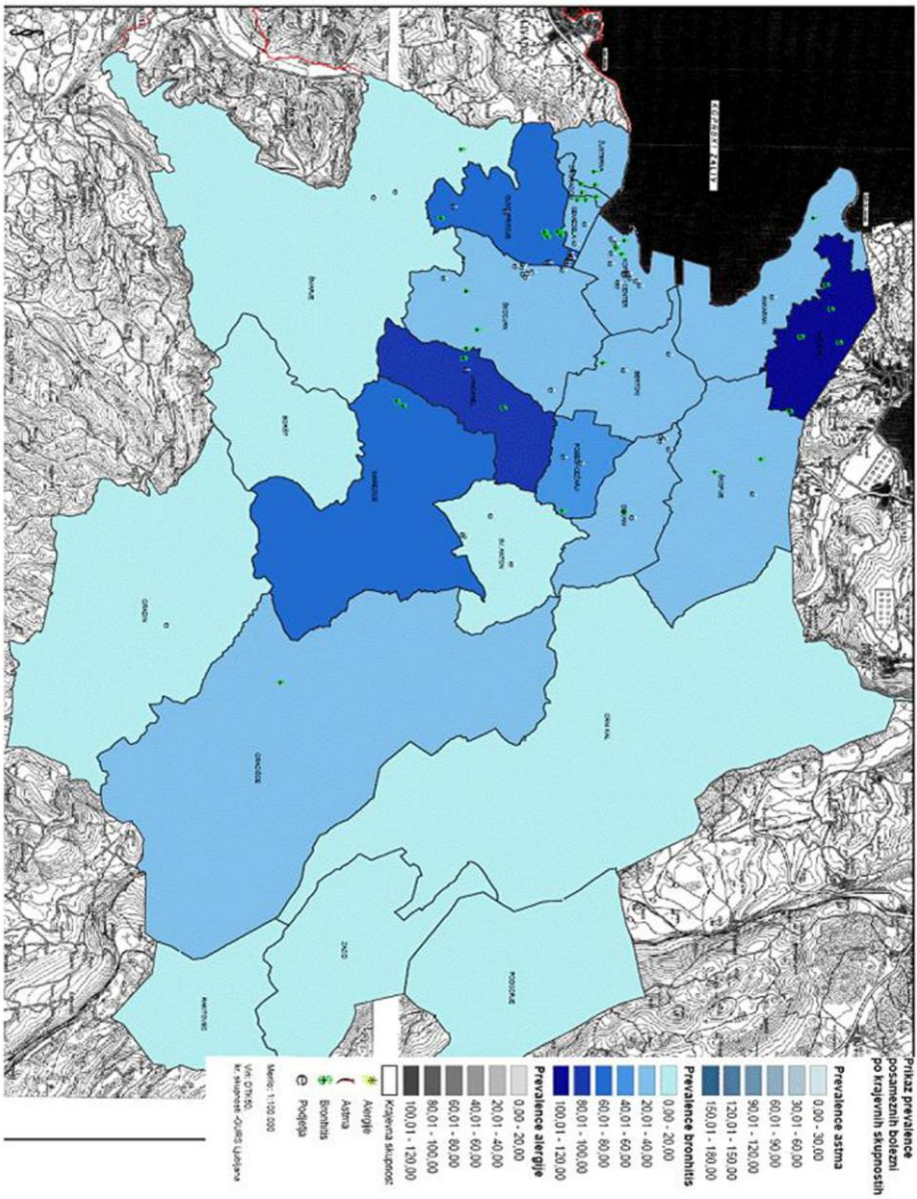


Figure 12. Geographic distribution of permanent residence addresses (corresponding local community) of children that were in year 2002/2003 in fourth class of primary school and had chronic bronchitis (37).

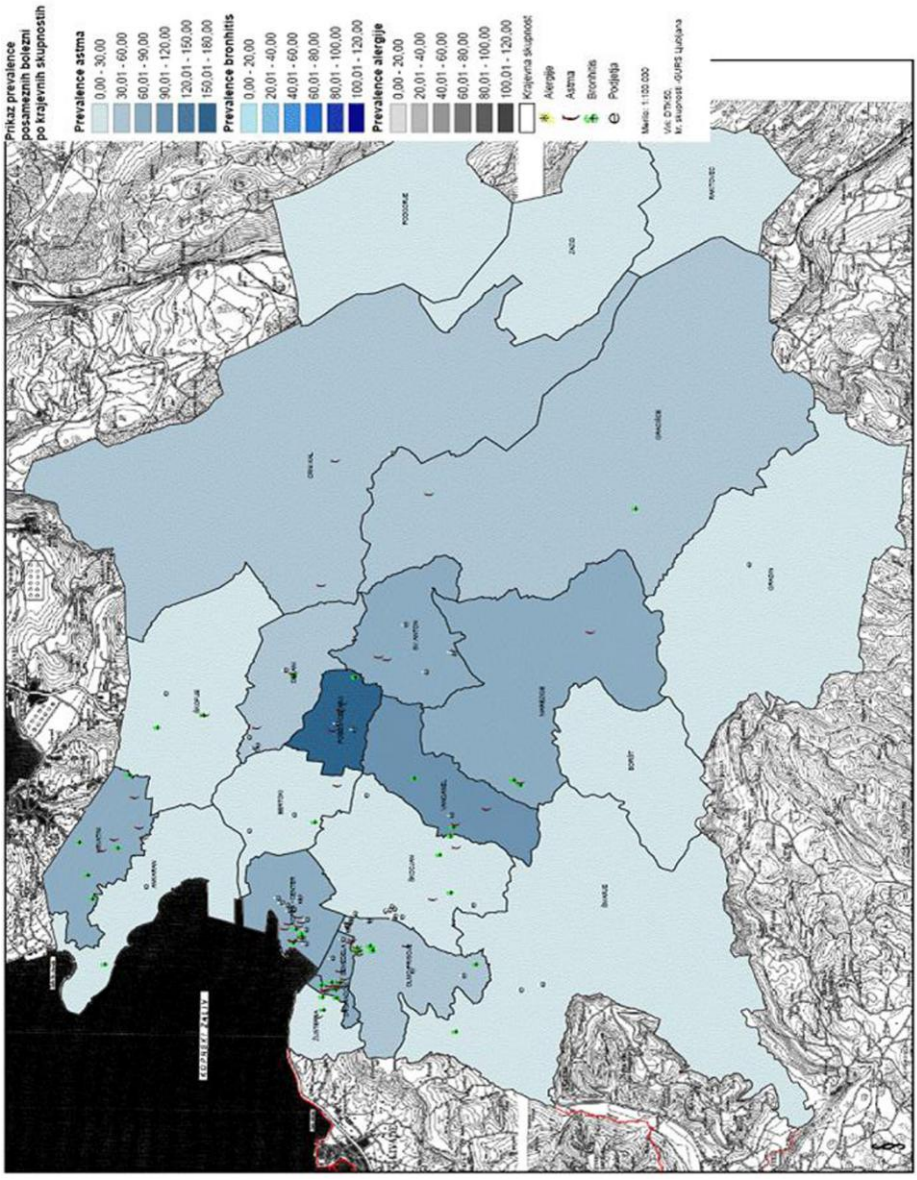


Figure 13. Geographic distribution of permanent residence addresses (corresponding local community) of children that were in year 2002/2003 in fourth class of primary school and had asthma or chronic bronchitis (37).

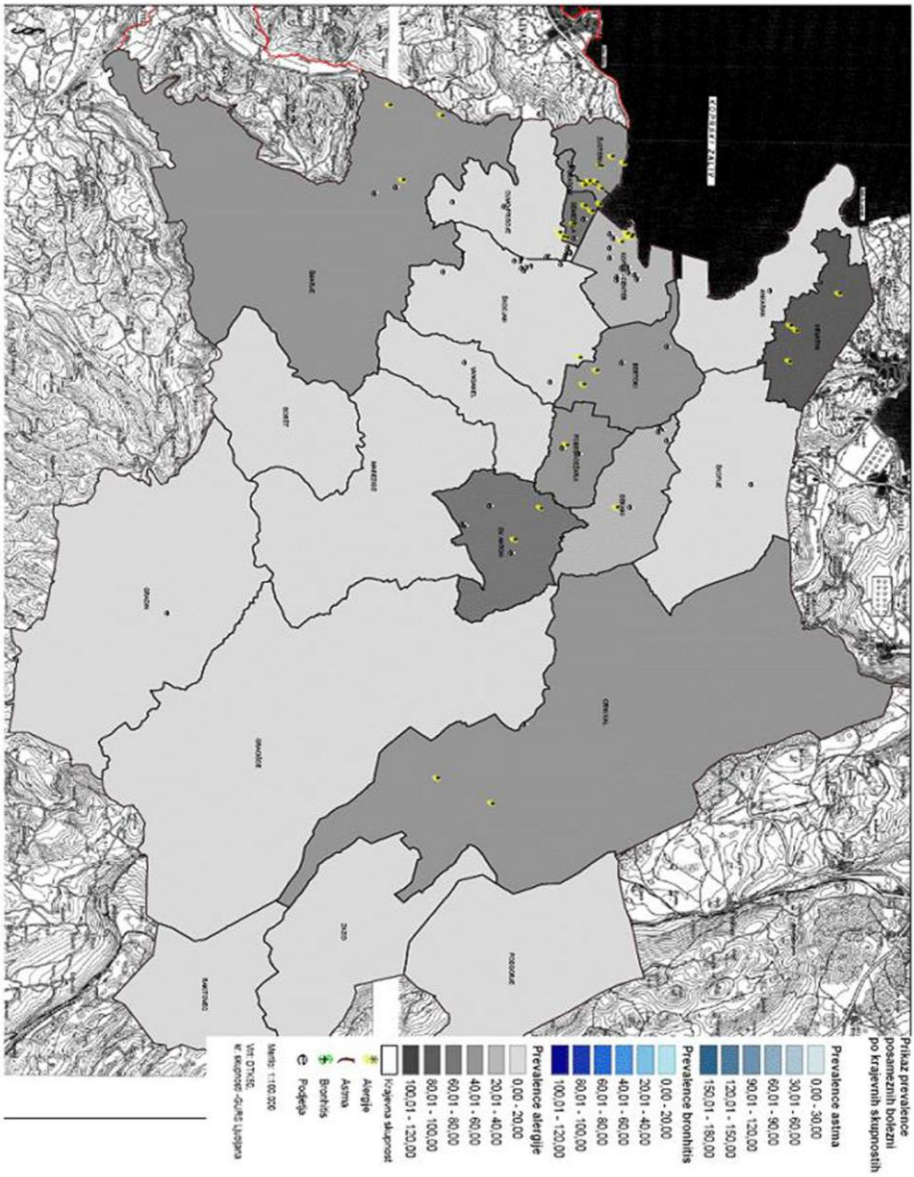


Figure 14. Geographic distribution of permanent residence addresses (corresponding local community) of children that were in year 2002/2003 in fourth class of primary school and had allergy (37).

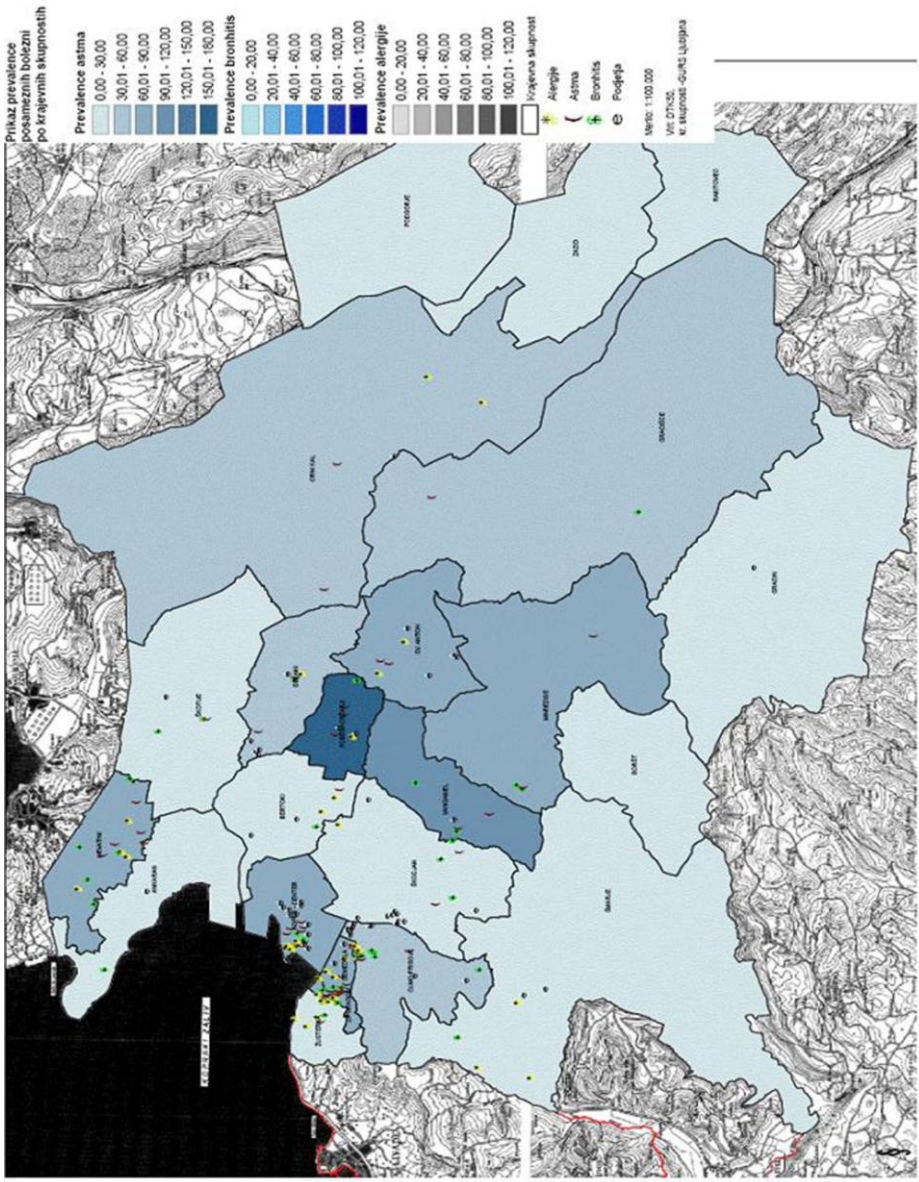


Figure 15. Geographic distribution of permanent residence addresses (corresponding local community) of children that were in year 2002/2003 in fourth class of primary school and had chronic lung disease (37).

EXERCISE

Task 1

Carefully read the theoretical background of this module and discus with other students about traffic air pollution and pulmonary diseases. Please, answer the following questions:

- Which pollutants have the major impact on human health?
- Can pollutants be reduced?
- What is the problem of reducing them?
- Is GIS model useful as a tool in epidemiology?
- What are its advantages and disadvantages?

Task 2

Find out if in your country data and tools for accomplishing a task like presented in our case study are available.

Task 3

Given you have access to the tools and the data necessary for accomplishing the task, presented in our case study, repeat the steps presented in this case study.

REFERENCES

1. Jensen SS. Mapping human exposure to traffic air pollution using GIS. J Hazard Materials. 1998;61:385-92.
2. Gulliver J, Briggs JD. Time-space modeling of journey-time exposure to traffic-related air pollution using GIS. Environ Research. 2005;97:10-25.
3. Cromley EK, McLafferty SL. GIS and Public Health. New York, NY: The Guilford Press; 2002.
4. National Environmental Research Institute. AIRGIS [homepage on the Internet]. Available from URL: http://www.dmu.dk/International/Air/Models/AIRGIS/. Accessed: February 15, 2008.
5. Tang UW, Wang ZS. Influences of urban forms on traffic-induced noise and air pollution: Results from a modelling system. Environmental Modelling & Software. 2007;22:1750-64.
6. Jarup L. The role of geographical studies in risk assessment. In: Elliot P, Wakefield J, Best N, Briggs D, editors. Spatial epidemiology: Methods and applications. Oxford: Oxford University Press; 2006. p.415-33.
7. Lomborg B. The sceptical environmentalist. Measuring the real state of the world. Cambridge: Cambridge university press; 2001.
8. Mahboob A, Makshoof A. Air pollution due to traffic, air quality monitoring along three sections of National Highway N-5, Pakistan. Environ Monit Assess. 2008;136:219-26.
9. World Health Organization, Regional Office for Europe. Health risk of particulate matter from long-range transboundary air pollution Copenhagen: World Health Organization, Regional Office for Europe; 2006. Available from URL: http://www.euro.who.int/document/E88189.pdf. Accessed: February 15, 2008.
10. World Health Organization, Regional Office for Europe. Health aspects of air pollution results from the WHO project »Systematic review of health aspects of air pollution in Europe«.

Copenhagen: World Health Organization, Regional Office for Europe; 2004. Available from URL: http://www.euro.who.int/document/E83080.pdf. Accessed: February, 15 2008.
11. Ostro B. Outdoor air pollution: Assessing the environmental burden of disease at national and local levels. Environmental Burden of Disease Series, No.5. Geneva: World Health Organization; 2004. Available from URL: http://www.who.int/quantifying_ehimpacts/publications/ebd5.pdf. Accessed: February 15, 2008.
12. Lah A. Remodeling of transport systems (in Slovene). In: Traffic and environment (in Slovene). Ljubljana: Svet za varstvo okolja Republike Slovenije; 2002.
13. Kongtip P, Thongsuk W, Yoosook W, Chantanakul S. Health effects of metropolitan traffic-related air pollutants on street vendors. Antmospheric Environment. 2006;40:7138-45.
14. Energy Information Agency (EIA). Petroleum [homepage on the Internet]. Available from URL: http://www.eia.doe.gov/oil_gas/petroleum/info_glance/petroleum.html. Accessed: February 15, 2008.
15. Woodcock J, Banister D, Edwards P, Prentice AM, Roberts I. Energy and transport. Lancet. 2007;370:1078-88.
16. Wang LK, Williford C, Chen WY. Desulfurization and emissions control. In: Wang LK, Pereira NC, Hung YT, editors. Handbook of Environmental Engineering, Volume 2: Advanced Air and Noise Pollution Control. Totowa, NJ: Humana Press, Inc.; 2005. p.35-92.
17. Holman C. Sources of air pollution. In: Holgate S, Samet JM, Koren HS, Maynard RL, et al, editors. Air pollution and health. London: Academic Press; 1999. p.115-48.
18. Fogg PGT, Sangster JM. Chemicals in the atmosphere: Solubility, sources and reactivity. Chichester: Wiley; 2003.
19. Dervent RG. Atmospheric Chemistry. In: Holgate S, Samet JM, Koren HS, Maynard RL, et al, editors. Air pollution and health. London: Academic Press; 1999. p.63-51.
20. Colls J. Air Pollution. 2th. Ed. London: Spon Press; 2002.
21. Fenger J. Urban air quality. Atmospheric Environment. 1999;33:4877-4900.
22. Garling T, Steg L. Environmental effects of urban traffic. In: Garling T, Steg L. Threats from car traffic. The quality of urban life: problems, causes, solutions. Amsterdam: Elsevier Science; 2007. p.11-33.
23. Anonymous. Sustainable consumption and production (in Slovene**)**. In: Bernard-Vukadin B, editor. Environment on hand (in Slovene). Ljubljana: Agencija Republike Slovenije za okolje; 2007. Available from URL: http://nfp-si.eionet.europa.eu/News/NEWS1203675146/Dokumenti/pdf/OND07/OND1_tpp.pdf. Accessed: March 14, 2008.
24. Schultz MG, Feichter J, Leonardi J. Climatic impact of surface transport. In: Hester RE, editor. Transport and the environment. Cambridge: Royal Society of Chemistry; 2004. p.111-26.
25. U.S. Environmental Protection Agency. Ground-level Ozone [homepage on the Internet]. Available from URL: http://www.epa.gov/air/ozonepollution/. Accessed: March 14, 2008.
26. Bailey D, Solomon G. Pollution prevention at ports: clearing the air. Environ Impact Assess Rev. 2004;24:749-74.
27. Schnelle BK, Brown AC. Air pollution control technology handbook. London: CRC Press; 2001.
28. Wright DA, Welbourn P. Environmental toxicology. Cambridge: Cambridge University press; 2002.
29. Health & Environment Impact Project. "Moving Forward" Conference 2007. University of Southern California, 2007 [homepage on the Internet]. Available from URL: http://www.scribd.com/doc/907226/2-Air-Pollution-Health-Effects-2007. Accessed: February 15, 2008.
30. Mishra V. Health Effects of Air Pollution, 2003 [homepage on the Internet]. Available from URL: http://www.mnforsustain.org/climate_health_effects_of_air_pollution_mishra_pern.htm. Accessed: February 15, 2008.

31. Yeh TJ, Chen WY. Control of NO_x During Stationary Combustion. In: Wang LK, Pereira NC, Hung YT, editors. Handbook of Environmental Engineering, Volume 2: Advanced Air and Noise Pollution Control. Totowa, NJ: Humana Press, Inc.; 2005. p.113-24.
32. Chaloulakou A, Mavroidis I, Gavril I. Compliance with the annual NO_2 air quality standard in Athens. Required NOx levels and expected health implications. Atmospheric Environment. 2008;42:254-65.
33. New Hampshire Department of Environmental Services (NHDES). Air Resources Division [homepage on the Internet]. Available from: URL: http://des.nh.gov/organization/divisions/air/index.htm. Accessed: February 15, 2008.
34. Grivas G, Chaloulakou A, Kassomenos P. An overview of the PM_{10} pollution problem, in the Metropolitan Area of Athens, Greece. Assessment of controlling factors and potential impact of long range transport. Science Total Environ. 2008;389:165-77.
35. Borrego C, Tchepel O, Costa A M, Martins H, Ferreira J. Urban population exposure to particulate air pollution induced by road transport. New York: Springer US; 2007.
36. Garçon G, Dagher Z, Zerimech F, et. al. Dunkerque city air pollution particulate matter-induced cytotoxicity, oxidative stress and infammation in human epithelial lung cells (L132) in culture. J Toxicol Vitro. 2006;20:519-28.
37. Erlih S. Geographic Information System (GIS) as a tool in geographical epidemiology – A case study. In: Cvahtetovi dnevi 2007. Zbornik prispevkov. Ljubljana: 28.september 2007. Medicinska fakulteta, Katedra za javno zdravje; 2007. p.48-58.
38. Fenger J. Urban-scale air pollution. In: Hewitt NC, Jackson AV, editors. Handbook of atmospheric science: Principles and applications. Chichester: Wiley-Blackwell; 2003. p.401-38.
39. Environment Protecting Act. Official Gazette of the Republic of Slovenia No. 41/2004, 17/2006, 20/2006, 28/2006, 39/2006, 49/2006, 66/2006, 112/2006, 33/2006.
40. Blanken PD, Dillon J, Wismann G. The impact of an air quality advisory program on voluntary mobile source air pollution reduction. Atmospheric Environment. 2001;35:2417-21.
41. Decree on sulphur dioxide, nitrogen oxides, particulate matter and lead in ambient air. Official Gazette of the Republic of Slovenia No. 52/2002, 18/2003, 121/2006.
42. Decree on nation emission ceilings for atmospheric pollutants. Official Gazette of the Republic of Slovenia No. 24/2005, 92/2005.
43. Decree on benzene and carbon monoxide in ambient air (RS) No. 52/2002. Operational programme on national emission ceilings for atmospheric pollutants in ambient air. Official Gazette of the Republic of Slovenia No. 24/2005.
44. Decree on sulphur dioxide, nitrogen oxides, particulate matter and lead in ambient air. Official Gazette of the Republic of Slovenia No. 52/2002, 18/2003, 121/2006.
45. World health organization. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Geneva: World health organization; 2005. Available from URL: http://whqlibdoc.who.int/hq/2006/WHO_SDE_PHE_OEH_06.02_eng.pdf. Accessed: February 15, 2008.

RECOMMENDED READINGS

1. Jarup L. The role of geographical studies in risk assessment. In: Elliot P, Wakefield J, Best N, Briggs D, editors. Spatial epidemiology: Methods and applications. Oxford: Oxford University Press; 2006. p.415-33.
2. Jensen SS. Mapping human exposure to traffic air pollution using GIS. J Hazard Materials. 1998;61:385-92.
3. National Environmental Research Institute. AIRGIS [homepage on the Internet]. Available from URL: http://www.dmu.dk/International/Air/Models/AIRGIS/. Accessed: February 15, 2008.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**BASIC OCCUPATIONAL HEALTH INDICATORS ON SICK LEAVE**
Module: 2.1.6	**ECTS (suggested): 0.15**
Author(s), degrees, institution(s)	**Marjan Bilban**, MD, PhD, Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Marjan Bilban** Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia Zaloška 4, Ljubljana, Slovenia E-mail: marjan.bilban@zvd.si
Keywords	Occupational health, health indicators
Learning objectives	After completing this module students should: • know the definition and characteristics of occupational health indicators on sick leave; • know where to find data for calculating these indicators; • be able to independently calculate these indicators.
Abstract	Monitoring of health related phenomena at workplaces is of utmost importance in the process of making workplaces safer and healthier. There exist several different occupational health indicators. A group of occupational health indicators on sick leave or sickness absence is a group of major indicators which provides information on the health status of the persons in employment. In this module four basic indicators, indicating occupational health on general, are presented, being sick leave percentage, frequency index, severity index, and inactivation index.
Teaching methods	An introductory lecture gives the students insight in characteristics of basic occupational health indicators on sick leave. The theoretical knowledge is illustrated by a case study. After introductory lectures students calculate sick leave percentage, frequency index, severity index, and inactivation index indicators by themselves using available data sources.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Calculation of indicators using given data. Only the correct procedure and result of the calculation is considered as positive.

BASIC OCCUPATIONAL HEALTH INDICATORS ON SICK LEAVE

Marjan Bilban, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Monitoring of health related phenomena at workplaces is of utmost importance in the process of making workplaces safer and healthier. Data are needed to determine the magnitude of work-related diseases and injuries, identify workers at greatest risk, and establish prevention priorities, since this kind of morbidity can be prevented. Data are also necessary to measure the effectiveness of prevention activities, and to identify workplace health and safety problems that need further investigation.

Definitions of basic terms

Occupational health

According to World Health Organization (WHO) (1), Several definitions of occupational health and safety and occupational health services have been produced by professional bodies, international organizations such as WHO and IL0-International Labour Organization, and different national bodies and authorities. When summarizing these definitions, occupational health is defined as (1):

- protection and promotion of the health of workers by preventing and controlling occupational diseases and accidents and by eliminating occupational factors and conditions hazardous to health and safety at work;
- development and promotion of healthy and safe work, work environments and work organizations;
- enhancement of physical, mental and social well-being of workers and support for the development and maintenance of their working capacity, as well as professional and social development at work;
- enablement of workers to conduct socially and economically productive lives and to contribute positively to sustainable development.

Health indicator

According to A Dictionary of epidemiology (2), health indicator is defined as a variable, susceptible to direct measurement, that reflects the state of health of persons in a community. Health indicators are measures which help to compare health status among different populations or population groups.

Occupational health indicator

An occupational health indicator is a specific measure of a work-related disease or injury, or a factor associated with occupational health, such as workplace exposures, hazards, or interventions, in a specified population (3).

There exist several different occupational health indicators. They could be general, indicating overall occupational health of a population, or specific, indicating occupational health of a specific population group (e.g. males and females, different age groups, etc.).

In this module four basic indicators, indicating occupational health on general, are presented, being sick leave percentage, frequency index, severity index, and inactivation index (4). All four indicators are indicators of sick leave or sickness absence.

Sick leave

According to Merriam-Webster Online Dictionary, sick leave**Napaka! Zaznamek ni definiran.**, in a meaning, used in this module, is an absence from work permitted because of illness (5).

Basic occupational health indicators on sick leave

A group of occupational health indicators on sick leave or sickness absence is a group of major indicators which provides information on the health status of the persons in employment. Sick leave figures are often used for example to reveal the need for preventive activities if absence rates are high. At a national level, absence rates are usually examined according to economic sectors to determine what action is necessary. It is also common to consult absence rates at company level in order to determine which departments should be targeted by health promotion activities. The effectiveness of health promotion activities is then often evaluated by the changes in sickness absence rates (6).

Sick leave percentage

Sick leave percentage (SL%) is a percentage of lost days because of a sick leave on one persons in employment. By sick leave percentage, daily percent ob absence is expressed (7). In calculation, calendar days, working days, or effective hours on one person in employment could be used.

When sick leave percentage for calendar days is calculated, number of persons in employment in denominator of the equation should be multiplied by average number of days in a year, being 365 (Equation 1a) (4,7):

$$SL\%_{CD} = \frac{N_{lost\ calendar\ days}}{N_{persons\ in\ employment} \times N_{days\ in\ a\ year(average)}} \times 100 \qquad \textbf{Equation 1a.}$$

SL = sick leave
CD = calendar days
N = number

When sick leave percentage for working days is calculated, number of persons in employment in denominator of the equation should be multiplied with average number of working days, being 312 or 260 for Slovenia (Equation 1b).

$$SL\%_{WD} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment} \times N_{working\ days\ (average)}} \times 100$$ **Equation 1b.**

SL = sick leave
WD = working days
N = number

When sick leave percentage for effective hours is calculated, number of persons in employment in denominator of the equation should be multiplied with number of all effective hours, available for one person in employment (Equation 1c).

$$SL\%_{EH} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment} \times N_{all\ effective hours}} \times 100$$ **Equation 1c.**

SL = sick leave
EH = effective hours (available for one person in employment)
N = number

Frequency index

Frequency index (FI) is number of persons in employment, absent from work because of illness (number of cases) on 100 persons in employment. By frequency index, relative incidence of absence from work because of illness is expressed (7). This indicator could be calculated using following equation (Equation 2) (4, 7):

$$FI = \frac{N_{cases}}{N_{persons\ in\ employment}} \times 100$$ **Equation 2.**

FI = frequency index
N = number

Severity index

Severity index (SI) is number of lost days on one person in employment, absent from work because of illness (number of cases). By severity index, average duration of one absence from work (one person absent from work because of illness) is expressed (7). This indicator could be calculated using following equation (Equation 3a) (4,7):

$$SI_{CD} = \frac{N_{lost\ calendar\ days}}{N_{cases}}$$ **Equation 3a.**

SI = severity index
CD = calendar days
N = number

Usually, severity index is calculated for calendar days, but it could be expressed in terms of working days as well (Equation 3).

$$SI_{WD} = \frac{N_{lost\ working\ days}}{N_{cases}}$$ **Equation 3b.**

SI = severity index
WD = working days
N = number

Inactivation index

Inactivation index (II) is number of lost days on one person in employment. By inactivation index, average duration of absence from work (average number of lost days) on one person in employment is expressed (7). This indicator could be calculated using following equation (Equation 4a) (4,7):

$$II_{CD} = \frac{N_{lost\ calendar\ days}}{N_{persons\ in\ employment}}$$ **Equation 4a.**

II = inactivation index
CD = calendar days
N = number

Usually, inactivation index is calculated for calendar days, but it could be expressed in terms of working days as well (Equation 4b).

$$II_{WD} = \frac{N_{lost\ working\ days}}{N_{persons\ in\ employment}}$$ **Equation 4b.**

II = inactivation index
WD = working days
N = number

CASE STUDY: OCCUPATIONAL HEALTH INDICATORS ON SICK LEAVE FOR SLOVENIA

Data sources for calculating basic occupational health indicators on sick leave

Data on sick leave

In Slovenia, data on sick leave are collected on the national level for all employees in all business sectors routinely.

Routine data for the data-base are being recorded according to the Healthcare Databases Act (8). The data-base is entitled »Temporary or permanent sickness absence due to illnesses, injuries, nursing, escort or other causes« (data base code in the Healthcare Databases Act: IVZ 3.) (8). They are recorded by medical doctors, authorized for recording of these data. Authorisation is conceded by the Health Insurance Institute of Slovenia (HIIS). The course of the data is as follows:

- the records, recorded by authorised practitioners are transmitted forward to nearest Regional Public Health Institute (there are 9 regional public health institutes in Slovenia), not later than 8th day every month;
- data, aggregated for insurance and social medicine analyses at this level, are transmitted forward to the Institute of Public Health of the Republic of Slovenia (IPHRS), not later than 15th day every month;
- data, aggregated at this level, are transmitted forward to the HIIS, not later than 20th day every month;
- annual data, aggregated at the regional level are transmitted from regional health institutes forward to the IPHRS not later than February 15 every year.

Number of persons in employment

In Slovenia, for calculation indicators on sick leave, number of employees according to the HIIS data-base is used. These data are available not later than June 30 every year.

Very similar data are available in the Statistical Yearbook, issued by the Statistical Office of the republic of Slovenia (9). These data will be used also in our case study[12].

Reports

All four indicators are routinely calculated by IPHRS. Annual and current three-month reports issued by the IPHRS are published at web-page of this institution (4).

[12] Data on all persons in employment except farmers are obtained from the Statistical Register of Employment (SRDAP). Data on farmers are obtained with the Labour Force Survey (LFS). SRDAP covers persons in paid employment who have employment contracts and self-employed persons who have compulsory social insurance (pension, disability and health insurance, parental protection insurance and unemployment insurance). Persons working under copyright contracts, contracts for work/service and citizens of the Republic of Slovenia working in Slovenian enterprises, on construction sites, etc., abroad are not covered.

Basic occupational health indicators on sick leave calculation for 2006

Data, needed for calculation of indicators

Data on elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index are presented in Table 1[13].

Table 1. Elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index for Slovenia in total for the year 2006.

Element	Value
Number of cases*	695,288
Number of lost calendar days*	13,026,763
Number of persons in employment#	824,839

* Data source: National Public Health Institute of Republic of Slovenia (4)

Data source: Statistical Yearbook 2007, Statistical Office of the republic of Slovenia (Table 30.17) (9)

Sick leave percentage

For calculation of SL%, IPHRS is using Equation 1a. In Equation 5, the procedure of calculation of indicator is presented:

$$SL\%_{CD} = \frac{13{,}026{,}763}{824{,}839 \times 365} \times 100 = 4.33 \qquad \textbf{Equation 5.}$$

Frequency index

For calculation of FI, IPHRS is using Equation 2. In Equation 6, the procedure of calculation of indicator is presented:

$$FI = \frac{695{,}288}{824{,}839} \times 100 = 84.29 \qquad \textbf{Equation 6.}$$

Severity index

For calculation of SI, IPHRS is using Equation 3a. In Equation 7, the procedure of calculation of indicator is presented:

[13] Owing to slight difference in number of persons in employment between different data sources, values of indicators on sick leave calculated in case study in this module slightly differ from those published by National Public Health Institute of Republic of Slovenia.

$$SI_{CD} = \frac{13{,}026{,}763}{695{,}288} = 18.74$$ **Equation 7.**

Inactivation index

For calculation of II, IPHRS is using Equation 4a. In Equation 8, the procedure of calculation of indicator is presented:

$$II_{CD} = \frac{13{,}026{,}763}{824{,}839} = 15.79$$ **Equation 8.**

EXERCISE

Task 1

Using internet, find sources of data, used in calculation of indicators in Equations 5-8, given in Table 1. Use internet links listed in References.

Task 2

In Table 2, you will find data for calculation of indicators on sick leave for the year 2006 for twelve statistical regions of Slovenia separately. Calculate all four indicators on sick leave for every of twelve statistical regions. Compare results for twelve specific population groups with the results of total population and discuss them with your colleagues.

Table 2. Elements for calculation of sick leave percentage, frequency index, severity index, and inactivation index.

Statistical region		Number of cases*	Number of lost calendar days*	Number of persons in employment#
1	Pomurska	31,983	608,766	42,253
2	Podravska	102,745	1,899,782	120,513
3	Koroška	21,093	446,621	26,389
4	Savinjska	78,538	1,778,667	104,145
5	Zasavska	7,811	282,156	13,486
6	Spodnjeposavska	21,443	363,319	23,539
7	Jugovzhodna	46,618	894,618	54,044
8	Osrednjeslovenska	196,301	3,836,227	260,344
9	Gorenjska	59,270	1,078,349	71,459
10	Notranjsko-kraška	20,110	307,717	17,557
11	Goriška	59,896	675,091	47,136
12	Obalno-kraška	49,361	852,776	43,975

* Data source: National Public Health Institute of Republic of Slovenia (4)

Data source: Statistical Yearbook 2007, Statistical Office of the republic of Slovenia (Table 30.17) (9)

Task 3

Find appropriate data for calculation of indicators on sick leave for years 2004 and 2005. Compare results for three successive years and discuss them with your colleagues.

REFERENCES

1. World Health Organization. Global Strategy on Occupational Health for All. The Way to Health at Work. Geneva: World Health Organization; 1995. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. Accessed: April 23, 2009.
2. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
3. Council of State and Territorial Epidemiologists. Putting Data to Work: Occupational Health Indicators from Thirteen Pilot States for 2000. Atlanta GA: Council of State and Territorial Epidemiologists, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention; 2005. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. www.cste.org/pdffiles/newpdffiles/cste_ohi.pdf. Accessed: April 23, 2009.
4. National Public Health Institute of Republic of Slovenia. Sick leave (in Slovene). Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. http://www.ivz.si/index.php?akcija=novica&n=122. Accessed: April 23, 2009.
5. Merriam-Webster Online Dictionary. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. http://www.merriam-webster.com/dictionary/sick%20leave. Accessed: April 23, 2009.
6. Kreis J, Bödeker W. Indicators for work-related health monitoring in Europe. Essen: BKK Bundesverband; 2004. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. www.enwhp.org/fileadmin/downloads/memberdocs/580_WORKHEALTH_report_English.pdf. Accessed: April 23, 2009.
7. Bilban M. Statistical indicators on sick leave (in Slovene). In: Teržan M, Pušnik S, Dodič-Fikfak M, Vuga V, Dšuban G, Molan M, Gspan P, Fatur-Videtič A, editors. Occupational medicine for family medicine practitioners. A handbook (in Slovene). Ljubljana: ZZD-Slovensko zdravniško društvo; Združenje za medicine dela, prometa in športa; 2002. p.99-109.
8. Healthcare Databases Act. Ur L RS. 2000;65. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. http://www.uradni-list.si/1/objava.jsp?urlid=200065&stevilka=2969. Accessed: April 23, 2009.
9. Statistical Office of the republic of Slovenia. Labour force. Statistical yearbook of the Republic of Slovenia. Available from URL: http://www.stat.si/eng/pub_letopis_prva.asp. Accessed: April 23, 2009.

RECOMMENDED READINGS

1. Bilban M. Statistical indicators on sick leave (in Slovene). In: Teržan M, Pušnik S, Dodič-Fikfak M, Vuga V, Dšuban G, Molan M, Gspan P, Fatur-Videtič A, editors. Occupational medicine for family medicine practitioners. A handbook (in Slovene). Ljubljana: ZZD-Slovensko zdravniško društvo; Združenje za medicino dela, prometa in športa; 2002. p.99-109.
2. Kreis J, Bödeker W. Indicators for work-related health monitoring in Europe. Essen: BKK Bundesverband; 2004. Available from URL: http://www.who.int/occupational_health/publications/globstrategy/en/index.html. www.enwhp.org/fileadmin/downloads/memberdocs/580_WORKHEALTH_report_English.pdf. Accessed: April 23, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**WORKPLACE RISK ASSESSMENT**
Module: 2.1.7	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Elisaveta Stikova,** MD, PhD**,** Professor National Public Health Institute and Faculty of Medicine, University "Ss. Cyril and Methodius", Skopje, Republic of Macedonia **Neda Milevska-Kostova,** MSc, MCPPM Centre for Regional Policy Research and Cooperation "Studiorum", Skopje, Republic of Macedonia **Petar Bulat**, MD, PhD**,** Professor University of Belgrade, School of Medicine, Occupational Health Department, Belgrade, Serbia **Doncho Donev**, MD, PhD**,** Professor Institute of Social Medicine, Institutes, Faculty of Medicine, University "Ss Cyril and Methodius", Skopje, Republic of Macedonia **Neda Jocić**, MD, PhD, Professor Faculty of medicine, Novi Sad, Serbia
Address for Correspondence	**Elisaveta Stikova** Faculty of Medicine, University "Ss. Cyril and Methodius", Skopje, R. Macedonia, 50 divizija 6, 1000 Skopje, R. Macedonia Tel: +389 70 230 183 e-mail: estikova@ukim.edu.mk or estikova@gmail.com
Keywords	Workplace, risk assessment, risk analysisis, risk management
Learning objectives	After completing this module students and public health professionals should be able to: • understand and define the terms risk, risk analysis - risk assessment, risk communication and risk management; • understand the concept and elements of workplace risk assessment (WRA); • understand the team work for WRA; • perform WRA using elaborated techniques.
Abstract	WRA is a new approach in the traditionally oriented occupational health. Current techniques for recognising and managing occupational safety and health risks were mainly oriented and defined as "workplace analysis" and they lack a risk management and risk quantification component. This module introduces a modern approach to risk quantification that incorporates recent scientific and legal advances in occupational safety and health, consistent with international regulations**.**
Teaching methods	Students are first introduced with the concept of WRA. In continuation they are required to assess the risk at a given workplace, while referring to the recommended readings. The findings are presented and shared with other students in the class.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: LCD projection, flip chart, access to the Internet and bibliographic data-bases; • training materials: recommended readings and matrices/chart tables; • target audience: master degree students according to the Bologna process.
Assessment of students	Multiple choice questionnaire, seminar paper, oral exam.

WORKPLACE RISK ASSESSMENT

Elisaveta Stikova, Neda Milevska-Kostova,
Petar Bulat, Doncho Donev, Neda Jocić

THEORETICAL BACKGROUND

Introduction - workplace risks

According to the estimations of the International Labour Organization (ILO), every year there are approximately 2.2 million deaths, or on average 6,000 die every day as a result of work-related accidents or diseases. Also, every year, there are about 270 million reported injuries at the workplace that lead to absences from work for 3 days or more, and about 160 million incidents of work-related disease, as a result of exposure to hazardous physical, chemical or biological agents, as well as improper working conditions. About 30-40% of them become chronic and irreversible conditions, and 10% lead to complete and permanent loss of working ability (1,2,3).

Hazardous substances kill about 438,000 workers annually, and 10% of all skin cancers are estimated to be attributable to workplace exposure to hazardous substances. There are more than 150 registered physical and chemical agents, classified as occupational cancerogens, and to which about 20-30% of the male and 5-20% of the female professionally active population are exposed (4).

Occupational malignant diseases are cause of about 146,000 deaths or 1.4 million disability adjusted life years (DALYs). Occupational exposure is cause of death for 10% of all cases of malignant diseases of the lungs, bronchi and trachea, and to 24% of all cases of leukaemia. Asbestos alone claims about 100,000 deaths every year and the figure is rising annually. Although global production of asbestos has fallen since the 1970s, increasing numbers of workers are now dying from past exposure to asbestos dust (5).

Another occupational disease, silicosis - a fatal lung disease caused by exposure to silica dust - still affects tens of millions of workers around the world.

There are 35 million health workers in the world. Significant occupational risk for them represents the percutaneous injury with sharp objects with which they work, such as needles, scalpels, peans, etc. The analyses show that every year nearly 3 million injuries at work are registered, where health professionals have come in contact with contaminated blood. In 2 million of the cases, they have been exposed to Hepatitis B, in 0,9 million of the cases with Hepatitis C, and in 170,000 cases they have been exposed to the HIV virus. As a consequence, every year, there are 85,000 cases of professionally contracted infectious diseases in the health professionals' community (6).

The analysis of the influence of the professional exposure in the total burden of disease shows that 5-18% of all cases of asthma and 14% of cases of Chronic Obstructive Pulmonary Disease (COPD) are related to work. Every year, 234,000 deaths are reported due to COPD with occupational ethiology, which represents 0.4% of the total number of deaths from this cause. This is related to 3 million disability adjusted life years (DALYs).

Occupational exposure to noise is a cause for hearing loss of 16% of those that have hearing impairment. In the total burden of disease, the professionally acquired deafness accounts for 0.3% or 4.2 millions of DALYs (7).

Although these figures represent basis for serious concern, the problem is additionally aggravated, if we take into account the large number of occupational diseases that remain undiagnosed or unaccounted for.

The direct costs for diagnosis, treatment and rehabilitation of the work-related injuries and occupational diseases represent a significant burden to the healthcare system. The indirect costs related to the absence from work, loss of working ability or decreased productivity represents a further burden to the national budgets of each country.

Occupational safety and health is one of the European Union's most important social policy areas: every three and a half minutes, somebody in the EU dies from work-related causes; this means almost 167,000 deaths a year as a result of either work-related accidents (7500) or occupational diseases (159,500).

According to European Statistics on Accidents at Work (ESAW), about 4 million accidents at work resulting in more than three days off work occurred in EU-15 in 2004. If accidents causing no absence from work or an absence of up to 3 days are added, the estimated total number rises to more than 7 million every year. It means that every four-and-a-half seconds, a worker in the EU is involved in an accident that causes three or more days absence from work (8,9).

According to the data from the European Agency for Safety and Health at Work (EU-OSHA), occupational diseases and work-related injuries that are consequence of improperly organized system for protection at work cost each country between 2.6-3.8% of GDP.

In the EU about 1 250 million working days are lost each year due to health problems in general. Within the structure by causes (Figure 1) about 210 million days are lost due to accidents at work and 340 million due to work-related diseases (10,11).

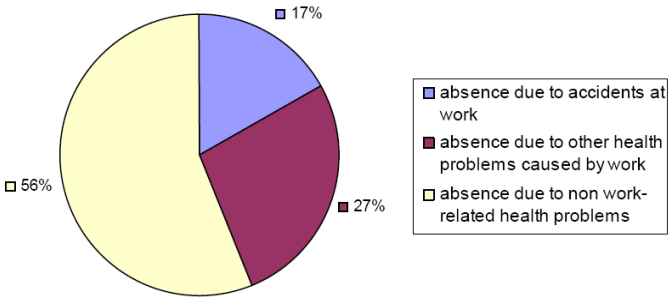


Figure 1. Causes for the sickness absence in EU.

The consequences of accidents at work and of work-related ill health are multiple and complex. The total costs of accidents at work and work-related ill health

are not equally divided between the different players. The total labour costs attributable to accidents at work in EU-15 in the year 2000 were estimated at around 48 billion euros. Other costs were estimated at around 6.6 bilion euros (12).

Thus, the safety and health of workers represents not only a legal requirement and obligation, but also substantial need for improvement of the productivity and economic progress of each country. At the same time, that means a possibility for establishing a balance between the economic interest, productivity and the health and working ability of the working population. The occupational safety and health is economically oriented category, which is aimed at protection of the interests of the workers, companies, health insurance companies, health workers, and finally of the country and European community as a whole (13).

Workers' safety and health is protected in Europe by an approach based on assessing and managing the risk. In order to carry out effective workplace risk assessement, all those involved require a clear understanding of the legal context, the roles of the main actors involved in the process, as well as the concepts and the process of assessing the risk (14,15).

Legal context and the role of the main players

The legal context of the risk assessment is set out in the EU Framework directive and other related so-called sisters' directives (16).

All European Union countries have legislation that set out measures to protect worker safety and health in order to improve quality in safety and security in the workplace. These laws are based on European Directives that lay down minimum obligations for employers and workers, covering the prevention of all types of risks and each activity or sector where risks exist. This legislation is in place to protect workers against occupational accidents and diseases and to assist in the prevention of occupational hazards (17).

Employers have a general duty to ensure the occupational safety and health in every aspect related to work and carry out a risk assessment. The EU framework directive highlights the key role played by risk assessment and sets out basic provision and measures that must be followed by every employer.

Every employer is responsible for taking the appropriate measures such as ensuring the availability of sufficient resources, carrying out a risk assessment, preventing occupational risk, informing, training and consulting workers.

The employer has a duty to:

- organise risk assessment;
- select the person(s) to carry out risk assessment and ensure they are competent;
- assess the risk and implement protective measures;
- consult the employees or their representatives about the organisation of the risk assessment;
- ensure all affected workers are informed of any hazard and any harm to which they may be at risk and the all protective measures taken to prevent such harm.

The employer has a duty and responsibility to make a final decision and to designate the people who carry out risk assessment. They can be:

- the employer;
- employees designated by the employer;
- external assessor and/or services if there is a lack of competent person within the company (16).

Among the most important characteristics of designated or nominated person/company should be competence. They don't need to be occupational safety or health experts, but they should demonstrate their competencies. They should have competencies to understand the general approach to the risk assessment and capabilities to apply this understanding, knowledge and skills to the workplace. But, they should have ability to identify the situations that overwhelmed their competencies in terms to ask for proper help and further assistance. They should demonstrate their competencies by showing they would be able to:

- identify safety and health problems;
- assess and prioritise the need for action;
- suggest options available to eliminate or minimise risk;
- evaluate their effectiveness, and
- promote safety and health good practices.

The last but not the least point in the risk's chain are workers and their representatives. They have a right to be consulted for organisation of the risk assessment, to participate in the risk assessment, to alert the designated person for perceived risk and to be informed about the risk on the work place and their safety and health.

What is risk

Risk represents a qualitative or quantitative estimation of the likelihood of adverse effects that may result from exposure to specified health hazards or from the absence of beneficial influences (18). Risk is a combination of the likelihood of an occurrence of a hazardous event and the severity of injuries or damage to the health of people that might be caused by this event.

Thus, the term professional risk embeds the probability of occurrence of negative effects on the workers' health and life, as a result of presence of certain hazards at work or exposure of the workers to hazardous agents in the working environment (17,19).

Occupational risk is expected frequency of occurrence of occupational diseases and work-related injuries or death, that are consequential to the exposure to different hazardous events or substances, and physical, chemical or biological factors of the working environment. A risk is a chance, higher or low, that somebody may be harmed by the hazard. A hazard can be anything -work materials, equipment, work methods or practices that have potential to cause harm.

Bearing in mind that it is practically impossible to eliminate the probability of occurrence of hazardous agents in the working environment, it is clear that the risk is always present. From expert point of view, it is inevitable to assess the actual and real level of risk, and furthermore important - to define the acceptable levels of risk (20)

Acceptable risk level

Acceptable risk level means that the likelihood of occurrence of hazardous health effects can justifiably be considered as insignificant, and that the occurring effects are so minimal that cannot be further minimized, neither through increased regulatory mechanisms nor through further investment for risk elimination.

Usually, acceptable risk level is agreed between all participants in the decision-making process - scientists, public health professionals, policy makers, administrative and legislative structures.

Risk analysis

Risk analysis is a process that enables insight into the correlation between the exposure, dose and effects from one side, and the measures for mitigation and elimination of the risk and the ways of informing the stakeholders.

Risk analysis consists of 3 consecutive and interrelated processes, i.e. procedures:

- risk assessment,
- risk management and
- risk communication.

Phases of risk analysis and interrelatedness of its segments are presented in the Figure 2 below.

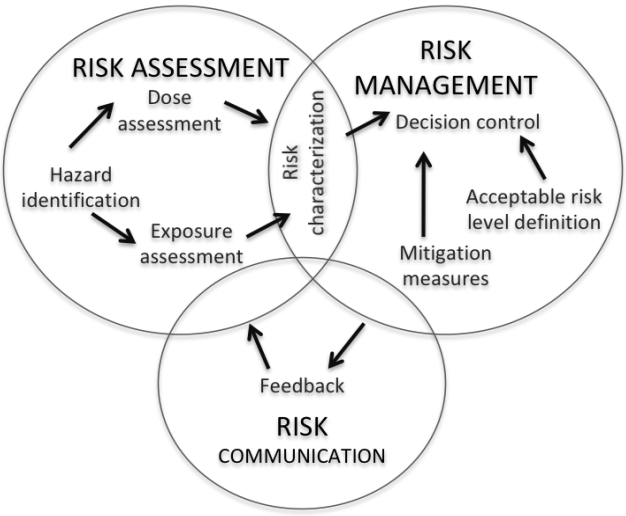


Figure 2. Interrelatedness of the phases of risk analysis

The starting key point in the risk analysis is the risk assessment.

Risk assessment

Risk assessment is quantitive measure of the likelihood of occurrence of hazardous health effects as a result of existing hazards at the workplace (21,22).

Workplace risk assessment is the process of evaluating risk to workers' safety and health from occupational hazards. It is a systematic examination and evaluation of all aspects of work that consider:

- what could cause injury or harm,
- whether the hazards could be eliminated, and if not,
- what preventive or protective measures are, or should be, in place to minimise and control the risk.

Workplace risk assessment means quantification of the likehood of an occurrence of hazardous event at the workplace and related negative effects (injuries at work or work related diseases).

Risk management

Next phase in the risk analysis process is the risk management. The risk management represents assessment of the successfulness of the preventive and mitigation measures (23).

Risk management includes the following activities:

- control of the decision made during the risk assessment,
- determination of the acceptable risk level,
- implementation of mitigation (corrective) measures and
- monitoring of the management.

Risk management is a process in which direct responsibility falls with those that participate in the decision-making process and creating policies for safety and health - from the level in the company where the risk assessment is made to the highest national structures (15). The elements for decision making are taken from the scientific research and expert opinions, standards and good practices established on national or international level (24,25).

Risk communication

Risk communication represents interactive exchange of information and opinions/forecasts related to the confirmed risk. The information is exchanged between those that participate in the risk assessment and risk evaluation, those that manage the risk and those responsible for making the decisions related to the whole process (26).

Information obtained through monitoring of the working environment in the process of health professional risk assessment are combined with the information from other sources, such as epidemiological studies for the impact of a certain professional exposure, referent values for the maximum allowed concentrations in the workplace, etc.

The strategy for risk communication should be determined, established and agreed at the beginning of each risk analysis. That should be a continuous process in which at first, the stakeholders as well as key responsible persons and methods of

communication are identified. At the same time, the most appropriate ways of communication should be defined. The most often used ways of risk communication are:

- public gatherings,
- technical and expert meetings,
- written and published documents,
- written and published results of related research and
- establishment of website.

Risk communication represents a legal obligation for the employers to inform the employees about the identified and confirmed level of risk to which they are exposed during their work as well as about the measures and procedures that they both, employer and employees, have to comply with in order to achieve successful risk management.

The interrelatedness of the processes and activities of risk analysis and risk assessment, as well as risk management are shown on the Figure 2.

What is risk assessment

As already shown, the starting and key point of the risk analysis is the risk assessment.

Risk assessment is an important step in the overall process for protecting the workers and business as well as complying with the low. It should be focused on the level of risk with potential to cause harm (17,27). In addition, risk assessment is quantitive measure of the likelihood of occurrence of hazardous health effects as a result of existing hazards at the workplace.

Risk assessment is based on the quantification of the probability that the hazardous event will occur during the working process and on the assessment of the severity of the hazardous effects that those events will have on the health of the exposed workers (28). Risk assessment represents obligation of the employers, and has the aim to enable them to:

- make easier identification of the dangers and hazards at the workplace,
- determine measures for occupational safety and health,
- check the efficiency of the measures undertaken and
- manifest responsibility in front of the authorities and workers in undertaking necessary preventive measures.

Risk assessment consists of several groups of consecutive activities (Figure 3) that have to be undertaken in the proper order, in magnitude and timeframe that are consistent with the determined objectives (29,30). These activities are:

- hazard identification,
- exposure assessment,
- dose-response assessment, and
- risk characterization.

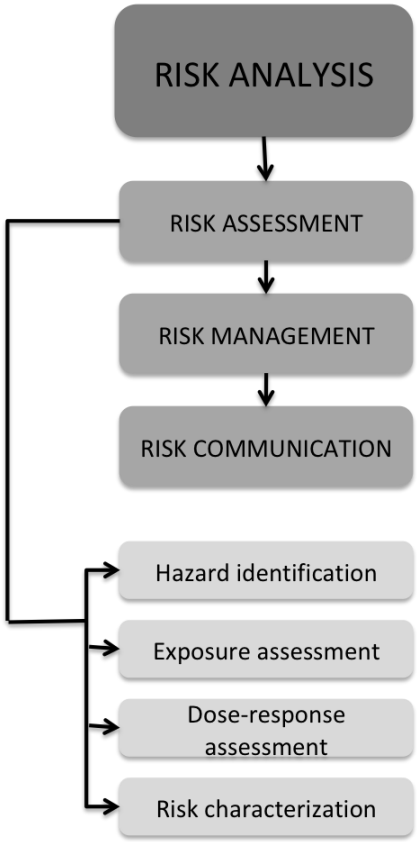


Figure 3. Group of activities that form the risk assessment.

Hazard identification

Hazard identification is a procedure of recognition of hazardous factor in the working environment and defining its physical, chemical and other properties. In the course of hazard identification, the "fate" of hazardous agent is determined, either in the external environment or in the organism, but also its properties to cause negative health effects or disease.

Exposure assessment

The assessment of the exposure means determination of the level of hazardous substances that are present in the working environment and their internalization/intake into the organism of the professionally exposed workers.

According to the intake route of the substance into the human organism, the exposure can be:

- direct and

- indirect.

Dose-response assessment

Dose response assessment is process of making a conection between the data of the exposure assessment and occurence of the work-related diseasses and other harm of the health of the exposed workers.

Risk characterization

The last step in the risk assessment process means characterisation of the actual risk on the examined workplace. It is essential and most visible phase of the quantification of the percived risk. It means establismnet of the interrelation among all information, data and results obtained from the previous steps. For practical use of the whole risk assessment proces in this phase the quantification of the probability of the occurence of the hazard and the severity of the harm caused by respective hazard is made.

Methodological approach to risk assessment

Preconditions for commencement of risk assessment are:

- existence of appropriate legal framework (obligation, responsibility and conditions),
- existence of national standards, guidelines, technical guides, agreed best practices, determined levels of maximum allowed concentrations (levels), manufacturer's guidelines, professional associations' standards, etc. and
- existence of legally and professinally defined hierarchy for risk prevention:
  - avoidance of hazards (modernization and automatization of the technology process),
  - replacement of hazardous with harmless or less hazardous substances (substitution),
  - combating hazards at their source (separation, hermetization),
  - applying collective protection measures,
  - adjustment to the new technological knowledge and their appliance into practice and continuous promotion of the care for safety and health at work.

Workplace risk assessment begins with the decision of the employer for commencement of the procedure and appointment of expert person from the pool of employees that will be responsible for the overall process of risk assessment (31,32). At the same time, depending on the size of the company, the employer appoints representative or representatives of the workers that are participating in the process of risk assessment, or the Committee for safety and health at work might be established.

The flow chart of the risk assessment and control consists of several procedures, as follows:

- preparation of Action Plan and monitoring of its implementation,
- data collection about the working environment and workplace (access, safety of working surfaces, safety of equipment and tools, microclimate conditions, lighting, chemical and biological hazards, etc.),

- identification of working processes and operations with description and/or identification of hazards and dangers,
- data collection on occurence of hazards and/or dangers during the working process,
- data collection on the employees,
- analysis of the recorded effects and/or diseases related to the working conditions,
- quantification of the likelihood of occurence of hazardous effects,
- quantification of the severity of the expected hazardous effects and/or diseases and injuries at workplace,
- defining priorities for prevention, reduction and elimination of workplace-related hazards,
- implementation of the control measures and assessment of their effectiveness,
- review of the risk assessment in pre-determined conditions (change of the technological process, occurence of severe, death or group injuries, etc.).

Starting point in the risk assessment is preparation of the action plan, which preparation, implementation and monitoring is under the responsibility of the employer.

The action plan should include:

- organization and coordination of the risk assessment,
- appointment of competent persons that will work on the risk assessment,
- appointment of representatives of the workers that will be involved in the risk assessment process,
- defining the flow of information and conditions under which the necessary data for risk assessment will be used (confidentiality),
- defining the monitoring of the preventive and/or corrective measures and undertaking responsibility for their implementation by signing written act,
- informing the employees about the identified conditions during the risk assessment and the preventive and corrective measures undertaken.

According to the European directives, and depending on the conditions in one company, the risk assessment can be made by (16):

- the employer,
- employed expert personnel appointed by the employer or
- external professional services.

The defining criterion regarding the selection of the responsible party that will undertake the risk assessment is the size of the company, working conditions and the competence of the appointed person for risk assessment in the company. In practice, very often a need is imposed for engaging a team of different specialties from various fields (outsourcing).

The persons involved in the risk assessment process are confirming their competencies by showing evidence of formal education in the field (if they have such education), as well as by showing competency and capability to:

- understand and be familiar with the general concept of risk assessment;
- have and apply their professional knowledge in particular conditions and workplaces for which the risk assessment is conducted.

Such approach requires that the persons involved in the risk assessment have knowledge and information based on which they will be able to:

- recognize the problems related to the occupational safety and health,
- prioritize accordingly the activities related to the occupational safety and health,
- predict the potential dangers and hazards, as well as consequences that can be expected in case of their occurence, based on relevant evidence,
- identify conditions that surpass their competencies, in order to seek additional expert opinions, and
- know the methodologies for risk assessment, including as much as necessary the complex logic analyses and technical simulations.

Key elements in risk assessment

Risk assessment is in fact reconsideration of all aspects of the work and work process, in order to identify and quantify the hazards that can inflict injury, professional disease or injury at workplace, for the purpose of reducing or removing them, or if possible, to introduce as much as justifiable, preventive and protection measures with which the risk will be eliminated or decreased and controlled.

The risk assessment should embed all predictable dangers and hazards at work in all aspects of work process. It is conducted with participation and consultation of all those involved in the work process - from workers to their employers.

Risk assessment is applicable to all workers and work positions that are in direct or indirect relation and interaction with the hazards present in the working environment, and special attention should be paid to the particularly vulnerable groups of workers - young people, pregnant and lactating women, disabled, etc.

Risk assessment is a set of procedures during which an analysis of the available information is undertaken. During the risk analysis, the appointed person or team does not conduct any measurements, analyses or medical check-ups, but rather makes overview and analysis of the existing situation and available documentation (33,34).

The information necessary for conducting risk assessment can be obtained through (but not limited to):

- analysis of the work opreations and procedures,
- filling stadndardized or adjusted check lists and questionnaires for workers,
- consultation and conversation with workers and their representatives (interviews),
- overview of the technical specifications and guidelines provided by the manufacturers or distributers of the equipment, tools and substances,
- databases of relevent information for safety and health at work and other scientific and technical literature,
- guidelines, directives and good practices given by national referent bodies in the field of standardization, scientific institutes or professional assocaitions,
- data for recorded accidents and injuries at work,
- data for the recorded morbidity (professional and other work-related diseases),
- data from undertaken mesurements of the factors in the working environment and
- national standards, criteria and lists of maximum allowed concentrations and biological exposure tests, etc.

The official and recorded data are used available from the records of the company, and, in the same way, all of the procedures would be recorded as part of the risk assessment.

Methodology for risk assessment can be semi-quantitative or numerically quantitative, and that depends on the national legislation, as well as of the type of certain working process at the workplace.

Regardless of the methodology chosen for undertaking risk assessment, each of the methodologies is based on defining the potential spectrum, i.e. severity of the potentially hazardous effects and the probability for their occurrence.

The spectrum, i.e. the severity of potentially hazardous effects on the health of professionally exposed workers falls into the following categories:

- insignificant consequences,
- accedent without injury,
- light injury,
- severe injury, acute poisoning or chronic professional disease,
- death,
- death of more workers at the same time.

The probability (likelihood) for occurrence of these hazardous effects as a result of professional exposure can be:

- very unlikely (practically improbable);
- likely (moderate probability);
- very likely (high probability).

In combination of these possibilities and with the use of risk assessment matrix, the result is semi-quantitative risk assessment (21). In this case the result can be low, medium or high, or a numerical scale can be used (for example, scale with values from 1-10 with pre-defined criteria about the interpretation of the meaning of each of the values on the scale).

In Figure 4 a sample for semi-quantitative matrix for risk quantification is presented. This matrix is recommended by European Agency for safety and Health at Work (35). In this matrix both, likelihood and severity of damage, have three levels (Figure 4):

1. Likelihood (frequency of occurrence) levels:
    - very unlikely - not expected to occur during the whole working lifetime,
    - likely - it can occur several times during the working lifetime,
    - very likely - it can occur repetitively during the working lifetime.

2. Severity of damage (severity of the expected negative health effects):
    - low severity - injuries and diseases that do not cause prolonged/long-term health problems (e.g. small injuries, eye irritation, headache, etc.),
    - moderate severity - injuries or disease that cause moderate but prolonged/long-term and periodically repetitive health problems (e.g. wounds, simple fractures, burns of second degree to a limited part of the body, skin allergies, etc.),
    - high severity - injuries or diseases that cause severe and permanent health problems or death (e.g. amputations, complex fractures with disability, cancer, burns of 3rd degree covering large part of the body, etc.).

	Severity of damage/level of risk		
Likelihood	**Low**	**Moderate**	**High**
Very unlikely	Low risk (1)	Low risk (1)	Moderate risk (2)
Likely	Low risk (1)	Moderate risk (2)	High risk (3)
Very likely	Moderate risk (2)	High risk (3)	High risk (3)

Figure 4. Likelihood for severity of damage/level of risk in 3x3 matrix (35).

In Figure 5, a more complex quantification of the risk assessment by using the risk matrix 5x5 is presented (36).

			CONSEQUENCES				
			1 Insignificant	**2 Minor**	**3 Moderate**	**4 Major**	**5 Catastrophic**
LIKELIHOOD	**A**	Almost certain to occur in most circumstances	High (H)	High (H)	Extreme (X)	Extreme (X)	Extreme (X)
	B	Likely to occur frequently	Moderate (M)	High (H)	High (H)	Extreme (X)	Extreme (X)
	C	Possible and likely to occur at some time	Low (L)	Moderate (M)	High (H)	Extreme (X)	Extreme (X)
	D	Unlikely to occur but could happen	Low (L)	Low (L)	Moderate (M)	High (H)	Extreme (X)
	E	May occur but only in rare and exceptional circumstances	Low (L)	Low (L)	Moderate (M)	High (H)	High (H)

Figure 5. Likelihood for severity of damage/level of risk in 5x5 matrix (36). LEGEND of consequences: 1-Insignificant: Dealt with by in-house first aid, etc.; 2-Minor: Medical help needed. Treatment by medical professional/hospital outpatient, etc.; 3-Moderate: Significant non-permanent injury. Overnight hospitalisation (inpatient); 4-Major: Extensive permanent injury (e.g. loss of finger/s). Extended hospitalisation; 5-Catastrophic: Death. Permanent disabling injury (e.g. blindness, loss of hand/s, quadriplegia).

How to prioritise the risk rating

Once the level of risk has been determined the following table may be of use in determining when to act to institute the control measures. In Table 1 risk rating and prioritization of measures is presented, while in Table 2 Hierarchy of control in order to minimize occupational exposure is presented.

Table 1. Risk rating and prioritization of measures

Risk rating	Immediate measures	Interim measures
Extreme	Act immediately to mitigate the risk. Either eliminate, substitute or implement engineering control measures.	Remove the hazard at the source. An identified extreme risk does not allow scope for the use of administrative controls, even in the short term.
High	Act immediately to mitigate the risk. Either eliminate, substitute or implement engineering control measures. If these controls are not immediately accessible, set a timeframe for their implementation and establish interim risk reduction strategies for the period of the set timeframe.	An achievable timeframe must be established to ensure that elimination, substitution or engineering controls are implemented. NOTE: Risk (and not cost) must be the primary consideration in determining the timeframe. A timeframe of greater than 6 months would generally not be acceptable for any hazard identified as high risk.
Medium	Take reasonable steps to mitigate the risk. Until elimination, substitution or engineering controls can be implemented, institute administrative . …or personal protective equipment controls. These "lower level" controls must not be considered permanent solutions. The time for which they are established must be based on risk. At the end of the time, if the risk has not been addressed by elimination, substitution or engineering controls a further risk assessment must be undertaken.	Interim measures until permanent solutions can be implemented: • develop administrative controls to limit the use or access. • provide supervision and specific training related to the issue of concern.
Low	Take reasonable steps to mitigate and monitor the risk. Institute permanent controls in the long term. Permanent controls may be administrative in nature if the hazard has low frequency, rare likelihood and insignificant consequence.	

Table 2. Hierarchy of control in order to minimize occupational exposure

Step in a control	Description
Elimination	Eliminate the hazard.
Substitution	Provide an alternative that is capable of performing the same task and is safer to use.
Engineering Controls	Provide or construct a physical barrier or guard.
Administrative Controls	Develop policies, procedures practices and guidelines, in consultation with employees, to mitigate the risk. Provide training, instruction and supervision about the hazard.
Personal Protective Equipment	Personal equipment designed to protect the individual from the hazard.

There are many other methodologies for quantification of the likehood of occurrence of the hazardous event at the workplaces as well as a consequences of those events to the workers' health, like Austrian Social Insurance for Occupational Risks-AUVA's methodology or Nohl and Thiemecke modification of risk approach (37), but it seams that matrix's approach is the most applicable one.

The documentation which is a part of the process of risk assessment should be made available to:

- experts responsible for implementing the measures for occupational safety and health
- persons - representatives of workers, involved in the process of occupational safety and health.

It is especially important to predict the time and conditions under which every employer is obliged to initiate the revision of the once undertaken risk assessment, but also there should be a legal obligation established for revision of risk assessment at any occasion when there are changes in the technological process or collective injury/death at work.

CASE STUDY: PROCESS OF RISK ASSESSMENT

There are a lot of resources in which it is possible to find a case studies and examples for good practices (24,38). The usefulness and applicability of such case study or good practice depend of the sector, size of the company, age, gender and other characteristics of workforce in the company, technological improvement, management and organization of the work etc. Many case studies could be found on the web-site of the EU-OSHA (39), or other agencies.

Process of risk assessment

Theoretical approach to the risk assessment seems to be very complicate and inapplicable. But, it is so simple process if basic tools and essentials are accepted (35). According to the European Agency for Occupational Health and Safety, workplace risk assessment consists of 5 simple steps (Figure 6):

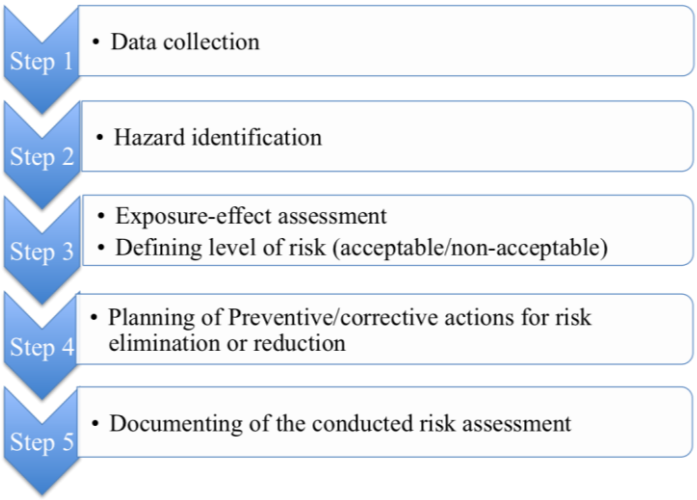


Figure 6. Five steps to the workplace risk assessment (35).

Step 1 - What information should be collected and how

What information is collected

For the workplace risk assessment it is necessary to have the following information at hand:

- location of the workplace,
- characteristics of the equipment, materials and processes,
- description of work operations,
- identified hazards and their sources,
- potential consequences,
- protection measures,
- number of exposed workers,
- number of workers from vulnerable groups (pregnant women, young workers, workers with disabilities),
- occurrence of professional diseases, work-related diseases and injuries at work (33,40).

How information is collected

Possible sources of information are:

- technical data on the equipment, materials and substances that are used in the production process,
- technological procedures and work protocols,
- results of measurements about the hazards at the workplace (workplace analysis),
- specification of the properties of chemical substances (registry),
- legal framework and technical standards,
- scientific and technical literature,
- observing the working environment
- observing the task performed at workplace,
- interviewing the workers,
- data on workplace injuries and professional diseases (registry), (34).

Step 2 - How hazard can be identified

The identification of the hazards

The identification of the hazards in all aspects of work should be approached by:

- walking around the workplace and looking at what could cause harm,
- consulting workers and/or their representatives about any problems they have encountered. Often the quickest and surest way to identify the details of what really happens is to ask the workers involved in the activity being assessed. They will know what process steps they follow, whether there are any short cuts, or ways of getting over a difficult task, and what precautionary actions they take,
- examining systematically all aspects of the work,
- considering long-term hazards to health, such as high levels of noise or exposure to harmful substances, as well as more complex or less obvious risks such as psychosocial or work organisational risk factors,
- looking at company accident and ill-health records.

The hazard identification is usually performed by using hazard specific check lists. These are adjusted to the need and aim of the investigation/assessment itself. EU-OSHA on its website gives one general check list that can be used as model for design of own check lists for every particular company or workplace for which a risk assessment needs to be conducted (39). But, at the same time other additional check lists exist that are related to the most common hazards and dangers that occur at the workplace and for which a risk assessment is required and necessary. The check lists can be of a general type and specific type, where particular and specific professional hazards and very particular technological processes and workplace are assessed. In general, the check lists are aimed at fast and simple orientation in the process of hazard identification, frequency of its occurrence and severity of the damage that can be caused by it.

Usually, the check lists are prepared in cooperation with all workers involved in the working process in question (41). Besides the *general*, there are also *specific check lists* for the hazards and sectors (Table 3).

Table 3. Check lists for workplace hazards and related sectors*.

Check lists - hazards	Check lists - sectors
Uneven slippery surface	Administrative (office) work
Vehicles and moving machines	Food production
Machines with movable parts	Wood processing
Electric installations	Agriculture
Fire	Small and surface mining
Explosion	
Chemical substances	
Noise	
Vibrations	
Lighting	

LEGEND: * - Additional info and specific check lists are available at: www.osha.europa.eu (39)

The identification of all those who might be exposed to the hazards

For each hazard it is important to be clear about who could be harmed; it will help in identifying the best way of managing the risk.

Account should be taken of workers interacting with the hazards whether directly or indirectly, e.g. a worker painting a surface is directly exposed to solvents, while others workers in the vicinity, engaged in other activities, are inadvertently and indirectly exposed. This doesn't mean listing everyone by name, but identifying groups of people such as 'people working in the storeroom' or 'passers-by'. Cleaners, contractors and members of the public may also be at risk.

Particular attention should be paid to groups of workers who may be at increased risk or have particular requirements: workers with disabilities, migrant workers, young and old workers, pregnant women and nursing mothers, untrained or inexperienced staff and temporary and part-time workers.

Step 3 - How to assess the risk araising from identified hazard

The next step is evaluating the risk arising from each hazard. This can be done by considering:

- how likely it is that a hazard will cause harm;
- how serious that harm is likely to be;
- how often (and how many) workers are exposed.

Based on the likelihood of the hazard or danger to occur and on the severity of the expected consequences on the health of professionally exposed workers, a decision on the level of professional risk should be performed. As previously said in the part on methodology, the matrices 3×3 and 5×5 (Figures 4 and 5) can be used for this purposes; or other semi quantitative or numerical matrices (for risk quantification).

The next very important step is to decide whether risk arising from a hazard is acceptable or unacceptable. In general:

- high risk is unacceptable,
- small and medium risk is acceptable,
- if legal requirements are not complied with, risk is not acceptable.

The risk assessment should always be carried out with the employees' active involvement. When deciding on the acceptability of risk, bear in mind their input, and take into account gender, age, and also health of the employees for whom assessment is conducted.

A straightforward process based on judgement and requiring no specialist skills or complicated techniques could be sufficient for many workplace hazards or activities. These include activities with hazards of low concern, or workplaces where risks are well known or readily identified and where a means of control is readily available. This is probably the case for most businesses (mainly small and medium-sized enterprises - SMEs). Risks should then be prioritised and tackled in that order.

Step4 - How the actions to eliminate or reduce the risk arising from hazard can be planed

It is essential that any work to eliminate or prevent risks is prioritised. How should the activities for risk elimination or reduction be prioritised? The basic principles are:

- if the risk is high and assessed as unacceptable, the measures for its reduction should be undertaken immediately,
- if the risk is moderate and assessed as acceptable, it is recommended to plan the activities for reduction of the risk level,
- if the risk is low and assessed as acceptable, it is essential to ensure that it will be reduced or at least remain on the same level.

When preventing and controlling risks, the following general principles of prevention have to be taken into account:

- avoiding risks (elimination of hazard/risk) by modernization and automatization of the technical process,
- substituting the dangerous by the non-dangerous or the less dangerous substances,
- combating or isolating the risks at source (separation, hermetization),
- minimization of the hazard/risk with organizational measures,
- minimization of the hazard/risk with collective protection measures,
- minimization of the risk through appropriate personal protective means,
- applying collective protective measures rather than individual protective measures (e.g. control exposure to fumes by local exhaust ventilation rather than personal respirators),
- adapting to technical progress and changes in information,
- seeking to improve the level of protection.

The next group of activities is to put in place the preventive and protective measures. It is important to involve the workers and their representatives in the process. Effective implementation involves the development of a plan specifying:

- the measures to be implemented;
- who does what and when;
- when it is to be completed.

Step 5 - How to document, monitor and review the process

The risk assessment for each workplace should be documented in the appropriate record for risk assessment. The record should contain the following sections of information:

- basic information: name of the company and address, name of the workplace, number and names of workers on the assessed workplace, date of the assessment and name of the person conducting the assessment,
- name of the hazard if identified.

For every identified hazard should be noticed:

- preventive/protection measures used for alleviation and limitation of the risk,
- results of risk assessment/evaluation,
- planned activities for reduction of the risk.

The risk assessment has to be reviewed regularly depending on the nature of the risks, the degree of change likely in work activity or as a result of the findings of an accident or 'near miss' investigation. Risk assessment is not a once-and-for-all activity.

Risk assessment tools are available at the web-page of the EU-OSHA and they should be used and modified in accordance with specific condition and needs of the company. The translation and modification of the check list and final report is necessary (35).

EXERCISE

Task 1

Individually, carefully read the theoretical part of this module and recommended readings. Try to make distinction between risk analysis and risk assessment and to explain all phases of risk assessment. Discuss risk assessment in the light of EU regulation. Explain how risk assessment should be used for prevention of occupational diseases and injuries at work.

Task 2

Using the information given in case study the students should define the 5 steps for risk assessment and explain:

- which data do they need for workplace risk assessment,
- how these data could be collected,
- how hazards would be identified and how check list can be used,

- how they would identify exposed workers,
- how they would classify the frequency of the occurrence of some hazard and severity of consequences,
- how they would measure the level of risk.

Task 3

Course students should be divided in two groups and using the results from Task 1 and Task 2, they should calculate the level of risk using two different risk matrixes. On the basis of given results they have to discuss:

- how they would use these results for prevention of the injuries at work and occupational diseases
- how they should use these results in the frame of hierarchy of control.

REFERENCES

1. International Labour Organization. Global estimates of fatal work related diseases and occupational accidents. Programme on Safety and health at work and the environment; world bank regions. Geneva: International Labour Organization; 2005. Available from URL: http://www.ilo.org/public/english/bureau/inf/download/sh_background.pdf. Accessed: July 07, 2008.
2. International Labour Organization.Yearbook of Labour Statistics. Volume 59. Geneva: International Labour Organization; 2000.
3. Hamalainen P, Takala J, Saarela K L. Global estimates of occupational accidents. Safety Science. 2006;44:137-56.
4. Leigh J, Macaskill P, Kuosma E, Mandryk J. Global burden of disease and injury due to occupational factors. Epidemiology. 2005;10:626-31.
5. Driscoll T, Nelson DI, Steenland K, Leigh J et al. The global burden of disease due to occupational carcinogens. Am J Industrial Med. 2005;48:463-71.
6. Verschuren R, De Groot B, Nossent S. Working conditions in hospitals in the European Union. Luxembourg: Office for official publications of the European Communities; 1995.
7. Schulte PA.Characterizing the Burden of Occupational Injury and Disease. J Occupat Environ Med. 2005;47:607-22.
8. The European Statistics on accident at work. ESAW 2003 Database. Available from URL: http://www.av.se/dokument/statistik/english/Accidents2003_EU15.pdf. Accessed: September 09, 2008.
9. European Agency for safety and Health at work. Healthy Workplaces. Good for you. Good for business. A European Campaign on risk assessment. Campaign Summary. Luxembourg: European Agency for safety and Health at work; 2008. Available from URL: http://osha.europa.eu/en/campaigns/hw2008/campaign/campaignsummary. Accessed: September 09, 2008.
10. Falkdal AH, Edlund C, Dahlgren L. Experiences within the process of sick leave. Scand J Occup Ther. 2006;13:170-82.
11. Hansen A, Edlund C, Branholm IB. Significant resources needed for return to work after sick leave. Work. 2005;25:231-40.
12. Eurostat. Statistical analysis of socio-economic costs of accidents at work in the European Union. Luxembourg: Office for Official Publications of the European Communities; 2004.

13. Occupational Medicine in Europe: Scope of competencies. EUR/ICP?EHBI 02 02 04. Available from URL: http://www.who.int/occupational_health/publications/eurmedicine/en/index.html. Accessed: October 26, 2006
14. Kolluru RV, Bartell SM, Pitblado RM, Stricoff RS. Risk assessment and management handbook: For environmental, health, and safety professionals. New York: McGraw-Hill; 1996.
15. Eduljee GH. Trends in risk assessment and risk management. Science Tot Environ. 2000;249:13-23.
16. European Council. Council Directive 89/391/EEC of 12 June 1989 on the introduction of the measures to encourage improvements in the safety and health of workers at work. Luxembourg: European Council; 1989.
17. European Legislation. Occupational safety and health. Available from URL: http://osha.europa.eu/legislation. Accessed: October 24, 2006.
18. Last JM, editor. A dictionary of Epidemiology. 4th edition. New York: Oxford University Pres; 2001.
19. Covello VT, Merkhofer MW. Risk assessment methods, approaches to assessing the health and environmental risk. New York: Plenum Publishing Corporation; 1993.
20. Carrivick PJW, Lee A, Kelvin K. Effectiveness of a workplace risk assessment team in reducing the rate, cost, and duration of occupational injury. J Occupat Environ Med. 2002;44:155-9.
21. Hertz-Picciotto I. Epidemiology and quantitative risk assessment: A bridge from science to policy. Am J Pub Health. 1995;85:484-91.
22. Kromhout H. Design of measurement strategies for workplace exposures. Educat Occupat Environ Med. 2002;59: 349-54.
23. Jardine C, Hrudey S, Shorteed J. et all. Risk management frameworks for human health and environmental risks. J Tox Environ Health B Crit Rev. 2003;6:569-720.
24. World Health Organization, Regional Office for Europe. Towards good practice in health, environment and safety management in industrial and other enterprises. Copenhagen: World Health Organization, Regional Office for Europe; 1999. Available from URL: http://www.who.int/occupational_health/publications/goodpractice/en/index.html. Accessed: October 24, 2006.
25. Franco G. Evidence-based decision making in occupational health. Occup Med. 2005;55: 15.
26. Cuny E, Jorgensen V, Galligan J. Comprehensive training for hazard communication. Compend Contin Educ Dent. 2005;26(Suppl 3):28-31.
27. Scholz JT, Gray WB. OSHA enforcement and workplace injuries: A behavioral approach to risk assessment. J Risk Uncertainty. 1990;3:283-305.
28. Concha-Barrientos M, Nelson DI, Driscoll T et all. Comparative Quantification of Health Risks - Selected occupational risk factors. Geneva: World Health Organization; 2004.
29. Stikova E. Risk and risk analysis (in Macedonian). Skopje: Republicki Zavod za Zdravstvena Zastita; 2008.
30. Stikova E. Health Ecology (in Macedonian). Skopje: Univerzitet Sv, Kiril i Metodij; 2006.
31. Nelson DI, Concha-Barrientos M, Driscoll T et al. The global burden of selected occupational disease and injury risks: Methodology and summary. Am J Ind Med. 2005;48:400-18.
32. Maroni M, Faneti AC, Metruccio F. Risk assessment and management of occupational exposure to pesticides in agriculture. Med Lav. 2006;97:430-7.

33. Bugajska J. Occupational data cards as a source of information for health prevention purposes. Med Pr. 2000;51:551-61.
34. International Labour Organization. Recording and notification of occupational accidents and diseases. ILO code of practice. Geneva: International Labour Organization; 1996.
35. European Agency for Safety and Health at Work. Safety and health at work is everyone's concern - checklists [homepage on the Internet]. Available from URL: http://osha.europa.eu/en/campaigns/hwi/ra_tools_checklists/. Accessed: July 07, 2008.
36. National Health Servicde Worcestershire. Future consequence or level of impact resulting from accident or incident. 5X5 scoring matrix [homepage on the Internet]. Available from URL: http://www.worcestershirehealth.nhs.uk/EXTRANET_Library/LINDAMARRIOTT/County%20Risk%20Matrix%202.pdf. Accessed: July 07, 2008.
37. Hohl J, Thiemecke H. Systematik fur Durchfuhrung von Gefahrdungsanalysen, Tail I und II. Schriftenreihe der Bundesanstalt fur Arbeitschutz Fb NR 536; Fb NR 542. Dortmund, 1988. In: Gruber H, Mierdel B, editors. Leitfaden fur die Gefahrdungsbeurteilung. Bochum: Verlag Technik and Information; 1997.
38. European Agency for Safety and Health at Work. Guidelines on the collection, evaluation and disemination of good practice information on the internet. Bilbao: European Agency for Safety and Health at Work; 2002. Available from URL: http://osha.europa.eu/en/good_practice/. Accessed: July 07, 2008.
39. European Agency for Safety and Health at Work [homepage on the Internet]. Available from URL: www.osha.europa.eu. Accessed: July 07, 2008.
40. Money C, Margary SA. Improved use of workplace exposure data in the regulatory risk assessment of chemicals within europe. Ann Occupat Hygiene. 2002;46:279-85.
41. European Agency for Safety and Health at Work. Safety and health at work is everyone's concern. Risk assessment essentials. Available from URL: http://osha.europa.eu/en/topics/riskassessment. Accessed: July 07, 2008.

RECOMMENDED READINGS

1. European Agency for Safety and Health at Work. Safety and health at work is everyone's concern. Risk assessment essentials. Available from URL: http://osha.europa.eu/en/topics/riskassessment. Accessed: July 07, 2008.
2. European Agency for safety and Health at work. Healthy Workplaces. Good for you. Good for business. A European Campaign on risk assessment. Campaign Summary. Luxembourg: European Agency for safety and Health at work; 2008. Available from URL: http://osha.europa.eu/en/campaigns/hw2008/campaign/campaignsummary. Accessed: September 09, 2008.
3. Leigh J, Macaskill P, Kuosma E, Mandryk J. Global burden of disease and injury due to occupational factors. Epidemiology. 2005;10: 626-31.
4. World Health Organization, Regional Office for Europe. Towards good practice in health, environment and safety management in industrial and other enterprises. Copenhagen: World Health Organization, Regional Office for Europe; 1999. Available from URL: http://www.who.int/occupational_health/publications/goodpractice/en/index.html. Accessed: October 24, 2006.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**SURVEILLANCE**
Module: 2.2.1	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Irena Klavs,** MD, PhD**,** Assistant Professor Institute of Public Health of the Republic of Slovenia
Address for correspondence	**Irena Klavs** Institute of Public Health of the Republic of Slovenia Trubarjeva 2, 1000 Ljubljana, Slovenia Irena.Klavs@ivz-rs.si
Keywords	Surveillance, behavioural risk factor surveillance system, population surveillance, sentinel surveillance
Learning objectives	After completing this module students will: • understand the definitions of surveillance (SU); • understand the aims and objectives of SU and uses of SU information; • understand the elements of SU systems and steps involved; • be able to identify the strengths and limitations of SU systems and consider possibilities for improvements.
Abstract	SU is ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and improve health. Monitoring trends is the cornerstone objective of SU systems. Information on characteristics of individuals with health problems permits identification of groups at highest risk of disease. SU information can provide a documentation of the success of an intervention or indicate the need for one. Information from SU systems can contribute to setting of public health priorities, advocacy, health planning and making decisions regarding the allocation of available resources, monitoring and evaluating of public health programmes. Evaluating public health SU systems is to ensure that problems of public health importance are being monitored efficiently and effectively.
Teaching methods	An introductory lecture gives students an overview of SU (the theory and a case study) is followed by an individual work (recommended readings, literature search and preparation for presenting an example of published results from a national SU system including critical analysis of strengths and limitations and suggestions for improvements).
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 35%/65%; • facilities: a lecture room, a computer room; • equipment: LCD projection, computers, access to the Internet and bibliographic data-bases; • training materials: recommended readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire (MCQ); presentation of an example of published results from a national SU system with critical analysis of strengths and limitations and suggestions for improvements.

SURVEILLANCE

Irena Klavs

THERORETICAL BACKGROUND

Definition of surveillance

In 1968, the 21st World Health Assembly described surveillance as the systematic collection and use of epidemiological information for the planning, implementation, and assessment of disease control; in short, surveillance implied "information for action".

According to Last (1), surveillance is systematic ongoing collection, collation, and analysis of data and the timely dissemination of information to those who need to know so that action can bet taken. He distinguishes surveillance from monitoring by the fact that it is continuous and ongoing, whereas monitoring is intermittent or episodic.

At greater length, the Centers for Diseases Control and Prevention (CDC) in the US defines surveillance as the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and improve health (2).

Surveillance methods were originally developed as part of efforts to control infectious diseases. Later, the basic concepts of surveillance have been applied to all areas of public health (3).

Aims and objectives of surveillance and uses of surveillance information

The ultimate aim of a surveillance system is the application of surveillance information for prevention and control of diseases. Surveillance information is meant to be and should be used for public health action.

In 1988, the Institute of Medicine defined three essential functions of public health that emphasized the central role of surveillance:

- assessment of the health of communities, which depends largely on surveillance;
- policy development based on the "community diagnosis" and prognosis established through surveillance; and
- assurance that necessary services are provided, using surveillance as one measure of the impact of programmes (4).

CDC (2) identified the following potential uses for the information resulting from public health surveillance systems:

1. guide immediate action for cases of public health importance;
2. measure the burden of a disease (or other health related event), including changes in related factors, the identification of populations at high risk, and the identification of new or emerging health concerns;
3. monitor trends in the burden of a disease (or other health-related event), including the detection of epidemics (outbreaks) and pandemics;

4. guide the planning, implementation, and evaluation of programmes to prevent and control disease, injury or adverse exposure;
5. evaluate public policy;
6. detect changes in health practices and the effects of these changes;
7. prioritize the allocation of health resources;
8. describe the clinical course of disease; and
9. provide a basis for epidemiologic research.

The list of principal objectives of surveillance provided by Haden and O'Brien (5) indicates similar uses of information generated by surveillance systems:

1. give early warning of changes of incidence;
2. detect outbreaks early;
3. evaluate the effectiveness of interventions;
4. identify at-risk groups; and
5. help set priorities for resource allocation.

Berkelman, Stroup and Buehler (3) list the following purposes of public health surveillance:

1. to define public health priorities;
2. to characterize disease patterns by time, place and person;
3. to detect epidemics;
4. to suggest hypotheses;
5. to identify cases for epidemiological research;
6. to evaluate prevention and control programmes; and
7. to facilitate planning, including projection of future trends and health care needs.

Surveillance systems generally provide descriptive information regarding when and where health problems are occurring and who is affected – the basic epidemiology parameters of time, place, and person.

Monitoring trends

Monitoring trends is the cornerstone objective of most surveillance systems. The detection of an increase in adverse health events can alert public health agencies to the need for further investigation.

For example, after reported incidence rates of early syphilis in Slovenia gradually declined through 1970s and 1980s, in 1994, an 18-fold increase in the annually reported incidence rate of early syphilis (to 1.81 cases per 100,000 population) in comparison to the preceding year was observed (Figure 1). Consequently, the epidemiologist from the Communicable Diseases Centre at the Institute of Public Health of the Republic of Slovenia together with the dermatovenerologist working at the Central Dermatovenerological Clinic in Ljubljana started an investigation. They reviewed the medical records of all notified syphilis cases in 1994. Available information included information on sexual partners, country of probable source of infection, and patients' occupation. The results of this initial investigation indicated that a majority of cases was acquired abroad, in the Russian Federation or in newly independent states (NIS).

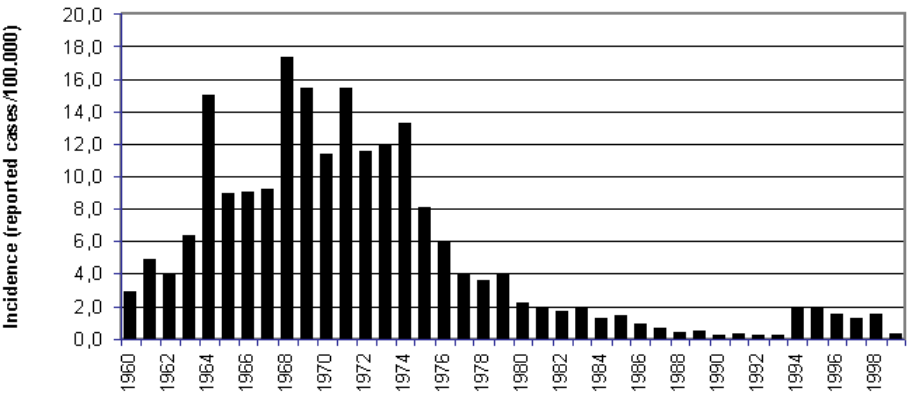


Figure 1. Reported incidence rates of early syphilis, Slovenia, 1960-1999

The annually reported incidence rates remained elevated for a period of five years. In 1999, the surveillance data were analysed for the whole period 1994-1998 (6). More than half of all reported cases (62%) during this period were directly or indirectly linked to a source of infection abroad. Among these the majority of cases (73%) were linked to the Russian Federation or NIS, where a major syphilis epidemic was evolving. Of these, 68% cases occurred in males, with a high proportion of long-distance lorry drivers.

Detection of outbreaks through monitoring trends in reported incidence rates of diseases is often cited as one of potential uses of surveillance information. In practice, however, very often outbreaks are first detected and reported by astute clinicians, before information on case reports is received and analysed in a public health agency.

Identifying populations at high risk

Information on characteristics of individuals with health problems permits identification of groups at highest risk of disease.

For example, in 2007, 36 cases of new human immunodeficiency virus (HIV) infection diagnoses were notified in Slovenia. Among these 36 cases, 34 were men and 29 of these men were known to be men who have sex with men (MSM). This information clearly indicated that MSM were the most affected population group in Slovenia.

Information on specific exposures and behaviours may provide insight into aetiology and modes of transmission. Thus, surveillance information can guide prevention activities even before the aetiology of a disease is known.

For example, in the early 1980s, surveillance of the acquired immunodeficiency syndrome (AIDS provided information on the behavioural characteristics as well as medical histories of AIDS patients. Nearly all had an identified sexual, injecting drug-use or transfusion exposure. This information was

sufficient to calm the public that the disease was not transmitted through ordinary social interaction. The surveillance information, together with initial epidemiologic investigations, defined the modes of HIV transmission even before HIV was discovered.

Evaluating interventions

Evaluation of the effect of public health interventions is complex. Full-scale evaluation, for example through randomised, placebo controlled trials or community randomised trials, may not be feasible. Decision makers and health planners often need to make decisions based on best available information. Trends in the reported incidence rates of a disease identified through surveillance can sometimes provide a convincing documentation of the success of an intervention.

For example, the introduction of childhood vaccination against pertusis in Slovenia in 1959 resulted in a dramatic decline in the reported incidence of pertusis (Figure 2) (7).

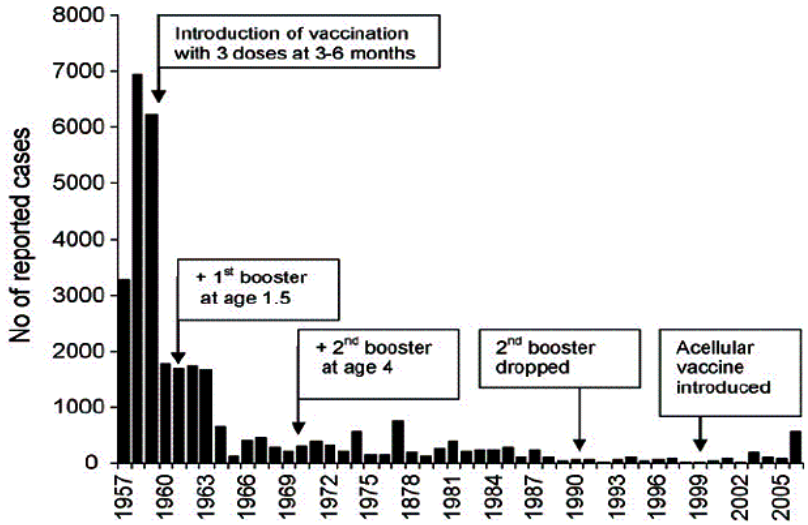


Figure 2. Reported cases of pertusis, Slovenia, 1957-2006.

Conversely, trends in reported incidence can indicate a need to introduce an intervention or modify an intervention. In addition, charting the characteristics of affected individuals, surveillance may provide a comparatively inexpensive and sufficient assessment of the impact of an intervention.

For example, despite the high vaccination coverage against pertusis in Slovenia for many decades, the reported incidence has increased slightly after 2002 and further in 2006 to the level of 27.5 per 100,000 population which represented a 6.5-times increase in comparison to the previous year (Figure 2) (7). In addition,

marked shift in age distribution among reported cases was observed with the highest age-specific incidence among 9 and 10 years old in 2006 (Figure 3). This indicated that a booster dose at school entrance or latest at the age of 8 years should be introduced to decrease the transmission of disease among school children and to further reduce the burden of disease among infants.

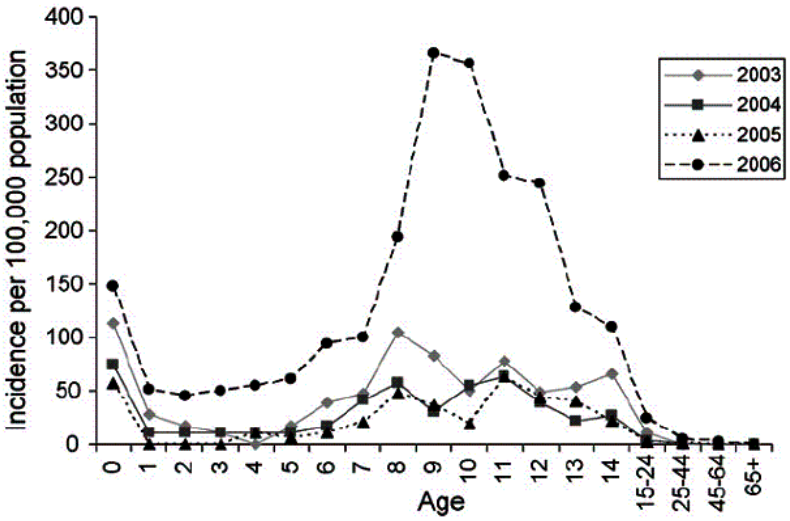


Figure 3. Age specific incidence rates of pertusis, Slovenia, 2003-2006.

Setting public health priorities and help in allocating resources

Information from surveillance systems can contribute to setting of public health priorities, advocacy, health planning and making decisions regarding the allocation of available resources, as well as monitoring and evaluating of public health programmes. In addition to political leaders, the information from surveillance systems educates the public, the media, and health care workers directly responsible for providing health care.

For example, in Slovenia in 2007, a dramatic increase in reported incidence of newly diagnosed HIV infection cases among MSM was documented (Figure 4). The Institute of Public Health of the Republic of Slovenia at that time presented the results to the Ministry of Health and advocated for the increase in the funding allocated to the prevention of HIV infection and other sexually transmitted infections (STI) among MSM. This resulted in increased funding allocated to selected non-governmental organisations of MSM, which were judged to be best capable to reach this population group with preventive interventions tailored to their specific needs.

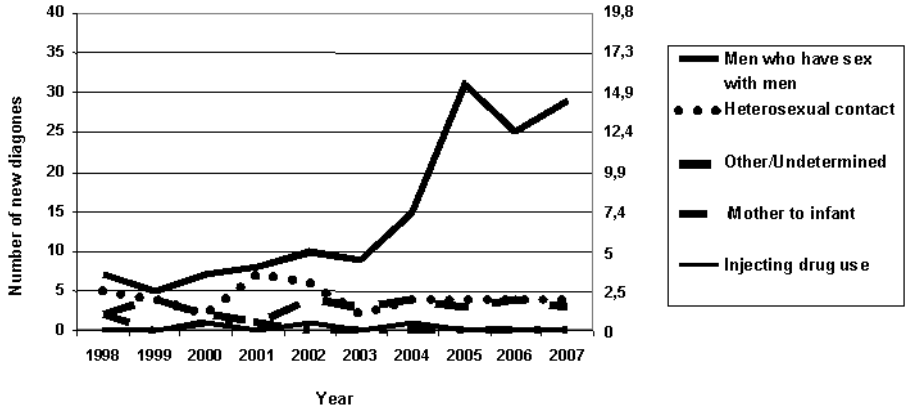


Figure 4. Newly diagnosed HIV cases according to transmission, Slovenia, 1998-2007.

Links to services

At the community level, surveillance is often an integral part of the delivery of preventive and therapeutic services by public health institutions. This role is particularly true for infectious diseases where interventions are based on known modes of disease transmission, where therapeutic or prophylactic interventions are available, and thus the receipt of a case report triggers a specific response.

For example, a report of invasive meningococcal disease caused by *Neisseria meningitidis*, which is spread by respiratory droplets, triggers an epidemiologic investigation with the aim of identifying close contacts at home, in school or elsewhere, who would benefit from post-exposure prophylactic therapy. Thus, at the local level, surveillance information can serve to initiate individual preventive actions.

Elements of surveillance systems, steps involved

Haden and O'Brien (5) identify four steps involved in all surveillance systems:

1. data collection;
2. data collation;
3. data analysis; and
4. data dissemination.

Data collection

There are many potential sources of surveillance data. A few of the major ones used in Slovenia are:

1. mortality data (8);

2. out of hospital care information system (8,9); (includes data on preliminary diagnosis for patients' visits in primary health care and on final diagnoses in the secondary or tertiary level outpatient care);
3. hospital care information system (includes data on hospitalisations and admission diagnoses) (8,10);
4. perinatal information system (11)
5. specific diseases registers (e.g. cancer register (12,13));
6. notifications of communicable diseases (14-19);
7. specific national administrative systems (e.g. information system with data on vaccinations, information system on data about adverse vaccination events) (20, 21).

The data collection methods must be defined precisely and understood well by those who collect the data (e.g. clinicians). This prevents misclassification of cases and ensures that more accurate and complete information is collected.

The case definition is fundamental to any surveillance system. It ensures that the same measure for a health-related event under surveillance is used across geographical area and through time.

For example, in a laboratory surveillance of sexually transmitted *Chlamydia trachomatis* infection, the data providers in participating laboratories should know well the laboratory surveillance case definition (e.g. at least one of the following three criteria should be met:

- isolation of *Chlamydia trachomatis* from a specimen of the ano-genital tract or from the conjunctivae;
- demonstration of *Chlamydia trachomatis* by DFA test in a clinical specimen;
- detection of *Chlamydia trachomatis* nucleic acid in a clinical specimen).

The desire to collect too detailed information must be tempered by the need to limit data to items that can be readily and consistently collected over long periods of time. For example, in a laboratory surveillance of sexually transmitted *Chlamydia trachomatis* infection, the data providers in participating laboratories could report the following data items (variables):

- specimen type;
- personal identifier (to eliminate duplicates);
- date of birth (to assist in eliminating duplicates and to compute age at diagnosis);
- gender;
- municipality;
- date of collection of specimen;
- reporting laboratory;
- type of clinical service.

This would provide good enough information to monitor reported rates of new diagnoses of sexually transmitted infection with *Chlamydia trachomatis* in different age groups of men and women in different municipalities and within different types of clinical services.

In order to identify trends over time, the importance of consistency in the on-going standardised data collection methods (means and data elements) is crucial. To be able to compare surveillance results between different geographical regions or different populations, data collection methods must be standardised.

Data collation

On the national and regional level surveillance data is often collated in computerised databases, which facilitates routine and more sophisticated analyses and production of regular as well as special surveillance reports. On the local level, it is often not necessary that surveillance data are in a computerized database. The data can be collated in a paper-based system or in a spreadsheet.

Surveillance data may be in the form of individual patient records or aggregate counts and tabulations. Individual cases data permit more flexibility of analysis than aggregated data.

Data analysis

Most often, the analysis of surveillance data does not call for sophisticated data manipulation. Simple descriptive analyses according to basic epidemiological parameters of time, place, and person - showing time trends and distribution of characteristics of cases are usually sufficient for surveillance reports.

A case report may include dates, such as those of the onset of disease, diagnosis, and report to local, regional or national public health authority. Analysis can be base on any of these dates, however, if there are long delays between dates of diagnosis and report, analyses of trends based on dates of diagnosis may be unreliable for the most recent period.

Dissemination

Timely dissemination of information to political leaders, the public, the media, and health care workers directly responsible for providing health care is essential. It is also important that the original providers of data to the surveillance system are given regular feedback. In addition, if resources permit, it is a courtesy to accommodate ad-hoc enquiries from data contributors.

Limitations of surveillance systems

Haden and O'Brien (5) identify four potential shortcomings of surveillance systems, which are similar for all information sources:

1. completeness;
2. accuracy;
3. relevance and / or representativeness; and
4. timeliness.

Completeness (sensitivity)

The importance of completeness, or sensitivity of surveillance information depends on the specific objectives of the particular disease surveillance. When surveillance

information is used to detect clusters to trigger an intervention and possible outbreak investigation, surveillance systems are quite tolerant to incompleteness.

For example, a few reported cases of hepatitis A in a primary school can reasonably be expected to be incomplete, as most cases will be asymptomatic. However, such information is good enough to trigger action to prevent the further spread and an outbreak investigation may ascertain the full number of cases.

If all people with the condition under surveillance in the target population are detected by a surveillance system, then its sensitivity is 100%.

For example, the information on reported AIDS cases in Slovenia is assumed to be relatively complete. AIDS is a serious disease, the individual presumably seeks health care and is very likely to be diagnosed correctly and when so, the case is most likely notified. In contrast, HIV infection is asymptomatic for many years and can be diagnosed very late. Thus the information on reported newly diagnosed HIV cases can reasonably be expected not to reflect the true incidence of HIV infection.

For comparing surveillance data over time and geographical areas, it is important that the degree of completeness is consistent between all data providers and through time. Otherwise, trends will be distorted and conclusions about the differences or similarities in the disease burden between different geographical regions will be flawed.

Accuracy

Accuracy is most critical for surveillance of diseases with very low incidence, where misdiagnosis or misclassification can generate a pseudo outbreak and trigger inappropriate action.

Representativeness

Representativeness is a measure of how well reported cases in a population reflect all cases that actually occurred in the population. Surveillance reporting is rarely complete, and cases that are reported may differ from those unreported in terms of demographic, behavioural, and risk exposures characteristics, geographic location or use of health-care services.

Timeliness

Surveillance information should be available in a timely manner. The judgement about what is timely may vary according to the health condition under surveillance and its potential consequences for the health of the community. For example, a case invasive meningococcal disease caused by *Neisseria meningitidis* should be reported immediately, while newly diagnosed cases of breast cancer will not require an immediate public health response and are often reported with longer reporting delays. When very rapid access to surveillance information is important, the system of data collection, the process of data management and the approach to analysis and interpretation should be kept as simple and as easy as possible. Especially when emergency action needs to be taken, for example in a communicable disease outbreak, the real time information based on surveillance data is important. The importance of timeliness may outweigh the need for completeness of surveillance data.

Whatever periodicity is used it should be specified and adhered to by all participants in all phases of the surveillance system loop. Timeliness refers to the entire surveillance cycle, from how quickly cases are reported to the distribution of surveillance reports. With increasing computerization and internet use, reporting at the time of case identification is becoming a reality in some countries.

Evaluating public health surveillance systems

The purpose of evaluating public health surveillance systems is to ensure that problems of public health importance are being monitored efficiently and effectively. In 2001, Centers for Diseases Control and Prevention published the updated guidelines for evaluating public health surveillance systems (2). The following tasks involved are described:

1. engage the stakeholders in the evaluation;
2. describe the surveillance system to be evaluated:
    - describe the public health importance of the health-related event under surveillance,
    - describe the purpose and operation of the surveillance system,
    - describe the resources used to operate the surveillance system;
3. focus the evaluation design;
4. gather credible evidence regarding the performance of the surveillance system:
    - indicate the level of usefulness,
    - describe each system attribute,
        - simplicity,
        - flexibility,
        - data quality,
        - acceptability,
        - sensitivity,
        - predictive value positive,
        - representativeness,
        - timeliness,
        - stability;
5. justify and state conclusions, and make recommendations;
6. ensure use of evaluation findings and share lessons learned.

Conclusion

Effective public health response depends on reliable, continuous flow of information provided by numerous and often complex surveillance systems. Adequate resources needed for surveillance systems and their regular evaluation should be allocated.

CASE STUDY: HIV INFECTION SURVEILLANCE IN SLOVENIA

Second generation HIV surveillance

HIV infection surveillance information has been crucial in understanding the evolvement of the pandemic and for generating evidence based global public health

response. Information about who is infected and who is at risk of infection can help targeting prevention efforts to slow the spread of HIV. Perhaps the most useful information for targeting HIV prevention is behavioural surveillance information. Surveillance information also helps assessing and forecasting treatment and care needs for those affected.

The national HIV surveillance system should be tailored to the pattern of the epidemic. According to the World Health Organization (WHO) and the Joint United Nations Programme on HIV/AIDS (UNAIDS), strengthened HIV surveillance systems, dubbed "second generation surveillance systems" aim to concentrate resources where they will yield information that is most useful in reducing the spread of HIV and providing treatment and care to those affected (22-24).

The goals of the second generation HIV surveillance systems are (22):

- better understanding of trends over time;
- better understanding of behaviours driving the epidemic in a country;
- surveillance more focussed on sub-populations at highest risk of infection;
- flexible surveillance that moves with the needs and state of the epidemic; and
- better use of surveillance data to increase understanding and to plan prevention and care.

According to WHO and UNAIDS, approach to national HIV surveillance that will provide most useful surveillance information should results from different data collection mix which should depend on the HIV epidemic state (classified as either low-level, concentrated or generalised) (22).

Low level HIV epidemic is defined as epidemic, where HIV infection may have existed for many years, but has never spread to significant levels in any sub-population and is largely confined to individuals with higher risk behaviour, e.g. sex workers, injecting drug users (IDU), MSM. Numerical proxy is that HIV prevalence has not consistently exceeded five percent in any defined sub-population. In such a low-level HIV epidemic situation, which is the case for Slovenia, the recommended components of the second generation HIV surveillance systems are (22):

- HIV and AIDS case reporting;
- HIV surveillance in sub-populations at risk;
- surveillance of STI and other biological markers of risk;
- tracking of HIV in donated blood; and
- cross-sectional surveys of behaviour in sub-populations with risk behaviour.

Approach to HIV surveillance in Slovenia

HIV infection surveillance information has been crucial in understanding the evolvement of the epidemic in Slovenia and for generating evidence based public health response.

The different components of the Slovenian national HIV surveillance system are:

1. HIV and AIDS case reporting;
2. monitoring HIV prevalence change in selected sentinel sub-populations at different behavioural risk;

3. collating information about the results of testing of all donated blood units for blood safety purposes;
4. behavioural surveillance in two sentinel sub-populations at highest behavioural risk, MSM and IDU.

We also use the information generated from the STI surveillance system (19).

HIV surveillance information is also complemented by the results of epidemiological studies that are relevant for HIV epidemiology.

HIV and AIDS case reporting

Methods

In Slovenia, notification of HIV infection and AIDS is mandatory according to the Communicable Diseases Law (14) and the Law on Health Care Data Bases (16).

We regularly collect, analyse, interpret and publish information about mandatory reported cases of newly diagnosed HIV infections, AIDS cases and deaths in patients diagnosed with AIDS. We use the European AIDS case definition (25). To be able to better interpret this information we also monitor overall national HIV diagnostic testing trends with annual collection of information about the number of HIV tests performed in laboratories.

Results

During the last ten years (1998-2007), the reported incidence rates of newly diagnosed HIV cases varied from the lowest 6.5 per million population in 1999 and 2000 to the highest 19.0 per million population in 2005 (Figure 5).

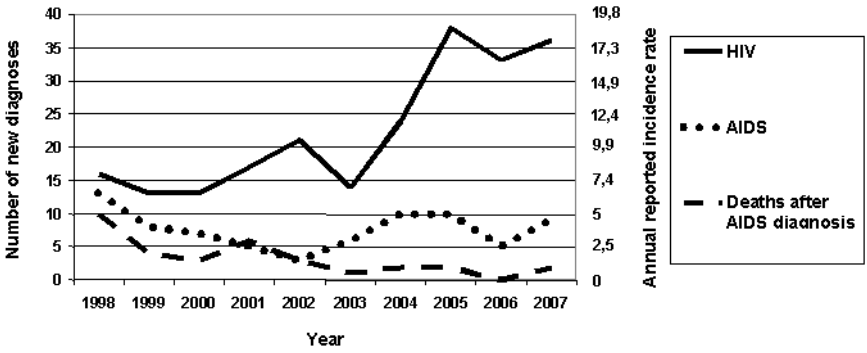


Figure 5. Newly diagnosed cases of HIV, AIDS, and deaths after AIDS diagnosis, Slovenia, 1998-2007.

This is rather low in comparison to most countries in the EU. The dramatic increase in reported HIV incidence after 2004 is due exclusively to the increase in new

diagnoses of HIV infection among MSM (Figure 4). Correspondingly, the great majority of new HIV diagnoses during this period occurred among men, the biggest proportion in 30-39 years old men, indicating that HIV prevention efforts should not only be targeted to very young MSM, but also to older (Figure 6). The last new HIV diagnosis among IDU was reported in 2001 and the last new diagnosis in a child born to mother with HIV infection in 2004.

Annually reported rates of AIDS cases per million population have not risen above 5 per million population since 1999 and deaths among AIDS patients remained below 5 per million population after 2001, which also reflects the effect of universal access to highly active anti-retroviral therapy (Figure 5).

For appropriate interpretation of HIV case reporting surveillance results it is important to understand the extent of HIV testing. In general, more extensive HIV testing results in reduced proportion of undiagnosed HIV infections in the population. Promotion of HIV testing for earlier diagnosis is one of the most important public health interventions. It aims also towards more timely interventions to prevent further transmission among people with newly diagnosed HIV infection.

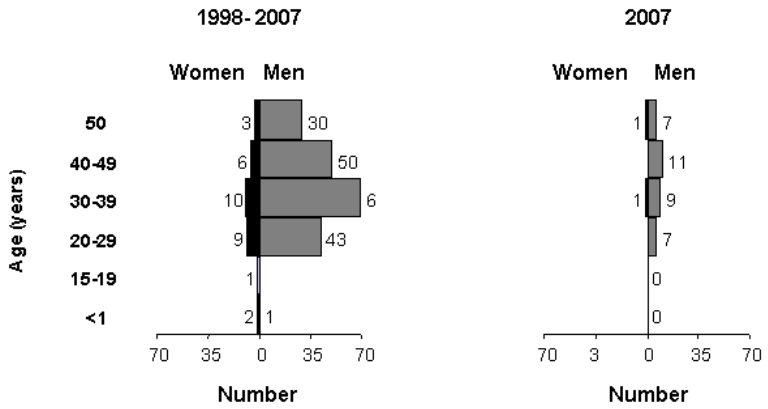


Figure 6. Newly diagnosed cases of HIV by gender and age, Slovenia, 1998-2007.

In Slovenia, overall HIV diagnostic testing rate is relatively low in comparison to most European countries (26), but has been increasing slowly (Figure 7). In 2007, 1.5 diagnostic HIV tests were performed per 100 Slovenians, a 19 percent increase in comparison to preceding year.

If HIV infection is diagnosed late, the opportunity for timely and very effective therapy that reduces the risk of early development of AIDS and AIDS related death has been missed. In Slovenia, a small proportion of HIV new diagnoses remains to be very late (at the same time as AIDS is diagnosed or when CD4 cells count has already fallen below 200/mm^3), however this proportion has been decreasing during recent years (Figure 8).

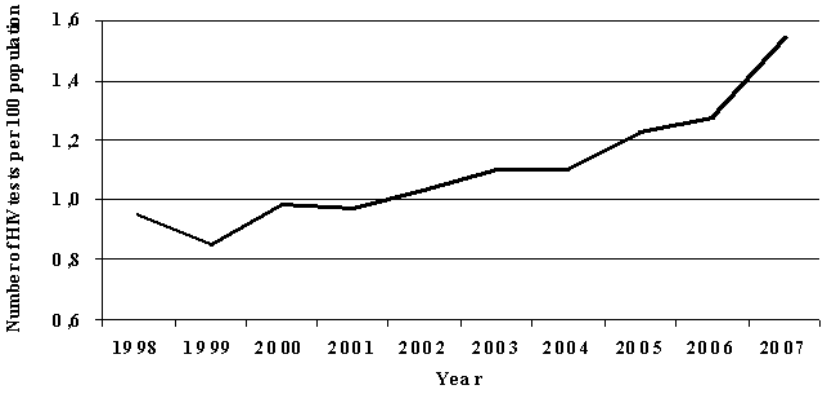


Figure 7. HIV diagnostic testing rates, Slovenia, 1998-2007.

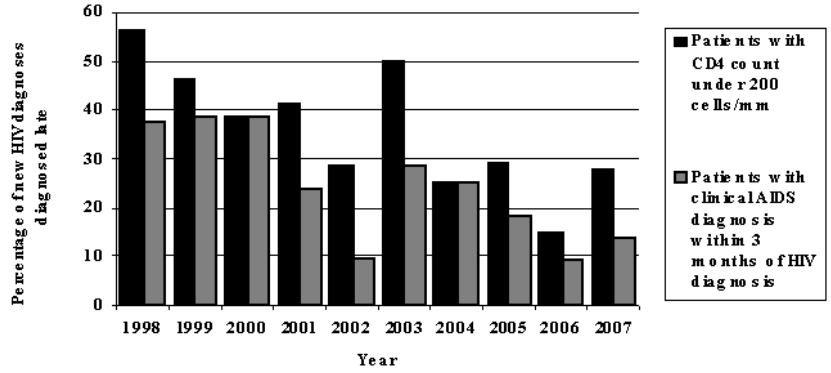


Figure 8. Late diagnoses of HIV infection, Slovenia, 1998-2007.

Monitoring HIV prevalence change in sentinel sub-populations

To complement surveillance information on HIV/AIDS case reporting we monitor HIV prevalence with unlinked anonymous HIV testing in several easily accessible sentinel groups at higher behavioural risk (IDU, MSM clients of STI clinics tested for syphilis) and also in one low-risk population group (pregnant women who are screened for syphilis) (27).

Methods

Residual sera from sera specimens obtained from STI patients and pregnant women for syphilis serology are sampled continuously in several laboratories since 1993. Since 1995, saliva specimens are continuously voluntarily obtained from IDU entering substitution therapy programmes at one or two sites and for the last three years also for two months per year from convenient samples of IDU clients at one needle exchange harm reduction programme in Ljubljana and one in Koper. Voluntary confidential testing for HIV is offered to all IDU included into the sentinel surveillance sample. Since 1996, once per year saliva specimens are voluntarily obtained from a convenient sample of MSM at a community venue in Ljubljana. Safer sex is promoted and information about access to voluntarily confidential and also anonymous HIV testing and counselling is provided to all MSM included into the sentinel surveillance sample. In addition to the information about the type of sentinel population, sampling period and sentinel site, all specimens are labelled only with sex and age group, frozen and stored at -20^0 C. Figure 9 presents the sentinel sites locations for all sentinel populations.

At the end of each sampling period, after each sampling year, all serum specimens are tested in pools of 12 for the presence of anti HIV-1/0/2 antibodies using third generation enzyme immuno assay (EIA). Individual sera from reactive pools are re-tested using the same assay. Saliva specimens are tested individually for the presence o anti HIV-1/2 antibodies, using EIA. All EIA repeatedly reactive individual specimens are supplementary tested using HIV Western Blot or Immuno Blot for anti HIV-1 or anti HIV-2 antibodies.

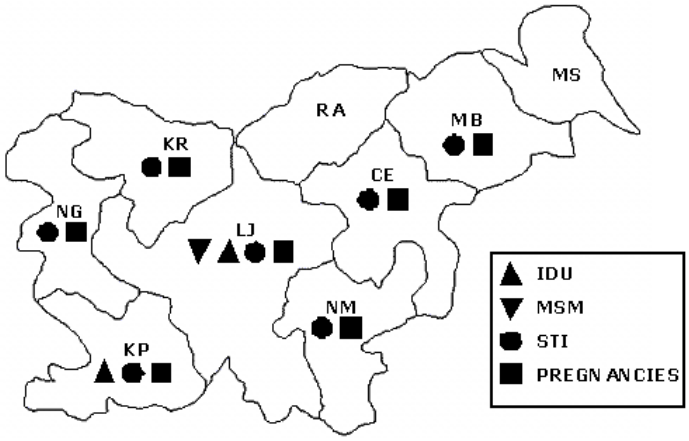


Figure 9. Sentinel sites and sentinel populations – unlinked anonymous HIV prevalence monitoring sentinel surveillance programme, Slovenia, 2003-2007. LEGEND: Health regions: CE-Celje; KP-Koper; KR-Kranj; LJ-Ljubljana; MB-Maribor; MS-Murska Sobota; NG-Nova Gorica; NM-Novo mesto; RA-Ravne; IDU- injecting drug users; MSM-men who have sex with men; STI-sexually transmitted infections.

Results

The results for the period of last five years (2003-2007) are presented in Table 1. MSM are clearly the most affected sub-population. The proportion of HIV infected is also high among male STI patient, the group that contains a disproportionally high proportion of MSM. The rapid spread of HIV infection among IDU in Slovenia has not started yet. The prevalence of HIV infection among pregnant women, a population at a very low behavioural risk, remains very low. The results for the period 1993-2002 have been published previously (27).

Table 1. HIV prevalence among sentinel populations of injecting drug users, men who have sex with men, patients with sexually transmitted infections and pregnancies, Slovenia, 2003-2007.

	Year	Number of sentinel sites	Number of tested		Number of HIV infected		Propor tion of HIV infected		Prevalence range*	
			Men	Women	Men	Women	Men	Women	Men	Women
Injecting drug users	2003	2	253	79	0	0	0%	0%		
	2004	3	173	59	0	0	0%	0%		
	2005	3	137	57	0	0	0%	0%		
	2006	3	125	35	0	0	0%	0%		
	2007	3	130	44	0	0	0%	0%		
Men who have sex with men	2003	1	101		1		0.9%			
	2004	1	79		2		2.5%			
	2005	1	82		3		3.7%			
	2006	1	94		2		2.1%			
	2007	1	124		3		2.4%			
Patients with sexually transmitted infections	2003	7	267	200	1	0	0.4%	0%	0–0,5%	
	2004	7	328	148	5	0	1.5%	0%	0–2,2%	
	2005	7	403	170	1	1	0.2%	0.6%	0–0,6%	0-2.6%
	2006	7	419	211	10	0	2.4%	0%	0–2,9%	
	2007	7	484	257	11	0	2.3%	0%	0–3,3%	
Pregnancies	2003	8		7544		0		0%		
	2005	8		8008		1		0.01%		0-0.5%
	2007	8		8963		0		0%		

* prevalence range between different sentinel sites

This component of the Slovenian national HIV surveillance system is relatively modest in terms of numbers of tested specimens for all higher-risk behaviour sub-populations, but, the results are informative and provide for crude monitoring of trends in HIV prevalence and early warning.

When inferring about the distribution and spread of HIV infection in different population groups in Slovenia we should be cautious, as these easily accessible

sentinel groups are not representative of all IDU, MSM, patients with STI, and women of reproductive age.

Collating information about the results of testing all donated blood units

All donated blood and blood components have been tested for blood and blood products safety reasons since 1986. Blood donors are a sub-population at a very low risk for HIV infection. We regularly collate information about the results of this mandatory testing for HIV. The proportion of HIV infected donated blood units has remained at a level of approximately one positive result per 100 000 donations. This indirectly indicates a very low level of HIV infection in the general population of Slovenia.

Behavioural surveillance in high-risk groups

In Slovenia, we have managed to develop a very basic behavioural surveillance system in two groups at high behavioural risk for HIV infection, MSM and IDU. The methods and some results are presented for the most affected sub-population in Slovenia, MSM.

Methods

In 2000, we have attached the behavioural data collection to the HIV prevalence monitoring with unlinked anonymous testing among MSM described previously (27). These are small scale, annually repeated one-day cross-sectional surveys in a community setting, just one sentinel site – MSM venue in Ljubljana. MSM themselves organize consecutive sampling of MSM attending an event. In addition to saliva specimens collection for HIV unlinked anonymous testing, all participants are invited to anonymously complete a very short self-administered questionnaire. This information is not linked to the saliva specimen of the same individual and the HIV testing result, which cannot be reported to the individual who provided the specimen. Safer sex is promoted and information about access to voluntarily confidential, and also anonymous HIV testing and counselling is provided to all MSM included into the sentinel surveillance sample. We collect information on:

- condom use at anal and oral sex (last occasion and during the preceding year);
- number of anal and oral sex male partners last year;
- number of female partners and condom use with female partners (the bridging to the general population indicator);
- having paid for and having been paid for homosexual sex last year;
- having tested for HIV last year and age in two broad categories (under 25, 25+).

Results

Figure 10 presents the variation in proportion of MSM who have reported different frequencies of condom use at anal sex wit men during the year preceding the survey.

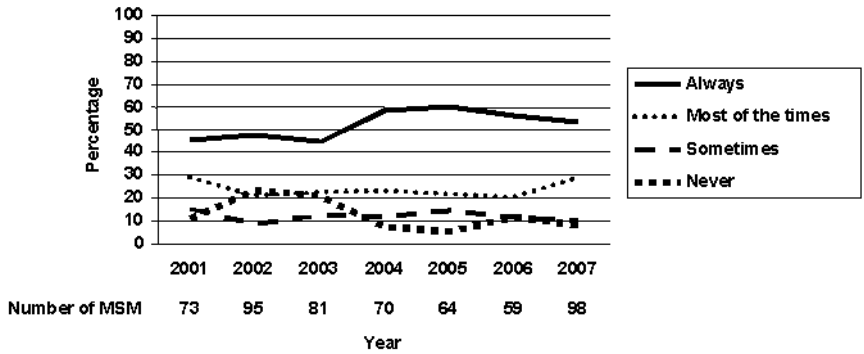


Figure 10. Condom use at anal sex of men with men during last year, Ljubljana, 2001-2007

The proportion of those who reported to have always used condoms has increased from below 50% during 2001-2003 to above 50% during 2004-2007. Conversely, the proportion of those who reported never to have used condoms has decreased during 2004-2007 in comparison to 2002 and 2003. These results suggest that safer sex behaviour among MSM in Ljubljana has not deteriorated substantially during recent years. Regretfully, the convenient samples sizes have been very small. However, this very simple and crude monitoring of behaviour change among MSM provides informative results with very modest resources.

Information complementing surveillance information

Surveillance information is often triangulated with or complemented with information from ad hock epidemiological studies.

To formulate appropriate and effective sexual health and reproductive health policies, including prevention of HIV, it is crucial to understand sexual behaviour of the population. An example of such a study in Slovenia was the first national sexual behaviour survey conducted in a probability sample of the general population.

Two broadly defined objectives were:

- to describe the patterns of sexual behaviour and identify demographic, social, and behavioural determinants of higher risk behaviour patterns, and
- to describe the distribution of sexually transmitted *Chlamydia trachomatis* infection.

Several results have been published (28-31). Information about condom use at first heterosexual intercourse is of particularly great relevance to HIV prevention (32). The information presented was similar to surveillance results, as it reflected changes through time (32).

Methods

Methods details were published previously (33). In brief, we included Slovenian citizens 18-49 years old. We used stratified two-stage sampling. The sampling frame

was designed using the list of enumeration areas and information from the Central Population registry.

The data was collected in 1999-2001 by means of face-to-face interviews in combination with anonymous self administration of more sensitive questions in the presence of interviewers (pencil and paper). We adapted the data collection methods used in the first National Sexual Attitudes and Lifestyles survey conducted in 1990 and 1991 in Britain (34). At the end of interview each respondent was invited to provide first void urine specimen to be confidentially tested for sexually transmitted *Chlamydia trachomatis* infection.

Analyses were conducted using the methods for complex survey data in statistical package STATA. All estimates were weighted.

Results

Figure 11 presents one result which is especially relevant for HIV prevention. Separately for men and women, the variation in the proportion of respondents who reported condom use at first heterosexual intercourse according to the time period of the event is shown. Among those who experienced first heterosexual intercourse during 1970-1974, only 2% of men and 6% of women reported to have used a condom. In contrast, among those who experienced first heterosexual intercourse during 1995-1999, already 72% of men and 64% of women reported condom use. The increase in condom use was most marked during the late 1980s and 1990s, the period during which the effect of "exposure" to AIDS awareness and condom use promotion was increasing.

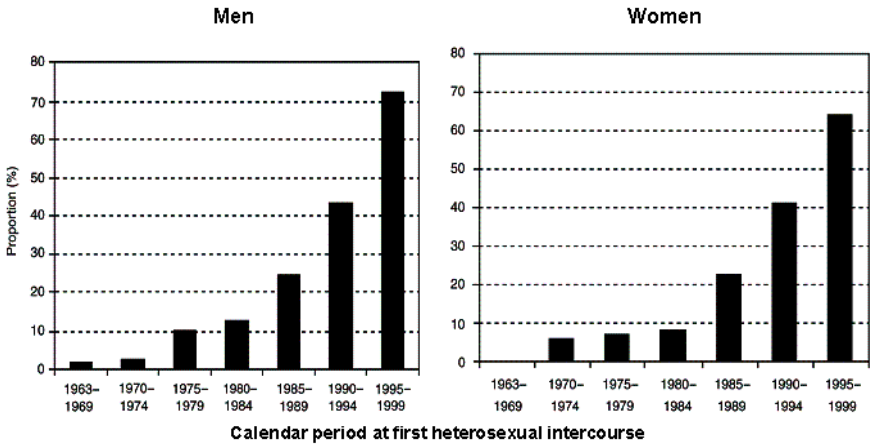


Figure 11. Proportion of men and women reporting condom use at first heterosexual intercourse according to calendar period the sexual debut occurred, Slovenia.

In Slovenia, we have not missed the window of opportunity for safer sex promotion including condom use promotion targeted to the general population and youth.

Strengths and limitations of the Slovenian HIV surveillance system

The Slovenian HIV surveillance system generates information that provides for fairly good understanding of the national burden of HIV, the identification of sub-populations most affected and most at risk as well as monitoring trends in the evolving HIV epidemic and underlying risk behaviours. Based on HIV surveillance information priorities for public health interventions have been identified and resources allocated correspondingly. Surveillance information has provided insight into the impact of public health interventions.

The HIV surveillance system is fairly simple, provides information of fair quality, and seems to be acceptable to participating persons and organizations. The data collection methods have been relatively stable over time, providing for consistency in monitoring trends. HIV surveillance reports are published annually, which is timely enough (19). In addition short quarterly reports are also published to provide for early warning in case of drastic increases in the numbers of newly diagnosed cases of HIV. The resources used to coordinate the surveillance system used at the Institute of Public Health of the Republic of Slovenian have been very modest.

The information on reported AIDS cases is relatively complete. AIDS is a serious disease, the individual presumably seeks health care and is very likely to be diagnosed correctly and when so, the case is most likely reported. In contrast, HIV infection is asymptomatic for many years and can often be diagnosed very late. Thus the information on newly diagnosed HIV case most often can not reflect the true incidence of HIV infection.

The information obtained from monitoring HIV prevalence with unlinked anonymous testing in sentinel groups at different behavioural risk (the sentinel surveillance component) has been informative and provided for crude monitoring of trends and early warning. The strengths of such sentinel surveillance include: feasibility, consistency, minimal participation bias, anonymity, and no need for additional invasive procedures to obtain biologic specimens. The limitations include: non-representativeness for all IDU, MSM, patients with STI and pregnant women (low geographical coverage and selection bias), too small sample sizes among MSM and IDU to be able to reliable monitor smaller changes in HIV prevalence, and non-availability of additional risk information (e.g. it is not known whether a MSM or a patient with and STI also had a history of IDU).

The information obtained from the very small scale and simple behavioural surveillance among MSM provides informative results with very modest resources. Since preventing the spread of HIV infection depends mostly on preventing high-risk behaviour and supporting behavioural change, monitoring behaviour is a necessary component of any HIV surveillance system and provides the information for evidence based targeting of prevention interventions and monitoring their impact. The validity of self-reported information can always be questioned, but, if it does not change with time, such approach is good for monitoring trends. The strengths of such sentinel behaviour surveillance include: feasibility, consistency, anonymity that may contribute to the validity of self-reported information. The limitations include: non-representativeness for all MSM (low geographical coverage and selection bias), participation bias, too small sample sizes to be

able to reliable monitor smaller changes in behavioural patterns, and limited amount of behavioural information collected.

Main challenges to improve the Slovenian HIV surveillance system

The existing HIV surveillance system fails to answer many relevant questions. Some examples of such questions are:

- How to prevent the spread of HIV among MSM effectively?
- Are there smaller, but important changes in the high-risk behaviour patterns among MSM?
- What is the proportion of undiagnosed HIV infection in the whole population and in the identified groups at higher behavioural risk?
- What is the proportion of diagnosed HIV infections not reported in the whole population and in the identified groups at higher behavioural risk?
- What is the uptake of HIV testing in different population groups at higher behavioural risk?
- What is the level of HIV testing among patients with diseases indicating high-risk behaviours or diseases indicating HIV infection?
- Do smaller changes in HIV prevalence occur that can not be detected by the behavioural surveillance because of relatively small numbers of participants from high-risk groups?
- Do changes in HIV prevalence occur among MSM and IDU in parts of the country that are not covered with the sentinel HIV surveillance system?
- How many individuals with HIV infection die from AIDS and how many from other causes of death and from which other causes?
- How many individuals with HIV infection have access to highly-active antiretroviral therapy?
- How many individuals with HIV infection are infected with HIV strains resistant to certain anti-retroviral drugs?

In addition to sustaining the existing Slovenian second generation HIV surveillance system challenges include:

- information collected through HIV and AIDS case reporting should include information on HIV therapy and resistance to anti-retroviral drugs;
- information collected through reports of deaths among HIV infected should include more detailed information on causes of death;
- the timeliness and completeness of HIV case reporting should be improved by the introduction of mandatory reporting of diagnoses of HIV infection from laboratories;
- laboratory based HIV surveillance should provide information on the uptake of HIV testing in different population groups at higher behavioural risk and on the level of HIV testing among patients with diseases indicating high-risk behaviour or indicating HIV infection. HIV prevalence monitoring in sentinel high-risk behavioural groups should be improved by increasing the geographical coverage and increasing sample sizes to provide for the detection of smaller changes in HIV prevalence;

- behavioural surveillance among MSM should be improved by regularly repeating large scale in-depth behavioural surveys with integrated biological markers (e.g. other STI).

Finally, formal evaluation of the Slovenian HIV surveillance system should be conducted using the guidelines published by the Centers for Disease Control and Prevention (2). The evaluation findings should be used to improve the efficiency and effectiveness of the surveillance system.

Conclusions and recommendations for prevention, treatment and care

Slovenia is a low HIV epidemic country with les than one individual living with HIV infection per 1000 population. MSM are the most affected sub-population. Rapid spread of HIV infection has not started yet among IDU and their sexual partners. We have not missed the window of opportunity for safer sex promotion including condom use promotion targeted to the general population and youth. We have also managed to target HIV prevention at groups at highest behavioural risk, particularly MSM.

Prevention and control of HIV infection within the broader frame of promoting sexual and reproductive health remains an important public health priority.

As MSM are the most affected population in Slovenia, promotion of responsible and safer sex including condom use is especially important in this group and should implemented by MSM non-governmental organizations. Sufficient resources for good coverage of all MSM with preventive interventions should be available and the interventions should be monitored for quality and impact.

In addition prevention and harm reduction interventions should be targeted to other population groups at higher behavioural risk such as IDU, sex workers and their client, prisoners, travellers to countries with generalised epidemics and immigrants from these countries to Slovenia.

Voluntary confidential testing for HIV should be promoted in population groups at higher behavioural risk, especially MSM.

Health-care provider initiated testing should be routinely offered to all patients with conditions indicating high risk behaviour (e.g. STI, hepatitis C) or indicating HIV infection (e.g. tuberculosis, one of the AIDS indicator diseases) (35).

Health care case management of all STI should include counselling for safer sex, notification and treatment of contacts and routine offer for voluntary confidential testing for HIV.

To contain the mortality and morbidity among people living with HIV on a low level, universal access to high quality treatment including highly-active anti-retroviral therapy should sustained.

EXERCISE

Task 1

Students read the theoretical background of the module, the case study, and recommended readings.

Task 2

Students search for other published surveillance results from their own country (e.g. in Slovenia: Biomedicina Slovenica; COBISS-Cooperative Online Bibliographic System of Slovenia; Institute of Public Health of the Republic of Slovenia) and from other European countries (e.g. on web-sites of other national public health institutes such as Health Protection Agency in the UK; Eurosurveillance website; and from international bibliographic data-bases such as Medline and PubMed) and reads a few examples.

Task 3

Each student presents an example of a national surveillance system and the whole group discusses the objectives of the presented surveillance systems, methods used, results and uses of surveillance information, characteristics of the surveillance system, its strengths and limitations as well as proposals for improvements.

REFERENCES

1. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
2. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50 (No. RR-13):1-36. Available from URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm. Accessed: July 19, 2008.
3. Berkelman RL, Stroup DF, Buehler JW. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press: 2002. p.759-78.
4. Institute of Medicine. The future of public health. Washington: National Academy Press; 1988.
5. Hadden F, O'Brien S. Assessing acute health trends: surveillance. In: Pencheon D, Guest C, Meltzer D, Gray JAM, editors. Oxford handbook of public health practice. Oxford: Oxford University Press; 2005. p.14-9.
6. Grgič-Vitek M, Klavs I, Potočnik M, Rogl-Butina M. Syphilis epidemic in Slovenia influenced by syphilis epidemic in the Russian Federation and other newly independent states. Int J AIDS & STD. 2002;13(Suppl 2):2-4.
7. Grgič-Vitek M, Klavs I, Kraigher A. Re-emergence of pertusis in Slovenia: Time to change immunization policy. Vaccine. 2008;26;1874-8.
8. Institute of Public Health of the Republic of Slovenia. Health statistics annual report 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2007. Available from URL: http://www.ivz.si/index.php?akcija=novica&n=1420. Accessed: July 19, 2008.
9. Institute of Public Health of the Republic of Slovenia. Out-of hospital care statistics (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2008. Available from URL: http://www.ivz.si/index.php?akcija=novica&n=964 (Accessed: 19 July 2008)
10. Institute of Public Health of the Republic of Slovenia. Health-care information system on hospitalizations (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2008. Available from URL: http://www.ivz.si/index.php?akcija=novica&n=668. Accessed: July 19, 2008.

11. Institute of Public Health of the Republic of Slovenia. Perinatal information system of the Republic of Slovenia (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2008. Available from URL: http://www.ivz.si/index.php?akcija=novica&n=929. Accessed: July 19, 2008.
12. Institute of Oncology Ljubljana. Cancer Registry of Slovenia (in Slovene) [homepage on the Internet]. Available from URL: http://www.onko-i.si/sl/dejavnost/epidemiologija_in_register_raka/. Accessed: July 19, 2008.
13. Institute of Oncology Ljubljana. Cancer Incidence in Slovenia 2005. Ljubljana: Institute of Oncology Ljubljana; 2008. Available from URL: http://www.onko-i.si/uploads/LP_2005.pdf. Accessed: July 19, 2008.
14. Slovenian Parliament. Act in communicable diseases (in Slovene). Offic J RS. 1995;69.
15. Ministry of Health. Regulation on notification of communicable diseases and special activities for prevention and control (in Slovene). Offic J RS. 1999;16.
16. Slovenian Parliament. Act on health care data bases (in Slovene). Offic J RS. 2000;65.
17. Institute of Public Health of the Republic of Slovenia. Surveillance of communicable diseases in Slovenia, 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2007. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/798-Epidemiolosko_spremljanje_NB_2006.pdf. Accessed: July 19, 2008.
18. Klavs I, Bergant N, Kustec T, Kastelic Z. HIV infection in Slovenia, 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2008. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/640-Leto_2007.pdf. Accessed: July 19, 2008.
19. Klavs I, Kustec T, Bergant N, Kastelic Z. Sexually transmitted infections in Slovenia, 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia, 2008. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/930-Leto_2007.pdf. Accessed: July 19, 2008.
20. Kraigher A, Grgič-Vitek M, Učakar V, Sevljak-Jurjevec M. Analysis of the implementation of the immunisation programme in Slovenia in 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia, 2007. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/1409-Analiza_izvajanja_imunizacijskega_programa_v_Sloveniji_2006.pdf. Accessed: July 19, 2008.
21. Kraigher A, Grgič-Vitek M, Učakar V, Simonovič Z, Sevljak-Jurjevec M. Adverse vaccination events in Slovenia in 2005 and 2006 (in Slovene). Ljubljana: Institute of Public Health of the Republic of Slovenia; 2007. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/1411-Nezeleni_ucinki_2005_in_2006.pdf. Accessed: July 19, 2008.
22. World Health Organization and Joint United Nations Programme on HIV/AIDS (UNAIDS). Guidelines for the second generation HIV surveillance. Geneva: World Health Organization, UNAIDS; 2000. Available from URL: http://www.who.int/hiv/pub/surveillance/pub3/en/. Accessed: July 19, 2008.
23. World Health Organization and Joint United Nations Programme. Initiating second generation HIV surveillance systems: practical guidelines. Geneva: World Health Organization and Joint United Nations Programme; 2002. Available from URL: http://www.who.int/hiv/pub/surveillance/guidelines/en/. Accessed: July 19, 2008.
24. World Health Organization. Guidelines for effective use of data from HIV surveillance systems. Geneva: World Health Organization and Joint United Nations Programme, 2004; Available from URL: http://www.who.int/hiv/strategic/surveillance/hivpubsurveillance/en/. Accessed: July 19, 2008.

25. European Centre for the Epidemiological Monitoring of AIDS. 1993 revision of the European AIDS surveillance case definition. AIDS surveillance in Europe. Quarterly Report No.37. Paris: Institute de Medicine et D'Epidemiology Africaines; 1993.
26. EuroHIV. HIV/AIDS Surveillance in Europe. End-year report 2006. Saint-Maurice: Institute de Veille Sanitaire; 2007.
27. Klavs I, Poljak M. Unlinked anonymous monitoring of HIV prevalence in high and low-risk groups in Slovenia, 1993-2003. Croat Med J. 2003;44:545-9. Available from URL: http://www.cmj.hr/2003/44/5/14515410.pdf. Accessed: July 19, 2008.
28. Klavs I, Rodrigues LC, Wellings K, Keše D, Hayes R. Prevalence of genital Chlamydia trachomatis infection in the general population of Slovenia: serious gaps in control. Sex Transm Infect. 2004;80:121-3.
29. Grgič-Vitek M, Švab I, Klavs I. Prevalence of and risk factors for self-reported sexually transmitted infections in Slovenia in 2000. Croat Med J. 2006;47:722-9. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2080463. Accessed: July 19, 2008.
30. Klavs I, Rodrigues LC, Weiss HA, Hayes R. Factors associated with early sexual debut in Slovenia: results of a general population survey. Sex Transm Infect. 2006;82:478-83.
31. Klavs I, Hamers FF. Male circumcision in Slovenia: results from a national probability sample survey. Sex Transm Infect. 2008;84:49-50.
32. Klavs I, Rodrigues LC, Wellings K, Weiss HA, Hayes R. Increased condom use at sexual debut in the general population of Slovenia and association with subsequent condom use. AIDS. 2005;19:1215-23.
33. Klavs I, Keše D, Švab I. Slovene national survey of sexual lifestyles, attitudes and health, 1999-2001: data collection methods. Zdrav Vars. 2007;46:1-8.
34. Johnson AM, Wadsworth J, Wellings K, Field J, editors. Sexual Attitudes and Lifestyles. Oxford: Blackwell Scientific publications; 1994.
35. World Health Organization. Guidance on provide-initiated HIV testing and counselling in health care facilities. Geneva: World Health Organization; 2007.

RECOMMENDED READINGS

1. Berkelman RL, Stroup DF, Buehler JW. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press: 2002. p.759-78.
2. Centers for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50 (No. RR-13):1-36. Available from URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm. Accessed: July 19, 2008.
3. Hadden F, O'Brien S. Assessing acute health trends: surveillance. In: Pencheon D, Guest C, Meltzer D, Gray JAM, editors. Oxford handbook of public health practice. Oxford: Oxford University Press; 2005. p.14-9.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**OUTBREAK INVESTIGATION**
Module: 2.2.2	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Maja Sočan**, MD, PhD, Assistant Professor National Institute of Public Health, Ljubljana, Slovenia
Address for correspondence	**Maja Sočan** National Institute of Public Health, Trubarjeva 2, 1000 Ljubljana, Slovenia tel: 0386 1 2441 522 e-mail: maja.socan@ivz-rs.si
Keywords	Communicable diseases, outbreak investigation
Learning objectives	After completing this module students should: • recognise the ongoing outbreak; • be aware of public health importance and the urgency of immediate public health action to prevent the spread of communicable diseases; • gain knowledge in procedures for outbreak investigation (OI); • be capable to perform field investigation and analysis of data collected; • increase knowledge in measures for the control of communicable disease outbreak; • understand the need of follow-up of control measures.
Abstract	The module is describing principals of OI with implementation and follow-up of control measures. The phases in OI are: defining the problem, appraising the existent data - including case identification, clinical observation, tabulation of data collected, identification of microorganism causing communicable disease, formulating the hypothesis, testing the hypothesis, drawing the conclusions and preparation of control measures.
Teaching methods	Teaching methods include introductory lectures on phases of OI. In continuation, students have to prepare a detailed work plan of OI with operations at central level, field investigation, analysis of investigation data and measures for the control of outbreak. The work plan is at first discussed in a small group, then presented to other students. At the end students prepare a complete written report on OI. Students have to find the published reports on similar outbreaks and compare the solutions proposed for their containment.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire; structured essay; seminar paper; case problem presentations.

OUTBREAK INVESTIGATION

Maja Sočan

THEORETICAL BACKGROUND

In the investigation of an ongoing outbreak of an infectious disease, which is an epidemic, limited to localized increase in the incidence of an infectious disease in an area, quick reaction is essential (1-3). When there is a possibility of increased number of communicable disease cases, it is necessary to carry out a thorough outbreak investigation which is the responsibility of public health (4,5,7). The success of the investigation will depend on the methodical organization of operations both at central level and in the field.

Crucial steps in an epidemiological investigation

The crucial steps in epidemiological investigation are (4):

1. defining the problem;
2. appraising the existent data - including case identification, clinical observation, tabulation and spot maps, if necessary, and identification of microorganism;
3. formulating the hypothesis;
4. testing the hypothesis;
5. draw conclusions and prepare control measures.

Operations at central level

As soon as the initial information on an outbreak reaches the public health institution, the epidemiologist must verify the information. If the existence of an outbreak is confirmed (or even pending confirmation) the epidemiologist must analyze the situation and initiate the decision-making processes at the central level (3).

The epidemiological team(s) for on-the-spot investigations should be set with no delay. The team of epidemiologist must have very clear instructions.

Checking of initial information on an outbreak

The information on an outbreak may come initially from different sources:

1. routine or emergency reports from medical facilities;
2. epidemiological surveillance;
3. the early warning system;
4. from other sources, e.g. veterinary services, laboratories, or, as frequently happens, from rumours disseminated rapidly by the media or by people who senses that a problem exists.

The epidemiologist checks carefully the validity of the information on the outbreak. Comparison of the information obtained from a number of different sources will indicate whether the initial reports are reliable. The situation may also require a rapid site visit by a competent person who should have both clinical

and epidemiological experience of the suspected disease. The epidemiologist has to be aware of different agents involved in the outbreak and collect appropriate laboratory specimens to confirm the tentative clinical diagnosis. The broad differential diagnosis must be kept in mind before a preliminary conclusion is reached (2).

Preliminary analysis of the situation

As a first step, it is necessary:

1. to establish a clear definition of the disease (case definition);
2. to formulate initial hypotheses as to nature of the agent and the cause of the outbreak;
3. to formulate the objectives and strategy of the outbreak investigations.

Case definition

The case definition is the crucial for outbreak investigation as it will serve as a guide in case finding. The case definition should be precise but at the same time not too exclusive. The provisional case definition will be based on the examination of the earliest cases seen and will be changed as soon as more precise clinical and epidemiological data become available. A final case definition should contain the following (2,3):

1. the name of the disease (syndrome);
2. the most frequent and the occasional signs and symptoms in both mild and severe cases;
3. the circumstances associated with the occurrence of cases (e.g. time, place, habit, contact);
4. a suggestive or confirmatory laboratory test;
5. the criteria for "confirmed", "presumptive", and "suspect" case, and index case, "primary" or "secondary" position in the transmission chain.

Since not all cases will be investigated with the same degree of thoroughness, a grading system is necessary to indicate the certainty with which the diagnosis is made. The criteria for confirmed, presumptive and suspected (or probable) cases are defined.

The following is an example of a case definition in an outbreak of a suspected food-borne disease to which the name dysentery-like syndrome was given. Initial definition (Example 1):

Example 1.

A person having diarrhoea and one or more of the following signs and symptoms: fever, nausea, vomiting, abdominal cramps, tenesmus, blood in the stools.

Final definition is made in the end of the investigation and can be graded:

1. by severity: a severe case is one with fever above 38ºC and bloody diarrhoea, with or without other signs and symptoms as indicated above, a mild case is one fitting the initial definition, but without high fever and bloody diarrhoea;

2. by level of certainty: a confirmed case is one from which a strain of Shigella has been insolated and identified, with or without clinical signs or symptoms, a presumptive case is one where no agent has been isolated but where the faecal exudate is rich in polymorphonuclear leukocytes or macrophages, a suspect case is one with a compatible clinical picture, but without positive laboratory findings;
3. by epidemiological associations: a primary case is someone who ate at a certain food with other primary cases and became ill.

Formulation of hypotheses

Initial hypotheses should be formulated as to:

1. the nature of the disease;
2. the origin of the outbreak;
3. the mode of transmission.

Such hypotheses are necessarily based on incomplete information but also essential as a guide to further investigations. The hypotheses are subjected to modification, refinement, or total change as the study proceeds. The approach by syndrome to the etiological diagnosis of disease causing outbreak, facilitates a comprehensive review of the different microorganisms that should be considered in the differential diagnosis (7).

Equipment and logistic support of outbreak investigation

Rapidity and efficacy are essential to successful outbreak investigation and most appropriate measures taken. The contingency planning should be prepared in advance to help in making available in the shortest possible time all the equipment and logistic support needed by field teams. The types of equipment that may be required will depend on local conditions (3).

Maximum protection, including respiratory, may be necessary in examining highly contagious patients, and in high-risk operations, such as post-mortems or the processing of dangerous laboratory specimens. Gowns, gloves and surgical masks can be effective in preventing disease transmission.

Laboratory support

Laboratories must be identified within the country, which are capable of diagnostic procedures. In decision process to which laboratory, or laboratories, specimens should be sent, the following should bet taken into consideration (1,2):

1. the nature of the suspected microorganism;
2. the level of expertise required and available;
3. the types of protective equipment available;
4. the facilities for the shipment of specimens;
5. the delays expected in receiving results.

Organization of field investigation

The organization of field investigation includes selection of investigation team, equipment and logistic support and depends on the time required to find and examine the cases, the time required to collect laboratory specimens, the extend of the outbreak and the time required for completing emergency control measures.

Outbreak investigation team should be given precise instructions covering their activities, including safety precautions, methods of case finding, contact tracing, special investigations, and collection and shipment of laboratory specimens (8-10).

Case finding

During outbreak investigation, the procedures used are aimed to identify as many cases as possible. The use of a carefully designed questionnaire form is necessary to ensure accurate and rapid investigations. Records of cases should be based on precise definitions and the findings systematically validated.

Case description and recording

A standard case investigation form is needed to ensure that complete information is obtained for each case. The type of information needed is the same for all diseases outbreaks, but specific details must be adapted to the individual disease and to the unique circumstances of each outbreak situation and each location. Despite the pressure of time, great attention must be given to patient identity, the serial numbers allocated to cases, forms and specimens, and the labeling of specimens. Important data are too often rendered meaningless or misleading because of errors.

In some cases, data should be collected from hospitals and community health centres. A retrospective survey should be conducted using records of inpatients, outpatients and laboratory results going back over the previous three months or so. Special attention should be given to cases that might have been misdiagnosed. In any case, an increase in the number of consultations or admissions should arouse suspicion (11,12).

Community survey is indicated not only to discover suspected cases but also to investigate the epidemiological factors that may have contributed to the spread of the causative agent in the community. The community may be a city, district, village or camp, depending on the circumstances. A case-finding strategy should be developed in the community concerned, using one or more of the following procedures:

1. visit to local health facilities;
2. interviews with doctors, pharmacists, nurses, veterinarians, and veterinary health personnel etc.;
3. interviews with a random statistical sample of the population or with the population at points where people gather together;
4. visit to hospitals known contacts of inpatients and outpatients who reported to hospitals and health centres;
5. systematic house-to-house visits (extensive or to a random statistical sample).

Search for source of infection and contact tracing

The primary of searching for the source of infection, either of an individual case or of the entire group of cases, is to eliminate, terminate, or isolate the source so that similar circumstances do not occur again or are less likely to do so in the future (13-15).

The methods used for identifying sources and tracing will differ according to whether an individual case or an outbreak is being investigated, whether the relevant infection is transmitted from person to person or by common source origin. However, the steps to bet taken and the order in which they are taken will remain the same:

1. identify the date (or time) of disease onset;
2. ascertain the range of incubation periods for the disease in question,
3. look for a source of infection in the time interval between the maximum and minimum incubation periods.

Incubation periods vary from an little as a few hours (e.g., salmonellosis), to days (e.g. influenza), weeks (e.g., hepatitis A), and even months (e.g. rabies).

Outbreak can be:

1. with continuous person-to-person transmission;
2. with a common source of infection.

When the disease under study may be transmitted by some other mechanism, a search should be made at the times indicated by the incubation period for the presence of vectors, reservoir animals, environmental contamination, or whatever the appropriate source of infection is.

If the analysis provides grounds for suspecting that a common source was involved, the search for that source is basically the same as in person-to-person transmission. The period of time during which the common exposure might have taken place is determined by the range of the incubation period and this in turn defines the period to be covered by the search. The common source may have been an individual and the mechanism of transmission person-to-person. Information on exposure history is obtained from the index case. As previously mentioned, this section of the case investigation form must be specially designed for a particular disease outbreak.

Continued transmission

The purpose of prospective or potential case contact tracing is the identification of new cases that may already have occurred or may still result from contact with the source of infection. Whether the mechanism of spread is person-to-person, via vectors, or contaminated food or the environment, the investigation should be based on the infective microorganism. As the infective period may be brief (e.g. influenza), a few weeks (e.g., hepatitis A), a year or more (e.g. malaria), or lifelong (in chronic carriers of hepatitis B) the maximum duration of the infective period defines the time during which a search for new cases should be conducted.

It is important to recognize that not only does the duration of the period of communicability vary, but also that this period may or may not begin before the appearance of signs and symptoms, and may or may not continue for a variable time after they have appeared.

Forward (prospective) tracing of contacts has two main purposes:

1. to identify continuing chains of infection and/or contamination, in order to interrupt them;
2. to locate new cases, so that they can be treated and further spread of infection discontinued.

It is particularly important if the patients have travelled outside his own community, and may therefore have introduced infection some distance away from his home (16-21).

Special investigations for common-source infections

Whenever there is some evidence of a common source of infection, special investigations must be carried out. Such sources may include arthropods, vertebrate animals, food, beverages and the environment.

Food-borne diseases

A food-borne disease will obviously be suspected when a number of persons who have eaten a meal together fall ill (22-25). Finding the infected dish responsible is more difficult and all those who eat the meal should be classified into subgroups according to the dishes that they consumed (26-28). Tracing the source of infection is even more difficult if the incriminated food has been eaten intermittently in different places, or if the contaminated product is mixed with different kinds of food or beverages (29,30). Food contamination may originate from infected animals, food handlers, flies and the environment (31-33).

Investigations may have to be carried out on:

1. the conditions under which the food concerned is grown, produced and consumed;
2. the handling and storage of foodstuffs, with particular reference to conditions known to be potentially hazardous;
3. the sanitary condition of restaurants, hotel etc., and their use by the community;
4. the possibility of cross-contamination from raw food to cooked food, the packages, bags, or containers in which the food has been transported, and the cooking utensils and working surfaces associated with its preparation;
5. the health status and hygienic practices of food handlers.

The assistance of the veterinary services may also be needed here (34,35).

Arthropod-borne diseases

Numerous blood-sucking insects are able to transmit diseases from man or from animal reservoirs to a respective person: ticks, mosquitoes, fleas body, lice mites, sand flies etc. They are most often specialized in the diseases that they can transmit and this knowledge, plus their abundance at the time of the outbreak, can lead to a particular insect being incriminated as the mode of transmission. The possible role of insects is best confirmed by a specialists entomologists-epidemiologist, who will also know how best to collect specimens for laboratory examination.

Zoonoses

Zoonoses may be transmitted from vertebrate animals to man both by arthropod bites and by contamination of food and the environment. Direct contact with sick domestic or wild animals or healthy carriers, such as horses, sheep, cattle, goats, pigs, dogs, cats, poultry, monkeys, rodents, and birds, may cause the epidemic diseases. Proof that infection in an animal species is related to the human disease may require the assistance of the veterinary services (34,35).

Originating in the environment

Common-source-infection may originate in the environment if water, soil or air is contaminated; the source of such contamination may be man or animals. In addition, certain agents of mycotic diseases may be present in soil. Several modes of transmission may be involved:

1. water (water-borne diseases): transmission via drinking water, beverages, contaminated food, or from bathing in recreational waters;
2. soil: transmission by direct contact, or contact of dust with mucous membranes (respiratory tract, eyes);
3. air: inhalation of drops or aerosols.

Proof of causation should be carefully established, and this may be very difficult. It may be necessary to investigate water-treatment systems, sewage systems, possible contamination of water and soil by human faeces and animal dugs, use of soil as fertilizer, water recycling, and air conditioning systems.

Collection and shipment of laboratory specimens

As has been noted in previous sections, laboratory support is essential in clinical and epidemiological investigations. The value of the results obtained will depend on:

1. correct sampling of appropriate specimens;
2. correct storage, packaging, and shipment;
3. appropriate formulation of requests for laboratory examinations;
4. the speed with which the laboratory responds to such requests.

If possible, therefore, the field team should include a microbiologist, or seek the advice of one whenever necessary (3).

Analysis of investigation data

After the field outbreak teams have finished their investigations, clinical, epidemiological and laboratory data is compiled and analyzed as soon as they become available. The data collected are used to arrive at a probable clinical diagnosis of the disease, define the epidemiological characteristics of the outbreak, confirm the identity of the causative agent, and identify the appropriate control methods (4).

Clinical data, i.e., the signs and symptoms recorded for each patient, are tabulated, the more precise picture of the disease thus obtained enabling the provisional case definition to be revised and providing a clinical approach to an etiological diagnosis (5). A disease is generally described either in terms of the

relative frequency of the various signs and symptoms that have been observed, or by drawing the way that these frequencies change during the course of the disease. If the number of person examined is large enough, the data for suspected, presumptive, and confirmed cases, or for mild and severe forms of the disease, can be tabulated separately (2).

Without microbiological confirmation, a syndrome can be identified (e.g. febrile rash, hemorrhagic fever, febrile lymphadenopathy, febrile neurologic diseases etc.) and possible causative agents listed.

Epidemiological data are required to gain knowledge who is affected, what is the size of the outbreak and to formulate a hypothesis as to the causative agent, the mode of transmission, and the probable progress of the outbreak (2,3).

Data on the numbers of persons affected are relatively meaningless for the purposes of epidemiological analyzes without knowing the number of persons exposed or at risk. Rates must be therefore calculated but only if numerators and denominators are known.

Attack rates and case-fatality rates are calculated by personal characteristics e.g. age, sex and occupation. Spot map and distribution of cases in time are equally important indicators. The distribution of cases in time is best shown by a graph (histogram).

Contact tracing usually points out to one of the transmission patterns (person-to-person etc.). A hypothesis of causation is deduced and confirmed by statistical analysis.

General measures for the control of outbreaks

An outbreak of communicable disease may be controlled by:

1. eliminating or reducing the source of infection;
2. interrupting transmission;
3. protecting persons at risk.

It may take some time before the exact nature of the causative agent is known and this will delay the application of specific control measures, such as the immunization of persons at risk or the treatment of carriers (2-4). In an emergency, therefore, the first step must be to try to interrupt transmission, since the epidemiological investigations will quickly provide some indication of the possible mode of transmission involved. This may be:

1. person-to-person transmission, whether direct or indirect;
2. common source infection;
3. a combination of both.

General measures to be taken in various types of outbreak are described. In emergency conditions, control measures may require a degree of improvisation whenever the necessary equipment is not immediately available; this is not difficult when the principles to be followed are well understood.

Measures in outbreaks of diseases with person-to-person transmission

Measures may be necessary in respect of patients, their contacts and the community.

Patients

The health personnel participating in medical care, specimen collection, laboratory examination, post-mortems, and field operations during outbreak investigations will all require protection. Immune personnel (after immunization or natural infection) should be employed if possible. However, when the agent is unknown or if there is no vaccine, general precautions are indicated which must be adapted to the degree of contagiousness of the disease; they should be reliable but not excessive, so as to avoid waste of time and money.

The most effective general precaution is careful hand-washing after any contact with a patient, or with a suspected case. Protective measures may be divided into four categories, depending to the degree of communicability of the disease and its mode of transmission, as determined by the outbreak investigations.

Duration of precautions or isolation has to be determined. The infective or contagious period is known for most communicable diseases. When the agent is unknown, the period of contagiousness can be determined from the data of infective contacts collected during the outbreak investigations, which may fit one of a number of different patterns.

Disinfection

Safe disposal of excreta infectious material (vomit, urine, secretions, discharges, dressings and bedding) is recommended, and may be mandatory, depending on the mode of transmission of the disease; this may be achieved by using disinfectants or by incineration. If contaminated material is to be transported, the double-bagging procedure must be used (31).

Contacts

Persons who are in contact with a patient with communicable disease during the contagious period may be at risk of becoming infected and therefore of becoming in their turn a source of infection. However, the magnitude of this risk is not the same for all disease and for all persons, and must therefore be assessed and preventive measures adapted accordingly (15,17,18).

Assessment of the risk of communicable disease

The following factors influence the risk of infection: the time of contact, and in particular whether it falls within the period of contagiousness, the degree of contagiousness of the disease, the closeness of contact and the routes of transmission to which the person may have been exposed and the specific and non-specific immunity of the person concerned (12,13,18).

During control operations, the time of contact and the closeness of contact are the essential factors in determining the measures to be taken. Two types of contact may be distinguished:

1. A close contact.

   A close contact is a person who has had occasional face-to-face contact, has given personal care without protection measures, or has shared the same meal or room during the period of communicability, or handled the patient's belongings (if indirect transmission is involved);

2. A possible contact.
   A possible contact is a person who may have been exposed either:
   - at some distance away from a highly contagious case during the period of communicability in circumstances not satisfying the above criteria, e.g. in public transport, in the next bed in a hospital, or in the same workplace; or
   - thought close contact with a patient outside the period of communicability, particularly if there is some doubt about its duration.

Quarantine

This is used to restrict the contacts of a well person who has been exposed to a patient with a communicable disease during the communicability period. Quarantine must be adapted to the risk to which the person concerned was exposed and the risk that he represents for the community. The restrictions imposed should not be excessive from either the humanitarian or economic point of view. A large number of contacts may have to be dealt with in a few days. They should therefore be divided into "possible contact" and "close contact" groups, which should be dealt with separately. Each group should be divided into cohorts depending on the expected time of onset of the disease concerned; this will be determined by the range of incubation periods following the infective contact. When there are numerous contacts should be separated physically so as to avoid introducing new suspects into a group that has already completed part of the quarantine period and who should then be obliged to begin the whole period again.

The protection of patients and the isolation of their contacts in quarantine will considerably decrease the risk for the community. However, as it may not be possible to identify all patients and contacts, other methods also have to be considered.

Immunization

Emergency mass immunization is possible for a limited number of diseases. But there will inevitably be some delay before a large enough part of the population is protected by the vaccine; other methods may therefore be necessary during the interim period.

Chemoprophylaxis

Chemoprophylaxis may sometimes be used during outbreaks to protect persons who have been in contact with source of infection or with infected person (21).

Public health measures

Public health measures are as follows:

1. Restrictions on mass gatherings.
   Such restrictions may be indicated including the closure of schools and even of public places, but their effectiveness is generally limited.
2. Restrictions on travel.
   This may involve the establishment of a cordon sanitarire in order to isolate the epidemic focus or to prevent the entry of infectious persons into a country. There is, however, more justification for a cordon sanitaire when immunization is possible and the aim is to make sure that unimmunized persons do not travel and thereby carry the disease to other places. Before a cordon sanitaire around an epidemic focus can be established, it is first

necessary to define the boundaries of both the infected and the receptive areas. This is expensive, and requires close cooperation between the health services, the police, and the army, without which the measure may be ineffective. Furthermore, considerable economic loss and inconvenience may be caused to individuals.

3. Strengthening of epidemiological surveillance.
   This has proved to be both more efficient and less expensive than the cordon sanitaire. Case finding, contact tracing, and prevention of transmission should all be strengthened in any group in which suspected cases have appeared.
4. Community participation.
   Keeping the community informed will reduce the risk of panic. If the community can be included to participate in the control measures, this will contribute considerably to their effectiveness.

Control of outbreaks caused by a common source of infection

Whenever an outbreak is caused by a common source of infection - whether by arthropods, rodents, direct contact with vertebrate animals, food, water, air, soil or a combination of any of these - control methods should be based on source reduction and interruption of transmission. The assistance of a specialist entomologist, veterinarian or sanitary engineer may be required.

Mosquito-borne diseases

Mosquitoes are capable of transmitting diseases to man belong to several species and their control raises technical problems that require the assistance of a specialized team. They constitute the most important group of insects vectors, transmitting malaria, filariasis and a number of arboviruses, including those causing outbreaks of yellow fever, dengue and dengue haemorrhagic fever, Japanese encephalitis, New Word equine encephalitides, and several dengue-like fevers. Only the females bite man. They lay their eggs in impounded water, selected according to the preference of the species. Mosquito control requires planning of strategy, logistic, and field operations. It should be noted that a patient with a mosquito-borne disease, e.g., dengue or yellow fever in Aedes aegypti infested areas or malaria in Anopheles-infested areas, should not be moved into an area where such mosquitoes are present; such movement may be subject to local health regulations (5).

Rodent-borne diseases

Rodents may be reservoirs of a number of epidemic diseases, including leptospirosis, plague, tularaemia, yersiniosis, lymphocytic choriomeningitis, Lassa fever, Junin and Machupo haemorrhagic fever with renal syndrome. Certain rodent-borne diseases may be passed from rodents to man by “direct” transmission, others thought arthropod vectors. "Direct" transmission occurs as a result of contamination of food and water by rodent urine and can thus also be regarded as indirect. The results of the outbreak investigations will determine which procedure(s) - environmental improvement, rodent-proofing, and domestic rodent extermination by rodenticidents - are to be used and in which order. In an

outbreak of plague, the first step in control operations is to use insecticides to kill rat fleat before using rodenticides to kill the rats.

Zoonoses

Different routes of transmission to man is possible, as follows:

1. direct;
2. thought arthropods and rodents;
3. thought food and the environment.

Direct transmission is mainly an occupational risk of veterinary personnel, farmers, and hunters, and may be more frequent in areas of poor hygiene. Control measures for outbreak resulting from direct contact with animals vary, depending on the diseases and circumstances.

Food-borne diseases

Food-borne diseases may be divided into intoxication (food poisoning) and infections. Outbreaks are most frequently caused by Salmonella, Clostridium perfringens, Staphylococcus aureus, Bacillus cereus, Campylobacter, Escherichia coli, Clostridium botulinum, and Yersinia enterocolitica. However, in many cases, the origin of food-borne outbreaks remains unknown and viral agents, including caliciviruses, astroviruses etc. A virus may be more frequently involved than indicated by present data (33,34).

Measures include: elimination of contaminated food and prevention of extension of the outbreak by withdrawal of suspected food from market, treatment of contacts, identification of infected food handlers, recommendations for good food handling and preparing practice (36-40).

CASE STUDY: OUTBREAK OF Q FEVER AMONG A GROUP OF HIGH SCHOOL STUDENTS IN SLOVENIA, MARCH-APRIL 2007

Introduction

Infection with Coxiella burnetii causes Q fever which occurs sporadically or in outbreaks in endemic areas (41-45). Q fever is an underdiagnosed and underreported communicable disease as majority of cases remains asymptomatic or manifest as a nonspecific flu-like febrile illness (46-49). According to ECDC report, there were 958 notified cases of Q fever in year 2005 in European Union, mainly from Germany, France and Spain. The number of notified Q fever cases in Slovenia is very low, with incidence rate varying between 0.15-0.25 cases per 100.000 (50-53).

A group of 33 veterinary students and two teachers contracted a laboratory-confirmed Q fever infection during a training course on a sheep farm in Slovenia in March 2007. The phases of outbreak investigation are described.

Outbreak

Outbreak investigation

On 17 April 2007 the Communicable Disease Centre at the National Public Health Institute was informed about a case of Q fever in an 18 year-old student of a veterinary high school. The patient had developed high fever and a severe headache

on 30 March 2007. Chest X-rays showed pneumonia. The student reported that her classmates in the same school year had been complaining about similar symptoms.

We suspected that the patient and her schoolmates might have been exposed to a common source of Q fever when attending a training course on a sheep farm (Farm A) located in the south-western part of Slovenia in March 2007. An outbreak investigation was launched involving all third grade students of the veterinary high school.

Epidemiological and clinical data

The patient's high school year has 66 students. As part of their training, 45 of the students spent several hours between 5 and 23 March 2007 on Farm A, together with three teachers. They were trimming sheep's feet, disinfecting wounds, and applying intramuscular vitamins and anti-helmintic injections as a preventive measure to healthy animals in a stable with approximately 500 sheep. Parturition time of the sheep on Farm A was from end of January to beginning of March this year.

We interviewed all individuals who had been at Farm A, and tested them for Q fever. In addition, 20 students who had not participated in the training course in March were included as a control group. One student from the control group had spent a short time on Farm A in autumn 2006. All interviewed individuals (68 altogether) had been in contact with different domestic animals at home and during school training courses at several locations.

Among 48 exposed individuals, there were 34 (71%) with high fever (38°C or more) with or without a headache. Four individuals (8%) had serious headaches only, but were not sure about fever, and three individuals (6%) reported symptoms of a common cold. Seven individuals (15%) were asymptomatic. The first person to develop symptoms of high fever and headache started to feel ill on 20 March 2007.

Among the 20 control subjects (including the student who had spent a short time on the sheep farm in autumn 2006), three students reported having a prolonged cough, and one had symptoms of a common cold. Sixteen (80%) students were completely asymptomatic.

Laboratory investigation

Serum samples were collected from 63 individuals (93% of the 68 interviewees) and sent to the Institute of Microbiology and Immunology, Medical Faculty in Ljubljana. They were tested by indirect immunofluorescence (FOCUS diagnostics) for the presence of IgG and IgM antibodies to C. burnetii phase I and II antigens.

A laboratory-confirmed case of acute Q fever was defined as an individual with IgM titres higher than 1:16 and/or IgG titres higher than 1:256.

Overall, the results confirmed acute Q fever in 36 individuals (57% of the 63 tested). Serum samples from 26 individuals (41%) were negative and the test result of one sample was inconclusive.

Of a total of 48 individuals who were exposed on Farm A in March 2007, 44 were tested and 35 of those (80%) were seropositive. They are shown in the Figure by date of onset of symptoms (Figure 1).

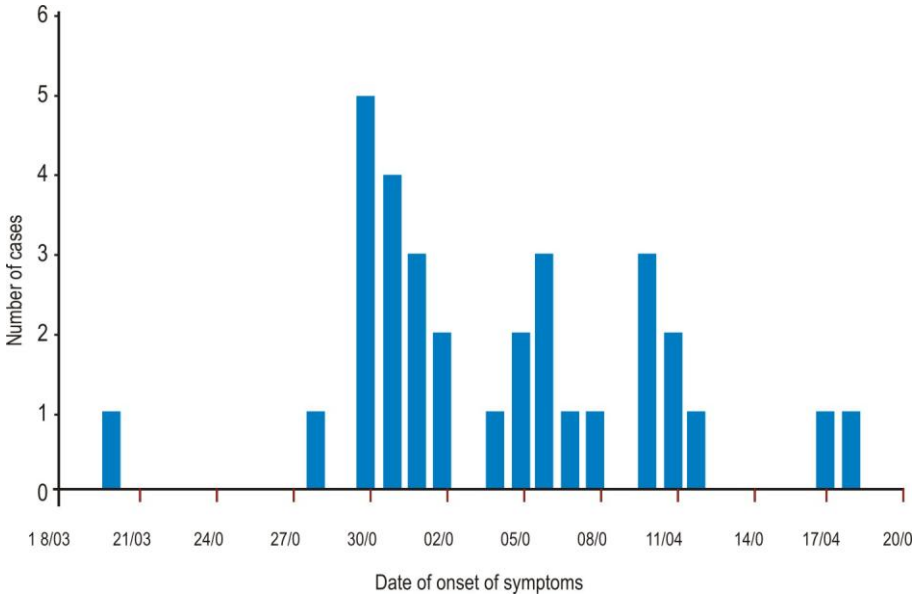


Figure 1. Laboratory-confirmed Q fever cases in Slovenia by onset of symptoms, March and April 2007 (N=32). All cases shown had visited farm A in March 2007. The date of onset of symptoms was known for only 32 of 35 cases.

Eight exposed individuals (18% of 44 tested) were not infected, and one had an inconclusive result. In the control group, none was seropositive, except for one student who visited the sheep farm in autumn 2006. The serology results of exposed versus non-exposed individuals are summarised in the Table 1.

Among the 36 laboratory-confirmed cases (35 exposed in March 2007 and one exposed in autumn 2006), four were asymptomatic, while the other 32 (89%) suffered from high fever, headache, chills, muscle aches, sweating and nausea. Twenty-five patients (69%) consulted a physician, and three (8%) developed radiologically confirmed pneumonia. Although 13 cases (36%) were given antibiotic treatment, only six (16%) received adequate antibiotic therapy with doxycycline, quinolones or macrolide antibiotic. At the time the treatment was initiated, the laboratory results indicating Q fever infection were not known.

Table 1. Laboratory confirmed cases of Q fever in individuals exposed or not exposed on a sheep farm.

	Exposed	**Not exposed**	**Total**
Seropositive	36	0	36
Seronegative	8	18	26
Inconclusive	1	0	1
Not tested	4	1	5
Total	49	19	68

The data analysis showed a significant statistical correlation between a positive serological test for C. burnetii and attendance of the training course on the sheep farm in March 2007 ($p<0,001$). There was no significant statistical correlation between a positive test and attendance of school training courses in other places or contact with domestic animals at the students' homes. We concluded that the sheep farm was the source of the Q fever outbreak.

Outbreak control measures

The first Q fever outbreak among humans in Slovenia was mentioned in 1954. Two outbreaks were described in 1985. The source of one was contact with a sheep herd; the other occurred among workers at a tannery, who were probably exposed through contact with sheep hides. Further outbreaks due to contact with infected sheep were reported in 1991 and 1992. Between 1996 and 2005, between zero and five Q fever cases were notified in Slovenia.

The Communicable Disease Centre at the National Public Health Institute was collaborating with regional epidemiologists from Ljubljana and Koper, and with the veterinary high school and faculty of Ljubljana on the outbreak investigation. Further collaboration with the Department for Infectious Diseases and the Health Inspectorate at the Ministry of Health, and the Veterinary Office at the Ministry of Agriculture resulted in a cascade of public health and veterinary measures.

Public health measures

Public health measures were as follows:

1. suspension of all student training courses and visits to the sheep farm;
2. ban on selling dairy products from the sheep farm;
3. improvement of sanitary conditions on the sheep farm;
4. serological testing of farm employees;
5. workplace risk reassessment (farm employees, forestry workers in the vicinity), in process;
6. serological testing of students who had been trained on the sheep farm in 2007, final results expected;
7. clinical and serological follow up of seropositive human cases, in process.

Veterinary measures

Veterinary measures are:

1. serological testing of animals at the sheep farm (as found later, 60% seropositive animals);
2. re-introduction of Q fever monitoring in small and large ruminants proposed by the Veterinary Chamber;
3. vaccination of animals considered, vaccine procurement in process.

The regions endemic for Q fever have never been determined in Slovenia. A joint research project is planned by public health and veterinary scientists to investigate the burden of disease in animals and humans, to determine endemic areas of Q fever in

order to develop a sufficient basis for workplace risk assessments, and to determine the risk for the general population.

EXERCISE

Task 1

Carefully read the part on theoretical background of this module. Critically discuss outbreak investigation with your colleagues.

Task 2

From domestic (e.g. Biomedicina Slovenica, and COBISS-Cooperative Online Bibliographic System of Slovenia in Slovenia), and/or international bibliographic data-bases (e.g. Medline, PubMed) find out the reports of outbreak investigation from your country.

Task 3

After finding an example, discuss the crucial steps of outbreak investigation. Analyze how these steps have been taken in Q fever outbreak presented here.

Task 4

Discuss on strength, limitations, and obstacles in different outbreak investigations.

REFERENCES

1. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
2. Morton RF, Hebel RJ. A study guide to epidemiology and biostatistics. Rockville, MA: Aspen Publication Inc.; 1984.
3. Friis R, Sellers TA, Epidemiology for public health practice. Gaithersburg, MA: Aspen Publication Inc.; 1999.
4. Baker DJO. Practical epidemiology. London: Churchill Livingstone Ltd.; 1973.
5. Last JM, editor. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
6. Premik M. Introduction to epidemiology (in Slovene). Ljubljana: Faculty of Medicine; 1998.
7. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998
8. Liu JW, Lu SN, Chen SS, Yang KD, Lin MC, Wu CC, Bloland PB, Park SY, Wong W, Tsao KC, Lin TY, Chen CL. Epidemiologic study and containment of a nosocomial outbreak of severe acute respiratory syndrome in a medical center in Kaohsiung, Taiwan. Infect Control Hosp Epidemiol. 2006;27:466-72.
9. Wang ShX, Li YM, Sun BC, Zhang SW, Zhao WH, Wei MT, Chen KX, Zhao XL, Zhang ZL, Krahn M, Cheung AC, Wang PP. The SARS outbreak in a general hospital in Tianjin, China - the case of super-spreader. Epidemiol Infect. 2006;134:786-91.
10. Porten K, Faensen D, Krause G. SARS outbreak in Germany 2003: workload of local health departments and their compliance in quarantine measures - implications for outbreak modeling and surge capacity? J Public Health Manag Pract. 2006;12:242-7.

11. Tan J, Mu L, Huang J, Yu S, Chen B, Yin J. An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. J Epidemiol Community Health. 2005;59:186-92.
12. Ho AS, Sung JJ, Chan-Yeung M. An outbreak of severe acute respiratory syndrome among hospital workers in a community hospital in Hong Kong. Ann Intern Med. 2003 7;139:564-7. Available from URL: http://www.annals.org/cgi/content/abstract/139/7/564. Accessed: July 30, 2008.
13. Fox JD. Respiratory virus surveillance and outbreak investigation. J Clin Virol. 2007;40(Suppl 1):S24-30.
14. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe.Euro Surveill. 2006;11:212-4.
15. Boivin G, De Serres G, Hamelin ME, Côté S, Argouin M, Tremblay G, Maranda-Aubut R, Sauvageau C, Ouakki M, Boulianne N, Couture C. An outbreak of severe respiratory tract infection due to human metapneumovirus in a long-term care facility. Clin Infect Dis. 2007;44:1152-8.
16. Spala G, Panagiotopoulos T, Mavroidi N, Dedoukou X, Baka A, Tsonou P, Triantafyllou P, Mentis A, Kyriazopoulou V, Melidou A, Tsiodras S. A pseudo-outbreak of human A/H5N1 infections in Greece and its public health implications. Euro Surveill. 2006;11:263-7.
17. Areechokchai D, Jiraphongsa C, Laosiritaworn Y, Hanshaoworakul W, O'Reilly M; Centers for Disease Control and Prevention. Investigation of avian influenza (H5N1) outbreak in humans - Thailand, 2004. MMWR. 2006;55(Suppl):3-6.
18. Morrow CB, Novick LF. A case exercise in public health preparedness: a community outbreak of influenza-like illness. J Public Health Manag Pract. 2005;11:306-10.
19. Balicer RD, Huerta M, Levy Y, Davidovitch N, Grotto I. Influenza outbreak control in confined settings. Emerg Infect Dis. 2005;11:579-83.
20. Young LC, Dwyer DE, Harris M, Guse Z, Noel V, Levy MH; Prison Influenza Outbreak Investigation Team. Summer outbreak of respiratory disease in an Australian prison due to an influenza A/Fujian/411/2002(H3N2)-like virus. Epidemiol Infect. 2005;133:107-12.
21. Horcajada JP, Pumarola T, Martínez JA, Tapias G, Bayas JM, de la Prada M, García F, Codina C, Gatell JM, Jiménez de Anta MT. A nosocomial outbreak of influenza during a period without influenza epidemic activity. Eur Respir J. 2003 Feb;21:303-7.
22. Black M, McAnulty J. The investigation of an outbreak of cryptosporidiosis in New South Wales in 2005. NSW Public Health Bull. 2006;17:76-9.
23. Dalben M, Varkulja G, Basso M, Krebs VL, Gibelli MA, van der Heijden I, Rossi F, Duboc G, Levin AS, Costa SF. Investigation of an outbreak of Enterobacter cloacae in a neonatal unit and review of the literature. J Hosp Infect. 2008;70:7-14.
24. Dalben M, Varkulja G, Basso M, Krebs VL, Gibelli MA, van der Heijden I, Rossi F, Duboc G, Levin AS, Costa SF. Investigation of an outbreak of Enterobacter in a neonatal unit and review of the literature. J Hosp Infect. 2008;70:7-14.
25. Murray RJ, Pearson JC, Coombs GW, Flexman JP, Golledge CL, Speers DJ, Dyer JR, McLellan DG, Reilly M, Bell JM, Bowen SF, Christiansen KJ. Outbreak of Invasive Methicillin-Resistant Staphylococcus aureus Infection Associated With Acupuncture and Joint Injection. Infect Control Hosp Epidemiol. 2008;29:466-72.
26. Centers for Disease Control and Prevention (CDC). Outbreak of multidrug-resistant Salmonella enterica serotype Newport infections associated with consumption of unpasteurized Mexican-style aged cheese--Illinois, March 2006-April 2007. MMWR. 2008;57:432-5. Available from URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5716a4.htm. Accessed: July 30, 2008.
27. Lim LS, Varkey P, Giesen P, Edmonson L. Cryptosporidiosis outbreak in a recreational swimming pool in Minnesota. J Environ Health. 2004;67:16-20.

28. Nygård K, Schimmer B, Søbstad Ø, Walde A, Tveit I, Langeland N, Hausken T, Aavitsland P. A large community outbreak of waterborne giardiasis-delayed detection in a non-endemic urban area. BMC Public Health. 2006;6:141. Available from URL: http://www.biomedcentral.com/1471-2458/6/141. Accessed: July 30, 2008.
29. Beaudeau P, de Valk H, Vaillant V, Mouly D, Ledrans M, Mannschott C, Tillier C. Lessons learned from ten investigations of waterborne gastroenteritis outbreaks, France, 1998-2006. J Water Health. 2008;6:491-503.
30. Fong TT, Mansfield LS, Wilson DL, Schwab DJ, Molloy SL, Rose JB. Massive microbiological groundwater contamination associated with a waterborne outbreak in Lake Erie, South Bass Island, Ohio. Environ Health Perspect. 2007;115:856-64.
31. Martinez A, Dominguez A, Torner N, Ruiz L, Camps N, Barrabeig I, Arias C, Alvarez J, Godoy P, Balaña PJ, Pumares A, Bartolome R, Ferrer D, Perez U, Pinto R, Buesa J; Catalan Viral Gastroenteritis Study Group. Epidemiology of foodborne norovirus outbreaks in Catalonia, Spain. BMC Infect Dis. 2008;8:47.
32. Outbreaks of Salmonella infections associated with eating Roma tomatoes--United States and Canada, 2004. Can Commun Dis Rep. 2005;31:225-8.
33. Smith KE, Stenzel SA, Bender JB, Wagstrom E, Soderlund D, Leano FT, Taylor CM, Belle-Isle PA, Danila R. Outbreaks of enteric infections caused by multiple pathogens associated with calves at a farm day camp. Pediatr Infect Dis J. 2004;23:1098-104.
34. Arshad MM, Asmar HA, Rahbar MH, Boulton ML, Wells E, Wilkins MJ, Saeed AM. Risk factors for Salmonella oranienburg outbreak in a nursing home in Michigan. J Am Geriatr Soc. 2006;54:715-7.
35. Talbot EA, Gagnon ER, Greenblatt J. Common ground for the control of multidrug-resistant Salmonella in ground beef. Clin Infect Dis. 2006;42:1455-62.
36. National Institute of Public Health of Slovenia, Centre for Communicable Diseases. News (in Slovene) [homepage on the Internet]. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/1311-CNB_maj.pdf. Accessed: July 30, 2008.
37. National Institute of Public Health of Slovenia, Centre for Communicable Diseases. Communicable diseases surveillance, 1985 (in Slovene). Ljubljana: National Institute of Public Health of Slovenia, Centre for Communicable Diseases; 1986.
38. National Institute of Public Health of Slovenia, Centre for Communicable Diseases. Communicable diseases surveillance, 1991 (in Slovene). Ljubljana: National Institute of Public Health of Slovenia, Centre for Communicable Diseases; 1992.
39. National Institute of Public Health of Slovenia, Centre for Communicable Diseases. Communicable diseases surveillance, 1992 (in Slovene). Ljubljana: National Institute of Public Health of Slovenia, Centre for Communicable Diseases; 1993.
40. National Institute of Public Health of Slovenia, Centre for Communicable Diseases. Communicable diseases surveillance, 2005 (in Slovene). Ljubljana: National Institute of Public Health of Slovenia, Centre for Communicable Diseases; 2006. Available from URL: http://www.ivz.si/javne_datoteke/datoteke/798-Epidemiolosko_spremljanje_NB_2005.pdf. Accessed: July 30, 2008.
41. Veterinary Office at Ministry of Agriculture, Slovenia [homepage on the Internet] (in Slovene). Available from URL: http://www.vurs.gov.si/si/javne_objave/mrzlica_q/. Accessed: July 30, 2008.
42. Cutler SJ, Bouzid M, Cutler RR. Q fever. J Infect. 2007;54:313-8.
43. van Woerden HC, Mason BW, Nehaul LK, Smith R, Salmon RL, Healy B, et al. Q fever outbreak in industrial setting. Emerg Infect Dis 2004;10:1282-9. Available from URL: http://www.cdc.gov/Ncidod/eid/vol10no7/03-0536.htm. Accessed: July 30, 2008.
44. Sukrija Z, Hamzić S, Cengić D, Beslagić E, Fejzić N, Cobanov D, Maglajlić J, Puvacić S, Puvacić Z. Human Coxiella burnetii infections in regions of Bosnia and Herzegovina, 2002. Ann N Y Acad Sci. 2006;1078:124-8.

45. Medić A, Dzelalija B, Punda Polić V, Gjenero Margan I, Turković B, Gilić V. Q fever epidemic among employees in a factory in the suburb of Zadar, Croatia. Croat Med J. 2005;46):315-9.
46. Carrieri MP, Tissot-Dupont H, Rey D, Brousse P, Renard H, Obadia Y, Raoult D. Investigation of a slaughterhouse-related outbreak of Q fever in the French Alps.Eur J Clin Microbiol Infect Dis. 2002;21:17-21.
47. Manfredi Selvaggi T, Rezza G, Scagnelli M, Rigoli R, Rassu M, De Lalla F, Pellizzer GP, Tramarin A, Bettini C, Zampieri L, Belloni M, Pozza ED, Marangon S, Marchioretto N, Togni G, Giacobbo M, Todescato A, Binkin N. Investigation of a Q-fever outbreak in northern Italy. Eur J Epidemiol. 1996;12:403-8.
48. Parker NR, Barralet JH, Bell AM. Q fever. Lancet. 2006 Feb 25;367:679-88.
49. Annual Epidemiological Report on Communicable Diseases in Europe. Stockholm: European Centre for Disease Prevention and Control; 2007.
50. Grilc E, Socan M, Koren N, Ucakar V, Avsic T, Pogacnik M, Kraigher A. Outbreak of Q fever among a group of high school students in Slovenia, March-April 2007. Euro Surveill. 2007;12(29):pii=3237. Available from URL: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=3237. Accessed: July 30, 2008.
51. Maurin M, Raoult D. Q fever. Clin Microbial Rev. 1999;12:518-53.
52. Aricau-Bovary N, Rodolakis N. Is Q fever an emerging or re-emerging zoonosis? Vet Res. 2005;36:327-49.
53. Marrie TJ, editor. Q fever. Vol. 1. The disease. Boca Raton, Fla: CRP Press Inc.; 1990.

RECOMMENDED READINGS

1. Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Oxford: Oxford University Press; 2002.
2. Beaglehole R, Bonita R, Kjellström. Basic epidemiology. Geneva: World Health Organization; 1998
3. Ofner-Agostini M, Wallington T, Henry B, Low D, McDonald LC, Berger L, Mederski B; SARS Investigative Team, Wong T. Investigation of the second wave (phase 2) of severe acute respiratory syndrome (SARS) in Toronto, Canada. What happened? Can Commun Dis Rep. 2008;34:1-11.
4. Areechokchai D, Jiraphongsa C, Laosiritaworn Y, Hanshaoworakul W, O'Reilly M; Centers for Disease Control and Prevention (CDC).Investigation of avian influenza (H5N1) outbreak in humans--Thailand, 2004. MMWR. 2006;28;55 Suppl 1:3-6.
5. Gilsdorf A, Kroh C, Grimm S, Jensen E, Wagner-Wiening C, Alpers K.Large Q fever outbreak due to sheep farming near residential areas, Germany, 2005. Epidemiol Infect. 2008;136:1084-7.
6. Porten K, Rissland J, Tigges A, Broll S, Hopp W, Lunemann M, van Treeck U, Kimmig P, Brockmann SO, Wagner-Wiening C, Hellenbrand W, Buchholz U. A super-spreading ewe infects hundreds with Q fever at a farmers' market in Germany. BMC Infect Dis. 2006;6:147.
7. Brouqui P, Badiaga S, Raoult D. Q fever outbreak in homeless shelter. Emerg Infect Dis. 2004;10:1297-9.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**ORAL HEALTH INDICATORS IN EUROPE**
Module: 2.3.1	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Barbara Artnik,** DMD, PhD, Teaching Assistant Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Barbara Artnik** Chair of Public Health, Faculty of Medicine Zaloška 4, SI-1000 Ljubljana, Slovenia E-mail: barbara.artnik@mf.uni-lj.si
Keywords	Oral health; dental health surveys; quality indicators; health status indicators
Learning objectives	The main two educational objectives of this module are: • to sensitise health professionals for developing an attitude about oral health indicators (OHI), and • to increase the awareness of health professionals about the positive effects of oral health surveys on oral health of a population; After completing this module students should be capable to: • assess the data currently available; • collect additional data according to the new OHI set; • analyse interpret and present the data; and • formulate a policy response to the results.
Abstract	The burden of oral diseases and the needs of populations have been changing rapidly over the past few decades. Therefore, oral health systems are required to adjust to the transition process. In order to meet these challenges effectively, public health care administrators and decision-makers need the tools, capacity and information to assess and monitor health needs, choose intervention strategies, design policy options appropriate to their own circumstances, and improve the performance of the oral health system. The aim of the EGOHIDP I (European Global Oral Health Indicators Development Project) started under the European Health Monitoring Programme was to develop a set of indicators for monitoring and describing oral health morbidity and different facets of oral health care systems. As the results a set of 40 indicators in oral public health were identified.
Teaching methods	The teaching programme is carried out as a discussion led by moderator. After every activity specific learning objectives are determined for every participant and until the next workshop their professional tasks are performed. Their achievements are reported and discussed with other participants at the next meeting.
Specific recommendations for teachers	Suggested: 2/3 of the work will be done under supervision, 1/3 individual public professionals' work. It is recommended that participants (group of 15 to 20) are all familiar with statistical package SPSS for Windows. A computer room should be provided.
Assessment of students	Attitude test for assessment of attitude changes. The questionnaires applied at the beginning of the first meeting and at the end of the course, or essay, discussing professional impact.

ORAL HEALTH INDICATORS IN EUROPE

Barbara Artnik

THEORETICAL BACKGROUND

Development of quality indicators

The main dimensions of quality of care include classification of quality into various levels: structure (organisational settings of care), process (health care treatment) and outcome (effects of care), known as indicators, which are divided up as follows (Table 1) (1).

Table 1. Donabedian's Dimensions of Quality of Care.

Structure indicators	Process indicators	Outcome indicators
• Resources	• Preventive care	• Health status
• Personnel	• Diagnosis	• Results of care
• Equipment	• Therapeutic care	• Patient wellbeing
• Facilities/Installations	• Rehabilitation	• Patient satisfaction
• Information systems	• Patient information and education	• Efficiency of resource utilisation

An indicator is defined as a variable or parameter, which can measure changes in a phenomenon directly or indirectly, in a valid, objective, sensitive and specific way.

Indicators are an essential component of all phases of health care: policy-making at the health authority level, and treatment and services at the health care provider level. Development of indicators is one of the most significant steps in any quality of care programme, and it is important that those who will primarily be using the indicators in their daily work be directly involved in the process.

Quality indicators are variables whose values indicate the level of quality. Ideally, indicators are related to the final (true) outcome but in some cases intermediate indicators of outcome must be employed. Differing from other methods for evaluating care, the use of true outcome indicators, intermediate outcome indicators and validated structure and process indicators, in the form of quality core data sets, puts the patient at the centre as the key to the successful outcome of care.

In the phase of developing quality indicators, it is important that the professional bodies be involved throughout the process, so that there will be agreement on the final selection. Without the support of the professions, the indicators will have little credibility or acceptability. It is easy to see why: if health care providers are to be motivated by their professional pride and satisfaction to improve quality of care, quality indicators will provide the basis of information regarding outcome of care, and providers must see these as relevant, valid and reliable.

In some cases, identifying or defining quality indicators is relatively easy because the literature contains evidence concerning effectiveness of interventions, etc. In other cases, it will be necessary to rely on less validated measures (2).

Sets of oral health indicators and goals for oral health

Based on this concept, and to enable the evaluation and monitoring of the results of health care, a number of different sets of indicators have been developed; for oral health in 1969, the World Health Organization (WHO) and the Fédération Dentaire Internationale (World Dental Federation) (FDI) agreed upon a basic outcome indicator (the number of decayed, missing, filled teeth = DMFT). The first global map with data on DMFT for 12-year-olds showed high prevalence of caries in industrialized countries and generally low values in the developing countries. A database was established and over a number of years an increasing number of epidemiological studies documented a pattern of change in caries prevalence, i.e. increasing levels of caries in certain developing countries and a decline in caries in many industrialized countries (3, 4). Several oral epidemiological studies have been carried out applying WHO methodology and criteria (5).

The caries decline observed in many developed countries (3, 4) was the result of a number of public health measures, coupled with changing living conditions, lifestyles and improved self-care practices. In some countries this positive trend could deter action to further improve oral health, or to sustain achievements. It might also lead to the belief that caries problems no longer exist at least in developed countries, resulting in precious resources currently available for caries prevention being diverted to other areas. However, it must be stressed that dental caries, as a disease, is not eradicated but only controlled to a certain degree.

In 1979 the most important goal ever to be formulated for global oral health was announced by WHO. By the year 2000, the global average for dental caries was to be no more than 3 DMFT at 12 years of age. At the World Health Assembly in 1979, this declaration was unanimously designated as being the overriding priority for WHO. In 1983 oral health was declared to be part of the Strategy for Health for All (resolution WHA36.14) (6). As a part of this goal, the FDI has decided to participate with "goals for oral health in the year 2000". The FDI already has many joint activities with WHO, serving as a link between that organization and the national dental member associations (7). The FDI is thus allowing for the fact that not all recommendations are applicable equally to all countries and populations. Appropriate differentiation is important. In 1989 WHO endorsed the promotion of oral health as an integral part of Health for All by the year 2000 (WHA42.39). In addition, World Health Day in 1994 was dedicated to oral health, which also reflects the importance attached to this issue (6).

In 1981, WHO and the FDI jointly formulated goals for oral health to be achieved by the year 2000, covering the following target age groups: 5-6 years old, 12 years old, 35-44 years old and 65 years old and over, as follows (8):

1. 50% of 5-6 year-olds to be free of dental caries;
2. the global average to be no more than 3 DMFT at 12 years of age;
3. 85% of the population should have all their teeth at the age of 18 years;
4. a 50% reduction in edentulousness among 35-44-year-olds, compared with the 1982 level;
5. a 25% reduction in edentulousness at age 65 years and over, compared with the 1982 level;
6. a database system for monitoring changes in oral health to be established.

The member countries of the WHO have decided to adopt a global strategy for achieving health for all in the year 2000 (8). Thirty-five European Member States have actually set up national programmes in oral health, and a number of these have reported achieving the target of ≤ 3 DMFT at age 12 (Albania, Bulgaria, Belgium, the Czech Republic, Denmark, Finland, France, Iceland, Ireland, Italy, the Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom) (4, 9).

This does not mean that it is possible to standardise concepts of care in all settings. Health care management and concepts vary from country to country and even from region to region within countries and indicators need to be adapted to the local situation. However, the basis remains the same no matter what the local conditions: it is a continuous process which works both through "top down" management and, in reverse, "bottom-up" engagement, which, taken together, can lead to continuous quality of care development (2).

For the new millennium, the WHO Regional Office for Europe specified updated oral health targets (target 8.5) as part of the Health21 policy (10). WHO, FDI and IADR jointly prepared new goals up for oral health in Europe by the year 2020; the indicator used to measure progress - DMFT at age 12 - read: no more than 1.5, and at least 80 % of children aged 6 years should be free of caries. New oral health goals were urgently needed not only to strengthen dental caries control and prevention activities, but also to address other significant components of the oral disease burden. The objectives and targets have been broadened in order to cover significant indicators related to oral health and care of population groups. The global goals are not intended to be prescriptive but the framework is primarily designed to encourage health policy makers at regional, national and local levels to set standards for oral health in relation to pain, functional disorders, infectious diseases, oro-pharyngeal cancer, oral manifestations of HIV-infection, noma, trauma, cranio-facial anomalies, dental caries, developmental anomalies of teeth, periodontal disease, oral mucosal diseases, salivary gland disorders, tooth loss, health care services, health information systems, and oral health-related quality of life. Such global goals for oral health will assist regions, countries and local health care planners to develop preventive programmes that are targeted at populations and high risk groups, and to further improve the quality of oral health systems.

Priority areas for the European oral health information systems

The burden of oral diseases and the needs of populations have been changing rapidly over the past few decades. Therefore, oral health systems are required to adjust to the transition process. In order to meet these challenges effectively, public health care administrators and decision-makers need the tools, capacity and information to assess and monitor health needs, choose intervention strategies, design policy options appropriate to their own circumstances, and improve the performance of the oral health system. Oral health systems play an important role in establishing optimum oral health by integrating oral health promotion and oral disease prevention into oral health services. Interdisciplinary and intersectional approaches to promotion of oral

health have the potential reorienting oral health services towards primary oral health care and that services may better diminish oral disease burdens (11, 12).

The global basis for an information system for surveillance of global trends in oral disease and risk factors was introduced by WHO in 1971 but still very few countries have so far prepared a comprehensive national plan for oral health care of the entire population including an information system for monitoring and evaluation. Therefore it is important to realize that the global goals for oral health in the year 2000 introduced by the WHO and the FDI (8) included a requirement to establish a national monitoring system for oral health. The WHO has developed a simple "Pathfinder" survey methodology (5) for production of baseline data for a national plan for oral health care. This methodology has been successfully applied in a number of countries. As a result the WHO Global Oral Health Data Bank (3) compiles valuable information for monitoring the global epidemiological picture and trends over-time in oral health (13).

The WHO/FDI goals for oral health by the year 2000 (8) urged Member States to establish oral health information systems, and this remains a challenge for most countries of the world. The WHO Oral Health Programme (14) was prepared to assist countries in their efforts to develop oral health information systems which include data additional to epidemiological indicators.

The information obtainable through a health information system may be usefully categorized into the following interrelated subsystems:

1. Epidemiological surveillance,
2. Service coverage of the population,
3. Service records and reporting,
4. Administration and resource management,
5. Quality of care provided,
6. Oral health programme monitoring and outcome evaluation.

Systematic evaluations of oral health systems are much needed and the WHO Oral Health Programme advocates a comprehensive model whereby input, processes, output and outcomes are measured (Figure 1).

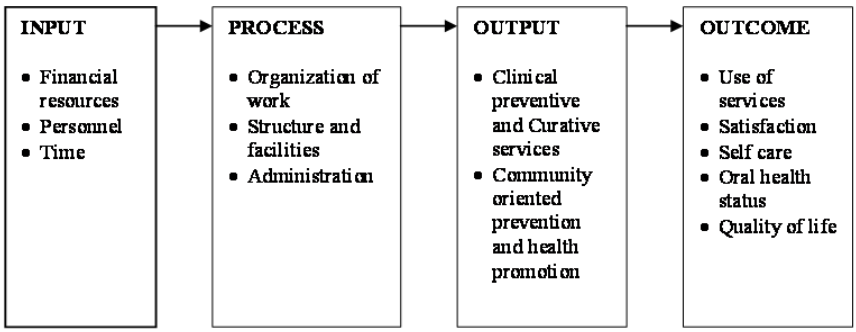


Figure 1. Oral health systems evaluation model. Source dr. Poul Erik Petersen, World Health Organization.

The WHO Oral Health Programme (14) has initiated integration of the existing database (3) with other WHO health databases and surveillance systems on risk factors. The main surveillance tool is called STEPS (STEPwise approach to surveillance), a simple approach which provides countries with core standardized methods but leaves them flexibility to expand tools by adding information relevant to the local situation (15, 16).

The WHO Oral Health Programme (14) provides modern global health information systems through several activities:

1. Revision of the WHO Oral Health Surveys Basic Methods (5), taking new oral disease patterns into account and allowing recording of risk factors to oral health (e.g. dental erosion and consumption of soft drinks);
2. Development of procedures for management and analysis of data based on the use of information technology;
3. Linking the Global Oral Health Data Bank (3) with the Country/Area Profile Programme information system (4);
4. Development of methodologies and approaches for evaluation of effectiveness of community oral health programmes with focus on health promotion and disease prevention. Such evaluation also includes process documentation in order to allow sharing of experiences from programmes.

European Global Oral Health Indicators Development Project

Numerous projects have been proposed by different teams from European countries within the framework of the Community Action Programme in the area of health surveillance (17). The European Commission Health Monitoring Programme (18) has as its main objectives to monitor the trends in the European community, to evaluate community programmes and actions and to provide Member States with appropriate health information to make international comparisons and to support their national health policies.

The development of national and international health surveillance systems has resulted in a great number of indicators overwhelming health services personnel in charge of epidemiological surveillance and evaluation of care programmes. The oral health sector is no exception. With in a context of a profusion of health indicators, operating a selection is not an easy task. The need for the necessary integration of the oral health sector within the national and European health information systems is an added challenge, considering that this should be done at all levels of the reference system. A challenge that this European public health project contributes is to meet with practical and decisive recommendations (18).

The EGOHIDP I (European Global Oral Health Indicators Development Project) (SPC 2002472) (12) has been developed in 2002 under the patronage of this Programme (18). The purpose was to establish priorities for a specifically European context in coordination with the existing programme and to make recommendations for improving health system information performance by the establishment of the major indicators of reference.

It was therefore to support the exchange of expectations and experiences among experts of oral health and their audience, policy makers in particular. The

terms of reference were also to conduct a systematic review and to outline a process for identifying a set of indicators of reference for oral health that will help national oral health public professionals and services to promote, improve and organize the global oral health promotion, quality of care and surveillance of people in Europe. Overall objectives were listed (12):

1. To support European Member States in their efforts to reduce the toll of morbidity, disability related to oral health diseases and especially to strengthen the ability at the local, national, regional levels to measure, compare and determine the effects of oral health services and use of resources on oral health;
2. To identify indicators of oral health - problems, determinants and risk factors related to lifestyle - of critical oral health care, its quality of care and of essential health resources; and
3. To identify the types of data generation and management problems within the health information system.

Principles for guiding the selection and use of oral health indicators

The major objective of this programme was to contribute to establish a community system for health surveillance (12). It embodied three specific objectives:

1. To develop community health indicators through a critical review of existing data and indicators;
2. To enable the realisation of a reliable communication system for data and health indicators transfer and sharing;
3. To define the necessary methods and instruments for analysis of activities and the production of reports on health status, trends, and policies' impact on health.

A high priority to identify indicators of reference was to encourage the development of standards for the design and implementation of computerized systems for the management of oral health systems. A goal was to seek a level of agreement sufficient to allow comparability of data that are conceptually equivalent and permit dear delineation of data.

The major principles for guiding the selection and use of oral health indicators focused on:

1. The identification of a list of priority oral health problems, populations and high risk group;
2. The definition of a table of essential indicators in the following areas: indicators of priority oral health problem, indicators of service delivery, quality of care and indicators of critical health resources;
3. The validation of the final long list of oral health indicators;
4. A common understanding of term s and criteria for selection of indicators; and
5. The recommendation of a short list of essential oral health indicators through a consultation process.

Before even starting to develop a list of existing indicators - a list that should be as comprehensive as possible - the following question should be raised (12) "Which of those indicators are we going to collectively retain on the final list?" As

soon as a selection process is engaged, a consensus should be reached on: "What will be the sorting criteria and their hierarchical order?"

This module is restricted to underline the main characteristics for a selection, in relation to the various reference areas: European Community health policies needs, scientific definition, usefulness and feasibility, ethical demand.

European Community health policy requirements

According to the European Community Health Indicators Project (12) the oral health indicator set should be:

1. Coherent in the sense of conceptual consistency, this implies that a shortlist should nevertheless cover the multidimensional aspect of oral public health surveillance, all areas usually included in the field of oral public health. This is indeed the fact for the long list already developed which is structured in the four main domains of reference;
2. Respond to oral health policy priorities, acknowledging the fact that these will be defined by each Member State and adjusted at local or regional levels;
3. Indicators should be scientifically valid, reliable and relevant.

Conceptual consistency

A set of indicators in oral public health, even restricted to a minimal essential list, has a time dimension and should cover the four major following dimensions (12):

1. Health status, morbidity and oral function status;
2. Determinants (behaviour, life habits);
3. Oral health system/promotion, prevention, access to care, quality care and system performance;
4. Outcomes and oral health quality of life.

The number of indicators in each area will vary mainly in relation to health policy priorities and to feasibility aspects of data collection and processing.

Methods used

Project members used the following procedure to select the indicators (12): a long list of over 600 possible indicators was drawn up after consultations within the group and with a wide range of relevant European clinical and scientific oral health organisations. Thirty-two group members were then asked to grade the possible indicators in order of importance. Again they were asked to confer widely before making their selections. A statistician then applied the Arrow Theorem (19) to the selections to aggretate the preferences and select 40, which were then discussed by the group.

Descriptions for all the indicators were then written using the structure described later on in this module.

Strengths and limitations of the descriptions of the indicators

The resulting indicators have been selected and described by a process of consensus between a group of decision-makers, clinicians, scientists, administrators and others. All contributors of the group can claim to be "experts" in some areas of oral health but none can claim to be "expert" in all areas. As the evidence base develops in the future and demography and epidemiology change, some of the selected indicators and

descriptions will need amendment. However, within the constraints of terms of reference of the European project (18), of time and other resources, the indicators and descriptions do provide a list, which should aid health planners in the future (12).

The issue of health policies

Increasingly European Member States or regions within Members States have formulated health priority areas or targets for health policies. There is a noticeable trend to broaden the spectrum of health objectives moving from simple morbidity measurements, or prevalence of specific diseases to objectives expressed in term s of quality of life improvements, reduction of health inequalities with reference to social policies enabling goals. For example health promotion and prevention tend to focus on specific population group s according to specific life-styles - specifically children or elderly -, goals are formulated for quality of care and access to care, or in terms of social life involvement of entire groups of population such as the aging population. For the oral health sector, this evolution implies a broader concept of the role of oral health professions and their contribution to general health. In addition, special attention should be given to the systematic integration of oral health indicators in any health surveillance system so that trends and changes in life-style and quality of life behaviour in relation to oral health can be monitored effectively. If there is a general move of health strategies towards health promotion and prevention, consideration should nevertheless be given to the fact that the situation varies considerably from country to country. There will be situations for example, where the information priority will be given to the organisation or the reorganisation of the health system for a better quality of care. Clearly health priorities are considerably variable in time and from country to country (12).

Scientific value, reliability and relevance of selected indicators

As short as the list may be, nevertheless, all selected indicators should have the four basic scientific qualities universally accepted. It was proposed to stick to the definitions given by the WHO health statistics programme (20):

1. validity: it is a true expression of the phenomena it is;
2. objectivity: it is able to provide the same result if measured by different people under similar circumstances;
3. sensitivity: it is capable of reflecting changes in the phenomena of interest;
4. specificity: it ret1ects changes in only the specific phenomena of interest.

If the WHO recommendations respond to the necessity of the scientific requirements it is also associated to a deep sense of pragmatism. An indicator that would be qualified “impeccable” scientifically but too expansive to collect or even impossible to use in a given practical situation would be totally useless. Therefore additional criteria should be considered relating to the actual use of the indicator and to the methodology used to collect the data (12):

1. The data required for the indicator are useful for case management or taking action in the community by the staff who originally recorded the data or the service unit from which the data originated;
2. It should be feasible to obtain as far as possible through routine service processes or through easily and rapidly executable surveys;

3. It should be simple and understandable, measuring one health condition or aspect of the service;
4. The indicator and the process of collecting and processing the relevant data are ethical.

Lastly, in the elaboration of the indicators selection process, quantitative principles should be considered as important criteria such as: the frequency of a given health problem, its total costs, its avoidable characteristic (prevention, promotion). This is particularly relevant for indicators of high oral health morbidity and indicators in the field or oral health determinants (12).

A flexible approach to a shortlist of oral health indicators

"STEPwise" approach developed by the WHO (15) is a practical example of a dynamic, multi-dimensional health data collection system, highly adaptable to the objectives and priority information required. In the same spirit, the European Community Health Indicators (ECHI) group proposed the concept of "user-windows" based on the selection of subsets of indicators taken from the comprehensive list of indicators developed (18). The specific user's perspective for selecting user-window could be:

1. Specific areas of health policy interest (prevention oriented, services oriented, intersectoral policies),
2. Specific thematic entries such as age-groups,
3. Specific disease groups with their determinants and costs, etc. This concept offers a more "natural" approach than that of the "core" as the number of possible windows is countless with expansion of information at any level.

Indicators and health objectives

Indicators are markers for health status, system performance and process or available resources. They are usually established to ensure follow-up and evaluation of progression towards health targets formulated by strategic programmes. They should not be confused with public health objectives expressed in terms of disease reduction or public health improvements. These are quantitative measurable achievements reached within a specific time-frame. It should be noted that oral health is broadly integrated within the health sector in the formulation of general targets as well as reflected in the list of proposed indicators. Oral health is considered as a full participative health sector, contributing not only to the promotion of oral health but also as a key actor to the promotion of general health (12).

It should keep in mind that beside their scientific qualities, the selected indicators should (12):

1. Respond to the priority needs of the community health strategies, national, local or regional, strategies for disease reduction and health promotion;
2. Be practically useful and easy to collect;
3. Be part of a highly adaptable information system, adaptable to the variety of needs and resourcesant to the evolution of scientific and economic contexts.

A set of 40 essential indicators of oral health

As the results of the European Global Oral Health Indicators Programme (2003-2005) (12) which has been supported by DG SANCO, 40 essential indicators of oral health

in Europe have been identified and harmonised. They concern problems, determinants and risk factors relating to lifestyle or critical oral health care, its quality of care and of essential health resources. The indicators have been grouped in 4 categories:

1. Children and adolescents (Part A.) (Table 2),
2. General population (Part B.) (Table 3),
3. Oral health systems (Part C.) (Table 4), and
4. Oral health quality of life (Part D.) (Table 5).

Table 2. Twelve indicators for monitoring the oral health of children and adolescents (Part A.).

Group	Indicator
Determinant	A.1. Daily Brushing with Fluoride Toothpaste
	A.2. Preventive Care-Seeking for Pregnant Women
	A.3. Mother's Knowledge of Fluoride Toothpaste for Child Tooth Decay Prevention
	A.4. Fluoridation Exposure Rates
Process	A.5. Preventive Oral Health Programs in Kindergartens
	A.6. Schools with Programs Centred on Daily Brushing with Fluoride Toothpaste
	A.7. Screening Oral Health Programme Coverage
	A.8 Protective Sealants Prevalence
	A.9 Orthodontic Treatment Coverage
Outcome	A.10. Early Childhood Caries
	A.11. Decay Experience in 1st Permanent Molars in Children
	A.12. Dental Fluorosis

Table 3. Eighteen indicators for monitoring the oral health of general population (Part B.).

Group	Indicator
Determinant	B.1. Tobacco Usage Prevalence
	B.2. Daily Intake of Food and Drink
Process	B.3. Geographical Access to Oral Health Care
	B.4. Access to Primary Oral Care Services
	B.5. Dental Contact within the Previous Twelve Months
	B.6. Reason for the Last Visit to the Dentist
	B.7. Reasons for not Visited the Dentist in the Last Two Years
	B.8. Tobacco Use Cessation
	B.9. Untreated Caries Prevalence
	B.10. Removable Denture Prevalence
	B.11.Periodontal Health Assessment
Outcome	B.12. No Obvious Decay Experience
	B.13. Dental Caries Severity
	B.14. Periodontal Diseases Severity
	B.15. Cancer of the Oral Cavity
	B.16. Functional Occlusion
	B.17. Number of Natural Teeth Present
	B.18. Edentulous Prevalence

Table 4. Five indicators for monitoring the oral health systems (Part C.).

Group	Indicator
Determinant	C.1. Cost of Oral Health Services
	C.2. Gross National Product Spent on Oral Health Care Services
Process	C.3. Dentists and Other Oral Care Clinical Providers
	C.4. Dentist Satisfaction with the Quality of Care Given
	C.5. Dentist Satisfaction with the Remuneration Provided

Table 5. Five indicators for monitoring the oral health quality of life (Part D.).

Group	Indicator
Outcome	D.1. Physical Pain due to Oral Health Status
	D.2. Psychological Discomfort due to Oral Health Status
	D.3. Psychological Disability due to Appearance of Teeth or Dentures
	D.4. Social Disability due to Oral Health Status
	D.5. Oral Disadvantage due to Functional Limitation

Indicators for monitoring the oral health of children and adolescents contain a priority list of these indicators which are specific to children and adolescents (Table 2). It must be appreciated that there are also a range of indicators in "Part B. Indicators for monitoring the oral health of general population" (Table 3), which may also be used to assess oral health in children.

In relation to oral health, "Oral health quality of life" indicators (Table 5) could be defined on the basis of performance indicators (Tables 2-4). For example, elderly people suffering from infected gums and tooth loss, will also suffer from diminished quality of life as they will not be able to eat properly and are more likely to be reluctant to socialise as they are embarrassed of their physical appearance.

As described in the WHO Catalogue of Health Indicators (1996) (21), each indicator description includes the following sections:

1. Title.
2. Rationale. Provides a brief description of the reasons why the indicator has been selected.
3. Definition of indicator textually or, in the case of proportions, rate s and ratios, by specifying the numerator and the denominator. The definition should be complete and leave no room for interpretation.
4. Definition of important temu, which may have specific meaning in the context of the indicator. Each term in the title of the indicator and its textual definition should be clear to administrative or technical staff not necessarily qualified oral health personnel. Clinical criteria, pathological terms may be defined under this section.
5. Common data sources, which could be either routine data collection, special survey or other sources. There may be a need to identify various types of data sources. This section could/should give an indication on how to collect the data (for example as part of community surveys) or where to find already existing information (for example access to databases, review of registers, of patient records etc.).

6. Recommended data collection methods which, for some indicators, are specially designed for the needs of the specific indicators.
7. Use of the indicator, which is an indication of how the indicator should be used at the facility level, and other level s of the health system. For example: to identify high-risk groups for implementation of preventive programme.
8. Recommended formats of presentation.
9. References providing primary sources of additional information about this and possibly other related indicators.

The report on selecting essential oral health indicators in Europe issued from the European Global Oral Health Indicators Project (12), supported by the Health and Consumer Protection Directorate-General, European Commission, called upon policy-makers, community leaders, private industry, health professionals, the media, and the public at large to affirm that oral health is essential to general health and well-being and to take action. Particular attention is given to the negative implications to health of changing diet, nutrition and unhealthy lifestyles related to tobacco and excessive use of alcohol. Moreover, priority should be given to the problems emerging within deprived communities or disadvantaged populations in Europe.

With the support from the European Commission, the expert contribution of the ministries of health, universities, regional and national dental associations, health professionals in the European member states, the European Global Oral Health Indicators Development Project will be instrumental to achieving the goals of improved performance of health systems and that oral health services in Europe may effectively match the needs of all population groups.

As a second step - (EGOHIDP II) (2006-2008) - methodological criteria are being defined for the collection of data to effectively promote and implement oral health indicators (22). The EGOHIDP I (12) has enabled a feasibility study (EGOHIDP II) (22), which is an essential part of an overall project, since it will allow Member States to evaluate their capability to use these indicators.

EXERCISE

For the purposes of this training programme four tasks will be executed (according to the learning objectives). The whole programme will be carried out as a discussion led by moderator. After every task specific learning objectives will be determined for every participant and until the next meeting their professional tasks should be performed. Their achievements will be reported (within 10 minutes) and discussed with other participants at the next meeting.

Task 1

Stimulating introduction at the first meeting will be led by moderator: key words will be used as a target to sensitise the participants that good oral health is essential for good general health. Discussion: The assessment process of the availability of data. Task 1 they have to achieve until the Meeting 2:

- to inventory the data that are already being collected and that can be used to assess the oral health status in different population groups;

- to assess the informative value of these data;
- to make provisions for generating new data.

Task 2

At the second meeting the reports should be presented by every participant. Discussion: Oral health indicators. The results of the first workshop will determine whether additional indicators need to be collected.

Task 2 they have to achieve until the Meeting 3 (if necessary):

- to add indicators to existing data sources;

Task 3

At the third meeting the reports should be presented by every participant. Methodological guidelines should be discussed and refined. It has to be decided:

- which oral health indicators will be used;
- the choice of an adequate level of analysis and the application of multilevel analysis.

Task 3 they have to achieve until the Meeting 4:

- to analyse differences in oral health;
- to interpreted the results carefully;
- to prepare the results for clear and understandable presentation.

Task 4

At the fourth meeting the results have to be presented clearly and understandably (e.g. to use graphical displays) by every participant. The discussion: Formulating a public health policy priorities to the results:

- to what extend has the state identified oral health as an important part of general health until now;
- what are the objectives for any interventions;
- who are the main groups with a concern for poor oral health;
- what are their interests, priorities, and commitments;
- what is the context within which interventions need to be considered, etc.

The formulated document should assure that public health policy satisfies identified needs and finally it should be submitted to policy-makers.

Follow up workshops on health policy development should be performed every six months.

REFERENCES

1. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260:1743-8.
2. World Health Organization, Regional Office for Europe. Joint ECE-WHO Meeting on Health Statistics, 14-16 Oct 1998. Available from URL: www.unece.org/stats/documents/ces/ac.36/1998/16.e.doc. Accessed: July 15, 2008.

3. World Health Organization. Global Oral Health Data Bank. Geneva: World Health Organization; 2002.
4. World Health Organization. WHO Oral Health Country/Area Profile [homepage on the Internet]. Available from URL: http://www.whocollab.od.mah.se/index.html. Accessed: July 15, 2008.
5. World Health Organization. Oral health surveys: Basic methods. 4th ed. Geneva: World Health Organization; 1997.
6. Petersen PE. The World Oral Health Report 2003. Continuous improvement of oral health in the 21st century - the approach of the WHO Global Oral Health Programme. Geneva: World health Organization; 2003.
7. Aggeryd T. Goals for oral health in the year 2000: cooperation between WHO, FDI and the national dental associations. Int Dent J. 1983;33:55-9.
8. World Health Organization/Fédération Dentaire Internationale. Global goals for oral health in the year 2000. Int Dent J. 1982;32:74-7.
9. World Health Organization. Health for All: list of indicators. Copenhagen: World Health Organization, Regional Office for Europe: 1993.
10. World Health Organization, Regional Office for Europe. Health 21. The health for all policy framework for the WHO European Region. European Health for All Series No. 6. Copenhagen: World Health Organization, Regional Office for Europe; 1999.
11. World Health Organization. Information systems. Oral health information systems, evidence for oral health policy and formulation of goals [homepage on the Internet]. Available from URL: http://www.who.int/oral_health/action/information/surveillance/en/index1.html. Accessed: July 15, 2008.
12. Bourgeois DM. Health surveillance in Europe. European Global Oral Health Indicators Development Project. SPC 2002472. Final report 2003-2005. Lyon: Community Action Programe on Health Monitoring, European Commision, Health and Consumer Protection Directorate-General; 2006. Available from URL: http://www.whocollab.od.mah.se/euro/eu/EUReportoralheath200305.pdf. Accessed: July 15, 2008.
13. Tala H. Information systems for oral health. Int Dent J. 1987;37:215-7.
14. World Health Organization. WHO oral health programme [homepage on the Internet]: Available from URL: http://www.who.int/oral_health/en/index.html. Accessed: July 15, 2008.
15. Bonita R, de Courten M, Dwyer T, Jamrozik K, Winkelmann R. Surveillance of risk factors for noncommunicable diseases: The WHO STEPwise approach. Summary. Geneva: World Health Organization; 2001.
16. Bonita R, Strong K. The SURF Report 1. Surveillance of risk factors related to noncommunicable diseases: Current status of global data. Geneva: World Health Organization; 2003.
17. Decision No 1786/2002/EC of the European Parliament and of the Council of 23 September 2002. Adopting a programme of Community action in the field of public health (2003-2008). Offic J Eur Commun. 2002; L 271/1, 9 October 2002. Available from URL: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:301:0003:0013:EN:PDF. Accessed: July 15, 2008.
18. ECHI. Design for a set of European Community Health Indicators. Final report of the ECHI project under the EU Health. Monitoring Programme. Bilthoven: RIVM; 2001.
19. Arrow KJ. Social Choice and Individual Values. 2nd ed., 8th printing. New Haven, London: Yale University Press; 1976..
20. World Health Organization. Development of health programme evaluation: Report by the Director-General. Geneva: World Health Organization; 1978.

21. World Health Organization. Catalogue of health indicators. A selection of important health indicators recommended by WHO Programmes. WHO/HST/SCI. Geneva: World Health Organization; 1996.
22. European Global Oral Health Indicators Development Phase II. Grant agreement 2005113. Kick off meeting. Lyon: Health and Consumer Protection Consumer Protection Directorate-General, Community Action Programme on Action Programme on Health Monitoring, European Commission; 2006. Available from URL: http://www.egohid.eu/Documents/Introduction%20presentation.pdf. Accessed: July 15, 2008.

RECOMMENDED READINGS

1. McKee M, Ryan 1. Monitoring Health in Europe: opportunities, challenges, and progress. Eur J Public Health. 2003;13:1-4.
2. Ochoa A, Imbert F, Ledesert B, Pitard A, Grimaud O. Health indicators in the European regions. Eur J Public Health. 2003;13:118-9.
3. Public Health Division. Report on the 1997 and 1998 NSW health surveys. Sydney: NSW Health Department; 2000.
4. U.S. Department of Health and Human Services. Healthy People 2010: Understanding and improving health. 2nd ed. Washington, DC: U.S. Government Printing Office; 2000.
5. World Health Organization, European Commission. International Compendium of Health Indicators (ICHI), version 1.1. Copenhagen: World Health Organization, Regional Office for Europe; 1999.
6. World Health Organization. WHO HFA indicators for the new health policy in Europe. Report on a WHO Expert Group Meeting, The Hague, Netherlands, 2-3 March 2000. Copenhagen: World Health Organization, Regional Office for Europe; 2000.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**HEALTH RELATED QUALITY OF LIFE AND GENERAL QUALITY OF LIFE - CONCEPTS AND MEASUREMENT**
Module: 2.4.1	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Gorka Vuletić Mavrinac,** BPsychol, PhD Andrija Štampar School of Public Health, Medical School University of Zagreb, Croatia
Address for correspondence	**Gorka Vuletić Mavrinac** Andrija Štampar School of Public Health, Medical School, University of Zagreb, Rockefellerova 4, Zagreb, Croatia Tel: +385 98 707 220; Fax: +385 1 4684441 E-mail: gorka.vuletic@gmail.com
Keywords	Quality of life, health status
Learning objectives	After completing this module students and public health professionals should: • understand the concept of quality of life in general (QOL); • differentiate health related quality of life (HRQOL) from general QOL; • increase knowledge on different measurement approaches in the field; • get to know various instruments for measuring HRQOL and QOL in general and learn how to measure HRQOL.
Abstract	Subjective health, assessed through self-reported measures of health status, is a standard in epidemiological and community-based survey research today. Their use reflects the importance of considering the people's point of view and the multidimensional nature of health. They often referred as health related quality of life measures. As distinct to this, general QOL concept and measurement was also discussed in regard to population health and well-being. As a case study, evaluation of Croatian version SF-36 is presented followed by normative data and population health profiles.
Teaching methods	Lectures, exercises, individual work through seminars and small group discussions.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 40%/60%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection equipment, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme, postgraduate students.
Assessment of students	Conducting a small research on health and quality of life using standardized instruments; writing and presenting seminar paper.

HEALTH RELATED QUALITY OF LIFE AND GENERAL QUALITY OF LIFE – CONCEPTS AND MEASUREMENT

Gorka Vuletić Mavrinac

THERORETICAL BACKGROUND

Literature on Quality of Life and related constructs is expansive. Due to the complex and philosophical nature of Quality of Life, there are many different ways it can be perceived, and many fields of study, and purposes, to which it can be applied. Although terms relating to Quality of Life are colloquially synonymous, the academic studies of Quality of Life require clear definitions that differentiate between these similar terms.

Term "Quality of life" was introduced by Medline as a heading in 1975, and accepted as a concept by Index Medicus in 1977. However, many researchers actually measures health related quality of life when they talking about QOL.

Quality of life in health care

The ultimate goal of health care is to maintain or improve the quality of life of people. Health is an important determinant of a person's quality of life although it is not the only one. Other factors such as culture, religion, environment, education and finance can also affect quality of life but they are often beyond the scope of health care.

Somebody's own perception of life needs to be recognized as valuable. Failure to consider the patient's viewpoint will underestimate disability and may also lead to failure to recognize worthwhile benefits of treatment. For example, successful heart failure treatment may have only modest demonstrable effects on cardiac function and measured exercise capacity but may enable basic everyday activities to be carried out with much greater ease and satisfaction. Quality of life in relation to one's state of health is the main concern for health care professionals and is becoming an important health outcome indicator (1). It is often referred to as "Health related quality of life" (HRQOL).

Concept of health related quality of life

The theoretical framework of health-related quality of life, then, is largely based on a multidimensional perspective of health as physical, psychological and social functioning and well-being, along the lines of the WHO's definition of health as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. The WHO has a working party on quality of life under its umbrella, which is undertaking a ten-country study of health-related quality of life. This is known as the World Health Organization Quality of Life Group (WHOQOL Group) (2).This group has provided a definition of quality of life which also takes individual perception and relationship to the environment into account: Quality of life

is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationships to salient features of their environment.

Ware, author of the widely known health status questionnaire (SF-36) (3), has argued that five health concepts are inherent in this definition: physical health, mental health, social functioning, role functioning and general well-being. He restricts his definition because the goal of health care is to maximize the health component of quality of life. Health status may influence quality of life without determining it.

In relation to previously mentioned, health status is increasingly referred to as quality of life, and, so as to narrow down its operationalisation in research studies, quality of life is increasingly referred to as health related quality of life. Health-related quality of life, like subjective health status, is patient based, but focuses more on the impact of a perceived health state on the ability to live a fulfilling life (3). From a health (or disease) perspective, quality of life has been said to refer to the social, emotional and physical well-being of patients following treatment, mirroring the World Health Organization's definition of health, and as the impact of disease and treatment on disability and daily functioning. It is a double-sided concept, incorporating positive as well as negative aspects of wellbeing and life, and it is multi-dimensional, incorporating social, psychological and physical health. It is also, ultimately, a personal and a dynamic concept for, as health status deteriorates, perspectives on life, roles, relationships and experiences change (3). Grant et al. (1990) also define quality of life as `a personal statement of the positivity or negativity of attributes that characterize one's life' (3). Taking these definitions into account, health related quality of life is defined here as optimum levels of mental, physical, role (e.g. work, parent, carer, etc.) and social functioning, including relationships, and perceptions of health, fitness, life satisfaction and well-being. It should also include some assessment of the patient's level of satisfaction with treatment, outcome and health status and with future prospects. It is distinct from quality of life as a whole, which is brother concept, including various aspects of somebody’s life (life domains) not just state of health and not only in relation to health of the individual.

In public health and in medicine, the concept of health-related quality of life often refers to a person or group's perceived physical and mental health over time. Physicians have often used health-related quality of life (HRQOL) to measure the effects of chronic illness in their patients to better understand how an illness interferes with a person's day-to-day life. Similarly, public health professionals use health-related quality of life to measure the effects of numerous disorders, short- and long-term disabilities, and diseases in different populations. Tracking health-related quality of life in different populations can identify subgroups with poor physical or mental health and can help guide policies or interventions to improve their health (4). At the present time, there is no single definition of HRQOL. Nevertheless, there is a broad consensus that it refers to the physical, psychological, and social functioning of patients and the impact of disease and treatment on their abilities and daily functioning (5).

Measurement of HRQOL

Health care researchers have developed numerous measures of HRQOL and QOL. Such instruments are widely used today in assessing the outcome of health care interventions. These measures include different indicators, from clinical symptoms, functional disabilities to subjective measures of life satisfaction and psychological well-being. Before using any of measures we should be aware of author's approach to QOL and underlying concept of QOL so we can interpret the results correctly.

HRQOL measuring instruments can be generic or disease specific. Generic instruments have the advantage of being applicable to all persons irrespective of their type or number of illnesses but they may not be sensitive to some problems unique to particular diseases. Instruments specific to a particular disease, for example Crohn disease, AIDS, or dermatologic disease such as melanoma, are more specific and sensitive but they allow comparison only between different patients with same disease. Furthermore, the results of disease specific instruments are difficult to interpret in persons with multiple diseases. The current trend favours the use of generic instruments for most patients, and if necessary, supplementing the assessment by disease specific instruments.

Gradually, definitions of health, like health-related quality of life, have moved away from a total disease model to one which incorporates health and well-being. The most widely used were measures of brother health status, or health-related quality of life, reflect this definition, and incorporate physical functioning, psychological well-being and social support and activity items. Those are in recognition of the emphasis on the positive as well as the negative aspects of health and consequences of illness.

There is large number of HRQOL instruments. One of the standardised generic instruments, internationally recognized and used is SF-36 health status questionnaire and it will be presented here.

The SF-36 Health Survey is a step forward in subjective health and HRQOL assessment. Firstly, the eight health concepts measured in the SF-36 were selected during two considerable empirical studies - the Medical Outcomes Study (MOS) and the Health Insurance Experiment (HIE), and represent the most frequently measured concepts in widely-used health surveys that have been shown to be affected by disease and treatment (6). Consequently, most SF-36 items have their roots in instruments that have been in use since the 1970s and 1980s, including well known General Psychological Well-Being Inventory and Health Perceptions Questionnaire (7). In addition to that, the International Quality of Life Assessment (IQOLA) Project gathered researchers from the whole world with the aim to translate the SF-36 Health Survey and to validate, norm, and document the translations as required for their use internationally (7).

There are also some other instruments that should be mentioned.

- WHOQOL group developed an instrument for measuring quality of life that can be used in a variety of cultural settings. The WHOQOL-100 and WHOQOL-BREF instruments are developed as a core instrument, but it can be also used in research on specific population groups (e.g. the elderly, chronic patients or cancer patients, etc.) Focusing on individuals' own views of their well being, those measures provide a new perspective on

disease, treatment and outcomes. It measures HRQOL multidimensional on five domains: Physical health, psychological health, social functioning, level of independence, and domain of environment. Those instruments can be used in particular cultural settings, but at the same time results are comparable across cultures. The WHOQOL instruments are now available in over 20 different languages and its development in further languages is progressing (8);

- EuroQol Quality of Life Scale (EuroQol 1990) - a measure that covers five dimensions - mobility, self-care, role (or main) activity, family and leisure activities, pain and mood - and expresses health status as a single index score.

Distinction in concepts

Health related quality of life should be distinguished from quality of life in general as we defined it earlier in that it concerns itself primarily with those factors that fall under the purview of health care providers and health care systems (9). Overall QOL comprised more than just health domain, therefore concept of QOL in general and will be discussed separately.

Concept of Quality of life - in general

Quality of life is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychological state, level of independence, social relationships, and their relationships to salient features of their environment (8). There is no one generally accepted definition of the QOL, and the meaning of the concept of quality of life is often dependent on the user of the term, his or her understanding of it, and his or her position and agenda in the social and political structure. According to Cummins, quality of life is both objective and subjective. We'll be focusing here on subjective QOL (10).

Subjective well-being (SWB) is another term that refers to self perceived overall QOL. It's the most global term to describe how people feel about their lives, and is based on both an emotional reaction and cognitive judgement. According to Diener (11), an assessment of subjective well-being involves the subjective experience of the individual, as the integrated judgment of the person's life both within specific domains and as a global assessment. There are two broad perspectives in conceptualisation of subjective QOL. In one, subjective QOL is a global, unitary entity (12), and in the other it is composed of discrete domains (13).

Measurement of QOL

Measurement of quality of life is increasingly being required and used in evaluative research and the planning of health services.

At the most simple form of measurement, subjective QOL can be measured as a single response to a question concerning satisfaction with 'life as a whole'. However, as single items do not specify the aspects of life for which satisfaction is important, it is more informative to measure QOL multidimensional, in areas or domains of life (11,13). Many researchers have adopted idea that quality of life can be conceptualised as a composite of satisfaction with specific domains. Despite disagreement on the definition of quality of life, there is considerable overlap among researches on relevant life domains for assessment. And considerable agreement exists on of quality of life as a multidimensional concept. According to Campbell, Converse, and Rodgers, there are five main domains: health, material well-being, intimacy, productivity, and emotional well-being (14). In addition to these, than Cummins has suggested the domains of safety and place in the community. These seven domains have been proposed on theoretical and empirical grounds based on results of previous research (13,15).

Today, there is agreement that QOL are multidimensional construct so it is suggested to be measured in that way, which was usually a set of questions measuring satisfaction within specific life domains (15,16).

It is recommended to measure subjective QOL with some of the standardised instruments (questionnaires). There are number of instruments, some of them are internationally standardised and widely used:

- the Personal Wellbeing Index (PWI) (Cummins, 2003) which measures subjective wellbeing and is designed as the first level deconstruction of 'satisfaction with life as a whole'. It comprises seven life domains as Standard of living, Health, Achievements in life, Personal relationships, Community connectedness, Safety and Future security. The Personal Wellbeing Index is an aggregate average score across the seven domains. It is standardised, psychometrically valid instrument have been extensively examined and internationally used and adapted to different languages including Croatian (17);
- the satisfaction with life scale (Pavot & Diener, 1993) is valid and reliable scale for measuring subjective QOL (18).

Assessment of subjective QOL (or SWB) gives us brother perspective on person's life despite of illness or health condition. Therefore it allows us comparison of the results between different group of patients or with the healthy (comparative) groups.

The argument in favour of measuring overall quality of life as an outcome of clinical interventions is: Quality of life indicators help to answer the question of whether the treatment leads to a life worth living, by providing a more patient-led baseline against which the effects of the intervention can be evaluated. Subjective well-being is suggested by several authors as a measure of perceived life quality irrespective to illness (19). Patient perceived symptoms as well as objective medical conditions may be added to this measurement for more comprehensive information.

CASE STUDY: PSYCHOMETRIC EVALUATION AND NORMING OF THE CROATIAN SF-36 HEALTH SURVEY: PROVIDING A FRAMEWORK FOR SUBJECTIVE HEALTH RESEARCH

Introduction

The example of generic instruments is SF-36, often used for health related quality of life measurement. The SF-36 Health Survey is a multi-purpose short-form health status questionnaire which contains 36 questions (20). It represents a theoretically based and empirically verified operationalisation of two general health concepts – physical and psychological health, and their two general manifestations – functioning and well-being. In accordance with this, the questionnaire contains four types of scales or four conceptually different measures of health. They refer to the following assessments or indicators of health:

1. functioning at the behavioural level,
2. perceived well-being,
3. limitations connected with social life and the realization of central life roles, and
4. direct personal perception of total health.

At the manifested level, each of the questionnaire items refers to one of the eight different indicators of health:

5. Physical Functioning – PF (10 items),
6. Role-Physical (limitations in performing important life roles due to physical health) – RP (4 items),
7. Bodily Pain – BP (2 items),
8. General Health – GH (5 items),
9. Vitality – VT (4 items),
10. Social Functioning – SF (2 items),
11. Role-Emotional (limitations in performing important life roles due to emotional problems) – RE (3 items),
12. Mental Health (absence of anxiety and depression) – MH (5 items).

Five questionnaire scales (PF, RP, BP, SF, and RE) define health as the absence of limitations and inability, so they represent continuing and one-dimensional health measures. The three remaining scales (GH, VT and MH) are bipolar, which means they measure a much wider range of negative and positive aspects of health. The scales PF, RP and BP refer to the general factor of physical health, and scales SF, RE and MH measure psychological health. Scales VT and GH are moderately connected with both factors. The total result is most often shown in the form of the profile defined with eight points that represent the measure of individual aspects of health transformed into a unique scale whose theoretical minimum is a score of 0, and the maximum a score of 100. On all scales higher results indicate better subjective health. Finally, the SF-36 contains one self-evaluated health transition item (five levels from “much better than one year ago” to “much worse than one

year ago"), which is not used in scoring the scales but is useful in estimating average change in health status during the year prior to its administration.

The SF-36 Health Survey was developed in the United States, but for purpose of use in international research. Original English language instrument can be obtained from co-copyright holders of all SF-36®, SF-12® and SF-8™ Health Surveys: The Medical Outcomes Trust (MOT), Health Assessment Lab (HAL) and QualityMetric Incorporated, through the SF-36 Community web site (21).

Translating, validating, and norming the SF-36 in 14 countries, including Austria, Belgium, Canada, Denmark, France, Hong Kong, Italy, Japan, the Netherlands, Norway, Spain, Sweden, the United Kingdom, and the United States (Chinese translation) were coordinated under the International Quality of Life Assessment (IQLA) Project. The results were published in the special issue of the Journal of Clinical Epidemiology in 1998 (22). In the second wave, 40 Western and Eastern European (Croatia among them) and non European countries were included in the project. Results generally support cross-cultural validity of the SF-36 Health Survey, making possible comparisons of health-related quality of life results across countries and encouraging a wide use of this instrument.

SF-36 Croatian version

The Croatian version of SF-36 questionnaire was licensed to Andrija Štampar School of Public Health. Croatian version of the questionnaire can be obtained for the research purpose; contacting the Andrija Stampar School of Public Health, Medical School University of Zagreb.

After the standard procedure of translation, a pilot study was carried out in 1998 (23). Trained interviewers applied the Croatian version of the SF-36 survey to the sample of adult population (n=5,048). Collected data provided the very first preliminary results related to metric characteristics of the Croatian version of the SF-36 (23,24). The scales showed good internal consistency, and obtained descriptive statistics were within the expected ranges. The mean values were similar to those from other European countries and provided expected profile in terms of its height and shape.

The Croatian version of SF-36 survey was incorporated in the 2003 Croatia Adult Health Survey (CAHS), the questionnaire which covers a wide range of health related variables. Multi-stage stratified sample design was adopted to define a representative sample of general adult population. The survey targeted persons aged 18 years or older living in private dwellings in the Republic of Croatia. The 2001 Croatian Census was used to select a representative sample of households to be included in this survey. The questionnaire was administered face-to-face to respondents by trained community nurses. Out of these selected households a response was obtained from 9,070 individuals, which resulted in the overall response rate of 84.3% (24). Results were shown in Table 1, and Figure 1.

More articles on perceived health in Croatian population will be published in special issue of the journal: Collegium Antropologicum, publication due to April, 2009.

Table 1. Descriptive statistics and features of score distribution for the SF-36 scales administered to a representative sample of Croatian general adult population (N=9,070).

Reported data	Physical Functioning	Role Physical	Bodily Pain	General Health	Vitality	Social Functioning	Role Emotional	Mental Health
M*	76.1	69.0	69.6	58.2	59.8	75.5	78.5	66.1
SD*	28.1	42.4	29.4	21.6	23.0	39.9	25.3	19.7
No. of items	10	4	2	5	4	2	3	5
Cronbach alpha	0.94	0.94	0.91	0.78	0.85	0.78	0.93	0.87

LEGEND: * Weighted values accordingly to the number of persons represented by the respondent for the entire population.

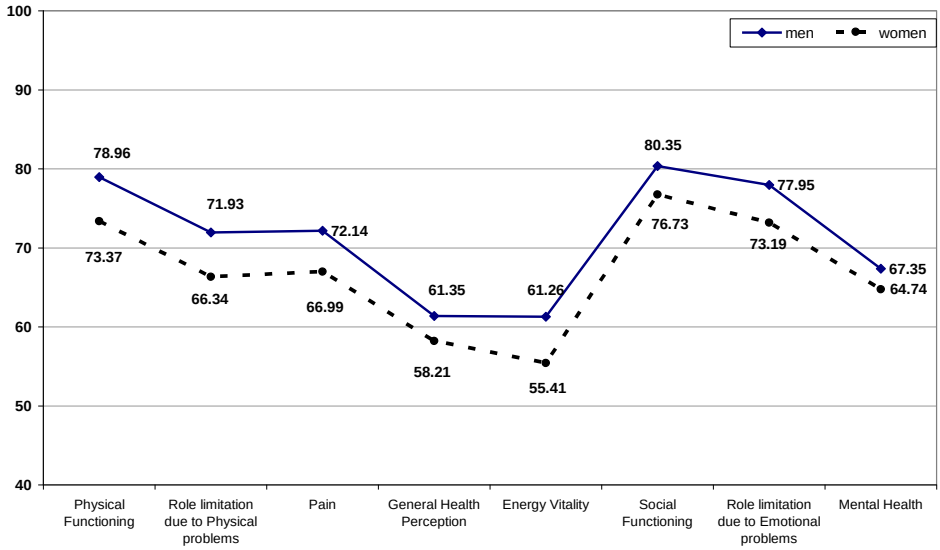


Figure 1. SF-36 Health Profile according to gender

Psychometric characteristics of the SF-36 Croatian version

Psychometric characteristics of the SF-36 Croatian version were as follows:

1. Metric characteristics.
   The Croatian version of the SF-36 had generally good metric characteristics including its construct validity - internal consistency of the scales ranged from 0.78 to 0.94; the reliability of the eight scales and two summary measures was estimated using Cronbach alpha coefficient of internal

consistency. In all cases its value exceeded the minimum standard of 0.70 and ranged from 0.78 (for General Health and Social Functioning) to 0.94 (for Physical Functioning and Role Physical Four scales - Physical Functioning, Role Physical, Bodily Pain, and Role Emotional had higher reliability value than 0.90, which is suggested as a minimum score for analyzing an individual patient's result.

2. Ceiling and floor effects.
   The highest floor effects were observed for the Role Physical and Role Emotional (30% and 26% respectively). These scales also had substantial ceiling effects (Role Emotional 63%; Role Physical 53%). Both ceiling and floor effects (percent of individual's results corresponding to the theoretical maximum or minimum) for the three bipolar scales (General Health, Vitality, and Mental Health) were minimal. These results were expected and they accord to data found with the original, U.S. version of SF-36. Obtained values for skewness, as well as for the floor and ceiling effects, were even lower in this study (24).
3. Bivariate correlations.
   Pearson's bivariate correlations showed moderate associations between SF-36 scales, and ranged from 0.45 (between Mental Health and Physical Functioning) to 0.68 (between General Health and Vitality). The exception was a somewhat higher correlation between scale results on Mental Health and Vitality, which counted 0.75. Factor analysis using principal component extraction method provided one latent dimension underlying all SF-36 scales. Extracted factor explained 63.3% of the variance.

Croatian data for open adult population, the mean scores and standard deviations for each dimension, were shown in Table 1.

Presented population norms for the Croatian version of SF-36 Health survey could be used in future research as a frame for comparisons across disease and condition groups in searching for better information about what works best in health care.

Croatian normative data for open adult population showed gender differences in health profiles (Figure 1). In general women scored lower than men with the most prominent differences in old age. As it was expected average health status declined with age for both gender.

EXERCISE

The purpose of the exercise is to provide students with basic skills necessary to explain, classify and accept the main concepts and knowledge on QOL in general and in relation to health of the individual or group by using different source of information (publications, online resources and free online journals in the field), which are listed in references and recommender readings.

Task 1

Carefully read the part on theoretical background of this module. In small groups critically discuss the concepts and measurement of QOL and HRQOL.

Task 2

Carefully read the following two documents:

- Bowling A. Measuring health: a review of quality of life measurement scales. Fourth Edition. Philadelphia: Open University Press, 1994.
- Cummins RA, Lau ALD, Stokes M. HRQOL and subjective well-being: non-complementary forms of outcome measurement. Expert Review of Pharmacoeconomics & Outcomes Research August 2004. Vol. 4. No. 4. Pages 413-420.

Task 3

Make groups of three to four students and fill the two chosen questionnaires (one for HRQOL and one for subjective QOL).

Task 4

Analyse the results in a small groups and than discuss the results of your group with results of other groups.

REFERENCES

1. McKee M, Ryan J. Monitoring health in Europe: opportunities, challenges, and progress. Eur J Public Health. 2003;13:1-4.
2. WHOQOL Group 1993 WHOQOL Group. Measuring Quality of Life: The Development of the World Health Organization Quality of Life Instrument (WHOQOL). Geneva: WHO; 1993.
3. Bowling A. Measuring health: a review of quality of life measurement scales. Fourth Edition. Philadelphia: Open University Press; 1994.
4. Centers for Disease Control and Prevention. Health-related quality of life. Available from URL: http://www.cdc.gov/hrqol/. Accessed: April 3, 2009.
5. Leplege A, Hunt S. The problem of quality of life in medicine. JAMA. 1997;278:47-50.
6. Stewart AL, Ware JE. Measuring Functioning and Well-Being: The Medical Outcomes Study Approach. Durnham, NC: Duke University Press; 1992.
7. Gandek B, Ware JE, editors. Translating functional health and well-being: International Quality of Life Assessment (IQOLA) Project studies of the SF-36 Health Survey. J Clin Epidemiol. 1998; 51:891-1214.
8. World Health Organization. Measuring Quality of Life. Geneva: World Health Organization; 1997.
9. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. JAMA. 1995;273:59-65.
10. Cummins RA. Objective and subjective quality of life: and interactive model. Soc Indicat Res. 2000;52:55-72.

11. Diener E. Subjective wellbeing. Psychological Bulletin. 1984;95:542-75.
12. Andrews FM, Withey SB. Social Indicators of Wellbeing: American's Perceptions of Life Quality. New York: Plennum Press; 1976.
13. Cummins RA. The domains of life satisfaction: an attempt to order chaos. Social indicators research 1996; 38:303-332.
14. Campbell A, Converse PE, Rodgers WL. The Quality of American Life: Perceptions, Evaluations and Satisfactions. New York: Russell Sage, 1976.
15. Cummins RA, Eckersley R, Pallant J, van Vugt J, Misajon R. Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. Soc Indicat Res. 2003;64:159-190.
16. Cummins R. Assessing quality of life. In R.J. Brown, editor. Assessing Quality of Life for People with Psychiatric Disabilities. Cheltenham, England: Stanley Thornes Publishers; 1997.
17. Cummins RA. Personal wellbeing index. Sydney: Australian Center for Quality of Life; 2003. Available from URL: http://acqol.deakin.edu.au/instruments/wellbeing.index.htm. Accessed: April 3, 2009.
18. Pavot W, Diener E. Review of the Satisfaction with Life Scale. Psychol Assess. 1993;5:164-172.
19. Cummins RA, Lau ALD, Stokes M. HRQOL and subjective well-being: noncomplementary forms of outcome measurement. Expert Rev Pharmacoeconom Outcom Res. 2004;4:413-420.
20. Ware JE, Kosinski MA, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. Lincoln: Quality Metric Incorporated, 2003.
21. SF-36.org. A Community for Measuring Health Outcomes using SF Tools [homepage on the Internet]. Available from URL: http://www.sf-36.org/. Accessed: April 3, 2009.
22. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) project. J Clin Epidemiol. 1998;11:903-12.
23. Jureša V, Ivanković D, Vuletić G, Babić-Banaszak A, Srček I, Mastilica M, Budak A. The Croatian Health Survey-SF-36: General quality of life assessment. Colleg Anthropolog 2000;24:69-78.
24. Maslić-Seršić D, Vuletić G. Psychometric evaluation and establishing norms of Croatian SF-36 Health Survey: Framework for subjective health research. Croat Med J. 2006;47:95-102.

RECOMMENDED READINGS

1. Bowling A. Measuring health: a review of quality of life measurement scales. Fourth Edition. Philadelphia: Open University Press; 1994.
2. Møller V, Huschka D, Michalos AC, editors. Barometers of quality of life around the globe: How are we doing? Social Indicators Research Series No.33. Heidelberg: Springer; 2009.
3. Gullone E, Cummins, RA. The universality of subjective well-being indicators. Social Indicators Research Series No.16. Berlin: Kluver academic Publishers; 2002.
4. Cummins RA, Eckersley R, Pallant J, van Vugt J, Misajon R. Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. Soc Indicat Res. 2003;64:159-190.
5. Diener E, Suh EM, Lucas RE, Smith HL. Subjective wellbeing: Three decades of progress. Psychol Bull. 1999;125: 276-302.
6. Centers for Disease Control and Prevention. Measuring Healthy Days. Atlanta, GE: U.S. Department of health and human services, Centers for Disease Control and Prevention; 2000.

7. O'Connor R. Measuring Quality of Life in Health. London: Churchill Livingstone, Elsevier; 2004.

ON-LINE RESOURCES

1. The Australian Centre on Quality of Life. Available from URL: http://acqol.deakin.edu.au/index.htm. Accessed: April 3, 2009.
2. Centers for Disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion. Health-Related Quality of Life. Available from URL: http://www.cdc.gov/hrqol/. Accessed: April 3, 2009.
3. Quality of Life Research (Journal). Available from URL: http://www.springer.com/medicine/journal/11136. Accessed: April 3, 2009.
4. Applied Research in Quality of Life (Journal). Available from URL: http://www.springer.com/social+sciences/quality+of+life+researc/journal/11482. Accessed: April 3, 2009.
5. Journal of Happiness Studies (Journal). Available from URL: http://www.springer.com/social+sciences/quality+of+life+researc/journal/10902. Accessed: April 3, 2009.
6. World Health Organization, Division of Mental Health and Prevention of Substance Abuse. WHOQOL: Measuring quality of life. Geneva: World Health Organization, Division of Mental Health and Prevention of Substance Abuse; 1997. Available from URL: http://www.who.int/mental_health/media/68.pdf. Accessed: April 3, 2009.
7. World Health Organization. WHO Quality of Life-BREF (WHOQOL-BREF). Available from URL: http://www.who.int/substance_abuse/research_tools/whoqolbref/en/. Accessed: April 3, 2009.

Chapter 3

METHODS OF PUBLIC HEALTH INTERVENTIONS

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**THE POTENTIAL OF PUBLIC SERVICE ADVERTISING IN THE FIELD OF PUBLIC HEALTH**
Module: 3.1.1	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Dobriana Sidjimova**, PhD, Assistant Professor Faculty of Public Health, Medical University, Sofia, Bulgaria **Mariana Dyakova,** PhD, Associate Professor Faculty of Public Health, Medical University, Sofia, Bulgaria **Zaharina Savova**, PhD, Associate Professor Faculty of Public Health, Medical University, Sofia, Bulgaria
Address for correspondence	**Dobriana Sidjimova** Faculty of Public Health, Medical University, Sofia, Bulgaria 8 "Bialo more" str., 1527 Sofia, Bulgaria e-mail: dsidjimova@mail.bg
Keywords	Public relations; advertising; social marketing; health information; health campaign
Learning objectives	After completing this module students and public health professionals should: • increase their knowledge for the potential of public service advertising and social marketing in public health; • differentiate public service advertising and PR activities; • increase knowledge of legal framework of advertising activities; • understand the social aspects of medical products' advertising.
Abstract	Advertising is a product of public demands, which determine its engagement in public, political, economic and cultural environment it functions in. From healthcare education and culture point of view the advertising potential should be directed to the promotion of healthy lifestyle, to demonstration of tolerant attitude to people with mental and physical disabilities, to respect and protection of the environment etc. A significant positive effect of the mass popularization of medicinal products via advertising is the increase of the general educational level of consumers.
Teaching methods	Teaching methods include lectures, interactive group discussions, case studies, exercises, internet searches.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on the group work, seminar papers, and case-problem presentations.

THE POTENTIAL OF PUBLIC SERVICE ADVERTISING IN THE FIELD OF PUBLIC HEALTH

Dobriana Sidjimova, Mariana Dyakova, Zaharina Savova

THEORETICAL BACKGROUND

Main functions of advertising

There are many definitions of the term "advertising". They differ depending on the emphasis placed in them, i.e. on the content and form of the advertising message and the target audience. The common parameters that contain in almost every interpretation are related to the fact that advertising is a purposeful activity in the field of the persuasive communication addressed to certain target group, selected based on socio- demographic indicators. The essence of advertising is expressed in introducing or enlarging the knowledge of a certain product, service or idea, formation of positive attitude to it, retaining in consumers' memory and inciting the consumers to undertake certain actions. The messages are spread through the means of communication and funded by a famous source.

Advertising is multifunctional. Some of its main functions are the following:

1. economic function – relates to trade promotion;
2. information function – relates to mass distribution of information about goods, services and ideas;
3. social function – promotes and boosts ideas that are significant for the public, thus forming the public opinion, helps for communication processes, outlines concrete behaviour patterns and promotes certain values and standards;
4. educational functions –in relation to propaganda of novelties in different fields of life;
5. aesthetic function – it is intended to create in the consumers taste for the exquisiteness and beauty. Most of the ads are made by talented designers, artists, directors, copywriters and representatives of other creative professions.

Olivero Toscani defines the advertising as the "new journalism" and on the opinion of Paul Humphrey the advertising should provoke the addressee in intellectual aspect and to make him/her start thinking about problems concerning him/her personally and concerning the society as a whole (1).

According to data from research made by I. Vladimirova (2) for a period of one year the time Bulgarian students spent in front of the TV time equals to 30 days. For students in the age group between 11 and 15 years, the factor of television is the second most important, just after the communication and interaction with their coevals. The time spent in front of the TV set is at the expense of preparation for school, practicing sport and hobbies. Such unhealthy daily lifestyle inevitably reflects on the nourishing habits, level of general knowledge as well as on the language and grammatical knowledge. In this context the role and responsibility of ads which are permanently present on the TV are of

key importance for making children (who, in many cases, are addressees of the messages) healthy, intelligent and harmoniously developed personalities.

Nature and specifics of public service advertising

The definition of public service advertising (or public service announcement) given in the marketing Terms Dictionary of the American Marketing Association is that it is created in order to educate or motivate certain target group with the aim of provoking socially significant behaviour. As part of the social marketing the public service advertising (PSA) strives to educate and motivate the audience to accept or change certain attitudes and behaviour. The appeals could be at local, national or world scope.

PSA is initiated in relation to concrete socially significant problems. Its role is to show the problems and the possible ways to overcome them. However, the personal decision of addressees for getting influenced by the advertising messages remains a matter of one's choice. Here comes the skill of advertising creators to select the necessary facts in order to give arguments for the idea they maintain, to succeed in motivating the recipients or to change their attitudes. That is why the advertising messages dealing with social issues are often personalized and put the emphasis on the role of every single individual in solving given global problem.

The advertising impact on PSA is achieved by applying psychological methods that influence consciously and subconsciously. The audience is provoked through the power of language as the texts contain many epithets, inclusion of scenes from the real life in TV spots, selection of appropriate music, etc.

Every year the American Advertising Federation awards premiums to the best projects in the field of PSA in the Public Service Category.

Public service advertising or public relations

In their book "The Fall of Advertising and the Rise of PR" Al and Laura Ries (3) note that wherever we look we can see the dramatic change from advertising based marketing to PR based marketing. According to the authors the advertising should come after PR not only in relation to time but also in relation to topics. The advertising is a continuation of public relations carried out with other means and it starts when the PR campaign has been exhausted. The process effectiveness is ensured provided that the advertising program repeats the perceptions set by PR. Al and Laura Reis use the metaphor that the advertising cannot light a fire. It only blows it on once it has been lighted (with PR means and methods).

One of the prominent Bulgarian PR specialists G. Ivanov notes that the advertising and PR are both components of a broader function – business communication, and they often overlap and complement each other, but he also outlines some major differences between them (4). We shall try to present these differences in a table (Table 1).

Table 1. Differences between advertising and PR

Advertising	Public relations
Marketing communication function	Management function
It relates to sale of goods and services and is it in accordance with the market conditions.	It generates public understanding and approval of certain organization for whose prosperity it creates favourable environment.
It is carried out mainly through the media.	It makes use of different communication channels: lobbying, organizing special events, sponsorship, press announcement, speeches, presentations and others.
Basic question: „Will it be salable?"	Basic question: „Shall we find new friends?"

CASE STUDY

Health Informational Campaigns in Bulgaria

In the last two years in Bulgaria there is a strong trend for initiating many campaigns promoting healthy lifestyle or explaining the prevention of socially significant diseases. Main channel for broadcasting such type of information is the internet, but in some cases it is only one of the sources which is used together with the printed and electronic media.

Indicative example in this relation is the National informational anti-AIDS campaign with advertising slogan "*Don't say - It could not infect me. Be informed*", which campaign was under the patronage of Ministry of Health, Joint UN Programme and Anti-AIDS Campaign (5). The main aim of this campaign is to inform the public about the possibilities for prevention. One of the most loved Bulgarian actors (V.V. Zueka) was chosen to be the image in the advertising spots and blocks. In his typically comic style he explains the serious problem regarding the transmission of AIDS. In a series of TV spots and printed as Zueka warns the public that: "*You could prevent yourself from transmission of HIV if you always use condoms*", that "*You will not die if you share your food with a person who lives with the HIV virus*" or that "*Mosquito bite does not cause risk of HIV infection"*. The possible ways for transmission of HIV are outlined in the advertising campaign and an appeal is addressed to prevent personal health and to show sympathy to those who live with the virus.

Although most of the PSAs have commercial aims it should be noted that their social effect is significant and therefore the health culture gradually increases.

For example, statistical data for the abortions are presented during the campaign against abortions with the motto "*Abortion leaves invisible scars. It depends on you*!" (6). The medical and psychological consequences of this act are emphasized as well as the low birth rate and the fact that in Bulgaria 270 000 couples in fertility age have reproductive problems and that is serious social problem. There is a link to a site where all interested could find information and this site is part of the national information campaign "*Become a mother when you are ready*", carried out under the aegis of the Bulgarian Society of Obstetrics and Gynaecology". Detailed information is provided about the different methods for prevention of undesired pregnancy, the advantages of using contraceptives. A new

service was launched – receiving free of charge SMS reminding the ladies when to take their contraceptive pill.

In April 2008 national campaign against cervix cancer with the motto "*For you and those you love*" was launched. Every Bulgarian woman could support the campaign by sending SMS with a photo to a given site and thus join the photo-petition in support of the prevention from the insidious disease. On this site detailed and in depth information could be found on the problems related to this disease, its diagnosis and ways of prevention and continuous protection. The revealed data are edited by the chairman of the Bulgarian Cancer Association (7).

The campaign entitled "Because we love life" has a long tradition and is strongly supported by many famous ladies from different fields of public and cultural life in the country. Its main objective is to inform the Bulgarian woman about the necessity to regularly go to examinations and to help for more accessible prevention and thus to help for the early diagnosis of breast cancer which is absolutely treatable if medical help is given in time (8).

Other internet advertising campaign which along with its direct marketing objectives (indirect advertising of a concrete medicinal product), has the task to increase the health culture of public, is the initiative for promotion of respiratory infections (9). Data are presented under different headings about the adults' and kids' flu symptoms. The information is addressed to the ordinary citizens and medical specialists, to managers in order to take measures to reduce the influenza epidemic within the organizations they manage, because according to the statistics one of the main factors for the reduction of labour efficiency is the high level of influenza illness.

Socially significant advertising champagne is also the initiative called "Sport against osteoporosis" aimed at informing about the possibilities for prevention and reduction of the risk from this disease. The advertising messages published in the printed and electronic media are supplemented with free DVD containing advice, diets and sport exercises which could be practiced in the office (10).

Nationwide public service advertising in Bulgaria was implemented through EU and Poland funded campaign for promotion of carrot juices, without taking into account any concrete brand. The main advertising text is "*Drink carrot juices and nectars. That's the way for you to smile… fibres take care of your figure and pro-vitamin A helps you to keep your complexion fresh and bright. Everybody looks at you with admiration!".* In relation to the balanced nutrition an EU platform with motto "Nutrition, physical activity and health" (11) is established as one of the basic components of healthy lifestyle. National and international multi-sectoral actions are planned. Example is that the Union of European beverages associations - UNESDA declared its readiness to refrain from direct advertising and promotion to children under 12 years of age. One of the most famous fast food chains also made a commitment to include information about the food content on the packages of all its products throughout Europe. Fruit and vegetable producers undertake to promote on the European market a logo which encourages the children to consume more fresh products. A lot of educational programmes and events are planned in relation to healthy nutrition and physical activity.

Another favourable trend promoting healthy lifestyle through advertising is the advertising campaign of one Bulgarian pharmaceutical company which has

developed a medicine helping people to give up smoking. The campaign idea is to show in the most influential way the fatal consequences of smoking to human health. The advertising message reads as follows: „*Smoking causes statistics. Black statistics. Out of the statistics: ex smoker*".

Of great social importance is the European campaign "*HELP – For a life without tobacco*" (12), which was launched in Bulgaria in 2007. A multimedia approach is adopted in order the messages to reach maximum addressees. TV advertising campaign was implemented with thematic video clips. Internet advantages are used as web page is developed containing detailed information for health and social problems, useful advice, tests for active and passive smokers, etc. Brochures are handed out with contact details of medicinal cabinet throughout the country where one could get professional aid.

Another social initiative of similar nature is the campaign entitled "*Knowledge saves! Lack of knowledge kills! Be ahead of street lessons!*" (13) and aims at providing young people and their relatives with reliable and professional information about the drug problem and different programs and alternatives for rehabilitation of those who already abuse. Contact details are given of the special aid centres throughout the country – foundations and programs, therapeutic communities, parents' associations and others. This kind of public service advertising is performed through an attractive banner of frequently visited site, which aims at drawing the attention of the mass audience which is asked to show sympathy and to have an active civil stand regarding the problems that drug addicted face.

Positive public response has marketing campaigns that are bound with a cause. Example here is also the initiative of one Bulgarian bank and one foreign pharmaceutical company providing medicines for children suffering from significant growth delay.

Legal framework of advertising activities aiming at reduce smoking and spirits consumption in the Bulgarian society

As a legislative initiative in Bulgaria in the field of advertising, which has a marked social and health effects, we could note the provisions of Law on Health relating to spirits advertising on the TV and radio. Paragraph 1, section 17 and 18 from the Additional provision to the Law on Health stipulates that alcoholic drinks are spirits (liquids intended to consumption containing at least15 vol.% ethyl alcohol") (14). Beer and wine are not considered spirits according to this criterion.

An imperative ban for direct advertising of spirits is laid down in Article 55, paragraph 1 of the Law on Health. According to the text in paragraph 1, section 19 from the Additional Provision to the Law on Health the term "direct advertising" is interpreted as "any form of market message, announcement or recommendation which aims to promote alcoholic drinks and/or the consumption, production or distribution of alcoholic drinks" (14). The objective of direct advertising is to promote the alcoholic drinks as well as their consumption.

In paragraph 1, section 20 from the Additional Provision to the Law on Health the term "indirect advertising" includes "any form of market message, announcement, recommendation or action that uses the name or trademark of alcoholic drink as well as name or trademark of a producer of an alcoholic drink on goods and products that are not alcoholic drinks." The advertising message in indirect advertising is not necessarily addressed at promoting certain drink or its consumption. The indirect advertising of spirits and the advertising of beer and wine should meet the conditions set in article 55, paragraph 2 from the Law on Health which prohibit the advertising messages promoting alcoholic drinks to be addressed to persons under 18 years of age as well as the participation of such persons in ads in the printed and electronic media. Deceptive and manipulative advertising of alcoholic drinks is prohibited. According to article 55, paragraph 3 from the Law on Health the indirect advertising is not allowed to be broadcasted in the TV and radio channels before 10.00PM (14).

These legislative measures aim to restrict the unfavourable downward trend in the age limit of persons consuming alcohol, to reduce traumas and death caused by road accidents, to reduce the negative effects on the economy, to increase the public awareness on problems related to alcohol abuse.

According to the European Commission data 23 million European citizens are addicted and the alcoholism costs Europe 125 billion EURO or 1.3% of the European Gross Domestic Product. Alcoholism is the main reason for 7.5% of diseases and untimely death in Europe. For young people between 15 and 29 years of age these figures are respectively 10% for women and 25% for men. Every ninth teenager tried alcohol for the first time at the age of 12.5 years. The last Eurobarometer research shows that 19% of young people (in the age group 15-24) in the EU drink five or more glasses of alcohol every time they sit at the table. Alcohol is one of the main reasons for the violence and road accidents. According to the Health Commissioner Markos Kyprianou the media, advertisers, sellers and owners of clubs and food and beverage establishments should contribute to change the attitude and behaviour of young people. We could not afford to lose so many young human lives every year because of alcohol abuse (15).

The International Trade Chamber has drafted many international codesfor marketing and advertising self control in relation to advertising of alcohol and tobacco that is intentionally addressed to children and popularized via the internet. These codes are implemented at national level by national self regulating bodies and contain recommendations not to incite children to activities that could seriously harm their health. An organization called European Advertising Standard Alliance operates in Europe and it promotes different initiatives for regulation of non ethic advertising directed to children. The Internet site of this organization contains practical information on how to proceed if laws and self regulating rules are broken at national level in relation to advertising of goods for children (16).

The following initiatives may be pointed out as policies implemented in Bulgaria for reducing another serious bad habit, and namely smoking:

- on the November 07, 2005, Bulgaria ratified the Framework Convention of Tobacco Control and took part in the work of the First Conference of Signatories to the Convention, held in February, 2006 in Geneva,

- since of the January 01, 2006, the cigarette excise in Bulgaria is increased thus increasing the price of cigarettes with more than 40%. As a result we could note 31% reduction in consumption in March compared to the same period during 2005,
- the Ordinance for the terms and conditions allowing, as an exception, smoking in certain areas in indoor public establishments and indoor working premises enters into force on the January 01, 2005, adopted by the Council of Ministers' Decree № 329 of 8th of December 2004,
- the Law on tobacco and tobacco products stipulates the maximum admissible concentration of harmful substances in cigarettes and until the December 31, 2010, the contents of tar in a cigarette will reach the level of 10 mg,
- with the Ordinance for the requirements for labelling, marking and appearance of tobacco products and for establishing standards for compliance assessment of the contents of harmful substances in cigarettes, the requirements for obligatory and additional notices on the packaging of tobacco products enter into force since of the January 01, 2005,
- in 2005 the Law for Health was enforced, which stipulates that the Minister of Health and other competent governmental authorities, together with nongovernmental organizations, are responsible to establish conditions for reduction of smoking, alcohol abuse and not allowing use of narcotic substances. Article 53, paragraph 3 of the Law provides that one percent of the funds in the national budget from excises on tobacco products and spirits, shall be used for funding of national programs for reduction of smoking, alcohol abuse and not allowing the use of narcotic substances. This covenant of the law enforced on the January 01, 2006, provides the opportunity to actually provide sufficient funds for the implementation of the objectives and tasks of the three programs, one of them being the reduction of smoking (14).

Advertising – a mean to increase the audience's dental culture

The power of advertising effect is also used in the field of dental medicine. TV and radio spots, printed media blocks, i.e. the advertising through all kinds of information distribution channels are full with plenty of ads appealing for maintaining high level of mouth hygiene. This type of ads is one of the sources to increase the health culture of society on dental health issues. Via everyday multiple repetition of one of the most popular ads of specific toothpaste in the primetime, the Bulgarian consumer is already acquainted with the seven signs of strong teeth.

The role of advertising for formation of environmental awareness of people

Another topic of advertising, which is directly related to individual's and the society's health as a whole, is the advertising activities in the field of environment. In Bulgaria, the campaigns initiated by the world organization World Wildlife Fund

(WWF) are actively promoted. Several campaigns intended directly for the Bulgarian audience can be distinguished. The slogan *"A single piece of garbage does not pollute – You thing, so as 7,720,000 Bulgarians do"* – appeals to people's awareness to deposit their wastes to the relevant places. In 2007 the campaign entitled "*For Sale – Protect Protected Areas"* was awarded the golden prize of the first Bulgarian competition for advertising efficiency, Effie. The advertising concept finds the following expressions: outstanding panorama views from the mountains of Strandzha, Rila and Pirin are presented, with plate *"For Sale"* in the foreground, accompanied with different parameters of the mountains for the purpose to associate natural wealth with real estates (17). At present, the next action of WWF with motto "*Climate anomalies happen here and now. We can stop them"* is being popularized across the World Wide Web. The visual design of this advertising is presented by means of analogy with the roulette game, as instead of numbers, the high temperatures as a result of global warm up are written on the table pot. This advertising technique is aimed to demonstrate that society neglects environmental problems and does not consider them serious.

The advertising campaign launched by an organization active in the establishment of sustainable systems for divided deposit and use of packaging wastes also has significant contribution for the increase of health-environmental culture of our nation. Attractive video spots and printed ads appeal for the formation of environmental awareness and behaviour by means of demonstrating the advantages of divided disposal of wastes. Specific facts are pointed out, for example the recycling of one glass bottle ensures 4 hours of electricity, thus appealing the addressees to support the process of divided disposal of everyday wastes made of glass.

Social aspects of medical products' advertising

Educational function of medicinal products' advertising

A significant positive effect of the mass popularization of medicinal products via advertising is the increase of the general educational level of consumers. The significant use of terminology is determined by the nature of advertised products. In this case the main components, the active substances of the medicine, as well as the names of different types of diseases. As a result, the Bulgarian consumer deals with the medical and pharmaceutical terms better and better.

Manipulative tactics for production and promotion of messages promoting medicinal products

The main purpose of corporate supply of medicines is to increase profit through encouraging consumer demand of medicinal products. Thus unethical methods for advertising thereof are encouraged. Pharmaceutical companies give the fact that they provide valuable information and increase the healthcare culture of people, as the only argument of theirs. But in practice, encouraging the irrational use of medicines may adversely affect the health.

Manipulation is mainly applied by means of misleading or incorrect statements about medicinal products or by means of deliberate non disclosure of

specific risk factors or side effects of some medicines. Widely used approach applied by the pharmaceutical companies is the sponsorship of patients' organizations, on one hand, and on the other – provision of financial or material incentives for the physicians to prescribe the medicines manufactured by the relevant company, at the expense of their competitors. The websites with detailed information for medicines, which do not have authorization for advertising, are essential channel for distribution of information. From them the mass consumer obtains information which is not attested and has biased character. Another important adverse trend is the financing of health campaigns for popularization of medicinal products, instead of investing the funds in promotion of prophylaxis and healthy lifestyle.

The advertising market of medicinal products

TV advertising is still one of the most expensive advertising channels worldwide. In relation to medicines, until 1997 manufacturers were prohibited to announce medicinal products and their application in TV spots, as the mass effect of the message distributed via television was taken in consideration. At present, there are no such limitations and according to the monitoring of TV advertising carried out by Media Links Agency (18), both during the first quarter of 2006 and in 2007 the product categories with the highest advertising budgets at the Bulgarian TV advertising market, are for the telecommunication, cosmetics, pharmaceutical, beverages and washing and cleaning detergents sectors.

What is most characteristic for the category of medicinal products is the relatively larger number of advertisers in comparison with the other categories. The advertisers from the pharmaceutical sector have increased their activeness and thus such category is now among the top 5 of TV advertising budgets for the first quarter of 2007. The existing situation directly results in 50% of the patients in Bulgaria to go to the drugstores rather than to the doctor (18).

This trend is observed globally being also confirmed by the analysis carried out by a group of American scholars published in Fitness and Wellness Week (19). The authors of the research found out that the advertisement of medicines does not only popularize and recognize the relevant medicinal, but also encourage its use. To the researchers' opinion, the fact that patients have clearer and clearer idea about the application of different medicinal products makes them confident to apply self treatment during the first symptoms when they even can avoid the use of medicines.

The consumption of medicinal products without doctor's prescription is also encouraged by the fact that there are no statutory restrictions for advertising of drugstores which attract their clients through spots with the cheerful pharmacist who plays the role of confident friend and healthcare adviser and offers quick and efficient solution for the health problems of the client who has visited the polished drugstore.

Within the existing unfavourable situation, not only legislation, but also the civil liability of advertising creators, has to regulate the market of medicinal products, and restrict the use of medicines without prescription. We already witness the first attempts in this direction. This is the example with message of drugstore chain – *"Because we sell medicines, we know that they are not the single remedy".*

Attempts for restriction of unethical advertising of medicinal products worldwide

The World Health Organization has drafted global criteria for assessment of unethical offering of medicinal products, but they are rather a wish than an obligation thus such criteria does not cause any legal consequences.

The European Federation of Pharmaceutical Industries and Associations (EFPIA) is the representative authority of the pharmaceutical industry in Europe. The associations of local national industries of more than twenty European counties manufacturing medicinal products, and more than forty pharmaceutical companies are members of EFPIA. Realizing the importance of providing correct, just and objective information about medicinal products thus to take reasonable decisions for their use, EFPIA adopted a Code of Practices for Promotion of Medicines (Code of EFPIA) (20). It is aimed at maintaining an environment where the society can be sure that the choice in relation to medicines is made on the basis of the merits of each medicine and the health demands of each patient. The Code is fully compliant with the Council's Directive 2004/27/EC Relating to Medicinal Products for Human Use in the European Union Member States.

In the USA, Australia and Canada (21,22), medicines promotion is controlled on daily basis by means of national associations of manufacturers. During this process, mainly five aspects are analyzed related to the mechanism of recognition of aggressive behaviour, of responsible commissions, of sanctions in case of offences, of data about qualitative and quantitative accumulation of complaints, of procedure for complaints review.

In the United Kingdom the Code for Best Practices Application in the Field of Medicinal Products Promotion has been enforced since 1958. Pursuant to this Code a commission is established with the following members: an independent lawyer, 12 representatives of companies and two doctors, 1 representative of the patients' association. Since 1985 another member is one independent medical consultant who has the obligation to study the advertising blocks in printed materials. The conclusions of the commission are only recommendations (21).

The manufacturing companies in the pharmaceutical sector also take active position in the establishment of ethical and legal standards for the promotion of medicines. At present, in our country an Code of Ethics of the Association of Research-based Pharmaceutical Manufacturers in Bulgaria is enforced, which is based on the provisions of the Code of the European Federation of Pharmaceutical Industries and Associations, and the Code of the International Federation of Pharmaceutical Manufacturers and Associations for Pharmaceutical Marketing Practices providing also the advertising activities in the field of medicinal products (22-24).

The importance of advertising of medicinal products cuts both ways – on one hand, it educates the addressee in relation to the market of medicines, their variety and functions, contributes for the formation of consumers' culture in the field of pharmaceutical industry, but on the other hand, it leaves misleading impression in people who need treatment, that they are competent to take decision for their health on their own, neglecting the professional diagnosis and treatment to overcome the relevant health problem. Namely this specification of the medicinal products'

advertising contains our statement that it has ambiguous effect on the overall picture of public health.

From healthcare education and culture point of view the advertising potential should be directed to the promotion of healthy lifestyle, to demonstration of tolerant attitude to people with mental and physical disabilities, to respect and protection of the environment, etc.

Advertising is a product of public demands, which determine its engagement in public-political, economic and cultural environment it functions in. Advertising is essential stimulating and regulating factor of the impact on the general audience's mind and behaviour, and these specific features of it, in addition to its mass character, make advertising one of the major elements in the process of health though the creation of purposeful messages for different target groups of the population, and through demonstration of specific behaviour patterns.

Via public service advertising people's attitudes to a specific socially significant problem are changed, new values are being formed, which afterwards turn to be the basis for responsible behaviour to ourselves, to the other, and to the surrounding environment as a whole. Some of the most effective and affecting public service advertisings are related to the public health issues, thus their potential must be used to the highest extent in order to popularize the healthy lifestyle and the prevention of socially significant diseases.

EXERCISE

Task 1

Trainees must specify an example of successfully realized public service advertising promoting the prevention of specific socially significant disease.

Task 2

Trainees should compose a message for a problem relevant to the public health. They should develop the advertising arguments corresponding to the relevant target group set up on the basis of social demographic principle.

Task 3

Develop overall strategy for model campaign, including PR and advertising to promote healthy lifestyle.

REFERENCES

1. Jekov S. Only one world is not enough. Media & Advertising. 2001;21:18-20.
2. Vladimirova I. The media and the formation of aggressive behavioural inclination in the young. Autoreferat. Sofia; 2006.
3. Ries A, Ries L. The Fall of Advertising and the Rise of PR. New York: HarperCollins; 2004.
4. Ivanov G. Public relations and trust formation. Sofia: UNESCO; 1996.

5. UNAIDS. National informational anti AIDS campaign. Available from URL: http://www.unaids-bulgaria.org/index.php?magic=2.58.0.0. Accessed: May 20, 2008.
6. Abortion leaves invisible scars. It depends on you! [homepage on the Internet]. Available from URL: http://www.antiabort.info. Accessed: May 20, 2008.
7. For you and those you love [homepage on the Internet]. Available from URL: www.zateb.info. Accessed: May 20, 2008.
8. Because we love life [homepage on the Internet]. Available from URL: www.otlubov.com. Accessed: May 20, 2008.
9. Stop the flue! [homepage on the Internet]. Available from URL: http://www.sprigripa.info. Accessed: May 20, 2008.
10. Sport against osteoporosis [homepage on the Internet]. Available from: http://www.wcif-bg.org/news/0/159/New-news. Accessed: May 20, 2008.
11. Barrington-Leach L., M. Canoy, A. Hubert, F. Lerais. Bureau of European Policy Advisers (BEPA). Investing in youth: an empowerment strategy. Luxembourg: European Commission, Bureau of European Policy Advisers (BEPA), 2007. Available from URL: http://ec.europa.eu/dgs/policy_advisers/publications/docs/Investing_in_Youth_25_April_fin.pdf. Accessed: May 20, 2008.
12. HELP – For a life without tobacco [homepage on the Internet]. Available from URL: http://www.help-eu.com. Accessed: May 20, 2008.
13. Knowledge saves! Lack of knowledge kills! Be ahead of street lessons! [homepage on the Internet]. Available from URL: http://www.initiativeforhealth.org. Accessed: May 20, 2008.
14. Health Act. Newspaper of the state. Vol. 70, 10.08.2004.
15. Alcohol and Health Forum [homepage on the Internet]. Available from URL: http://www.eurocare.org/resources/special_topics/alcohol_and_health_forum. Accessed: May 20, 2008.
16. Advertising Education Forum (AEF) [homepage on the Internet]. Available from: www.aeforum.org. Accessed: May 20, 2008.
17. World Wildlife Fund [homepage on the Internet]. Available from URL: http://www.panda.org/bg/get_involvedd/campaign/index.cfm. Accessed: May 20, 2008.
18. Angelov L. The advertising stresses on cosmetics, pharmacy and beverages. Insight. 2007;18:23.
19. Anonymous. The advertising sells "unnecessary" medicines. Dnevnik. 2006;15:16.
20. Code of practices for promotion of medicines (Code of EFPIA). Available from URL: http://www.efpia.eu/Content/Default.asp?PageID=296. Accessed: May 20, 2008.
21. Mitkov T, Petrova G. Measuring the results from pharmaceutical promotion from pharmaceutical companies. 2008;2:90-97.
22. Mintzes B., Bassett, K., Lexchin J. An assessment of the health system impact of direct-to-consumer advertising of prescription medicines (DTCA). Volume IV: pills, persuasion and public health policies. Report of an expert survey on direct-to-consumer advertising of prescription drugs in Canada, the United States and New Zealand. Vancouver: Center for Health Services&Policy Research; 2001.
23. Code of Ethics of the research-based pharmaceutical industry in Bulgaria. Available from URL: http://www.arpharm.org. Accessed: May 20, 2008.
24. European Federation of Pharmaceutical Industries and Associations (EFPIA). european code of practice for the promotion of medcines. 2004 edition. Available from URL: http://www.efpia.eu/Objects/2/Files/Promomedicines2004.pdf. Accessed: May 20, 2008.

25. International Federation of Pharmaceutical Manufacturers and Associations (IFPMA). Code of pharmaceutical marketing practices. Revision 2006. Available from URL: http://www.ifpma.org/pdf/IFPMA-TheCode-FinalVersion-30May2006-EN.pdf. Accessed: May 20, 2008.

RECOMMENDED READINGS

1. Advertising Standards Authority website [homepage on the Internet]. Available from URL: www.asa.org.uk. Accessed: May 20, 2008.
2. Amended by Directive 2004/27/EC of the European Parliament and of the Council of 31 March 2004 amending Directive 2001/83/EC on the Community code relating to medicinal products for human use, Official Journal L 136, 30/4/2004 p. 34–57. Available from URL: http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/homev1.htm. Accessed: May 20, 2008.
3. Bass F. A Simultaneous Equation Regression Study of Advertising and Sales of Cigarettes. J Marketing Research, 1969;6:291–300.
4. Code of Ethics of the research-based pharmaceutical industry in Bulgaria. Available from URL: http://www.arpharm.org. Accessed: May 20, 2008.
5. Consolidated Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use as amended by Directive 2002/98/EC, Directive 2004/24/EC and Directive 2004/27/EC. Corrigendum (Official Journal L 235, 30/8/2006 p. 24 HU). Available from URL*:* http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/homev1.htm. Accessed: May 20, 2008.
6. Council Directive 92/28 of 31 March 1992 on the advertisement of Medicinal Product for Human Use. Official Journal L 113, 30/04/1992 p. 0013–0018. *Available from URL:* http://www.ikev.org/docs/eu/392L0028.pdf. *Accessed: May 20, 2008.*
7. Dikshit R, Dikshit N. Commercial source of drug information: comparison between the United Kingdom and India. Br Med J. 1994;309:990-1.
8. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use, Official Journal L 311, 28/11/2001 p. 67–128. *Available from URL:* http://ec.europa.eu/enterprise/pharmaceuticals/eudralex/homev1.htm. Accessed: May 20, 2008.
9. Dodd CH. Dynamics of intercultural communication. Boston, MS: McGraw Hill; 1998.
10. Fornell C. A national customer satisfaction barometer: The swedish experience. J Marketing. 1992;56:1-21.
11. ICC International Code of Advertising Practice (1997 Edition) Document No. 240/381 Rev. Available from URL: http://actrav.itcilo.org/actrav-english/telearn/global/ilo/guide/iccadv.htm. Accessed: May 20, 2008.
12. Jedidi K, Mela C, Gupta S. Managing advertising and promotion for long-run profitability. Marketing Science. 1999;18:1-22.
13. Little J, Lodish L. A media planning calculus. Operations Research. 1969;17:1-35.
14. Mooij M. Global marketing and advertising: understanding cultural paradoxes. Thousand Oaks, CA: Sage Publications, Inc.; 1998.
15. Zaltman G., R. Higie. Seeing the voice of the consumer: Metaphor based advertising research. J Advert Research. 1995;35:35-5.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**COMMUNICATION AND BEHAVIOUR IN DENTAL PRACTICES**
Module: 3.1.2	**ECTS (suggested): 0.25**
Author(s), degrees, institution(s)	**Zaharina Savova**, MD, PhD, Associate Professor Faculty of Public Health, Medical University, Sofia, Bulgaria **Dobriana Sidjimova**, MD, PhD, Assistant Professor Faculty of Public Health, Medical University, Sofia, Bulgaria
Address for correspondence	**Dobriana Sidjimova** Faculty of Public Health, Medical University, Sofia, Bulgaria 8 "Bialo more" str., 1527 Sofia, Bulgaria e-mail: dsidjimova@mail.bg
Keywords	Dental practice, interaction, communication, psychological behavior models, verbal techniques, non-verbal techniques, fear, anxiety, dental phobia
Learning objectives	After completing this module students should: • gain command of knowledge and skills concerning dentist-patient communication; • recognize models of dentist-patient relationship; • increase their knowledge in types of patients, psychological techniques and methods; • understand and differentiate the approach towards children and elderly people who have dental phobia; • distinguish fear from anxiety, coming from dental phobia and find the appropriate approach; • improve their practical communication skills for work with patients.
Abstract	Contemporary expectations of the patient from the stomatologist include expectations of professional and communication skills, and individual approach to the patient. Development of stomatologist's communication qualities could help to establish confidence, certainty and respect towards the therapeutic process as well as lead to a successful treatment. An individual does not visit the office only because of one's disease, but also with one's worries, fears, expectations, imaginations already accomplished, etc. The patient is not a diagnosis but a personality. The module is describing models of the patient–dentist relationship; characteristics of patient's behaviour; psychological techniques and methods in the dental practice.
Teaching methods	Lectures, interactive methods such as small focus group discussions, case studies.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment should be based on the group-work, seminar papers, and case-problem presentations.

COMMUNICATION AND BEHAVIOUR IN DENTAL PRACTICES

Zaharina Savova, Dobriana Sidjimova

THEORETICAL BACKGROUND

The contemporary understanding and expectations of the patient towards the dentist include not only expectations for the availability of professional treatment skills but also possession of communication and individual approach to the patient. The creation of dentist's communication qualities could help to establish confidence, certainty and respect towards the therapeutic process, and so would lead to a successful treatment. The individual does not visit the office only because of one's disease, but also with carrying his worries, fears, expectations, imaginations already accomplished, etc. The patient is not a diagnosis but personality.

The mastering of communication skills requires psychological preparation, which includes: knowledge of different models of the relationship patient- dentist, acknowledgement with the approaches for most effective communication between the dentist and the patient – attaining proficiency in verbal and non-verbal communication techniques and methods, and also distinguishing the different types of personalities with all their specific behaviours and the relevant suitable ways of influence, psychological and cognitive-behaviour approaches and techniques.

Models of the patient– dentist relationship

There are four models of patient-dentist relationship:

1. Model of agreement (Parsons 1951), (1)
   The individual accepts the role of the sick. This allows him to avoid any social responsibilities but yet he requires from the odontologist to confirm that he is seriously "ill". Such role would give him the possibility to divagate from social life and responsibilities (for example, a patient with terrible toothache; patient who decides that his jaw is unusually big and looks out of shape). Such patients could behave in a very passive manner (this model is appropriate for the respectable role of the professional).
2. Model of mutual dependency.
   Shazz and Hollander (1956), (1) upgrade Parson's model. They describe three basic models of the doctor-patient interaction, which are defined from the patient sensations, related with his disease. These three types of relationships vary depending on the level of stomatologist's participation in the process of understanding the patient's embarrassing situation:
   - activity-passivity type relates to patients, who cannot participate actively in interaction because they are either in status of unconsciousness or are not able to realize interaction (due to disabilities);
   - leading-cooperation type relates to the forms of heavy diseases;
   - mutual participation type relates to the forms of chronic diseases.

3. Model of conflict – Fridson (1970), (1).
   While monitoring the interaction between doctor and patient he found out a certain level of conflict. He rejected the assumption that the professional is always the active part and the patient – the passive. He stressed his attention to the effect that the difference of the cultural environment of the professional and the one of the patient could have on the interaction.
4. Model of conflict and control (1).
   The latest tendencies in the theory of behaviour take into account and unite the models of Shazz and Hollander, and Fridson. Thus Steward and Rotter (1989), (1) developed the model of conflict and control. This model takes into account cultural and other factors, and also the level and type of the disease. It refers to four different types of relations: paternal, mutual, customer's type, and type, where the relation is absent. Reciprocity is an aspect of Steward and Rotter's model and it is very important for the reduction of conflict and the proportional division of the control between the dentist and the patient. The lack of balancing of the control usually leads to unsatisfactory final results whereas referred to the paternal and customer's type.

Many patients avoid taking responsibility during the treatment process but yet this does not lead to higher effectiveness. Therefore, in order to have good results from treatment, it is necessary that both parties take control and responsibility.

Characteristics of patient's behaviour

The psychological reactions of the patient in the dentist's office are defined from a number of background factors:

- past experience from previous visits in the dentist's office and the reason for the current visit;
- attitude towards his own health, conviction in the importance of the oral health;
- life style, availability of stress and conflicts in the routine regime;
- socio-economic and cultural environment.

Influences to the good results of the treatment have not only the professionalisms and the skills of the doctor, but also the patient's behaviour. There are several specific typical peculiarities of patient's behaviour and his attitude towards treatment.

1. Activity on behalf of the patient is found supportive to the therapeutic process. The behaviour of the "active patients" is characterized with the questions they ask, explanations they seek, preferences they state and also their requirements for attention and understanding from the professional.
2. There are patients who hamper the therapeutic process with their behaviour. It is hard to establish communication with them because they are reticent (or too talkative), close, distrustful, have preliminary created opinions and understanding on the oral health.

3. ”Difficult patients are those, who never follow advices and instructions, have variable temper, those who come strongly anxious, frightened and practically it is impossible to establish contact, especially with people who suffer from dental phobia.
4. There are patients who set up requirements, which from dentist’s point of view are groundless and unnecessary. The doctor shall be able to refuse steady and persuasively, from professional position.
5. In situations when the patient is angry, furious and aggressive it is important that dentist remains calm and not be miss leaded by patient. He shall not feel offended by the patient and shall let him leave when he decides. If possible, it is appropriate to clarify the situation, and dissatisfaction and reaction that appeared.
6. While working with patients who consider them selves ugly-looking and disfigured, personnel shall be extremely delicate with their feelings. Communication difficulties are possible due to their low self-confidence or availability of depressive symptoms. Bradbury indicates several requirements for the medical personnel communication with such patients:
    - personal warmth;
    - self-confidence;
    - spontaneity;
    - unconditional positive support;
    - involvement;
    - sincerity;
    - not taking defensive position;
7. Behaviour of patients, who cooperate to treatment, manifests in several ways:
    - they have high health care culture and are responsible to their health, do not underestimate the necessity of prevention examinations and timely treatment;
    - they can be distinguished by their strict and conscious attitude and observation of the recommended by the dentist measures and behaviour;
    - thy are precise concerning their visits;
    - they show respect and good approach to the dentist;
    - they react adequately even in the norm of paid and discomfort;

The child as a patient

There are several behaviour characteristics of the work with children that support the work of the dentist and the process of treatment. Children, especially younger ones use to interpret some sensations in a more different way, frequently enduing a magic meaning to procedures and actions. Real world for them is still connected to stories and fantasies. Sometimes they can accept that the visit to the dentist is a kind of punishment for something. Especially important are the first visits because they can have a lasting imprint to their attitude to the dental treatment. On another hand, this is the time when the basis of a good health protection culture, in particular dental health, can be laid down. It is very important to demonstrate positive support during each stage of the

treatment. The frequent use of praises stimulates the child and creates correct behaviour. If the behaviour of the child impedes the treatment or it is dangerous for the child then very strictly formulated oral instructions shall be given. The child shall be explained what exactly it has to do and how it shall behave. Frequent and well phrased praises and instructions can bring good results and can encourage the child to get used to dental treatment. Tests for the loudness and the strength of the voice show that moderate voice is the most influent while working with children. Studies of Mellamed and Siegel (1) show that when treating children who feel fear and process runs difficult, showing movies with children in similar situation who manage their fear and finalize the treatment process, can be helpful. There are studies, which show that using pharmacological support when treating children is also helping the successful implementation of the process. Work with children requires more attention, patience and predisposal before the start and during the treatment.

Adult and aged people as patients

The percentage of aged people has increased not only in Bulgaria but also worldwide. There is a tendency that older people more rarely use to search dental help and take less care for their dental health. Reasons are different.

Ageing is characterized with specific physical and psychical changes that sometimes hamper the aged person to search for dental help. For example worsened health condition and difficult mobility could be the reason not to leave his home and practically going to the dentist's office is prevented. It is observed that some aged people do not feel worried about the status of their teeth. This observation is very frequent when there is a disturbed psychic state like depression, loss of cognitive abilities, dementia disease, Alzheimer, conditions which influence the ability to take care of teeth. Some chronic diseases could also be obstacle for maintaining labium hygiene or taking care of their dentures and lead to negligent attitude to teeth. Poor economic conditions of aged people in Bulgaria do not allow them to visit dental offices and pay for health care. Communication with aged people could be hampered due to their unsatisfactory health and psychic health as well as due to some suspicious attitude towards the change of the physician, the office and others. And last but not least, a stereotypical, sometimes discriminating attitude to older people and their capabilities and needs, exists sometimes. Due to one or another reason, aged and old people experience real difficulties in accessing dental care and health.

One of the basic psychological problems in dental practice and work with aged and old people is related with the loss of teeth and the functional limitations, which appear later on. Patients, who preserved their own natural teeth until very late age, may suffer very deep from the perspective to lose their teeth. Therefore, it is important for the dentist how he shall announce the news. Lloyd & Bor (1) propose a scheme for presenting bad news:

- firstly he should check what knows the patient;
- he shall listen to the patient and shall present him information;
- he shall evaluate patient's potential;
- he shall invite him for feed back information and shall revise patient's worries;

- he shall give realistic hopes;
- they shall book the next visit to the office.

Dental practice of work with aged and old people has its characteristics and therefore an individual approach to each patient is necessary. The main line of behaviour shall be guided by the idea that aged people have the right to live a life of full value, good attitude, attention and respect.

Verbal and non-verbal communication techniques

During the process of communication each individual reaches some aims, lead by particular motives. Intercourse as a process contains the communication as an element.

Communication is unity of three aspects: communicative (exchange of information); interactive (reciprocal action); perceptive (interpersonal perception). The means of communication are verbal and non-verbal.

The patients who look for dental aid usually experience a necessity for information and understanding of their emotional state. The abilities of the dentist to understand the patient's anxieties and problems and his capacities to explain them the provided procedures and treatment, support the implementation of maximally effective dental practice.

Gary Humphries and Margaret Lynch (1) propose a model of psychological care, which divides tasks into two categories: informational need and emotional needs.

Informational need – these necessities concern both patient and dentist t. Firstly it should be started with general questions, used to prepare the patient. Patient is encouraged to share how he feels, what the questions he is exited about are and what does he want to understand about the treatment and procedures; he is questioned whether any problems appeared since the last visit. Patient is asked whether he understands dentist's instructions and directions and he is proposed to choose certain approach during the process, clarifying if there are additional questions or need for information.

Emotional needs – dentist is interested whether the patient had any emotional anxieties before the examination or the treatment and replays have to be listened carefully; patient is asked how he feels about some specific procedures, dentist makes proposals, which could help the patient; he proposes ways for management of the emotional reactions and during the process looks for specific verbal signs related to manipulations; before the patient leaves the dental office his feelings concerning the visit are checked and questions about his next visit are to be asked, finally, dentist thanks to the patient for his visit.

Personal qualities and communication skills of dentist

Personal qualities and communication skills of a dentist are as follows:

- to possess empathic behaviour;
- to show warmth and amicability during the communication process;
- to have neat and clean outlook;
- not to look tired, bored or angry;

- to be able to interpret and assess the behaviour and the messages of the patient;
- if the patient (not depending on his age) would feel more comfortable if his relative presents in the office, dentist shall agree;
- to take into consideration the differences, resulting from the age, education, communicational culture of the patient;
- not to show impatience when patient wills to share how he feels or when needs more clarifications and information. This is a very specific feature for the aged patients;
- ability to motivate the patient to cooperate during the treatment and encourage him to improve his dental care culture;
- ability to listen - dentist shall be capable to listen actively and carefully the patient and monitor not only patient's verbal, but also non-verbal behaviour; he shall not accelerate his responses and conclusions;
- he shall not tolerate discussions, which do not concern the patient;
- he shall be skilled in making jokes and good humour in order to brighten the atmosphere;
- when detecting that patient breaths fast and perfunctory - the signal that patient is frightened, he shall calm him and then continue his work. Several times a slow and calm breathing (counting to five or six when inspiring and expiring) is very appropriate for patient's relaxation.

Verbal skills and techniques

Some analyses in behaviour sciences show that the impact of the message to the person against us is distributed as follows: words have 7% to 10% of the total impact; sounds - 20% to 30%, and body language from 60% to 80% of the total impact (2).

Verbal communication is realized through words (written or spoken). It is of high importance what is the voice that dentist speaks with and what are his skills to express correctly and clearly. The style of the speech, sonority and temp of speaking, his expressiveness and intonation, the pauses are very important. The influence on communication starts in the very first seconds of the meeting, and therefore it is important:

- to welcome the patient with cheerful smile;
- that dentist refers to the patient using his name;
- to chose the appropriate words when talking to the patient – without medicine terminology that could provoke anxiety or could be misunderstood;
- to use words, which do not frighten, for example instead of pain use discomfort, words like cut, bleed, blood aspiration, boring, etc., shall be avoided;
- to speak calmly and without panic;
- to master his intonation because it could change the message.

Dentist shall take in consideration the fact that strongly worried patients usually remember a poor volume of information. Therefore the important issues shall be communicated in the beginning and the end of the meeting, after attracting patient's attention. It is important to check whether the patient has understood correctly dentist's

directions, but is it also of great importance to confirm that the doctor himself has correctly understood the patient. So, unclear information could be corrected.

Non-verbal techniques in communication

Each professional should know that if there is no cross point of the verbal message and non-verbal behaviour, the patients feel anxiety, mistrust and confusion. The strength of the non-verbal message is sometimes stronger and with wider impact than the words to be said. Non-verbal communication encompasses:

1. Gestures and facial expressions.
   Through the non-verbal behaviour – gestures and facial expressions, messages and information might be exchanged in moments when verbal communication is impossible, for example: patient and dentist agree on giving signs with his hand during some manipulations, and this gives a significant relieve because the patient knows that he would be understood without talking and have the conviction that he can control the situation.
2. Touch.
   Touch can be used as means to calm the patient and also as a controlling tool.
3. The body position
   The body position (pose) is of great importance. The trust is created when the dentist is turned to the patient with his face and body, he looks straight into his eyes when speaking to him and he carefully monitors the patient's reactions. Thus, some reactions of the patient can be identified and also to transmit warmth, care and natural behaviour, to attest an attention. Closeness and distantness is transmitted when we speak with hands crossed before the chest while turned to another direction and our look is abstracted.
4. Facial expression.
   Facial expression confirms or not what dentist has said. The patient watches the doctor in order to receive more information. Facial expressions can calm, confirm, and turn attention to something important. Eyes emit not only the patient's, but also dentist's emotions. Anxiety, annoyance and indifference can be easily felt.
5. Physical proximity.
   The work of the dentist takes place in close physical proximity to the patient. During the dental process there is a physical contact, which in many cases is being rejected by the patients and which arise behaviour of resistance. People are different and there are patient which permit the proximity and physical attendance extremely hard. Therefore empathy, warmth and trust are exceptionally important for the work of the dentists. These are the ways to overcome the barriers to the proximity.
6. Cloths.
   Cloths cover 90% of the body and have a mighty impact on the way the other people evaluate the reliability, respect, social success, expertise and status of the individual. The outlook of the doctors has always been considered important for communications. It is approved as a custom form of respect to the patient and expression of self-respect.

7. Environment in the office.
   Environment in the office has strong impact as well. It includes not only the mandatory strict hygiene, comfortable furniture, paintings, health educating materials. Psychological impact of the colours, music and aromatherapy has been proven as well. Appropriately chosen and combined, they can have calming and relaxing influence and decrease the pressure and the fear.

Fear, anxiety and dental phobia

Notwithstanding the latest improvement of dental techniques, for a great number of people the necessity of dental treatment inspires fear. Fear and anxiety are assisting emotional conditions for significant percentage of patients. Those were dictated by the unknown and the expectation for pain during the treatment, or fear from physical damage or injury, fear to lose control, fear of helpfulness and dependency. Sometimes, depending on the momentary emotional state people are more or less sensitive. Fear and anxiety as experience are different from dental phobia because they are provoked by real source. Fear and anxiety can be influenced by passed experience, monitoring of someone else's reactions or rumours, some people are even afraid of instruments. The reasons for the appearance of the fear could be provoked by traumatic experiences, single events, information from relatives and friends, media information, individuality of the patient.

The fear of dental manipulation could be related to the expectation for loss of teeth, surgery intervention, noise, unpleasant smell, needles, and penetration in tissues. Reactions of fear and anxiety are different – emotional (tormenting feeling of tension and worry), vegetative (enlarged pupils, pale face, sweating, accelerated breath rhythm, accelerated pulse, changes of blood pressure, tachycardia), psychomotoric (anxious facial expression, psychomotoric agitation or vice-versa, allocation of the body on the dental chair – stiff, slightly turned aside, insufficient opening of the mouth, avoidance of eye contact, neurotic movements of hands and legs, clutching the elbow-rest or holding tight the hands).

Some anxious patients may become over detailed, speak more than necessary (this is a reaction of tension and fear); they could look for a touch with the dentist, which gives them perception of security and support. Among some of the patients could be observed more aggressive behaviour, hostile and not corresponding to the atmosphere and the attitude of the dentist. These reactions are not targeted to the doctor's personality as they represent reaction which is consequence from the fear and anxiety. As a result of the fear some patients could postpone the visit for the treatment, avoid situations that seem anxious for them.

Feldbau (3) proposes four rules of dentist's behaviour aiming at overcoming the fear, pain and stress:

- scrutiny evaluation of the level of anxiety and stress of the patient by the means of careful and precise interview. Uncontrolled worry and stress could lead to inadaptable situations, dangerous for the patient's life and of high medical risk. The most important strategy is prophylactics;

- based on the entire collected information on the health status and psychological particularities of the individual, correct methods for control of the fear and anxiety can be determined. This evaluation is dominant for the choice of the adequate treatment. Of great importance is also the control of the patient's reactions during the treatment process;
- to use of pharmaceutical means as auxiliary tools for increasing the positive effect, but not as a method of control. Medications fight the fear but do not eliminate conflicts. Mutual understanding and relationship between patient and dentist are always of great importance;
- to apply differentiated methods to the necessities of the patients. Application of single method to all patients can lead to the lack of success.;
- particularly good influencing on anxiety and fear has the relaxation techniques – progressive muscle relaxation and abdominal breathing.

Dental phobia is state of unconquerable fear from dental treatment, excessive manifestation of worry with lasting negative emotions each time when mentioned dental procedures, manipulations, treatment, medical offices, instrument, on the whole everything related to dental aid and practice.

Dental phobia is not determined by the age of the patient or his gender. It is not inborn but acquired during lifetime experience. Certain facts, events, previous experience, anatomic-physiological features of the face and jaw area, increased pain threshold, individual particularities of the central nervous system create preconditions for the appearance of dental phobia. Categorization, which determines four types of dental phobia, was made by Moore, Brodsgard &Birn (1):

- fear, which is conditioned by specific, painful or unpleasant stimulus, needle, dental machine, specific sounds or smells;
- anxiety, provoked by the appearance of psychosomatic reaction, for example allergic reaction to some local anaesthetic, panic attack.

Some of these characteristics are directly related to the state of agoraphobia, in which the individual reveals reaction of worry to any strange and unknown environment;

- disorders related to commonly manifested anxiety or multi-factorable phobia;
- lack of confidence in the personnel, caused by social phobia.

Dentist shall manifest understanding, respect, compassion to the pain and fears; he shall create good psychological climate of serenity, compassion and attention. Therefore, the good influence to the patient could be achieved when diagnosis, necessary manipulations, treatment and expected perceptions during the process are explained with calm and understandable language.

Psychological techniques and methods in dental practice

Psychological techniques and methods to manage the fears and anxiety of the patients are different and each dentist shall have individual approach for the choice.

In 1983 Dr. Spasimir Gazdov introduces the term "melotherapy" (3). He thinks that most of the people (despite their age) react negatively, with fear and anxiety, during the dental treatment process. He believes that music is universal tool that impacts in maximal positive way on the psycho-emotional background of the treatment process.

Especially good impact on the anxiety and fear, have relaxation techniques like:

- progressive muscle relaxation – through subsequent contraction of the muscle groups (three times for 20 seconds for each group) and repetition to the extend when the entire body grows heavier and wormer. Thus relaxation of the entire body is reached and pressure is decreased. After the estimation of the doctor this exercise can be experienced for specific group of muscles;
- abdominal breathing – patients put one of their hands on the chest and the other on the stomach, breathing is calm and slow – 8-10 inhalations by minute. Firstly, patient breaths with his stomach, and then with his chest by gradually limiting the movement of the latest. Such technique is appropriate for strongly anxious patients who suffer form nausea;
- another respiratory technique relates to the recovery of the alveolar carbon dioxide levels – known as breathing in paper bag; patient makes „cup" with his hands in front of his mouth and receives instructions to inhale his own exhaled air;
- distraction and re-direction of patient's attention to other direction, aiming at concentration on different stimulus.

Systematic desensibilization

Together with the patient dentist creates a hexarchy scale of his fears (in order of their strength). Afterwards the patient is educated how to use some of the relaxation techniques. The patient is proposed to imagine the first situation from the scale of his fears (the weakest one) and using the technique to succeed to manage in completely emotional plan and outlive it up to the level of desensibilization. Then he goes to the next situation of fear and worry, survive it the same way as previous and thus continue to the end of the scale. The time end the number of the sessions is individual for each patient.

The approach of the *cognitive restructuring* is used to change the negative thinking and wrong believes of the patient. The target is to change the manner of acceptance of situations, threatening for the patient, by the means of the arguments "for" and "against". He comes to conclusions that his thinking is irrational and fights the reasons that maintain his unreal fears.

Application of the hypnosis in dental practice.

According to A. Moss (3) the sphere of clinical hypnosis application is developed in two directions:

1. Therapeutic hypnosis – used for relaxation of the patient; elimination of his internal pressure, anxiety, fear from pain and discomfort; removal of resistance against ortodont or denture device after the patient has agreed on its mounting; maintenance of continuous comfort during the dental procedure, support the

patient to get accustomed to the dentures; elimination of some bad habits; problems of the work with children; bruxism (clenching or grinding of teeth, especially at night; tightly clamping top and bottom teeth together; the force of clenching causes stressful pressure on the muscles, tissues and jaw; jaw disorders, jaw pain, soreness, headaches, earaches, damaged teeth and other problems can result from bruxism), and brux-mania;

2. Operational hypnosis – avoidance of medication anaesthesia; amnesia of unpleasant perceptions; premedication on individual basis or as combination with general anaesthesia; elimination of intensified pharyngeal reflex; management of increased secretion of salivary gland; prevention of bleeding; post operational anaesthesia; avoidance of post operational complications.

Training in psychology of communication contributes to the complete, complex establishment of the professional appearance of the young specialist. The acquired knowledge creates possibilities for individual approach during the working process depending on the patient's personality. It guarantees high quality of the treatment process, which is completely coherent with the contemporary requirements of dental practice. Mastership in specific behaviour techniques and communication methods supports dentist in the following:

- to chose the appropriate model for relationship with the patient, taking into account his personal qualities;
- to process a calm treatment avoiding undesired conflicts and difficulties;
- to be aware of the specific features of work with children, aged and old people;
- to understand the essence and the meaning of communication process in the framework of the dental practice;
- to gain command of verbal and non-verbal communication techniques and to apply them in his daily practice;
- to be capable to organize good psychotherapeutic environment in his office;
- to be aware of patient's reasons and reactions of fear, to be able to identify the behaviour of anxious patients and those with dental phobia;
- to possess psychological – behaviour, cognitive techniques for over-living the fear and anxiety in patients;
- to be capable to strengthen in competent way the holistic approach in the health care. After their graduation students are expected to apply effective approaches for organization and management of dental care, which is based on their good psychological education being a necessity and requirement of the contemporary standards for medical care.

EXERCISE

Task 1

„Game of roles" – students divide into couples as one of them takes the role of the patient and the other – of dentist. The aim is that students, playing the role of the patient, understand what the perceptions are during the communication with the dentist in dependence of the communication model selected. Different behaviours and roles of

different types of individualities (patients) are exercised. Students in the role of the dentist have the opportunity to test different models of behaviour, to experiment various psychological approaches. Thus they are able to experiment and attain proficiency, in safely environment, in different techniques, methods and strategies. At the end of each game a feed back is exchanged.

Task 2

Solving cases, for example:

- woman of 40 have not visited dentist for 10 years. She ethers the dental office and says" My last dentist said that I am allergic to local anaesthetics. After the injection I fainted on the chair". How are you going to interpret this information? How are you going to proceed with this patient?
- man of 34 feels dental pain, but manifests strong fear. How are you going to overcome the fear and the anxiety? What would you say, what will be your behaviour? How are you going to implement the procedure?

Task 3

Themes for self education, for example:
"Students shall prepare and submit in written their concept for the organization of relevant psychological environment in the office of dental medicine."; "Psychological techniques for decrease of dental fear".

REFERENCES

1. Humphries G, Lynch M. Principles in the behavior of the dental practice. Sofia: IKTUS; 2008.
2. Peez A, Garner A. The languages of the body. Sofia: Ciela; 2000.
3. Stoev V. Psychological approach of the stomatologist towards the patient. Sofia: Ever -S. Penov; 2005.

RECOMMENDED READINGS

1. Ten Have H, Gordjin B, editors. Bioethics in a European perspective. Dordrecht: Clawer Academic Publishers; 2001.
2. Bonaiuto M, Maricchiolo F. La comunicazione non verbale. Roma: Carocci; 2007.
3. Borella V. La comunicazione medico/sanitaria. Milano: Franco Angeli; 2004.
4. Brener C. Psychology basic course. Sofia: FOB; 1993.
5. Desev L. Dictionary of psychology. Sofia: Bulgaric; 1999.
6. Derogates L, Wise T. Disturbing and depressive disorders of somatically ill patients. Sofia: LIK; 2000.
7. Delfino V. The patient-medical personnel communication. Torino: CSE; 2001.
8. Froid A. The ego and defensive mechanisms. Sofia: LIK; 2000.
9. Goldberg D, Benjamin S. Psychiatry in medicine practice. Sofia: FNP; 1997.

10. Hristozov H, Aleksiev A, Atanasov A, Milenkov K. Psychotherapy, methods and directions. Sofia: MF; 2002.
11. Hersen M, Hasselt V. Trattamenti psicologici nell'anziano. Milano; McGraw-Hill; 1998.
12. Ivanov V, Kokocharova A, Hristozov H, .Meckov K. Clinical psychology. Sofia: MF; 1976.
13. Karvasaski B. Classical psychology. Moskow: Piter; 2006.
14. Lindemann H. Autogenic training. Sofia: Kibea; 1999.
15. Marchi I. Autogenic training. Firenze: Demetra; 2002.
16. ICD-10. Sofia: NCKIC; 1998.
17. Milev V, Milev R. Psychopathology. Sofia: Balkan Press; 1994.
18. Mechkov K. Medical psychology. V.Tarnovo: PIK; 1995.
19. Magris C. Saper ascoltare, saper comunicare. Roma; PSE; 2005.
20. Petkov A. Psychology of personality. Sofia: NI; 1976.
21. Petrov R. Psychatry. Sofia: Veda Slovena; 1997.
22. Peels A. Questions and answers. Sofia: Ciela; 2006.
23. Rimann F. Basic forms of the fear. Sofia: LIK; 2002.
24. Solovyova S. Medical psychology. Moskow: Sova; 2006.
25. Shurtleff M. La fiducia e denaro. Milano: Franco Angeli; 1998.
26. Shurtleff M. The trust and the money. Milano: FrancoAngeli; 1998.
27. Todorkin K. Cognitive psychotherapy. V.Tarnovo: Faber; 2007.
28. Velichkov I. Hypnosis. Sofia: LIK; 2001.
29. Vodenicharov T, Popova S. Medical ethics. Sofia: Falkor Consult; 2006.
30. Vodenicharov T. Medicine ethics guidelines. Sofia; Goreks Press; 1997.
31. VerrastroV. Psicologia della comunicazione. Milano: FrancoAngeli; 2007.
32. VerrastroV. Psychology of communications. Milano: FrancoAngeli; 2007.
33. Wilkinson Gr. Stress. Sofia: Softpress; 2007.

Chapter 4

METHODS OF PLANNING AND EVALUATION

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASURING THE BURDEN OF DISEASE: DISABILITY ADJUSTED LIFE YEARS (DALY)**
Module: 4.1.1	**ECTS (suggested): 0.40**
Author(s), degrees, institution(s)	**Doncho Donev**, MD, PhD**,** Professor Institute of Social Medicine, Faculty of Medicine, University "Ss Cyril and Methodius", Skopje, Republic of Macedonia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Vesna Bjegović,** MD, PhD, Professor Institute of Social Medicine and Centre School of Public Health, School of Medicine, University of Belgrade, Serbia **Genc Burazeri**, MD, PhD Department of International Health, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
Address for correspondence	**Doncho Donev** Institute of Social Medicine, Faculty of Medicine, 50 Divizia 6, MKD-1000 Skopje, Republic of Macedonia E-mail: dmdonev@gmail.com
Keywords	burden of disease, summary measures of the population health, DALYs
Learning objectives	After completing this module students and public health professionals should be able to: • understand general approach in measuring the burden of diseases (BoD); • understand concept of disability adjusted life year (DALY); • perform DALY calculations using basic procedure on an individual as well as on population basis.
Abstract	The BoD concept is used in examining health issues internationally from the perspective of determining cost-effective interventions and priority setting for resource allocation. The DALY has emerged as a measure of the BoD in early 1990-ties. DALYs for a disease or health condition are calculated as the sum of the Years of Life Lost (YLL) due to premature mortality and the Years Lost due to Disability (YLD) of the health condition of less than full health.
Teaching methods	An introductory lecture gives the students insight into DALY concept and calculation procedure. In continuation they apply in small groups calculation according to examples and case study provided, discuss the results and problems.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers, LCD projection, access to the Internet, access to the bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	A multiple choice test and practical calculation of a given case study.

MEASURING THE BURDEN OF DISEASE: DISABILITY ADJUSTED LIFE YEARS (DALY)

Doncho Donev, Lijana Zaletel-Kragelj,
Vesna Bjegović, Genc Burazeri

THEORETICAL BACKGROUND

Burden of disease

Various diseases people suffering from put different amount of disease-burden on populations. In last decades this burden is increasingly measured across the nations with intention to compare it. We could define burden of disease (BoD) as the burden that a particular disease process has in a particular area as measured by cost, morbidity, and mortality. It is quantified by so called summary measures of population health.

Summary measures of population health

Summary measures of population health are measures that combine information on mortality and non-fatal health outcomes to represent the health of a particular population as a single number. Over the past 30 years or so, several indicators have been developed to adjust mortality to reflect the impact of morbidity or disability. These measures fall into two basic categories, health expectancies and health gaps (1-3).

1. Health expectancies.
   Health expectancies measure years of life gained or years of improved quality of life. In this group of measures, among others, following measures are classified:
   - active life expectancy (ALE),
   - disability-free life expectancy (DFLE),
   - disability-adjusted life expectancy (DALE),
   - health adjusted life expectancy (HALE),
   - quality adjusted life expectancy (QALE).

2. Health gaps.
   Health gaps measure lost years of full health in comparison with some 'ideal' health status or accepted standard. In this group of measures among others, following measures (indicators) are classified:
   - potential years of life lost (PYLL),
   - healthy years of life lost (HYLL),
   - quality adjusted life years (QALY),
   - disability adjusted life years (DALY).

Both approaches use time and multiply number of years lived (or, not lived in case of premature death) by the "quality" of those years. The process of adjustment of the years of healthy life lived is called "quality adjustment" (expressed as QALYs), and the process of adjustment of the years of healthy life lost is called "disability adjustment" (expressed as DALYs) (4,5). It means that QALYs represent a gain which should be maximised, and DALYs represent a loss which should be minimised. In the QALY approach the quality is

weighted (sometimes called "utility", as it is the case of cost-utility analyses) on a scale from 1 indicating perfect health and the highest quality of life, to 0 indicating no quality of life and is synonymous to death. In the DALY approach the scale goes in opposite way: a disability weighted zero indicates perfect health (no disability), and weighted 1 indicates death. The disability weighting is the most difficult and controversial part of the DALY approach (5). A typology of summary measures of population health is presented in Figure 1.

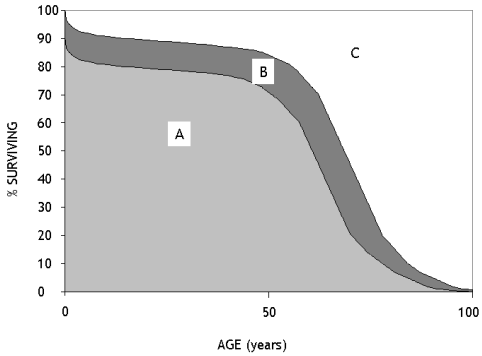


Figure 1. A typology of summary measures. Adapted from Murray CJL and Lopez AD (1). LEGEND: A=time lived in optimal health, B=time lived in suboptimal health, C=time lost due to mortality.

There are two lines (upper and lower) and three areas (A, B and C) (Figure 1). The upper line is the survivorship curve from a hypothetical life table population. The lower curve is a hypothetical curve of survivors to each "age x" in optimal health. Area A represents time lived in optimal health, area B time lived in suboptimal health, and area C time lost due to mortality. Total life expectancy at birth is given by the area under the upper curve (Equation 1):

$$LE = A + B \qquad \textbf{Equation 1.}$$

LE = total life expectancy at birth
A = time lived in optimal health
B = time lived in suboptimal health

Health expectancies are population indicators that estimate the average time that a person could expect to live in a defined state of health. In terms of Figure 1, health expectancy is given by following equation (Equation 2):

$$HE = A + f(B) \qquad \textbf{Equation 2.}$$

HE = health expectancy
A = time lived in optimal health
B = time lived in suboptimal health
f(B) = function that assigns weights to years lived in suboptimal health (optimal health has a weight of 1)

Health gaps measure the difference between actual population health and some specified standard or goal (Equation 3):

$$HG = C + g(B) \qquad \textbf{Equation 3.}$$

HG = health gap
B = time lived in suboptimal health
C = time lost due to mortality (premature death)
f(B) = function that assigns weights to health states lived during time B, but where a weight of 1 equals to time lived in a health state equivalent to death

One of the most employed measures of health gaps is DALY measure (1,2).

Disability-Adjusted Life Year (DALY) concept

Introduction

DALY is an indicator of BoD in a population. It takes into account not only premature mortality, but also disability caused by disease or injury. As a new single summary measure was introduced in a 1990 Global Burden of Disease Study (GBDS) (6), that represented a major step in quantifying global and regional effects of diseases, injuries, and risk factors on population health. It is worth to note that DALYs are an inverse form of the more general concept of QALYs (7).

Two dimensions of DALYs

The DALY is a time-based measure that combines years of life lost due to premature mortality and years of life lost due to time lived in health states less than ideal health. One DALY can be thought of as one lost year of "healthy" life, and the BoD can be thought of as a measurement of the gap between current health status and an ideal situation where everyone lives into old age, free of disease and disability (1,4,8-11). In other words, DALYs are the combination (more precisely the sum) of two dimensions: the present value of future years of lifetime lost through premature mortality, and the present value of years of future lifetime adjusted for the average severity (frequency and intensity) of any mental or physical disability caused by a disease or injury (10-12).

The years of life lost dimension

As a basis for the DALY measure, a "gold standard", or most desirable life, is defined as living in a completely healthy state until death at age around 80 years. Perfect health is 1 on the y-axis and death is 0 on the DALY diagram shown in Figure 2. The "ideal" life is quantified as the total area in the box, a combination of the number of years lived and the full quality of life without disability (5,8).

For each premature death[14], the number of years lost is counted up to the "standardised" maximum life span. The standardized maximum life span is 82.5 years for

[14] Here, a premature death is defined as one that occurs before the age to which a person could have expected to survive assuming a life expectancy at birth approximately equal to that of the world's longest-surviving population - Japan (1,8,9).

females and 80 years for males. It is taken from the country with the highest life expectancy in the world, Japan. Such a measure of premature death in number of years lost is known as "years of life lost" (YLL) (8).

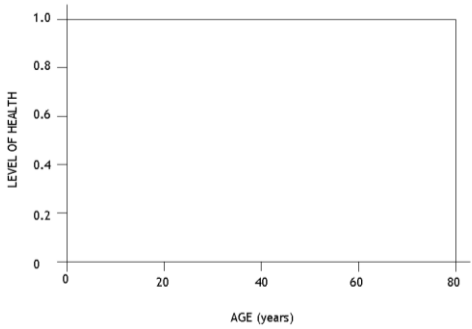


Figure 2. Graphical presentation of a life in full health until sudden death at the age of 80 years (8).

The following example illustrates how the YLLs could be calculated (Example 1).

Example 1.

A scenario: a man dies in a car accident at 30 years of age.

In terms of years of life lost, 50 years are lost due to this premature death (YLLs = 80-30 years). This could be illustrated as presented in a Figure 3. The gray area represents the time lost due to premature death.

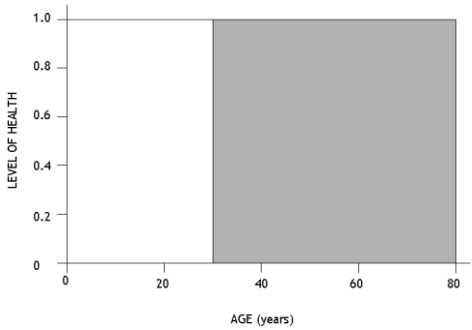


Figure 3. Presentation of a life in full health until sudden death at the age of 30 years. The gray area represents the time lost due to premature death.

The disability dimension

Injury and disease cause not only deaths but also varying time periods with morbidity and disability. The time period in years that is lived in states of poor health or disability due to each disease is another dimension of the DALY measure.

The disability is measured in length in years and in severity.

Severity weights have been appointed for each disabling condition on a scale from one to zero[15] (11,13,14). The disability severity weight for each disease reflects the average degree of disability a person suffers with each condition. Panels of healthy experts with knowledge about disease conditions have determined the weights. We will discuss the disability weights in details later in this module.

The severity weight is then multiplied by the average time a person is suffering from the disability from each disease (5,8). A measure of years lived in health states less than ideal health is known as "years lived with disability" (YLD) (8). Two examples from a person's life are presented in continuation (Examples 2 and 3).

Example 2.

A scenario: at the age of 30, a man gets a knee injury and his health is jeopardized with a weighted severity of 0.1. The injury is incurable and a man suffers until he dies at the age of 80 years.

In terms of years lost due to disability this man's health is only 0.9 of the maximum of 1.0 for the entire 50-year period. This could be illustrated (Figure 4). The grey area in Figure 4 represents his life years lost due to disability, and YLDs corresponds to 5 years (YLDs = 0.1×50 = 5 years).

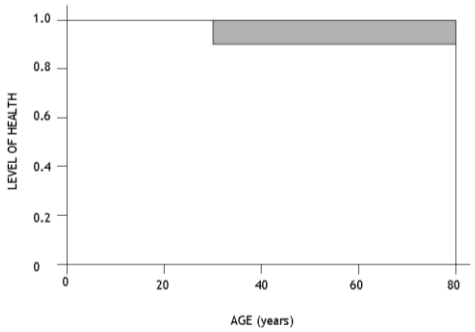


Figure 4. Illustration of life of a man who gets a knee injury at the age of 30. The injury is incurable and a man suffers until he dies at the age of 80 years. The grey area represents life years lost due to disability.

Example 3.

A scenario: at the age of 30, a man gets a knee injury and his health is jeopardized with a weighted severity of 0.1. During the years he suffers from the knee injury his health is only 0.9 of the maximum of 1.0. After At the age 35 he is successfully operated and recovers completely.

In terms of years lost due to disability this man suffers from the knee injury and his health is only 0.9 of the maximum of 1.0 for the 5-year period. This could be illustrated in a figure (Figure 5). The grey area in Figure 5 represents his life years lost due to disability, and YLDs correspond to 0.5 years (YLDs = 0.1×5 = 0.5 years).

[15] For example, schizophrenia was given a weighted severity loss of 0.8, whereas the common cold only causes a loss of 0.007.

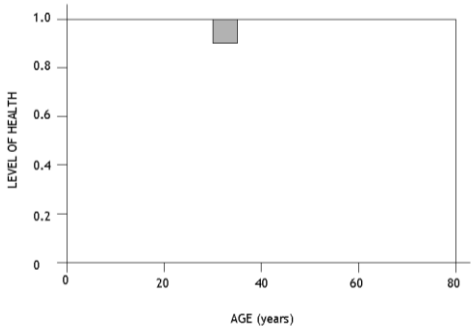


Figure 5. Illustration of life of a man who gets a knee injury at the age of 30. At the age 35 he is successfully operated and recovers completely. A man is healthy until he dies at the age of 80 years. The grey area represents life years lost due to disability.

Both dimensions combined

Usually, both dimensions are combined. Another two examples from a person's life are presented in continuation (Examples 4 and 5).

Example 4.

A scenario: at the age of 40, a man gets cancer which disables him for a certain amount but after a surgery he is in remission for 10 years. After 10 years he suffers from a progress of a disease which disables him substantially more. At the age of 60 years he dies.

In terms of years of life lost, 20 years are lost due to this premature death (YLLs = 80-60 years) (Figure 6, dark grey area).

In terms of years lost due to disability the health of this man is 0.9 for the 10-year period and after a progression is 0.6. YLDs in this case correspond to 5 years (YLDs = 0.1×10 + 0.4×10 = 1+4 = 5 years) (Figure 6, light grey area).

The total loss, that is DALYs, is 25 years (DALYs = YLDs+YLLs = 20+5 = 25)

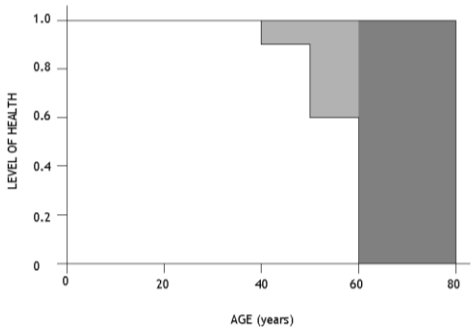


Figure 6. Illustration of a life of a man who gets a cancer at the age of 40. The disease disables him more and more. He dies at the age of 60. The grey area represents life years lost due to disability (light gray) and due to premature death (dark gray).

A scenario in this example is very similar as in Example 4. The difference is that the disease disables a man gradually more and more until he finally dies at the age of 70 years (Figure 7). **Example 5.**

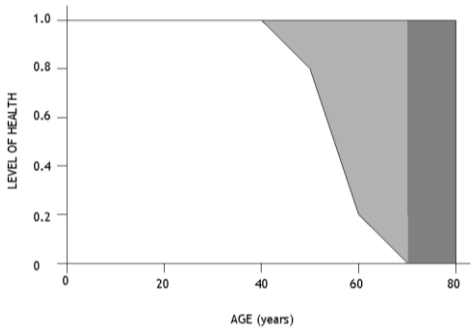


Figure 7. Illustration of a life of a man who gets a cancer at the age of 40. The disease disables him gradually more and more until he finally dies at the age of 70 years. The grey area represents life years lost due to disability (light gray) and due to premature death (dark gray).

The calculation in this case is more complex than presented in Examples 1-4.

Calculation of DALYs

The DALY measure is the sum of both dimensions/components just described - the sum of the YLLs and the YLDs (4,10,11,15-19) (Equation 4):

$$DALY = YLL + YLD \qquad \textbf{Equation 4.}$$

DALY = disability adjusted life years
YLL = years of life lost due to premature death
YLD = years lost due to disability

For YLLs we already know that they measure the number of years lost when a person dies prematurely. Consecutively, the younger is the age at which death occurs, the greater is the number of YLLs. The YLDs measure the number of years of healthy life lost due to temporary or permanent disability. Consecutively, the more severe disability or the longer the duration of this disability associated with a given health condition, the greater is the number of YLDs.

Another characteristic is that DALYs basically comprise so called social values/preferences. The two basic ones are:

- the sex differences which are built in the YLL component (standard life expectancies are separate for men and for women), and
- disability weighting which is built in the YLDs component.

But they could comprise additional social preferences, for example discounting and age weighting as well. According to this characteristic, DALYs could be computed:

- without considering age weighting and discounting,
- with considering age weighting, or
- with considering age weighting and discounting.

All procedures will be discussed in continuation but only the most simple, without considering age weighting and discounting will be presented in details since both other procedures are out of the scope of this module.

Basic procedure considering basic social preferences

Before presenting the basic procedure of calculation of DALYs the main characteristics of two basic social preferences built in the DALY calculation procedure need to be discussed in more details.

Two basic social preferences - sex differences and disability weighting

Major characteristics of basic social preferences are:

1. Sex differences.
   To assess premature mortality, a standard life table for all populations, with life expectancies at birth fixed at 82.5 years for women and 80 years for men is utilized. Life expectancy is calculated separate for men and for women because women live on average longer than men. This table could be accessed in full or abridged form from the publicly available World Health Organization (WHO) web page[16] (20,21). It is also presented in Appendix (Appendix, Table A1). The same table was used by Murray and co-workers (22). What this table tells us is illustrated in Example 6.

Example 6.

Standard life expectancy table tells that (23):

- *a male infant who dies shortly after birth would lose all 80.00 years of his life he would have been expected to live (Appendix, Table A1),*
- *a man who dies at age 50 would lose 30.99 years of expected life, because if he has already survived up to age 50 he has a life expectancy of 80.99 years (50 years + 30.99 years) (Appendix, Table A1),*
- *a man who dies at age 75 would lose only 10.17 years of expected life, because his life expectancy at age 75 is 85.17 years (75 years + 10.17 years) (Appendix, Table A1).*

[16] The GBD methods can be applied on a national or sub-national level and tailored to health values of that country. WHO provides practical tools for people undertaking a BoD study. A link to software to help create life tables and disease models, a manual, supplementary files to help with calculations, and examples of where these methods have been applied by researchers around the globe.

A standard life expectancy allows deaths at the same age to contribute equally to the BoD irrespective of where the death occurs. Alternatives, such as using different life expectancies for different populations that more closely match their actual life expectancies, violate this egalitarian principle. As life expectancy is rarely equal for men and women, a lower reference life expectancy could be used for both sexes because of "biological differences in survival potential", as well as men are higher exposed to various risks such as alcohol, tobacco, and occupational injury (1,5,9,24).

2. Disability weights.

On one side, there are health conditions that frequently cause significant disability or death, while on the other side there are those conditions that rarely cause death but may cause severe and/or prolonged disability. The other health conditions can cause severe disability but they occur late in life and they are of shorter duration (Example 7).

Example 7.

Here we give some examples:

- *Alzheimer's disease often cause severe disability but it typically occurs in older ages and consecutively lasts shorter,*
- *Down syndrome is a congenital anomaly that causes limited ability to perform most activities for entire life of a person,*
- *stroke is highly lethal disease etc.*

In order to quantify time lived with a non-fatal health outcome and assess disabilities in a way that will help to inform health policy, disability must be defined, measured, and valued in a clear framework that inevitably involves simplifying reality.

Disabilities were assigned severity weights ranging from zero, representing perfect health, to one, representing death. These weights were determined at a meeting of experts in international health who had not participated in the first GBDS. In order to reduce the number of weights to be assigned and to emphasize large differences in the severity of disability, each disability condition was assigned to one of six classes. In Table 1 definitions of each of these classes are presented (22).

In continuation, for disability, the incidence of cases by age, sex, and demographic region was estimated on the basis of community surveys or, failing that, on expert opinion; the number of years of healthy life lost was then obtained by multiplying the expected duration of the condition (up to remission or to death) by a severity weight that measured the severity of the disability in comparison with loss of life.

For the GBDS 2000 and 2004 some updates were made and diseases were grouped into seven classes of severity of disability. For example, class IV, which includes arm or below-the-knee amputation, deafness etc., was given a range of severity weight from 0.24-0.36, and class VI, which includes AIDS cases not on antiretroviral drugs, Alzheimer and other

dementias, blindness etc., was assigned a severity weight of 0.6 (range from 0.5-0.7) (Table 2) (11,14).

Table 1. Definitions of disbility weighting classes. Adapted from Murray (22).

Class	Description	Weight
1	Limited ability to perform at least one activity in one of the following areas: recreation, education, procreation or occupation	0.096
2	Limited ability to perform most activities in one of the following areas: recreation, education, procreation or occupation	0.220
3	Limited ability to perform activities in two or more of the following areas: recreation, education, procreation or occupation	0.400
4	Limited ability to perform most activities in all of the following areas: recreation, education, procreation or occupation	0.600
5	Needs assistence with instrumental activities of daly living such as meal preparation, shopping or housework	0.810
6	Needs assistence with activities of daly living such as eating, personal hygiene or toilet use	0.920

Table 2. Disability classes for the Global Burden of Disease Study (GBDS), with examples of long-term disease and injury sequelae falling in each class[1] (11,14).

Dysability class	Severity weights	Conditions[2]
I	0.00-0.02	Stunting due to malnutrition, schistosomiasis infection, long-term scarring due to burns (less than 20% of body)
II	0.02-0.12	Amputated finger, asthma case, edentulism, mastectomy, severe anaemia, stress incontinence, watery diarrhoea
III	0.12-0.24	Angina, HIV not progressed to AIDS, alcohol dependence and problem use, radius fracture in a stiff cast, infertility, erectile disfunction, rheumatoid arthritis, angina, low vision (<6/18, >3/60)
IV	0.24-0.36	Amputated arm, below-the-knee amputation, deafness, congestive heart failure, drug dependence, Parkinson disease, tuberculosis
V	0.36-0.50	Bipolar affective disorder, rectovaginal fistula, mild mental retardation, neurological sequelae of malaria
VI	0.50-0.70	AIDS cases not on antiretriviral drugs, Alzheimer and other dementias, blindness, paraplegia, Down syndrome,
VII	0.70-1.00	Active psychosis (schizophrenia), severe depression, severe migraine, quadriplegia, terminal stage cancer

[1] Based on average severity weight globally for both sexes and all ages in the GBDS 2004 update

[2] Conditions are listed in the disability class for their global average weight. Most conditions will have distributions of severity spanning more than one disability class, potentially up to all seven

Two methods are commonly used to formalize social preferences for different states of health. Both involve asking people to make judgements about the trade-off between quantity and quality of life. This can be expressed as a trade-off in time (how many years lived with a given disability a person would trade for a fixed period of perfect health) or a

trade-off between persons (whether the person would prefer to save 1 year of life for 1,000 perfectly healthy individuals or 1 year of life for perhaps 2,000 individuals in a worse health state). The DALY approach which is the basis of the GBDS currently in use has been much criticised because the method presupposes that life years of disabled people are worth less than life years of people without disabilities (1,5,9). The GBDS for example developed a protocol based on the person trade-off method. In a formal exercise involving health workers from all regions of the world, the severity of a set of 22-indicator disabling conditions - such as blindness, depression, and conditions that cause pain - was weighted between 0 (perfect health) and 1 (equivalent to death). These weights were then grouped into seven classes where class 1 has a weight between 0 and 0.02 and class VII a weight between 0.70 and 1 (Table 2). In the protocol, a life year for 1,000 healthy people has been set as equally valuable as one life year for (1,5):

- 9,524 people with quadriplegia;
- 2,660 blind people;
- 1,686 people with Down's syndrome without cardiac malformation;
- 1,499 deaf people;
- 1,236 infertile people.

WHO has announced changes to such approach.

Classification for estimating mortality and disability

Deaths were classified using a tree structure, in which the first level of disaggregation comprises three broad cause categories of diseases (1,9,14,16):

- Group I: communicable diseases, perinatal, and nutritional conditions;
- Group II: non-communicable diseases;
- Group III: injuries.

Each group was then subdivided into categories: for example, cardiovascular diseases and malignant neoplasms are two subcategories of group II. Beyond this level, there are two further disaggregation levels such that 107 individual causes from the ninth revision of the ICD (ICD-9) can be listed separately. Consistent with the goal of providing disaggregated estimates of BoD to assist priority setting in the health sector, estimates were prepared by age and sex and for eight broad geographic regions of the world: Established Market Economies, Formerly Socialist Economies of Europe, China, India, Latin America and the Caribbean, Middle-Eastern Crescent, Other Asia and Islands, and sub-Saharan Africa (1,9).

The disabilities in a particular class differ in kind (for example, blindness versus paralysis) but were considered to be of equal severity. Each participant then voted on the weight to be assigned to the entire class, not to individual disabilities, and the class was weighted according to the average vote. It is important to note that many disabling conditions lead to two or more distinct disabilities, which may be classified in more than one class of severity (9,11,24).

There is surprisingly wide agreement between cultures on what constitutes a severe or a mild disability. For example, a year lived with blindness appears to most people to be

a more severe disability than a year lived with watery diarrhoea, while quadriplegia is regarded as more severe than blindness. These judgements must be made formal and explicit if they are to be incorporated into measurements of BoD (1,9,14).

When one would like to perform calculation of YLD, he/she will need to find the actual list of disability weights. In the GBDS 2004 update publication is stated that the authors used for the weight factor the weights listed in Annex Table A6 of a publication of Mathers et al (7). The same table could be obtained as special WHO document available on the Internet (25), while more detailed table, including weights according to age, is available on special WHO web page on the Internet (20,26). For the purposes of this module an adapted table presenting average disability weights for diseases and conditions including cancers and injuries is available in Appendix (Appendix, Table A2).

The basic procedure

We will present the basic procedure for calculating DALYs by using an example (Example 8).

Example 8. *A scenario: a woman who had moderate depression since she was 20 commits a suicide at age 50.*

1. Calculating the YLLs.
   On a population basis the YLLs for a given age basically correspond to the number of deaths for that given age multiplied by the standard life expectancy at the age at which death occurs. The basic formula for calculation of YLLs on a population basis is the following (12,15,19) (Equation 5):

$$YLL = N \times L \qquad \textbf{Equation 5.}$$

YLL = years of life lost due to premature death
N = number of deaths
L = standard life expectancy at age of death in years

On an individual basis the YLLs for an individual person correspond to the standard life expectancy at the age at which death occurs. We have already worked out this situation in Example 6, but it is worked out again for Example 8. For the woman from Example 8 the YLLs are calculated as follows:

Example 8. Cont. *In terms of YLLs the woman from scenario presented in Example 8:*

- *would lose 33.99 years of expected life, because if she has already survived up to age 50 she has a life expectancy of 83.99 years (50 years + 33.99 years) (Appendix, Table A1),*
- *YLL =33.99*

2. Calculating the YLDs.
   Because YLLs measure the incident stream of lost years of life due to deaths, an incidence perspective is also taken for the calculation of YLDs.

To estimate YLDs for a particular cause in a particular time period, the number of incident cases in that period is multiplied by the average duration of the disease and a weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (death). The basic formula for calculation of YLDs on a population basis is the following (12,15,19) (Equation 6):

$$YLD = I \times DW \times L \qquad \textbf{Equation 6.}$$

YLD = years lost due to disability
I = number of incident cases
DW= disability weight
L = average duration of the case until remission or death (years)

On an individual basis the basic formula for calculation of YLDs is the following (Equation 7):

$$YLD = DW \times L \qquad \textbf{Equation 7.}$$

YLD = years lost due to disability
DW= disability weight
L = duration of the case until remission or death (years)

For the woman from Example 8 the YLDs are calculated as follows:

In terms of YLDs: **Example 8. Cont.**

- *her disability is weighted to 0.350 (disability weight for moderate depressive episode) (Appendix, Table A2),*
- *it lasts for 30 years (50-20 years),*
- *YLD = (20×0) + (30×0.350) = 10.50 years*

3. Calculating DALYs.
   At the end the YLLs and the YLDs are summed up according to Equation 4.
   For the woman from Example 8 the DALYs are calculated as follows:

DALY = 33.99 + 10.50 = 44.49 **Example 8. Cont.**

The burden of disease in this case in terms of DALYs is 44.49 years.

Procedures considering additional social preferences

Before presenting procedures of calculation of DALYs that consider additional social preferences their main characteristics are briefly discussed.

Additional basic social preferences – discounting and age weighting

Major characteristics of additional social preferences are:

1. Discounting.

   Discounting means that future gains and losses are counted less than if they had occurred today. The years lost in the future are discounted, so that years lost now are worth more than years lost in the future. This is a standard procedure and common practice in economics when it comes to valuing material goods, and in the DALY calculations a discount rate of 3% per year is used (5,8,9,16,27).

   The innovative GBDS for example calculated the total sum of the combined loss of all premature deaths that occurred in the world in 1990 and the loss of healthy life from disability in future years from specific diseases arising in that year (28). The study used all possible data sources of recorded causes of death and prevalence and incidence of disease, as well as expert judgment when data were not available (8).

   Individuals commonly discount future benefits against current benefits similarly to the way that they may discount future dollars against current dollars. Whether a year of healthy life, like a dollar, is also deemed to be preferable now rather than later, is a matter of debate among economists, medical ethicists, and public health planners, since discounting future health affects both measurements of disease burden and estimates of the cost-effectiveness of an intervention (3,12). There are arguments for and against discounting. In the GBDS, for example, future life years were discounted by 3% per year. This means that a year of healthy life bought for 10 years hence is worth around 24% less than one bought for now, as discounting is represented as an exponential decay function. Another effect is that it reduces the value of interventions, especially preventive, that provide benefits largely in the future, such as vaccinating against hepatitis B, which may prevent thousands of cases of liver cancer, but some decades later (1,9,11,12).

2. Age weighting.

   Another characteristic of the DALY measure is that years lost due to premature deaths or disability could give different values at different ages. These differences in values are introduced in DALY calculations by what is called "age weights". The age weight used in the DALY calculations is obtained from a scale where the value of a year lost rises steeply from zero at birth to a maximum at 25 years of age, and then decreases progressively in older ages (Figure 8). From the Figure 8 it could be seen that the relative value of a life year is below 1 for the children under 10, and for the persons more than about 55 years of age.

   Because of weighting the value of the lifetime the years of life in childhood and old age are counted less because of social roles vary with age and social value of that time may be different i.e. "young, and often elderly, depend on the rest of society for physical, emotional and financial support" (3,5,24). This means that if a newborn girl dies, 32.5 weighted years are lost, if she dies at age 30, 29 weighted years are lost, and at age 60, 12 weighted years are lost (8). For males, the above figures will be slightly lower, because their shorter life expectancy is taken into consideration. Another example is that living with disability, e.g. blindness, for a person aged 80 years is considered "less bad" than living with blindness for a 25-year old individual (1,5,9,16,24).

Figure 8. Relative values of a year of life. Adapted from the World Bank development report 1993 (9).

If individuals are forced to choose between saving a year of life for a 2-year old and saving it for a 22-year old, most prefer to save the 22-year old one. A range of studies confirms this broad social preference to weight the value of a year lived by a young adult more heavily than one lived by a very young child, or an older adult. Adults are widely perceived to play a critical role in the family, community, and society. It was for these reasons that the GBDS incorporated age weighting into the DALYs. It was assumed that the relative value of a year of life rises rapidly from birth to a peak in the early twenties, after which it declines steadily (1,3,8).

Calculation of DALYs with discounting

When 3% discounting and uniform age weights are considered, the formulas for calculating YLLs (Equation 8) and YLDs (Equation 9) on the population basis are the following (19,29):

$$YLL = \frac{N}{0.03}\left(1 - e^{-0.03L}\right) \qquad \textbf{Equation 8.}$$

YLL = years of life lost due to premature death
N = number of deaths
L = standard life expectancy at age of death in years

$$YLD = \frac{IDW\left(1 - e^{-0.03L}\right)}{0.03} \qquad \textbf{Equation 9.}$$

YLD = years lost due to disability
I = number of incident cases
DW= disability weight
L = average duration of the case until remission or death (years)

Detailed presentation of this procedure is out of the scope of this module.

Calculation of DALYs with discounting and age weighting

When non-zero discounting and age weighing are considered, the formulas for calculating YLLs (Equation 10) and YLDs (Equation 11) on an individual basis are the following (12,19,29):

$$YLL = \frac{KCe^{ra}}{(\beta + r)^2}\left[\begin{array}{l}\left[e^{-(r+\beta)(L+a)}\right] \\ \left[-(r+\beta)(L+a)-1\right] \\ \left[-e^{-(r+\beta)a}\right] \\ \left[-(r+\beta)a-1\right]\end{array}\right] + \left[\frac{1-K}{r}\left(1-e^{-rL}\right)\right] \qquad \textbf{Equation 10.}$$

YLL = years of life lost due to premature death
K = age-weighting modulation constant (e.g. K=1)
C = adjustment constant for age weights (GBDS standard value is 0.1658)
r = discount rate (GBDS standard value is 0.03)
a = age of death (years)
β = age-weighting constant (GBDS standard value is 0.04)
L = standard life expectancy at age of death (years)

$$YLL = DW\left\{\frac{KCe^{ra}}{(\beta + r)^2}\left[\begin{array}{l}\left[e^{-(r+\beta)(L+a)}\right] \\ \left[-(r+\beta)(L+a)-1\right] \\ \left[-e^{-(r+\beta)a}\right] \\ \left[-(r+\beta)a-1\right]\end{array}\right] + \left[\frac{1-K}{r}\left(1-e^{-rL}\right)\right]\right\} \qquad \textbf{Equation 11.}$$

YLD = years lost due to disability
DW= disability weight
K = age-weighting modulation constant (e.g. K=1)
C = age-weighting correction constant (GBDS standard value is 0.1658)
r = discount rate (GBDS standard value is 0.03)
a = age of onset
β = parameter from the age-weighting function (GBDS standard value is 0.04)
L = duration of disability (years)

Detailed presentation of this procedure is out of the scope of this module since it is very complicated and a separate module is needed to present it. For all who want to study this procedure in more details, a paper by Fox-Rushby and Hanson (12), is recommended where also the formulas in Excel programme for calculation of YLLs and YLDs are given.

These formulas have also been programmed into calculation spreadsheet templates for calculation of DALYs on a population basis (19,29), available at the WHO web site (20,30).

Sensitivity analysis

To gauge the impact of changing these social choices on the final measures of BoD, the GBDS assessments were recalculated with alternative age weighting and discount rates, and with alternative methods for weighting the severity of disabilities. Overall, the rankings of diseases and the distribution of burden by broad cause groups are largely unaffected by age weighting and only slightly affected by changing the method for weighting disability. Changes to the discount rate, by contrast, may have a more significant effect on the overall results. Changes in the age distribution of burden, in turn, affect the distribution by cause, as communicable and perinatal conditions are most common in children while non-communicable diseases are most common in adults. The most significant effect of changing the discount and age weights is a reduction in the importance of several psychiatric conditions (1,11,24).

However, sensitivity analysis has shown that the results of the GBDS are not greatly affected by these social preferences. Another problem is that the GBDS calculates DALY on data, which on some continents are of poor quality. Especially for the disability calculations, the data is of varied quality in different regions, e.g. Sub-Saharan Africa, and for different disease conditions, e.g. depression (8,11).

Controversies and criticism

Some critical articles on the DALY approach have questioned both the validity of the results as well as the underlying value-judgements (5,24,27). In the Journal of Health Economics Anand and Hanson argues that: "the conceptual and technical basis for DALYs is flawed, and that the assumptions and value judgements underlying it are open to serious question" (24).

According to some authors, the DALY concept has "the potential to revolutionize the way in which we measure the impact of disease, how we choose interventions, and how we track the success or failure of our intervention (31,32). Furthermore, DALYs are considered to be an "advancement" over other composite indicators, such as QALYs, because the value choices incorporated in the DALYs are made explicit: "The black box of the decision-maker's relative values is then opened for public scrutiny and influence" (24). Yassin (33), pointed out several advantages for using DALY in studies of health inequalities:

- the DALY is the only measure that can infuse information about non-fatal health outcomes into debates of health inequalities,
- DALY uncouples social and epidemiological assessment of health inequalities from advocacy,
- the DALY can measure the magnitude of premature death and non-fatal health outcomes attributable to proximal biological causes, including diseases and injuries or attributable to more distal causes such as poor living standards, tobacco use or socio-economic determinants,
- the DALY is a stable measure that can be used for purposes of comparisons either between different communities or between different points of time.

Jankovic (3), emphasized that DALY measurement of clinical outcomes and cost-effectiveness analyses allows existing or prospective interventions to be judged both in terms of cost-effectiveness, and their relative impact in reducing the BoD and ill-health. DALY as a composite indicator is a useful analytical tool for health policy-makers and analysts in priority setting and resource allocation in health systems providing unique and desirable health information on non-fatal health outcomes that is essential for determining appropriate health research priorities, too (3).

The DALY measure has been criticized because of the four built-in social preferences:

- different weights for sexes,
- different age weights,
- discounting future years lost, and
- severity weighting of disabilities.

Many argue that life years for men and women should be given the same weight. However, as has been described above the difference is small and only gives a slightly greater value for diseases that affect females. Some people argue that all years lost should be given the same value independently of the age at which the years are lost. Others argue that discounting years is wrong, because they value the current and the future years equally. In a complex measure like DALY, the built-in social preferences may conceal issues of inequity. The most difficult part of any approach combining data on quality of life and length of life is how to measure the quality of life. Many philosophical questions as well as questions regarding the limits of natural sciences arise. The first requirement of a valid measurement is that one knows what is being measured. The concept of quality of life is, however, vaguely defined, and different people as well as different cultures may have very different opinions on the main elements of a good life (5,16).

The approach has been criticized for violating the principle of treating people equally and for discriminating the young, the elderly, future generations (future health benefits), the disabled, and the women (5,16,24).

Use of DALYs

This approach increased the validity of comparisons of the burden of different diseases between world regions and countries over time. In fact, the World Bank and the World Health Organization (WHO) were the first institutions to use the DALY measures to compare the BoD in different regions of the world and thereby the value and effectiveness of different health interventions and changes in living conditions. It became possible to estimate and compare the cost of avoiding the loss of a DALY for each intervention (4,8,9).

Prior to the GBDS, which began in 1992, there had been no comprehensive efforts to provide comparable regional and global estimates and projections of the causes of loss of health and disease and injury burden in populations based on a common methodology and denominated in a common metric comparable across populations and over time (1,9).

One of the major goals of the GBDS was to facilitate the inclusion of non-fatal illness/conditions and their long-term health consequences/outcomes (mental

and musculoskeletal disorders, blindness etc.) into debates on international health policy, beside the causes of death and mortality data. In addition, there was a need to quantify the BoD using a measure that could then be used for cost-effectiveness analysis. The GBDS method quantifies not merely the number of deaths but also the impact of premature death and disability in a population, combining these measures into a single unit of measurement of the overall BoD in the population - the DALY. DALYs allow the losses or the BoD from the premature death and nonfatal consequences of over 100 diseases and injuries to be expressed in the same unit. The study also presented the first global and regional estimates of disease and injury burden attributable to certain risk factors for disease, such as tobacco, alcohol, poor water and sanitation, and unsafe sex. The method uses 107 diagnoses, covering all conceivable causes of death and 95% of all possible causes of disability (1).

The methods of the GBDS 1990 created a common metric system to estimate the health loss associated with morbidity and mortality. It generated widely published findings and comparable information on disease and injury incidence and prevalence for all world regions. It also stimulated numerous national studies of BoD. These results have been used by governments and non-governmental agencies to inform priorities for research, development, policies and funding. In 2000, the WHO began publishing regular GBDS updates for the world and 14 regions. These revisions were aided by methodological improvements and more extensive data collection that covered key aspects of the global BoD, including mortality estimation, cause of death analysis, and measurement and valuation of functional health status. Standardized concepts and approaches to comparative risk assessment were applied for over 25 risk factors (10,11).

Conclusion

Bearing its inherent weaknesses in mind, DALY is still a very useful measure, because it is the first comprehensive attempt made to summarise the world's burden of injury, disease and premature death. It has initiated a debate and new research to find even better complex indicators for global comparisons (34). The DALY measure is useful to describe the disease burden across the world and to make projections for the future. At present, many countries are exploring the possibility of using DALYs as a measure of trends in disease burden and as a tool for cost-effectiveness studies and priority setting (8,32).

In 1999 WHO has started to include the DALY measure in their annual reports. This allows for refinement of the results from the initial study because of new health data. From the year 2000, they also included a DALE, disability-adjusted life expectancy, which was renamed as the more cheery HALE, health-adjusted life expectancy, in 2002. This measure is based on life expectancy at birth, but includes an adjustment for the time spent in poor health. It is the equivalent of the number of years a newborn can expect to live in full health, based on current statistics of mortality and morbidity. In Japan, for instance, the HALE is 72 years, while in Afghanistan only 35 years. Many find this a measure that is instinctively easier to understand compared to DALY (3,8).

CASE STUDY

A scenario

In a heavy frontal car collision of two cars 7 people are involved. In the first car a 4-member family dies, while in the second car 3 young people are heavily injured. The sequelae of a car collision for every participant are presented in Table 3.

Table 3. The sequelae of a car collision for every participant with their weights.

	Participant	Sequela	Disability weight
	CAR 1		
1.	36 years of age father	dies	1
2.	29 years of age mother	dies	1
3.	7 years of age daughter	dies	1
4.	2 years of age son	dies	1
	CAR 2		
5.	27 years of age man	Injured spinal cord (lifelong)	0.725
6.	25 years of age woman	Intracranial injury - short term (½ year)	0.359
7.	22 years of age woman	Fractured ribs (¼ year)	0.199

Calculation (basic procedure)

In continuation we will calculate the DALYs for this car accident. In Table 4 is presented calculation of elements of DALYs - YLLs and YLDs - and the final result (DALYs).

Table 4. Elements of calculation of DALYs and final result.

Participant	YLLs	YLDs	DALYs
1.	44.58 years LE		
2.	54.25 years LE		
3.	75.97 years LE		
4.	78.36 years LE		
5.		53.49 years LE × 0.725 = 38.78	
6.		0.5 × 0.359 =0.18	
7.		0.25× 0.199 =0.05	
Total	253.16	39.01	292.17

LEGEND: DALY = disability adjusted life years; YLL = years of life lost due to premature death; YLD = years lost due to disability; LE = life expectancy

In summary, altogether 292.17 DALYs are lost in this heavy frontal car collision.

EXERCISE

Task 1

After introductory lecture, students carefully read the part on theoretical background of this module and corresponding recommended readings.

Task 2

Students consider following scenario: in the age 55 a previously healthy man is diagnosed a prostate cancer. After a surgery and radiotherapy he is in remission for 15 years. In the age of 70 he is diagnosed a metastatic disease. He dies in the age of 75.

In groups of 2-3 in a process of calculation of DALYs students should follow the following steps:

- make graphical presentation (a sketch) of a time horizon for this case,
- calculate DALYs[17],
- compare results to the results of other student groups.

Task 3

This task is based on the examples presented in the manual »Selecting an essential package of health services using cost-effectiveness analysis: a manual for professionals in developing countries« (35).

Let us imagine a 5-year old girl who falls sick with poliomyelitis at this age. Following scenarios are possible:

1. immediately after she contracts poliomyelitis at age 5, she dies,
2. after she contracts the disease at age 5 she lives unitil age 10 with a disability,
3. after she contracts the disease at age 5 she becomes permanently disabled over her entire life span (77.95 years),
4. after she contracts the disease at age 5 she is in an acute phase of the disease confined to bed for ½ month (disability weight 0.500). Afterwards she fully recovers.

In a process of calculation of DALYs students (this task is to be done by every student alone) should follow the following steps:

- make graphical presentation (a sketch) of a time horizon for each of these scenarios,
- calculate DALYs (be careful in calculating YLLs and YLDs; help yourself with graphical presentations) for each scenario[18],
- compare results to the results of other students.

[17] Final result: 31.93 years

[18] Final results:
Scenario 1: 77.95 years
Scenario 2: 74.835 years
Scenario 3: 28.764 years
Scenario 4: 0.021 years

Task 4

This task is continuation of the Task 3 and is basing on working paper of Homedes (36).

Let us imagine that in a particular community there are 20 girls who fall sick with poliomyelitis at age 5:

- 2 of them die immediately,
- 2 die at age 10 after a 5-year period of disability,
- 10 of them are permanently disabled, and
- 6 recover completely after the ½ month acute phase of the disease.

Calculate the total number of DALYs lost due to poliomyelitis in this community[19].

Task 5

Students first carefully study the data presented in Table 5.

Table 5. Morbidity and mortality from ishaemic heart disease (heart attack), by the age of getting heart attack and period of survival, in the population of the Region X in SEE, 2008.

Health outcome of the heart attack		Age group							
		< 35	35-39	40-44	45-49	50-54	55-59	60-64	65 +
Deaths	No. of cases	9	14	16	13	12	10	8	7
	YLL per case	40	35	30	25	20	15	10	5
Disability (weight 0.3)	No. of cases	9	14	16	13	12	10	8	7
	Years of survival	40	35	30	25	20	15	10	5

Afterwards they should consider following tips:

- they discuss what these data present,
- they calculate BoD (expressed as DALYs) using the data from the Table 5,
- they should use the basic formula for a simple way of calculation, without age-weighting and discounting,
- they should compare the results[20],
- if possible, students should calculate DALYs for various diseases and injuries causing disability and/or premature death, using official data from their own countries or local communities.

[19] Final result: 593.336 years
[20] Final result: 2,808 years

Task 6

This task is comprised of following steps:

- on PubMed Central students should try to find papers using DALYs as summary measure of BoD,
- they should analyze the ratio DALYs calculated according to basic method: DALYs calculated using discounting or/and age weighting,
- they should critically discuss the pros and contras of these different methods.

REFERENCES

1. Murray CJL, Lopez AD. Assessing health needs: the Global Burden of Disease Study. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford Textbook of Public Health - Fourth edition. New York: Oxford University Press Inc.; 2004. p.243-54.
2. Atanackovic-Markovic Z, Bjegovic V, Jankovic S, et al. The burden of disease and injury in Serbia. Serbian Burden of Disease study - an EU funded project managed by the European Agency for Reconstruction. Belgrade: Ministry of Health of the Republic of Serbia; 2003.
3. Jankovic S. Summary measures of population health and their relevance for health policy. In: Galan A, Scintee G, editors. Public Health Strategies. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2005. p.190-207. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
4. Tulchinsky TH, Varavikova EA. Measuring and evaluating the health of a population. In: Tulchinsky TH, Varavikova EA. The New Public Health. An Introduction for the 21st century. San Diego: Academic Press; 2000. p.113-68.
5. Arnesen TM, Norheim OF. Disability Adjusted Life Years - possibilities and problems. National Institute of Public Health, Oslo, Norway. Available from URL: http://iier.isciii.es/supercourse/lecture/lec2911/index.htm. Accessed: April 8, 2009.
6. World Health Organization. About the Global Burden of Disease (GBD) project [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/global_burden_disease/about/en/index.html. Accessed: August 14, 2009.
7. Mathers CD, Lopez AD, Murray CJL. The burden of disease and mortality by condition: data, methods and results for 2001. In: Lopez AD, Mathers CD, Ezzati M, Murray CJL, Jamison DT, eds. Global burden of disease and risk factors. New York, Oxford University Press; 2006. p.45-240. Available from URL: http://www.dcp2.org/pubs/GBD. Accessed: August 26, 2009.
8. Lindstrand A, Bergstrom S, Rosling H, Rubenson B, Stenson B, Tylleskar T. Global health - an introductory textbook. Copenhagen: Narayana Press; 2008.
9. The World Bank. The World Bank development report 1993. Investing in Health. New York, NY: Oxford University Press; 1993. Available from URL: http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&siteName=WDS&entityID=000009265_3970716142319. Accessed: August 14, 2009.
10. World Health Organization. Global burden of disease study. Study overview. Available from URL: http://www.who.int/healthinfo/global_burden_disease/GBD_2005_study/en/index.html Accessed: April 16, 2009.
11. World Health Organization. Global Burden of Disease 2004 Update. Geneva: World Health Organization; 2004. Available from URL:

http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf. Accessed: August 14, 2009.
12. Fox-Rushby JA, Hanson K. Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. H Health Economics and Financing Programme, Health Policy Unit, London School of Hygiene and Tropical Medicine, London, UK. Available from URL: http://heapol.oxfordjournals.org/cgi/reprint/16/3/326. Accessed: April 15, 2009.
13. Stibich M. Disability Adjusted Life Years - DALYs - Measuring Suffering. Available from URL: http://longevity.about.com/od/longevity101/g/dalys.htm Accessed: April 18, 2009.
14. Galan A. Disability-Adjusted Life Years: A Method for the analysis of the burden of disease. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.140-50. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
15. World Health Organization. Metrics: Disability-Adjusted Life Year (DALY) - quantifying the burden of disease from mortality and morbidity. Available from URL: http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/index.html Accessed: April 16, 2009.
16. Holden B. DALYs and disagreement. Available from URL: http://blog.givewell.net/?p=269 Accessed: April 20, 2009.
17. Somerford P, Katzenellenbogen J, Western Australian Burden of Disease Study: Disability-Adjusted Life Years: Technical overview. Perth: Department of Health; 2004. Available from URL: http://www.health.wa.gov.au/publications/documents/BOD/BOD9.pdf. Accessed: April 9,2009.
18. World Health Organization. Global Burden of Disease (GBD). Available from URL: http://www.who.int/healthinfo/global_burden_disease/en/index.html Accessed: April 12, 2009.
19. Prüss-Üstün A, Mathers C, Corvalan C, Woodward A. Introduction and methods: assessing the environmental burden of disease at national and local levels. WHO Environmental Burden of Disease Series, No. 1. Geneva: World Health Organization; 2003. Available from URL: http://www.who.int/quantifying_ehimpacts/publications/9241546204/en/index.html. Accessed: August 26, 2009.
20. World Health Organization. National tools [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/global_burden_disease/tools_national/en/. Accessed: August 26, 2009.
21. World Health Organization. National tools. National burden of disease supplementary files. Standard life table. Available from URL: http://www.who.int/healthinfo/bodreferencestandardlifetable.xls. Accessed: August 26, 2009.
22. Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ. 1994;72:429-445. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=8062401. Accessed: August 26, 2009.
23. Los Angeles County Department of Health Services, UCLA Center for Health Polocy Research. The burden of disease in Los Angeles County. Available from URL: www.lapublichealth.org/epi/reports/dburden.pdf. Accessed: August 26, 2009.
24. Anand S, Hanson K. Disability adjusted life years: a critical review. Journal of Health Economics, 1997;16:658-702.
25. World Health Organization. Global burden of disease 2004 update: disability weights for diseases and conditions. Geneva: World Health Organization; 2004. Available from URL:

http://www.who.int/healthinfo/global_burden_disease/GBD2004_DisabilityWeights.pdf. Accessed: August 26, 2009.
26. World Health Organization. National tools [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/bodreferencedisabilityweights.xls. Accessed: August 26, 2009.
27. World Health Organization. Global Burden of Disease (GBD) 2005 study [homepage on the Internet]. Available from URL: http://www.who.int/healthinfo/global_burden_disease/GBD_2005_study/en/index.html Accessed: April 12, 2009.
28. Murray CJL, Lopez AD, editors. The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2030. Global Burden of Disease and Injury Series, Vol. 1. Cambridge, MA: Harvard University Press; 1996.
29. World Health Organization, Global Program on Evidence for Health Policy. National burden of disease studies: a practical guide. Geneva: World Health Organization; 2001. Available from URL: http://www.who.int/healthinfo/nationalburdenofdiseasemanual.pdf. Accessed: August 25, 2009.
30. World Health Organization. DALY calculation template. Available from URL: http://www.who.int/healthinfo/bodreferencedalycalculationtemplate.xls. Accessed: August 26, 2009.
31. Foege W. Preventive medicine and public health. Journal of American Medical Association, 1994;271:1704-5.
32. Bobadilla JL, Cowley P, Musgrove P, Saxenian H. Design, Content and Financing of an Essential National Package of Health Services. The World Bank Group Private and Public Initiatives: Working Together in Health and Education. The World Bank, March 17,1995:1-10. Available at: http://www.worldbank.org/html/extdr/hnp/health/hlt_svcs/pack1.htm Accessed: Jan. 11, 2004.
33. Yassin K, Galan A. Inequalities in Health as Assessed by the Burden of Disease Method - Case Study. In: Bjegović V, Donev D, eds. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2004:163-9.
34. Murray CJ, Lopez AD. Progress and directions in refining the global burden of disease approach: a response to Williams. Health Economics. 2000;9:69-82.
35. Brenzel L. Selecting an essential package of health services using cost-effectiveness analysis: a manual for professionals in developing countries. Washington, DC, Boston, MA: The World Bank, Population, Health and Nutrition Department; Harvard School of Public Health, Department of Population Studies and International Health; 1993. Available from URL: http://www.hsph.harvard.edu/ihsg/publications/pdf/No-5.PDF. Accessed: August 31, 2009.
36. Homedes N. The disability-adjusted life year (DALY): definition, measurement and potential use. Human Capital Development and Operations Policy Working Paper. Washington DC: The World Bank; 1995. Avalible from URL: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/1996/07/01/000009265_3970311114344/Rendered/PDF/multi0page.pdf. Accessed: August 31, 2009.

RECOMMENDED READINGS

1. Jankovic S. Summary measures of population health and their relevance for health policy. In: Galan A, Scintee G, editors. Public Health Strategies. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing

Company; 2005. p.190-207. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
2. Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ. 1994;72:429-445. Available from URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=8062401. Accessed: August 26, 2009.
3. Murray CJL, Lopez AD. Assessing health needs: the Global Burden of Disease Study. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford Textbook of Public Health - Fourth edition. New York: Oxford University Press Inc.; 2004. p.243-54.
4. Prüss-Üstün A, Mathers C, Corvalan C, Woodward A. Introduction and methods: assessing the environmental burden of disease at national and local levels. WHO Environmental Burden of Disease Series, No. 1. Geneva: World Health Organization; 2003. Available from URL: http://www.who.int/quantifying_ehimpacts/publications/9241546204/en/index.html. Accessed: August 26, 2009.
5. World Health Organization, Global Program on Evidence for Health Policy. National burden of disease studies: a practical guide. Geneva: World Health Organization; 2001. Available from URL: Available from URL: http://www.who.int/healthinfo/nationalburdenofdiseasemanual.pdf. Accessed: August 25, 2009.

APPENDIX

Table A1. Standard life expectancy table. Adapted from World Health Organization (21).

	Sex			Sex			Sex	
Age	**Males**	**Females**	**Age**	**Males**	**Females**	**Age**	**Males**	**Females**
0	80,00	82,50	34	46,55	49,36	68	15,15	17,90
1	79,36	81,84	35	45,57	48,38	69	14,36	17,05
2	78,36	80,87	36	44,58	47,41	70	13,58	16,20
3	77,37	79,90	37	43,60	46,44	71	12,89	15,42
4	76,38	78,92	38	42,61	45,47	72	12,21	14,63
5	75,38	77,95	39	41,63	44,50	73	11,53	13,85
6	74,39	76,96	40	40,64	43,53	74	10,85	13,06
7	73,39	75,97	41	39,67	42,57	75	10,17	12,28
8	72,39	74,97	42	38,69	41,61	76	9,62	11,60
9	71,40	73,98	43	37,72	40,64	77	9,08	10,93
10	70,40	72,99	44	36,74	39,68	78	8,53	10,25
11	69,40	72,00	45	35,77	38,72	79	7,99	9,58
12	68,41	71,00	46	34,81	37,77	80	7,45	8,90
13	67,41	70,01	47	33,86	36,83	81	7,01	8,36
14	66,41	69,01	48	32,90	35,88	82	6,56	7,83
15	65,41	68,02	49	31,95	34,94	83	6,12	7,29
16	64,42	67,03	50	30,99	33,99	84	5,68	6,76
17	63,42	66,04	51	30,06	33,07	85	5,24	6,22
18	62,43	65,06	52	29,12	32,14	86	4,90	5,83
19	61,43	64,07	53	28,19	31,22	87	4,56	5,43
20	60,44	63,08	54	27,26	30,29	88	4,22	5,04
21	59,44	62,10	55	26,32	29,37	89	3,88	4,64
22	58,45	61,12	56	25,42	28,46	90	3,54	4,25
23	57,46	60,13	57	24,52	27,55	91	3,30	3,98
24	56,46	59,15	58	23,61	26,65	92	3,05	3,71
25	55,47	58,17	59	22,71	25,74	93	2,80	3,43
26	54,48	57,19	60	21,81	24,83	94	2,56	3,16
27	53,49	56,21	61	20,95	23,95	95	2,31	2,89
28	52,50	55,23	62	20,09	23,07	96	2,14	2,71
29	51,50	54,25	63	19,22	22,20	97	1,97	2,53
30	50,51	53,27	64	18,36	21,32	98	1,80	2,36
31	49,52	52,29	65	17,50	20,44	99	1,63	2,18
32	48,53	51,31	66	16,71	19,59	100	1,46	2,00
33	47,54	50,34	67	15,93	18,74			

Table A2. Average disability weights (ADW) for diseases and conditions including cancers and injuries. Adapted from World Health Organization[21] (25,29).

Disease/Sequela	ADW	Disease/Sequela	ADW
Diseases and conditions except cancers and injuries			
TUBERCULOSIS		DIARRHOEAL DISEASES	
Cases	0.271	Diarrhoeal diseases - episodes	0.105
SYPHILIS		PERTUSSIS	
Congenital syphilis	0.315	Episodes	0.137
Primary	0.015	Encephalopathy	0.452
Secondary	0.048		
Tertiary - Neurologic	0.283	POLIOMYELITIS	
		Poliomyelitis - Cases - lameness	0.369
CHLAMYDIA			
Ophthalmia neonatorum	0.180	DIPHTHERIA	
Cervicitis	0.049	Episodes	0.231
Neonatal pneumonia	0.280	Neurological complications	0.078
Pelvic inflammatory disease	0.329	Myocarditis	0.323
Ectopic pregnancy	0.549		
Tubo-ovarian abscess	0.548	MEASLES - EPISODES	
Chronic pelvic pain	0.122	Measles - Episodes	0.152
Infertility	0.180		
Symptomatic urethritis	0.067	TETANUS	
Epididymitis	0.167	Episodes	0.638
GONORRHOEA		BACTERIAL MENINGITIS,	
Ophthalmia neonatorum	0.180	MENINGOCOCCAEMIA	
Corneal scar - Blindness	0.600	Streptococcus pneumoniae - Episodes	0.615
Corneal scar - Low vision	0.233	Haemophilus influenzae - Episodes	0.616
Cervicitis	0.049	Neisseria meningitidis - Episodes	0.615
		Meningococcaemia without	0.152
Pelvic inflammatory disease	0.169	meningitis - Episodes	
Ectopic pregnancy	0.549	Deafness	0.229
Tubo-ovarian abscess	0.548	Seizure disorder	0.100
Chronic pelvic pain	0.122	Motor deficit	0.381
Infertility	0.180	Mental retardation	0.459
Symptomatic urethritis	0.067		
Epididymitis	0.167	HEPATITIS B AND HEPATITIS C	
Stricture	0.151	Hepatitis B - Episodes	0.075
		Hepatitis C - Episodes	0.075
HIV			
Cases	0.135	MALARIA	
AIDS cases not on ART	0.505	Episodes	0.191
AIDS cases on ART	0.167	Anaemia	0.012
		Neurological sequelae	0.471

[21] Only average disability weights are given in this table. Many of sequelas varies with age, and many varies also with treatment. For details please see special WHO document available on the Internet (25), and a special WHO web page on the Internet (20,26).

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
TRYPANOSOMIASIS		Cotemporaneous cognitive deficit	0.006
Episodes	0.350	Cognitive impairment	0.463
		Intestinal obstruction	0.024
CHAGAS DISEASE			
Infection	0.000	TRICHURIASIS	
Cardiomyopathy without congestive heart failure	0.062	High intensity infection	0.000
		Cotemporaneous cognitive deficit	0.006
Cardiomyopathy with congestive heart	0.270	Massive dysentery syndrome	0.116
Failure		Cognitive impairment	0.024
Megaviscera	0.240		
		HOOKWORM DISEASE -	
SCHISTOSOMIASIS		ANCYLOSTOMIASIS AND	
Infection	0.005	NECATORIASIS	
Advanced renal disease	0.104	High intensity infection	0.006
Advancerd hepatic disease	0.104	Anaemia	0.024
		Cognitive impairment	0.024
LEISHMANIASIS			
Visceral	0.243	LOWER RESPIRATORY	
Cutaneous	0.023	INFECTIONS	
		Episodes	0.279
LYMPHATIC FILARIASIS		Chronic sequelae	0.099
Hydrocele >15 cm	0.073		
Bancroftian lymphoedema	0.106	UPPER RESPIRATORY	
Brugian lymphoedema	0.116	INFECTIONS	
		Episodes	0.000
ONCHOCERCIASIS		Pharyngitis	0.070
Blindness	0.594		
Itching	0.068	OTITIS MEDIA	
Low vision	0.170	Episodes	0.023
		Deafness	0.229
LEPROSY			
Cases	0.000	MATERNAL HAEMORRHAGE	
Disabling leprosy	0.152	Episodes	0.000
		Severe anaemia	0.093
DENGUE			
Dengue fever	0.197	MATERNAL SEPSIS	
Dengue haemorrhagic fever	0.545	Episodes	0.000
		Infertility	0.180
JAPANESE ENCEPHALITIS			
Episodes	0.616	HYPERTENSIVE DISORDERS OF	
Cognitive impairment	0.468	PREGNANCY	
Neurological sequelae	0.379	Episodes	0.000
TRACHOMA		OBSTRUCTED LABOUR	
Blindness	0.581	Episodes	0.000
Low Vision	0.170	Stress incontinence	0.025
		Rectovaginal fistula	0.430
ASCARIASIS			
High intensity infection	0.000		

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
ABORTION		UNIPOLAR DEPRESSION	
Episodes	0.000	DISORDERD	
Infertility	0.180	Mild depressive episode	0.140
Reproductive tract infection	0.067	Moderate depressive episode	0.350
		Severe depressive episode	0.760
OTHER MATERNAL		Dysthymia	0.140
CONDITIONS			
Stress incontinence	0.025	BIPOLAR DISORDER	
		Cases	0.367
LOW BIRTH WEIGHT - ALL			
SEQUELAE		SCHIZOPHRENIA	
All sequelae	0.106	Cases	0.528
BIRTH ASPHYXIA AND BIRTH		EPILEPSY	
TRAUMA -		Cases	0.113
All sequelae	0.372		
		ALCOHOL USE DISORDERS	
PROTEIN-ENERGY		Cases	0.134
MALNUTRITION			
Wasting	0.053	ALZHEIMER AND OTHER	
Stunting	0.002	DEMENTIAS	
Developmental disability	0.024	Cases	0.666
IODINE DEFICIENCY		PARKINSON DISEASE	
Total goitre rate (G1 + G2)	0.000	Cases	0.351
Mild developmental disability	0.006		
Cretinoidism	0.255	MULTIPLE SCLEROSIS	
Cretinism	0.804	Cases	0.411
VITAMIN A DEFICIENCY		DRUG USE DISORDERS	
Xerophthalmia	0.000	Cases	0.252
Corneal scar	0.277		
		POST-TRAUMATIC STRESS	
IRON-DEFICIENCY ANAEMIA		DISORDER	
Mild	0.000	Cases	0.105
Moderate	0.011		
Severe	0.090	OBSESSIVE-COMPULSIVE	
Very severe	0.249	DISORDERS	
Cognitive impairment	0.024	Cases	0.127
DIABETES MELLITUS		PANIC DISORDER	
Cases	0.015	Cases	0.165
Diabetic foot	0.133		
Neuropathy	0.072	INSOMNIA (PRIMARY)	
Retinopathy - Blindness	0.552	Cases	0.100
Amputation	0.102		
		MIGRAINE	
		Cases	0.029

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
MILD MENTAL RETARDATION ATTRIBUTABLE TO LEAD EXPOSURE		APPENDICITIS	
Cases	0.361	Episodes	0.463
		NEPHRITIS AND NEPHROSIS	
GLAUCOMA		Acute glomerulonephritis	0.091
Low vision	0.170	End-stage renal disease	0.098
Blindness	0.600		
		BENIGN PROSTATIC HYPERTROPHY	
CATARACTS		Symptomatic cases	0.038
Low vision	0.170		
Blindness	0.570	SKIN DISEASES	
		Cases	0.056
HEARING LOSS, ADULT ONSET			
Mild	0.000	RHEUMATOID ARTHRITIS	
Moderate, untreated	0.120	Cases	0.199
Severe or profound, untreated	0.333		
		OSTEOARTHRITIS	
RHEUMATIC HEART DISEASE		Hip	0.126
Cases	0.253	Knee	0.129
HYPERTENSIVE HEART DISEASE		GOUT	
Cases	0.246	Cases	0.132
ISCHAEMIC HEART DISEASE		LOW BACK PAIN	
Acute myocardial infarction	0.439	Episode of limiting low back pain	0.061
Angina pectoris	0.124	Episode of intervertebral disc displacement or herniation	0.061
Congestive heart failure	0.201	Chronic intervertebral disc	0.121
CEREBROVASCULAR DISEASE		CONGENITAL ANOMALIES	
First-ever stroke	0.920	Abdominal wall defect - Cases	0.850
Long-term stroke survivors	0.266	Anencephaly - Cases	0.850
		Anorectal atresia - Cases	0.850
INFLAMMATORY HEART DISEASES		Cleft lip - Cases	0.050
		Cleft palate - Cases	0.103
All sequelae	0.252	Oesophageal atresia - Cases	0.850
		Renal agenesis - Cases	0.850
COPD		Down syndrome - Cases	0.593
Mild and moderate symptomatic cases	0.170	Congenital heart anomalies - Cases	0.323
Severe symptomatic cases	0.530	Spina bifida - Cases	0.593
ASTHMA		DENTAL DISEASES	
Cases	0.043	Dental caries - Cases	0.081
		Periodontal disease - Cases	0.001
PEPTIC ULCER		Edentulism - Cases	0.025
Cases not treated with antibiotic	0.042		
CIRRHOSIS OF THE LIVER			
Symptomatic cases	0.330		

Table A2. Cont.

Disease/Sequela	ADW	Disease/Sequela	ADW
Malignant neoplasms			
CANCERS (SITES)		Corpus uteri	0.081
Mouth and oropharynx	0.118	Ovary	0.084
Oesophagus	0.217	Prostate	0.124
Stomach	0.217	Bladder	0.086
Colon and rectum	0.217	Lymphomas and multiple myeloma	0.073
Liver	0.239	Leukaemia	0.098
Pancreas	0.269		
Trachea, bronchus and lung	0.146	CANCERS - METASTASIS	0.750
Melanoma and other skin	0.045	STAGE (ALL SITES)	
Breast	0.078		
Cervix uteri	0.071	CANCERS TERMINAL (ALL SITES)	0.809
Injuries			
POISONING		INTERNAL INJURIES	
Short term	0.609	Short term	0.208
FRACTURES		OPEN WOUND	
Skull - Short term	0.431	Short term	0.108
Skull - Long term	0.384		
Face bones	0.223	INJURY TO EYES	
Vertebral column	0.266	Short term	0.108
		Long term	0.354
INJURED SPINAL CORD			
Injured spinal cord	0.725	AMPUTATIONS	
		Thumb	0.165
FRACTURES		Finger	0.102
Rib or sternum	0.199	Arm	0.308
Pelvis	0.247	Toe	0.102
Clavicle, scapula, or humerus	0.153	Foot	0.300
Radius or ulna	0.180	Leg	0.300
Hand bones	0.100		
Femur - Short term	0.372	CRUSHING	
Femur - Long term	0.272	Short term	0.218
Patella, tibia, or fibula	0.271		
Ankle	0.196	BURNS	
Foot bones	0.077	<20% - Short term	0.156
		<20% - Long term	0.002
DISLOCATIONS		>20% and <60% - Short term	0.469
Dislocated shoulder, elbow, or hip	0.074	>20% and <60% - Long term	0.255
Other dislocation	0.074	>60% - Short term	0.469
		>60% - Long term	0.255
SPRAINS			
Short term	0.064	INJURED NERVES	
		Short term	0.078
INTRACRANIAL INJURY		Long term	0.078
Short term	0.359		
Long term	0.359		

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**TEST VALIDITY MEASURES AND RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYSIS**
Module: 4.2.1	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Jadranka Božikov**, PhD, Professor Andrija Štampar School of Public Health, School of Medicine, University of Zagreb, Croatia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Jadranka Božikov** Andrija Štampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: jbozikov@snz.hr
Keywords	tests, validity, sensitivity, specificity, false positive rate, false negative rate, positive predictive value, negative predictive value, ROC analysis
Learning objectives	After completing this module students should: • understand • understand measures of tests validity and differences between them; • know how to calculate these measures and be capable to calculate them by themselves; • understand principles of ROC analysis, and how basic measures of tests validity are related to ROC analysis.
Abstract	Diagnosis is based on the results of diagnostic tests. Most of them are imperfect instruments, and make errors in both directions - a healthy individual can be classified as diseased, and vice versa. Ability of each diagnostic test to correctly classify patients as diseased or healthy is called validity of test. Concept apply also in screening tests. There exist several nosological (sensitivity, specificity, etc.) and diagnostic measures (ositive predictive value, negative predictive value) to assess validity of a test with a binary outcome. In other tests ROC method could be used as a method of analysis.
Teaching methods	Teaching methods include introductory lecture, exercises, and interactive methods such as small group discussions. Students after introductory lectures first carefully read the recommended sources. Afterwards they discuss the issue of tests validity measures. In continuation, they in practice in groups of 2-3 students perform the procedure of calculation of all different measures of tests validity using the programme tool (e.g. MS Excel) on given data. At the end they compare and discuss their results.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a computer room; • equipment: computers (1 computer per 2-3 students), LCD projection, access to the Internet; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaire (MCQ) and case-study.

TEST VALIDITY MEASURES AND RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYSIS

Jadranka Božikov, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Introduction

Diagnosis is based on the results of diagnostic tests (in the broadest meaning of that term). Most of these tests are imperfect instruments, and make errors in both directions - a healthy individual can be classified as diseased, and diseased as healthy. Ability of each diagnostic test to correctly classify patients as diseased or healthy is called **validity of test.** Assessing validity of a tests is especially important at introduction of new diagnostic procedures (1-3).

Concept of test validity apply also in population studies, in the screenings of populations.

Test validity measures

Suppose that a diagnostic or screening test under observation provides us the binary outcome - the disease is present (usually referred as »positive«) or the disease is not present (usually referred as »negative«). Therefore, »positive« means a greater probability of disease, whereas »negative« a greater probability of absence of disease (1-4). Of course we are interested in:

- how well the patients with disease are recognized by the test, or
- how well the test indicates whether the disease is really present.

For answering these two questions it intuitively follows that the test results are compared with the actual situation. Relation of results produced by some diagnostic to the actual state is presented in 2×2 table of contingencies (Figure 1), also called the **decision matrix**.

		Disease	
		Present	Absent
Test	Positive	a	b
	Negative	c	d

Figure 1. Relation of results produced by a test to the actual state in the decision matrix.

Box "a" represents number of diseased examinees that are correctly recognized by the test as diseased. We say that the test recognize these examinees as **true positive** (TP). Box "b" represents number of healthy examinees that are incorrectly recognized by the test as diseased - **false positive** (FP). Box "c" represents number of diseased examinees that are incorrectly recognized by the test as healthy - **false negative** (FN). Box "d" represents number of healthy examinees that are correctly recognized by the test as healthy - **true negative** (TN). Figure 2 is presenting TP, FP, FN and TN classifications which are in fact absolute frequency measures of performance of a test.

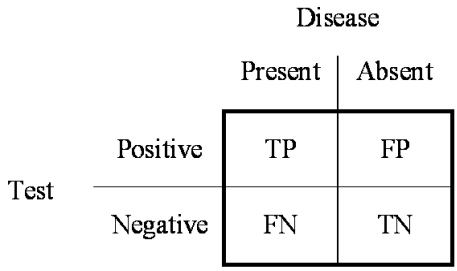


Figure 2. Relation of results produced by a test to the actual state with absolute measures of performance of a test in the decision matrix. Legend: TP – true positive classifications, FP – false positive classifications, FN – false negative classifications, TN – true negative classifications.

However, to answer the above raised two questions we need to form relative measures. Consequently, we need to complete the contingency table (the decision matrix) from Figure 1 and Figure 2 with marginal totals (Figure 3).

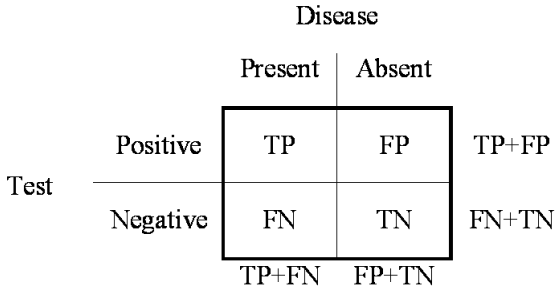


Figure 3. Complete decision matrix for calculating relative measures of performance of a test.

To answer to the question »how well the patients with disease are recognized by the test« so called »nosological test validity measures« need to be formed, while to

answer to the question »how well the test indicates whether the disease is really present« so called »diagnostic test validity measures«.

Nosological test validity measures and nosological probability

Nosological test validity measures are those validity measures where test results are compared to actual situation of the disease (1,2). There exist four nosological test validity measures:

- **sensitivity** or true positive rate (TPR) – **nosological sensitivity** of test is proportion of sick people that test correctly recognizes as diseased (test-positive) of the total number of really diseased (Equation 1).

$$sensitivity = \frac{TP}{TP + FN} \qquad \textbf{Equation 1.}$$

- **specificity** or true negative rate (TNR) – **nosological specificity** of test is proportions of healthy people that test correctly identified of the total number of really healthy (Equation 2).

$$specificity = \frac{TN}{TN + FP} \qquad \textbf{Equation 2.}$$

Sensitivity and specificity are two main nosological measures of the validity of the test. The other two proportions are false positive and false negative ratio

- **false positive rate (FPR)** – false positive rate is a proportion of healthy that test incorrectly classified as diseased (Equation 3). False positive rate is at the same time 1-sensitivity (Equation 3).

$$FPR = \frac{FN}{TP + FN} = 1 - sensitivity \qquad \textbf{Equation 3.}$$

- **false negative rate (FNR)** – false negative rate is a proportion of the diseased that test wrongly placed as healthy (Equation 4). False negative rate is at the same time 1-specificity (Equation 4).

$$FNR = \frac{FN}{TP + FN} = 1 - specificity \qquad \textbf{Equation 4.}$$

All these equations could be seen also from the probability point of view. Let's mark events in following way:

- B means the presence of disease (the event »be diseased«)

- B' is the absence of disease (the event »not to be diseased« i.e., »be healthy«)
- indicates the presence of features (symptoms) or a positive test result
- O' is absence of features (symptoms) or a negative test result

From definitions for measures of validity of test and the concept of conditional probability (module 1.1.2) obviously sensitivity, specificity, FNR and FPR are (Equations 5-8):

$$sensitivity = P(O \mid B) = \frac{P(O \cap B)}{P(B)} \quad \textbf{Equation 5.}$$

$$specificity = P(O' \mid B') = \frac{P(O' \cap B')}{P(B')} \quad \textbf{Equation 6.}$$

$$FNR = P(O' \mid B) = P(O \mid B) \quad \textbf{Equation 7.}$$

$$FPR = P(O \mid B') = P(O' \mid B') \quad \textbf{Equation 8.}$$

It should be noticed that it is necessary and sufficient to know two of four listed nosological probabilities (exactly two): for example, sensitivity, and specificity, because the other two are opposite probabilities.

Diagnostic test validity measures and diagnostic probability

Of course, medical doctors would be more interested in diagnostic test validity measures and hence diagnostic probabilities. Diagnostic test validity measures are those validity measures where actual situation of the disease is compared to test results (1,2). There exist two diagnostic test validity measures:

- **positive predictive value (PPV)** - positive predictive value represents proportion of really diseased of those who are positive on the test (Equation 9). PPV is known also as **diagnostic specificity**. This often causes confusion, especially because the attribute »diagnostic« often tends to be lost.

$$PPV = \frac{TP}{TP + FP} \quad \textbf{Equation 9.}$$

- **negative predictive value (NPV)** - negative predictive value represents proportion of real healthy individuals among individuals with negative test results. (Equation 10). NPV is known also as **diagnostic sensitivity**. Again, this often causes confusion, especially because the attribute »diagnostic« often tends to be lost.

$$NPV = \frac{TN}{FN + TN}$$ **Equation 10.**

From the probability point of view diagnostic probabilities are probabilities of presence/absence of disease for positive or negative test results. According to the concept of conditional probability and Bayes' theorem (module Probability - basic concepts) obviously PPV and NPV are (Equations 11 and 12):

$$PPV = P(B \mid O) = \frac{P(O \mid B) \cdot P(B)}{P(O \mid B) \times P(B) + P(O \mid B') \times P(B')}$$ **Equation 11.**

$$NPV = P(B' \mid O') = \frac{P(O' \mid B') \cdot P(B')}{P(O' \mid B) \times P(B) + P(O' \mid B') \times P(B')}$$ **Equation 12.**

These parameters are **influenced by disease prevalence in the observed population** (on nonconditional probabilities P(B)). Thus PPV (Equation 13) and NPV could be expressed also as (Equation 14):

$$P(B \mid O) = \frac{P(O \mid B) \times P(B)}{P(O \mid B) \times P(B) + P(O \mid B') \times P(B')} =$$ **Equation 13.**

$$= \frac{P(O \mid B) \times P(B)}{P(O \mid B) \times P(B) + (1 - P(O' \mid B')) \times (1 - P(B))} =$$

$$= \frac{sensitivity \times prevalence}{sensitivity \times prevalence + (1 - specificity) \times (1 - prevalence)}$$

$$P(B' \mid O') = \frac{P(O' \mid B') \times P(B')}{P(O' \mid B') \times P(B') + P(O' \mid B) \times P(B)} =$$ **Equation 14.**

$$= \frac{P(O' \mid B') \times P(B')}{P(O' \mid B') \times (1 - P(B)) + (1 - P(O \mid B)) \times P(B)} =$$

$$= \frac{specificity \times (1 - prevalence)}{specificity \times (1 - prevalence) + (1 - sensitivity) \times prevalence}$$

Absolute and relative test accuracy

With test evaluation also go terms **absolute test accuracy (ATA)** (Equation 15) **and relative test accuracy (RTA)** (Equation 16), among which the first is and the second is not under the influence of disease prevalence in the tested population.

$$ATA = \frac{TP + TN}{TP + FP + FN + TN} \qquad \textbf{Equation 15.}$$

$$RTA = \frac{\left(\frac{TP}{TP + FP} + \frac{TN}{FN + TP} \right)}{2} \qquad \textbf{Equation 16.}$$

Mathematical basis of the above measures (which are used also for evaluation of expert systems) can be found in probability theory.

Receiver Operating Characteristic (ROC) analysis

However, many screening and diagnostic procedures in medicine does not only have two possible outputs - the disease is present or the disease is not present, but several values that can be measured on ordinal or even continuous scale of values. In such kind of diagnostic/screening test, we must first put the cut-off point on the scale of values in which to put the decision of a positive or negative result of test. In this cut-off point the decision matrix is constructed (3,5-7).

Often it happens that we cannot just immediately put the best cut-off point. In this case we can put more cut-off points and in each of them we construct a decision matrix. In each, sensitivity (TPR), specificity (TNR), FNR and FPR are calculated. By varying the cut-off point these proportions change (Figure 4).

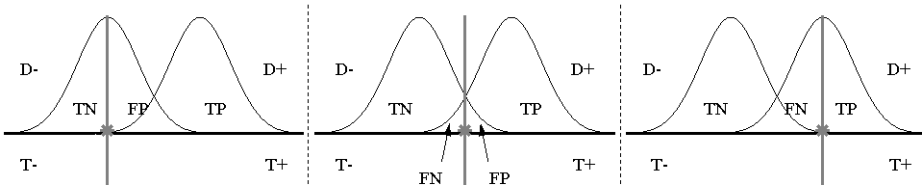


Figure 4. Change of the the proportions of TP, TN, FP and FN test results as a function of changing the cut-off point for decision. LEGEND: * = the position of cut-off point, TN = true negative test results, TP = true positive test results, FN = false negative test results, FP = false positive test results, D- = disease not present, D+ = disease present, T- = negative test result, T+ = positive test result.

Choosing the best possible cut-off point depends on the cost-benefit associated with the classification of patients with the disease among those who do not have the disease, compared with the classification of healthy as those with the disease:

- when discovering the disease, which, if untreated, is mortal, we will gladly accept a slightly higher FPR, because we thus ensure that the TPR will be closer to 100%. But we must also consider the fact that people who would be labeled as diseased, but in fact they would be not, could suffer intangible costs (for example stress),
- in less serious diseases, or very expensive treatments, we would be interested in lowering the value of TPR on account of minimizing the FPR.

However, one can support his/her decision about best possible cut-off point by calculating special measures. One of them will be discussed at the end of the theoretical part of the module.

ROC curve

Analysis of sensitivity and specificity of the test, depending on setting borders is known as the **Receiver Operating Characteristic (ROC) analysis**. To deal with these multiple pairs of sensitivity and specificity values, we can draw a diagram using the **sensitivities as the y coordinates** and the **FPRs** (**1-specificities) as the x coordinates** (Figure 5) (3,5-7).

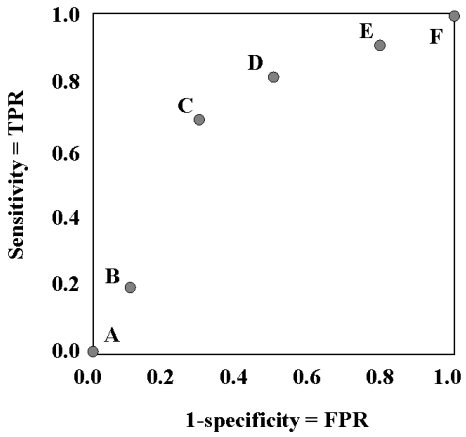


Figure 5. Receiver Operating Characteristic (ROC) diagram made of four cut-off points (A-F).
LEGEND: TPR = true positive rate, FPR = false positive rate.

The points in the ROC diagram could be fitted with a curve which can be smoothed (Figure 6).

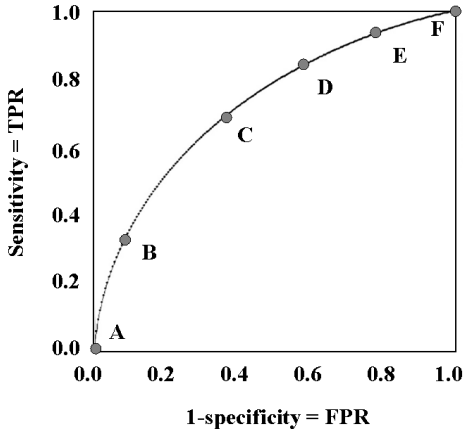


Figure 6. Fitted and smoothed Receiver Operating Characteristic (ROC) curve. LEGEND: TPR = true positive rate, FPR = false positive rate.

A good and very informative diagnostic/screening test is characterized by high sensitivity (TPR) values and low FPR values in all possible cut-off points. ROC curve of such a test is shifted to the left upper corner (Figure 7, curve W). A perfect test (Figure 7, curve W) has an area under the ROC curve of 1. By contrast, the ROC curve of poor diagnostic test is approaching the diagonal connecting the lower left corner of the image by right-upper one.

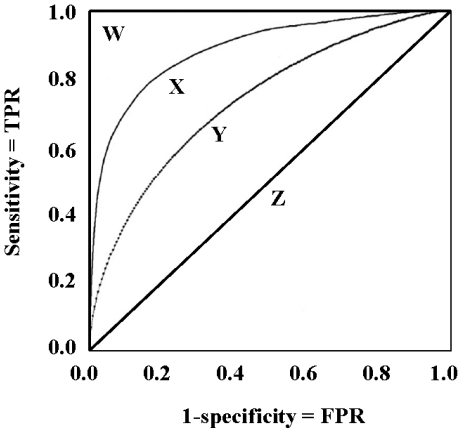


Figure 7. Fitted Receiver Operating Characteristic (ROC) curves for diagnostic tests, with varying degrees of informativity (W = very informative test, X and Y = more/less moderately informative tests, Z = noninformative test). TPR = true positive rate, FPR = false positive rate.

On this diagonal (Figure 7, curve Z - diagonal), the sensitivity (TPR) values and FPR values in each cut-off points are the same.

Area under ROC curve

All this is reflected in the most important summary measure related to ROC curves – the **area under the curve** (AUC) (7,8). AUC is measuring the overall performance (accuracy, informativity) of a diagnostic test and is interpreted as the average value of sensitivity for all possible values of specificity. The AUC is calculated using different methods, the simplest being the trapezoid rule (7). In this process AUC is being divided into several parts, depending on how many measurement points we have (Figure 8).

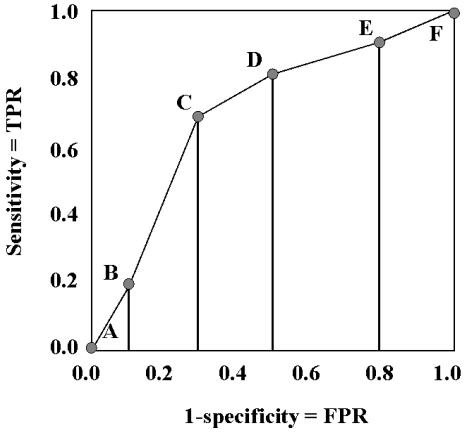


Figure 8. Graphical representation as the basis of calculating area under Receiver Operating Characteristic (ROC) curve using trapezoid rule.

All parts are added together at the end.

By using trapezoid rule, AUC, which is usually designated with greek letter θ, is in practice calculated as (Equation 17):

$$\theta = \theta_{1-0} + \theta_{2-1} + ... + \theta_{(i+1)-i} =$$
$$= \left[\left(\mathrm{FPR}_1 - \mathrm{FPR}_0 \right) \times \left(\frac{\mathrm{TPR}_0 + \mathrm{TPR}_1}{2} \right) \right] +$$
$$+ \left[\left(\mathrm{FPR}_2 - \mathrm{FPR}_1 \right) \times \left(\frac{\mathrm{TPR}_1 + \mathrm{TPR}_2}{2} \right) \right] + \ldots$$
$$\ldots + \left[\left(\mathrm{FPR}_{i+1} - \mathrm{FPR}_i \right) \times \left(\frac{\mathrm{TPR}_i + \mathrm{TPR}_{i+1}}{2} \right) \right]$$

Equation 17.

For the case presented in Figure 8 the Equation 17 reads as (Equation 18):

$$\begin{aligned}\theta &= \theta_{B-A} + \theta_{C-B} + ... + \theta_{F-E} = \\ &= \left[\left(\mathrm{FPR_B} - \mathrm{FPR_A}\right) \times \left(\frac{\mathrm{TPR_A} + \mathrm{TPR_B}}{2} \right) \right] + \\ &+ \left[\left(\mathrm{FPR_C} - \mathrm{FPR_B}\right) \times \left(\frac{\mathrm{TPR_B} + \mathrm{TPR_C}}{2} \right) \right] + ... \\ &...+ \left[\left(\mathrm{FPR}_F - \mathrm{FPR_E}\right) \times \left(\frac{\mathrm{TPR_E} + \mathrm{TPR}_F}{2} \right) \right]\end{aligned}$$

Equation 18.

The problem of this method is that it gives an underestimates estimate of AUC. It should be noted that the more points there are, the better estimate of AUC we get. Much better estimate one can get by fitting the data to a binormal model with maximum likelihood estimates (7). However, this method is out of the scope of this module.

AUC can take any value between 0 and 1 (or 0 and 100% respectively). The closer AUC is to 1 (or 100%), the better the overall informativity (diagnostic performance) of the test. In another words, an area of 1 represents a perfect test. On the contrary an area of 0.5 represents a worthless test. This would be a test in which the positive test results were equally likely in both groups of people under investigation - those with the disease and those without it. A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system (9):

- AUC: 0.90-1.00 = excellent performance
- AUC: 0.80-0.90 = good performance
- AUC: 0.70-0.80 = fair performance
- AUC: 0.60-0.70 = poor performance
- AUC: 0.50-0.60 = fail performance

Youden index as a support in decision about best cut-off point

The decision about best possible cut-off point could be supported by calculating special measures. One of them is Youden index (10). This index which is one of the oldest measures for diagnostic accuracy (11), is an index that is a function of nosological sensitivity and nosological specificity. The calculation procedure is rather simple (10) Youden index denoted as *J* maximizes the vertical distance from line of equality (Figure 7, curve Z - diagonal) to certain point at the ROC curve (Equation 19):

$$J = \text{maximum } \left\{ \left(\text{sensitivity} + \text{specificity} \right) - 1 \right\}$$

Equation 19.

This is shown also in Figure 9.

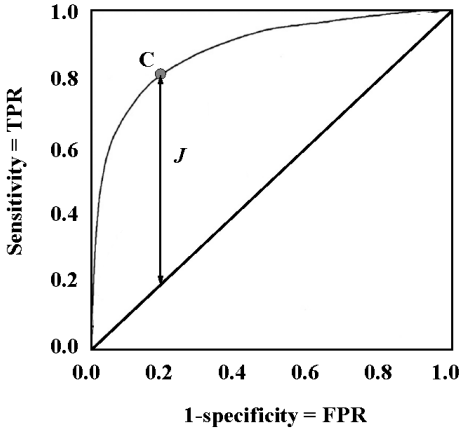


Figure 9. Finding best cut-off from the ROC curve by using Youden index *J*. LEGEND: C = the best cut-off point on the ROC curve.

The Equation 19 could be writtes also as (Equation 20):

$$J = \text{maximum } \{(\text{TPR} + \text{TNR}) - 1\} \qquad \textbf{Equation 20.}$$

We could calculate it also by using the quantities used in graphing ROC curve. In this case the equation reads as (Equation 21):

$$J = \text{maximum } \{\text{TPR} - \text{FPR}\} \qquad \textbf{Equation 21.}$$

The Youden index ranges between 0 and 1. values close to 1 indicate that the test's effectiveness is rather large, and values close to 0 that the test's effectiveness is rather small.

When we have a series of cut-off points, for each of them the the difference between TPR and FPR is calculated. Youden index *J* is the highest value of this difference. The point with the maximal difference s according to this criterion the best cut-off point.

CASE STUDIES

Case study 1: Test validity measures

Impact of disease prevalence on positive and negative predictive values

For illustration of the impact of disease prevalence in the population (sample) on PPV and NPV we will use two virtual sets of data representing two virtual

situations. In both of them sensitivity and specificity of the test are exactly the same. Let's suppose that **sensitivity is 0.90** (90%) and **specificity is 0.95** (95%).

The first case is the case of open population in which the observed **prevalence of disease is 0.01** (1%). In Table 1 the situation in which sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%) is presented in details.

Table 1. The case of open population in which the observed prevalence of disease is 0.01 (1%), and sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%).

		Disease		
		Yes	**No**	**Total**
Test	**Positive**	90	495	585
	Negative	10	9405	9415
	Total	100	9900	10000

The result of calculation of two nosological and two diagnostic measures of test validity for data set presented in Table 1 is presented in Table 2.

Table 2. The result of calculation of two nosological and two diagnostic measures of test validity in the case presented in Table 1.

Measure	**Calculation**	**Result**
Prevalence	100/10000	0.010 (1.00%)
Sensitivity	90/100	0.900 (90.0%)
Specificity	9405/9900	0.950 (95.0%)
Positive predictive value	90/585	0.154 (15.4%)
Negative predictive value	9405/9415	0.999 (99.9%)

The second case is the case of hospital population where disease prevalence is 0.60 (60%). In Table 3 the situation in which sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%) is presented in details.

Table 3. The case of hospital population in which the observed prevalence of disease is 0,60 (60%), and sensitivity is set to 0.90 (90%) and specificity to 0.95 (95%).

		Disease		
		Yes	**No**	**Total**
Test	**Positive**	54	2	56
	Negative	6	38	44
	Total	60	40	100

The result of calculation of two nosological and two diagnostic measures of test validity for data set presented in Table 3 is presented in Table 4.

The comparison shows that the PPV in the second case is much, much higher than in the first case, although the sensitivity and specificity are exactly the same.

Table 4. The result of calculation of two nosological and two diagnostic measures of test validity in the case presented in Table 3.

Measure	Calculation	Result
Prevalence	60/100	0.600 (60.0%)
Sensitivity	54/60	0.900 (90.0%)
Specificity	38/40	0.950 (95.0%)
Positive predictive value	54/56	0.964 (96.4%)
Negative predictive value	38/44	0.864 (86.4%)

Case study 2: Receiver Operating Characteristic (ROC) analysis

The case of the test of blood glucose level

Curves in Figure 10 show the distribution of blood glucose level measured two hours after meal in healthy people and people suffering from diabetes.

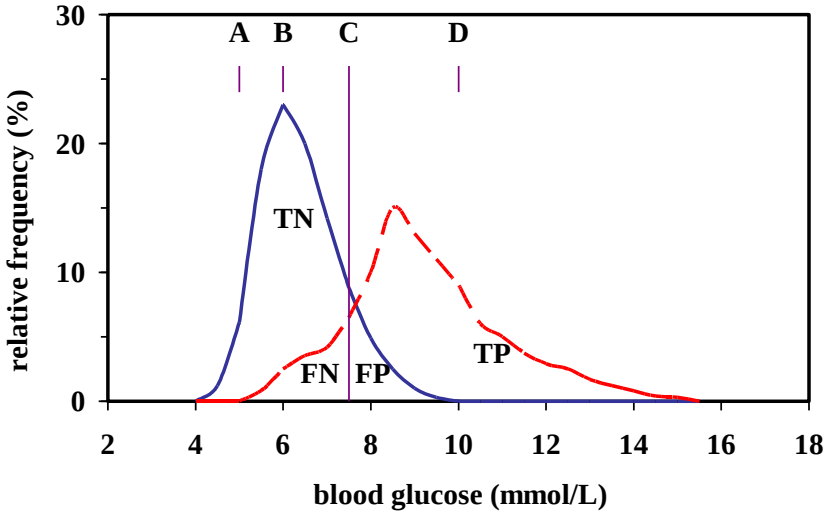


Figure 10. The distribution of blood glucose level measured two hours after meal in healthy people (blue solid line) and people suffering from diabetes (red dashed line). LEGEND: A, B, C and D = the position of four possible cut-off points, TN = true negative test results, TP = true positive test results, FN = false negative test results, FP = false positive test results.

Capital letters (A, B, C and D) in Figure 9 denote four possible cut-off points for final clinical decision. If, for example, in screening of population for diabetes limits are set very low (criterion A), the test will be very sensitive and detect all diseased people with very low specificity (many FP that request additional diagnostic tests).

With increase of the limit (criteria) sensitivity will decrease, and specificity will increase (1-specificity or FPR will decrease).

In continuation a ROC diagram made of four cut-off points with fitted ROC curve (Figure 11) is constructed.

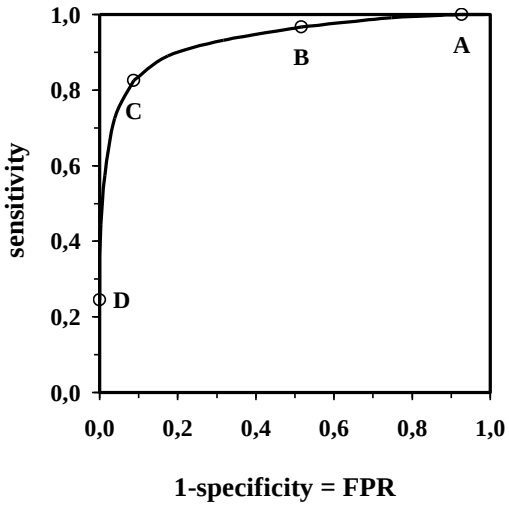


Figure 11. Fitted and smoothed Receiver Operating Characteristic (ROC) curve in the case of making decision in blood glucose level test. LEGEND: FPR = false positive rate.

In Table 5 TPRs (sensitivities) and FPRs (1-specificities) in several cut-off points are presented. They can be used to determine the Youden index to find best cut-off point. In the last column of this table, the differences between TPRs and FPRs (by using Equation 21) are presented as well. In cut-off points 9 and 10 the difference has the highest value, indicating the Youden index. One of these two cut-off points are the best cut-off point according to this criterion.

Table 5. TPR (sensitivity) and FPR (1-specificity) in several cut-off points of blood glucose level measured two hours after meal in healthy people and people suffering from diabetes.

Cut-off point	TPR (sensitivity)	FPR (1-specificity)	TPR-FPR
1	1	0,988	0,012
2	1	0,927	0,073
3	1	0,927	0,073
4	0,992	0,747	0,245
5	0,967	0,517	0,45
6	0,967	0,517	0,45
7	0,932	0,317	0,615
8	0,89	0,175	0,715

Table 5. Cont.

Cut-off point	TPR (sensitivity)	FPR (1-specificity)	TPR-FPR
9	**0,825**	**0,087**	**0,738**
10	**0,825**	**0,087**	**0,738**
11	0,725	0,038	0,687
12	0,575	0,013	0,562
13	0,445	0,003	0,442
14	0,335	0	0,335
15	0,245	0	0,245
16	0,245	0	0,245
17	0,135	0	0,135
18	0,135	0	0,135
19	0,098	0	0,098
20	0,069	0	0,069
21	0,044	0	0,044
22	0,027	0	0,027

The case of interpretation of radiocardiograms

Of course, in qualitative diagnostic procedures, we can have only a few discrete values. Perfectly described example can be found in the work of Malčić dealing with classification of findings of radiocardiograms - the graphic record produced by radiocardiography (12).

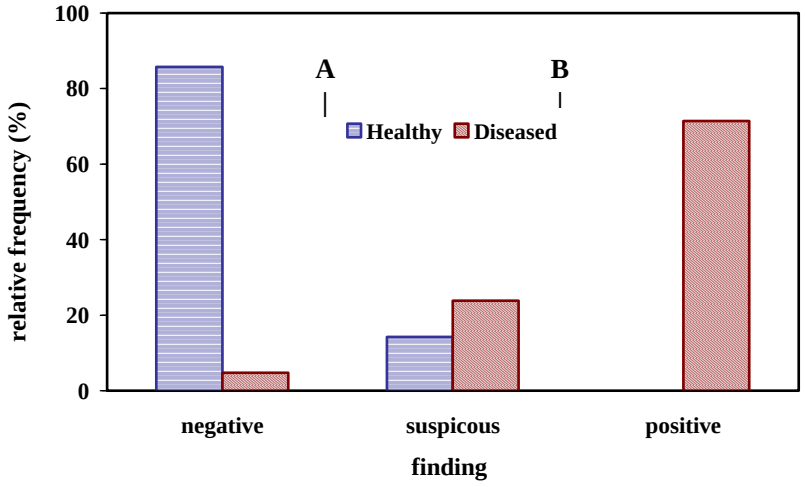


Figure 12. The distribution of findings of radiocardiography that are classified as negative, suspect or positive in relation to the proven existence of shunt (12). LEGEND: A, B = the position of two possible cut-off points.

In the work of Malčić, the value of radiocardiographic testing in diagnostic process of intracardial shunt from left to right is test whose results are compared with results of catheterisation and angiography as the criteria for the existence of shunt in work (12). Figure 12 shows the distribution of findings of radiocardiography that are classified as negative, suspect or positive in relation to the proven existence of shunt. Capital letters (A and B) in Figure 11 denote two possible cut-off points for final clinical decision. Criteria A places all positive and suspect findings as positive and has a sensitivity somewhat lower than 100% (because two of 42 patients have negative radiocardiography findings) and the specificity is 71.4% (4/28 healthy have suspect radiocardiography finding). Criteria B only positive radiocardiography finding consider as positive and it is clearly evident that its sensitivity is 100% and specificity is 85.7% because some diseased examinees were not recognize as diseased.

In continuation a ROC diagram made of four cut-off points with fitted ROC curve (Figure 13) is constructed.

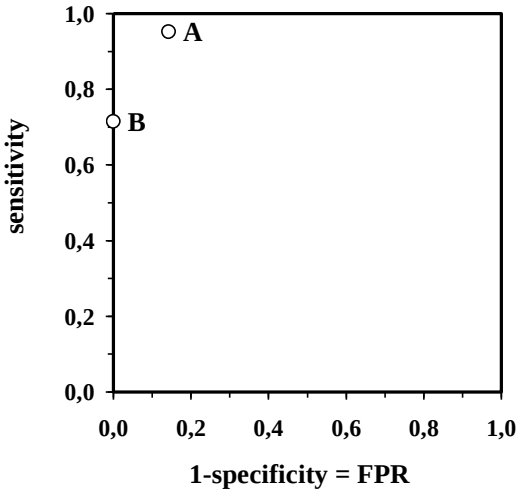


Figure 13. Fitted and smoothed Receiver Operating Characteristic (ROC) curve in the case of classification of findings of radiocardiograms in diagnostic process of intracardial shunt from left to right (12). LEGEND: FPR = false positive rate.

EXERCISE

Task 1

In Table 6, a data set on performance of test T with binary results in diagnosing disease D is presented[22].

[22] This exercise is based upon real data presented in a paper on results in the test for antibody to neutrophil cytoplasmic antigens as a diagnostic aid for Wegener's granulomatosis (13).

Table 6. Performance of test T with binary results in diagnosing disease D.

		Disease D		
		Yes	**No**	**Total**
Test T	**Positive**	18	3	21
	Negative	5	214	219
	Total	23	217	230

Carefully read the theoretical part and calculate:

1. Nosological sensitivity of the test T
2. Nosological specificity of the test T
3. Diagnostic specificity of the test T
4. Diagnostic sensitivity of the test T

Define:

1. Another name for diagnostic specificity
2. Another name for diagnostic sensitivity

Compare the results to the results of the teacher[23].

Task 2

In Table 7, a data set on performance of test T with several values of results in diagnosing disease D is presented.

[23] **Answers to Task 1:**

The results of calculating rates:

1. Nosological sensitivity:
   18/23 = 0.78 (78%)
2. Nosological specificity:
   214/217 = 0.99 (99%)
3. Diagnostic specificity:
   18/21 = 0.86 (86%)
4. Diagnostic sensitivity:
   214/219 = 0.98 (98%)

Replies to definitions:

1. Positive predictive value
2. Negative predictive value

Table 7. Performance of test T with several values of results in diagnosing disease D.

Number of symptoms	Diseased	Healthy
1	12	48
2	13	30
3	40	20
4	60	10
5	40	2

Sketch the ROC curve (point) if the positive test is considered as the presence of 3 or more symptoms of the disease[24].

Task 3

The group of researchers was interested how good is the performance of test T with several values of results in diagnosing disease D in three different groups of examinees. In Table 8 the results of TPRs and FPRs calculated in several cut-off points are presented.

Table 8. Performance of test T with several values of results in diagnosing disease D in three different groups of examinees.

	Group 1		Group 2		Group 3	
Cut-off point	**TPR**	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**
1	0.000	0.000	0.000	0.000	0.000	0.000
2	0.023	0.000	0.019	0.000	0.027	0.000
3	0.051	0.001	0.036	0.001	0.046	0.001
4	0.067	0.002	0.068	0.002	0.069	0.002
5	0.072	0.003	0.084	0.004	0.082	0.003
6	0.089	0.004	0.093	0.005	0.098	0.004

[24] **Answer to Task 2:**

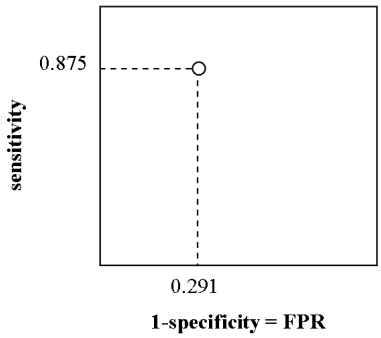


Table 8. Cont.

Cut-off point	Group 1		Group 2		Group 3	
	TPR	**FPR**	**TPR**	**FPR**	**TPR**	**FPR**
7	0.097	0.005	0.098	0.006	0.115	0.005
8	0.104	0.006	0.105	0.007	0.121	0.006
9	0.107	0.007	0.112	0.008	0.135	0.007
10	0.116	0.008	0.117	0.009	0.142	0.008
11	0.120	0.009	0.129	0.010	0.159	0.009
12	0.125	0.010	0.131	0.013	0.169	0.010
13	0.133	0.011	0.138	0.015	0.174	0.011
14	0.142	0.012	0.138	0.016	0.181	0.012
15	0.145	0.013	0.142	0.017	0.186	0.013
16	0.148	0.014	0.170	0.019	0.193	0.015
17	0.168	0.018	0.182	0.020	0.201	0.016
18	0.173	0.019	0.186	0.022	0.210	0.017
19	0.201	0.022	0.207	0.025	0.222	0.019
20	0.207	0.024	0.210	0.028	0.235	0.021
21	0.208	0.025	0.224	0.029	0.247	0.022
22	0.228	0.029	0.235	0.033	0.251	0.024
23	0.233	0.031	0.256	0.037	0.259	0.027
24	0.235	0.032	0.261	0.040	0.270	0.029
25	0.247	0.035	0.263	0.042	0.292	0.034
26	0.255	0.038	0.277	0.050	0.304	0.038
27	0.265	0.043	0.294	0.057	0.318	0.042
28	0.277	0.047	0.305	0.061	0.340	0.049
29	0.292	0.054	0.340	0.070	0.378	0.056
30	0.319	0.067	0.352	0.077	0.399	0.065
31	0.370	0.092	0.378	0.088	0.422	0.074
32	0.378	0.101	0.403	0.101	0.449	0.086
33	0.436	0.132	0.427	0.115	0.476	0.102
34	0.488	0.167	0.462	0.133	0.502	0.123
35	0.593	0.266	0.506	0.162	0.547	0.157
36	0.712	0.437	0.557	0.207	0.594	0.201
37	0.864	0.652	0.597	0.250	0.681	0.276
38	0.924	0.774	0.739	0.417	0.789	0.420
39	1.000	0.996	0.830	0.574	0.923	0.694
40	1.000	1.000	0.911	0.731	0.989	0.917
41			0.967	0.912	1.000	1.000
42			1.000	1.000		

From the data presented in Table 8[25]. make the following:

[25] This exercise is based upon real data used in analysis of the effectiveness of multiple regression, discriminant analysis and logistic regression to identify and evaluate predictors of premature birth (14).

1. Calculate area under ROC curve for all three groups of examinees
2. Interpret calculated AUCs according to rough guide for classifying the accuracy of a diagnostic test.

Compare the results to the results of the teacher[26].

Task 4

For all three groups of examinees presented in Table 8 find Youden index. Compare the results to the results of the teacher[27]

REFERENCES

1. Wulff HR. Kliničko prosuđivanje. Zagreb: Jumena; 1982.
2. Bradley GW. Evaluation of tests. In: Bradley GW, editor. Disease, diagnosis and decisions. Chichester: John Wiley & Sons; 1993. p.70-90.
3. McNeil BJ, Keeler E, Adelstein JS. Primer on certain elements of medical decision making. N Engl J Med. 1975;293:211-215.
4. Ivanković D. Osnove statističke analize za medicinare. Zagreb: Medicinski fakultet Sveučilišta u Zagrebu; 1989.
5. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978;4:283-298.
6. Brismar J. Understanding receiver-operating-characteristic curves: a graphical approach. Am J Roentgenol 1991;157:1119-1121.
7. Kallin Westin L. Receiver operating characteristic (ROC) analysis: valuating discriminance effects among decision support systems. Umeå: Deptartment of Computing Science, Umeå University; 2001. Available from URL: http://www8.cs.umu.se/research/reports/2001/018/part1.pdf. Accessed: March 29, 2009.
8. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
9. Tape TG. Interpreting Diagnostic Tests. Omaha, NE: University of Nebraska Medical Center [Internet book]. Available from URL: http://gim.unmc.edu/dxtests/Default.htm. Accessed: March 29, 2009.

[26] **Answers to Task 3:**

1. The results of calculating AUCs:
   Group 1: $\theta = 0.717$
   Group 2: $\theta = 0.732$
   Group 3: $\theta = 0.771$
2. Interpretation of the accuracy of a diagnostic test: in all three groups the performance of test is fair. However, the best is in Group 3.

[27] **Answer to Task 4:**

The results of calculating Youden index:
Group 1: Cut-off point 35: $J = 0.327$
Group 2: Cut-off point 36: $J = 0.350$
Group 3: Cut-off point 37: $J = 0.405$

10. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and corresponding Youden index to discriminate individuals using pooled blood samples. Epidemiology. 2005;16:73-81
11. Youden WJ. An index for rating diagnostic tests. Cancer.1950;3:32-35.
12. Malčić I. Dijagnostička vrijednost radiokardiografije u djece s lijevo-desnim intrakardijalnim shuntom - ROC analiza. Liječ Vjesn. 1985;107:500-6.
13. Harrison DJ, Simpson R, Kharbanda R, Abernethy VE, Nimmo G. Antibodies to neutrophil cytoplasmic antigens in Wegener's granulomatosis and other conditions. Thorax. 1989;44:373-7.
14. Zaletel-Kragelj L. Analiza učinkovitosti multiple regresije, diskriminantne analize in logistične regresije pri identifikaciji in vrednotenju napovednih dejavnikov prezgodnjega poroda: doktorska disertacija. Ljubljana: Medicinska fakulteta; 1997.

RECOMMENDED READINGS

1. Bradley GW, editor. Disease, diagnosis and decisions. Chichester: John Wiley & Sons; 1993.
2. Kallin Westin L. Receiver operating characteristic (ROC) analysis: valuating discriminance effects among decision support systems. Umeå: Deptartment of Computing Science, Umeå University; 2001. Available from URL: http://www8.cs.umu.se/research/reports/2001/018/part1.pdf. Accessed: March 29, 2009.
3. Wulff HR. Kliničko prosuđivanje. Zagreb: Jumena; 1982.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**BASIC CONCEPTS IN HEALTH ECONOMICS AND METHODS FOR ECONOMIC EVALUATION**
Module: 4.3.1	**ECTS (suggested): 0.15**
Author(s), degrees, institution(s)	**Doncho Donev**, MD, PhD**,** Professor Institute of Social Medicine, Faculty of Medicine, University "Ss Cyril and Methodius", Skopje, R. Macedonia
Address for correspondence	**Doncho Donev** Institute of Social Medicine, Faculty of Medicine, 50 Divizia 6, MKD-1000 Skopje, Republic of Macedonia E-mail: dmdonev@gmail.com
Keywords	Health economics, economic evaluation, efficiency, cost-minimization, cost-effectiveness, cost-utility, cost-benefit
Learning objectives	After completing this module students should be able to understand: • basic concepts in health economics theory and evaluation; • measuring costs and economic measures of health status; • economic evaluation terminology; • value of human life approaches; • cost analysis and economic evaluation of health care; • principles of cost-minimization, cost-effectiveness, cost-utility, and cost-benefit analyses;
Abstract	Health economics (HE) as a science of choice is an important element in health policy. HE analysis and evaluation provides useful tools in decision making and selection of priorities, resource planning and allocation. The "input-output" theory in HE provide useful measures for examining the potential costs and benefits of a specific health interventions or health programs and allows the decision makers to select the most suitable alternatives. This module describes four different types of economic evaluation: cost-benefit analysis, cost-effectiveness analysis, cost-utility analysis, and cost-minimization analysis.
Teaching methods	An introductory lecture gives the students insight into basic concepts of health economics and characteristics of economic evaluation techniques. After introductory lectures students should discuss differences among four types of economic evaluation, their value, and limitations in decision-making process. In group work, students should apply all four different types of economic evaluation in practical examples.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 50%/50%; • facilities: a lecture room, a computer room; • equipment: computers, LCD projection, access to the Internet and bibliographic data-bases (PubMed Central); • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice test.

BASIC CONCEPTS IN HEALTH ECONOMICS AND METHODS FOR ECONOMIC EVALUATION

Doncho Donev

Introduction

Health economics is the science of choice. It is an important element in the health policy, both at the strategic level (macroeconomics) dealing with overall financing and allocation of health resources, and tactical/operational level (microeconomics) which compares alternative approaches to dealing with specific health issues. Financial power is declining and monetary resources for health care are limited in each country to face increased demands of the population for higher quantity and quality of health care services. It is not easy to make decisions for proper allocation of resources as various incentives and disincentives affect the supply, demand, and ultimately the cost of health services. The amount of money spent in health care system is known but usually "health production" processes are unknown and the efficiency of the health care delivery process cannot be sufficiently controlled or influenced. Because of that there is a tendency to rationing of services and cutting down expenditures by different administrative means rather than to increasing productivity. Limited funds require careful selection of choices and distribution based on an objective assessment of costs and benefits of available options. Costs of health care (direct, indirect and intangible) in terms of resources used are compared to benefits in terms of reduced morbidity and mortality, as well as improved productivity and quality of life (1-4).

Health economic analysis and evaluation provides useful tools in decision making and selection of priorities, resource planning and allocation. Health economics sometimes conflicts with professional, ethical, and moral issues in solving the everyday problems of preventive and curative services. A balance between these issues is part of present day health care management and therefore of the New Public Health. The "input-output" theory in health economics provide useful measures for examining the potential costs and benefits of a specific health interventions or health programmes and allows the decision makers to select the most suitable alternatives (3,4).

Economic issues of health systems

Making balance between revenues and expenditures for health care is very intricate problem and source of serious concerns in practically all countries in the world. Permanent increasing of the number and proportion of elderly, patients with chronic diseases, use of expensive health technologies and their non-critical implementation, and some other factors are causing higher needs/demands and increase of the expenditures for health care. Big inequalities are recognized among countries in global health spending, ranging from under 3% to over 15% of GDP, and many health system reforms are characterized by transformation from central planning to market-based (5).

Higher needs and demands than resources available emphasize the need for rationing and allocation of resources due to defined priorities. There are two basic

economic problems in health affecting different countries in similar ways: underinvestment and overinvestment (overspending) or, more accurately, misallocation of available health resources. Overspending in one area (selected diseases, specific patient groups, provision of care like prevention vs. cure) very often goes along with under-spending in other areas. The World Bank's 1993 *World Development Report Investing in Health* addresses problems besetting health care systems, mainly, but not exclusively, in the developing countries. This report stresses that health is essential for productivity and economic growth, and that allocation of limited resources to costly, relatively unproductive services, such as excessive expenditures for armed forces, prevents many countries from meeting their basic health needs (6).

The *World Development Report, 1993* stresses the role of health in economic development, stating that a healthy population is not only a well-meaning social goal but is also essential to the development of a strong economy. Healthier populations are better workers and contributors to economic growth. Healthier children learn better in schools and, thus, also have prospects for contributing to economic development of their country.

This report addresses investment in health as follows:

1. Good health is a crucial part of well-being;
2. Spending on health can be justified on purely economic grounds;
3. Improved health contributes to economic growth:
    - it reduces production loss by worker illness;
    - it permits use of natural resources that have been inaccessible because of disease;
    - it increases the enrolment of children in school and makes them better able to learn;
    - it frees for alternative uses resources that would otherwise have to be spent on treating illness;
4. Sound policy in financing and resource allocation is essential to achieve good health.

Table 1. Essential, cost-effective health services for developing countries (6).

Public health interventions	**Clinical services**
1. Immunizations: DPT, polio, measles, hepatitis B, yellow fever, vitamin A, iodine supplements (EPI Plus)"	1. Short-term chemotherapy for TBC
2. School health program: deworming, micronutrient supplementation, health education	2. Management of the sick child: ARI diarrheal diseases, measles, malaria, malnutrition
3. Information on health, nutrition, family planning	3. Pregnancy care and delivery
4. Tobacco and alcohol control programs	4. Family planning
5. Disease monitoring, surveillance, vector control	5. Treatment of STDs
6. AIDS prevention program	6. Assessment, care of pain, trauma, infection etc. as resources permit

LEGEND: EPI = Expanded Programme of Immunization; ARI = Acute respiratory infection; STDs = Sexually transmitted diseases

The World Bank report calls on governments to increase spending on health and also to foster an atmosphere in which individual families do the same. It calls for competition and diversity in health care, based on an economic rationale, with health development based on "a basket of essential public health and clinical services" (7), (Table 1). This report establishes investment in health as an efficient tool for economic development, and thereby places health among the priorities for national and international financial investment (4,7-9).

Measuring health needs, demands, and utilization of health services

Need and demand for medical service are not necessarily the same:

- *Need* in medical care exists when an individual has symptoms, illness, or disability for which there may be an effective or acceptable treatment or cure from which the patient can benefit. Need also refers to nonmedical conditions, such as for designated preventive services such as immunization. Definitions of need vary depending on wide perception, interpretation and values are in play. Health needs can be addressed and met through treatment services, rehabilitation services, caring services, prevention at an individual level, community health promotion activities, environmental protection, and in many other ways. Health needs are thus most commonly set in the context of a need for care, but in the wider context of public health, need may be more inclusive.
- *Demand* for medical care exists when an individual considers that he or she has a need and is willing to spend resources of money, time, energy, loss of work, travel, and inconvenience to receive care.
- *Utilization* occurs when the individual actually acts on this demand or need and receives health services (4,10,11).

Health needs

We distinguish between different types of health needs:

1. Normative needs.
   Normative needs are those services determined by experts or by regulation to be essential for a specific need or for a specific population group. These include many examples of standard protocols for both clinical and preventive health care, such as prenatal care, immunization, child care for infants and toddlers, management of diabetes and hypertension, and screening for breast and prostate cancer. There are very often legitimate differences of opinion on professional issues in public health which are based on alternative interpretations of the available information, or where the evidence is incomplete. Furthermore, as scientific knowledge advances, new information may not be absorbed into decision making processes as rapidly as needed. Professional value judgments may be biased by trends in medical opinion, or influenced by advances in clinical, technological, and epidemiologic evidence. Such normative needs should be under continuous review by qualified professionals and academics representing clinical and public health services, as well as managers and consumers of health care. Other disciplines such as health

economics, sociology, health education, and planning add to the understanding of contributing factors to a disease and how to address it. Each field contributes to interpretation and decisions on standards for the health system.

The individual characteristics of people seeking care, or within the responsibility of a health system, including factors such as their age and sex, help determine the type and amount of health services needed. For example, a woman of 40 may not need a mammography as frequently as a woman over 50 years of age. An infant may need to be seen for preventive care assessment more often than a 3 year old. A male aged 45 needs his blood pressure checked more often than a 25 year old, and a teenager needs more attention paid to prevention of risk-taking behaviour than a 35 year old (4,11).

2. Felt (potential) needs.
   Felt need is the subjective view of the patient or the community, which may or may not be based on actual physiological needs. Though subjective, felt need is a prerequisite to whether a person actually wants and undertakes to seek care. There is a growing recognition of the importance of sharing health information with the population to increase the possibility that rational choices will be made (the health-belief model). Greater public knowledge is vital to acceptance of preventive programs such as immunization and compliance with treatment regimens for chronic diseases. Felt needs also affect health planning and allocating of scarce resources efficiently (allocative efficiency) in order to maximize benefits or maximize the needs met within the available resources. A community or donor may, for example, feel that a community needs a new hospital, whereas the same resources might better be spent on developing primary care, health educational services or building water supply system that have a greater impact on health of the population (4,11).
3. Expressed need.
   Expressed need/want/demand (equated with use) is a felt need that is acted on, such as in visiting a clinic or general practitioner. Felt needs may not be acted on because economic, geographic, social, or psychological barriers may inhibit a person from seeking or receiving care. Accessibility may be limited because the individual cannot afford to pay the fee. A service may be free, but not readily accessible due to such obstacles as distance, language or cultural barriers, difficulty in arranging an appointment, or a long waiting period. As a result, the person seeking care may not be able to receive it and may delay interfacing with the health system until a more urgent, and often more costly, problem arises. Distance, time, and cost of travel, inconvenience, and loss of wages affect the seeking of service, more so for preventive care than urgent surgical conditions, for example, even if the service is free of charge. Altering the supply, location and type of service availability can change these factors, thus improving equity of access. Elderly persons may sometimes avoid turning their felt needs into action as they may not feel comfortable with the fact that they are ill, or may not wish to become a burden. Religious, cultural, or even political factors may prevent a woman from practicing birth control even when further pregnancies may jeopardize her life. Lack of knowledge may also interfere with appropriate use of available clinical or preventive health care (4,10,11).
4. Comparative need.

Comparative need is a term that relates needs of similar population groups, as in two adjacent regions with the same age/sex/ethnic mix and socioeconomic status. One region may have a certain service, such as fluoridation of the community water supply, while the comparison community does not. The population of the second community is objectively in need of that service or service according to the best present professional and scientific evidence. There are no definable absolutes in the extent of health care, but there are accepted basic standards that are part of world standards at a particular point in time for health care, for prevention or health promotion. These are derived from trial and error as much as from science and must be continuously re-examined in light of new information, as well as the measurable benefits and costs derived from them (4).

Demands

Demand is willingness and/or ability to seek, use and, in some settings, pay for services. Demand is based on the individual and community expectations (Figure 1). The individual may feel that he needs a service, but expert opinion may say that this is not a reasonable demand. A patient may ask a physician for an antibiotic to treat a viral infection which would not help and may even cause harm. A patient may feel that denial of a referral by a doctor is not justifiable, but there may be a legitimate and ethical reason for the refusal. Doctors may wish to have the prestige and convenience of certain equipment locally, but economic and planning assessments may say that this is not justified on economic or medical grounds. Such equations, however, are not immutable; costs of a procedure may change as technology or clinical experience accumulates, so what was once not justifiable may become so. Such conflicts are unavoidably part of health care planning (4,10).

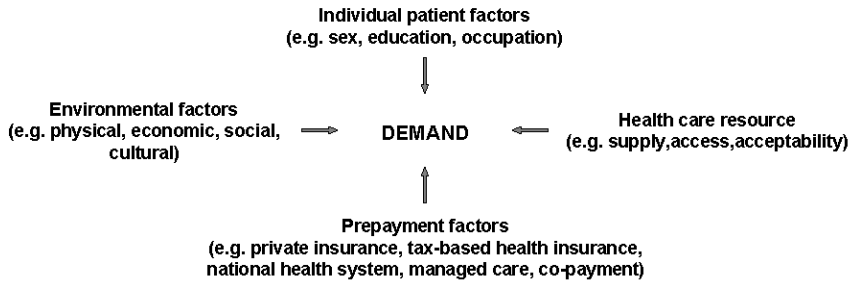


Figure 1. Factors in demand for health services (4).

Supply

Demand may also be induced by the supply or provision of care or by the method of payment to the provider (supplier induced demands). Making available more hospital beds may increase their use beyond justifiable need, or it may lead to an expectation or demand by patients or their families for an unnecessarily long stay in the hospital. Providing some services at no cost to patients may induce people to utilize those services more than they

really objectively require for health reasons according to current best standards. An inappropriate or excessively frequent use of a service may be promoted, and used by the upper middle class, while other important services may be lacking to serve the poor due to inequitable allocation of resources. Sometimes the interest of the health care providers is such that they may act to promote the use of services because payments are received for each service rendered (fee for service). This occurs, for example, in situations where a greater supply of surgeons results in unnecessary surgery being performed.

Measuring costs

The measurement of all effects of an intervention strategy or programme in terms of cost and outcome components (benefit, results, consequences) is based on the distinction between the input of resources used by the intervention on the one hand, and its positive and negative outcome effects on the other. Costs in health care can be analyzed in various ways. Costs could be classified in:

1. *Direct costs* to the patient related to the utilization of resources in the form of goods and services; costs to the insurer or sick fund on behalf of the patient; costs to the hospital or other provider;
2. *Indirect costs* refer to inputs and outputs outside healthcare system/industry, costs of illness to the patient, his family, and society, associated with a loss production due to sick leave, disability or premature death; and
3. *Intangible costs* (direct or indirect) are those that are incurred by patients and their families as a result of illness or intervention but are not measured in monetary terms (pain or grief levels associated with disability, morbidity or death), (1,3,4,10).

They can also be classified in:

1. Opportunity costs
   Opportunity costs refer to the resources used that could have been applied to other uses. It relates to the benefit foregone, or value of opportunity lost, by engaging resources in a service; usually quantified by considering the benefit that would accrue by investing the same resources in the best alternative manner. The concept of opportunity cost derives from the notion of scarcity of resources. Hospital land and building costs, for example, could be allocated for other purposes, such as primary health care facilities or facilities outside the health sector such as after school programs for children. Increasing the proportion of the GDP spent on health care may limit society's ability to spend money on education and other important social programs (1,4,10).
2. Social costs
   Social costs include indirect expenditures for health effects, such as the total value of lost production or costs of social support for a person whose health and work capacity has been impaired by illness.
3. Private costs
   Private costs include out-of-pocket expenditures that an individual makes to purchase health care plus related expenses such as payments for health insurance, loss of wages, purchase of pharmaceuticals, and copayments for health services.

Absence of (meaningful) market prices for many health care goods and services is a fundamental difficulty in the assessment of costs. Generally, true market prices are available only for (some) direct cost and outcome components, due to third party payment. The results or consequences of a medical intervention can be called its medical and economic outcome. The evaluation is based on a comparison of alternative treatments and the medical benefits are measured by different parameters, i.e. life expectancy and quality of life, progression of disease, patient compliance, frequency of complications and adverse events etc. (3).

Economic measures of health status

Measures of health status and health needs in many countries have been constructed to guide the allocation of health care resources. Economic analysis assesses not only input, as in costs and resources, but also output, as in morbidity, mortality, extension of years of life, reduction of disability and improving the health status in general. Greater functional levels of the individuals that improve the quality and quantity of life are output measures of health care. This should be part of an economic evaluation of the use of national or personal resources. Disability-adjusted life years (DALYs) and quality adjusted life years (QALYs) are measures of the total burden of disease as a guide to population health status (both of death and disability). Both are based on the concept of health-related utility which recognizes that health, as an output of health services and of other activities that promote health, is a function of both quantity of health and quality of health - or mortality and morbidity (4,11).

The QALY allows individuals, groups, or societies to 'trade-off' quantity of life against quality of life arguing, for example, that according to people's preferences 10 years living with a chronic condition which results in the individual being confined to his or her own home is equivalent to 8 years of full health. The implication is that the 'quality adjustment' for the chronic condition is 0.8. Furthermore, it is implied that intervening to cure the chronic condition would result in an improvement in quality of life of 0.2 per annum which over 10 years means 2 QALYs.

The DALY is a variant of the QALY. DALYs are calculated as the present value in years of disability-free life that might be lost as a result of premature death and disability occurring due to a disease in a particular year. The approach involves the measurement of health status (strictly lost health status) into a universal index of mortality and morbidity. Quality-adjusted life years (QALYs) measure life expectancy adjusted by changes in quality of life measured by assessing two or more aspects of health, such as pain, disability, mood, or capacity to perform self-care or socially useful activities such as paid employment or housework. DALYs and QALYs are constructed by using expert evaluation to estimate the degree of impairment (normal, impaired, or incapacitated) from specific diseases, These include impairments such as loss of ability to communicate, sleep disturbance, pain, depression, and sexual, eating, and mobility dysfunctions (4,6,11).

The value of the health status of an individual can then be assessed in terms of numerical values for comparisons. The values of the total scores are then added together and the overall score calculated out of a maximum value for comparison. This allows a measure of health status for purposes of comparison and may be used in comparing the effectiveness of alternative interventions. This is subjective depending on the perception of the assessor, and inter-observer variability may be high. While such measures do not

include all factors in determinants of disease, by pooling mortality and economic indicators they contribute to using the economic impact of disease as part of health planning (4,11).

DALYs and QALYs provide a common base for comparing mortality along with the dimensions of disability and quality of life as measures to compare different causes, settings, and change over time. They are used as proxy health status indicators to analyze different approaches to health policy and to justify specific interventions and determine priorities. The World Bank, WHO and other groups are examining alternative indicators to link health and its underlying determinants of the total burden of disease and disability, and to refine the process of establishing priorities for research and decision making for interventions (4,6,11).

The value of human life

One anticipated benefit of health intervention is the saving of human life. Placing an economic value on life is useful in calculating the benefits of specific interventions or perhaps for compensation to the family of a person who loses his or her life as a result of, for example, a negligent doctor or hospital manager.

More recently, economists have calculated the value of human life by human capital approach, willingness to pay for services, and other methods of quantifying the value of human life. Ethical and, indeed, political conflicts surround the issue of calculating the economic value of human life. A materialistic approach would evaluate human life based on the value of production that the individual might make to society. A humanistic approach would place virtually unlimited value on a human life according to the ethical value that saving of one life is as saving all human beings (sanctity of human life). By valuing human life as infinite, doctors may use precious resources to save one individual, without considering that this may be at the expense of other lives. For example, the cost of a heart transplant, which may add quality and years to one person's life, could be alternatively applied to a preventive program that might save many more lives through the prevention of heart disease. Should international agencies spend hundreds of millions of dollars to eradicate polio, a much feared, crippling, but nonlethal disease, while measles, thought to be a common benign disease, kills over 1 million children per year? The valuing of human life is not meant to fuel ethical argument, but rather to provide a measurement tool for the planning of priorities and litigation needs of health economics (4).

In health economics some arbitrary measures are used in order to demonstrate alternative ways of using limited resources. The implicit social value of life (ISV) rates a program by the lives it saves and assumes that, in a democratic society, all lives have the same intrinsic value.

Early economists valued life in terms of loss of net output to society, or the future loss of earnings minus the future loss in consumption resulting from the death of an individual. This human capital method is still widely used because of the simplicity of its calculations. However, it does not take into account the grief of the family. It places a negative value on the life of a pensioner who is no longer a "producer" in society, and gives no value to work done in the household, such as cooking, home maintenance, and rearing children. Nor does it give value to the intangible social and psychological benefits of the multigenerational family for all its members.

Another approach to valuation of life is based on court awards for compensation. It is a highly subjective method, often based on the court's interpretation of degree of contributory negligence, such as whether the injured person in a car accident was wearing a seat belt at the time of injury.

A major method is the willingness-to-pay approach, where valuations of life are based on what individuals are willing to pay for reductions in their probability of dying. For example, how much would persons pay for new tires on their car, or how much extra would they pay in order to travel on an airline with a better safety record? How much will a patient be ready to pay above his or her insurance coverage to have a world-famous surgeon operate on him or her as opposed to accepting the service available within the health care system. Such measurement is difficult and is often based on asking questions about hypothetical situations. Answers are also influenced by the income level of the respondent, by their attitude toward risk, and by the probability of death (4,12,13).

Basic concepts in health economics evaluation

All societies have limited resources and must, according to politically determined priorities, provide funds for health care in competition with funds for education, defence, agriculture, and others. The use of limited funds requires making choices. These choices reflect the overall political commitment to health and should, as far as possible, be based on an objective assessment of costs and benefits of available options.

There are three basic components of economic evaluation in health care (Figure 2):

1. Expenditure of resources (in terms of financial and human resources), both direct and indirect, are targeted to a
2. Health program with possible alternatives, which is expected to produce
3. Health benefits or utilities, which can also be both direct and indirect.

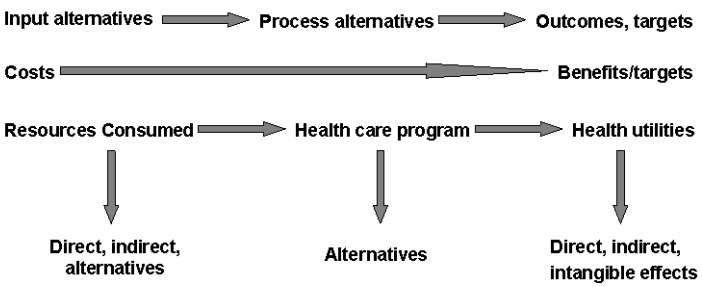


Figure 2. Economic models of health evaluation: resources-programs-benefits (4).

Health benefits may be expressed in terms of a direct reduction in morbidity and mortality, or as improved productivity and quality of life.

Measurement of both input and output is an essential part of health care management. Health inputs include expenditures on buildings, hospital or nursing home beds, equipment, personnel, home care, ambulatory care, and preventive programmes.

Other elements of health costs include patient's travel time, loss of work time for both, the patient and caregivers, loss of full functioning years of life, and loss of quality of life. The "input-output" theory of health economics may sound simplistic, but it provides a useful measure when examining the benefits and costs of a specific health intervention. Alternatives can be examined and analysis of their cost-effectiveness made, in order for decision makers to select the most suitable alternatives (3,4).

Health economics is a central discipline for the analysis what is effective in improving health. The core of health economics is *choice* and *decision making.* Health economics covers many aspects of health, from defining the meaning and value of health, to studies of how well resource utilization within health services function. Economic analysis leads us to define economic evaluation as the comparative analysis of alternative courses of action in terms of both their costs and consequences. It requires a thorough understanding of the complex context to do an economic analysis, especially in poor countries. However, when this is carried out it can contribute immensely to rational policy and resource allocation. Economic evaluation is potentially most useful where the resources are most scarce. To support decision making, information is needed on the desirability of the choice and the possible outcome in the future. The desirability (or anticipated satisfaction) of a good is described by its *value*. (2,3,12).

What does economic evaluation involve?

The basics of economic evaluation related to health care involve identifying, measuring, valuing, and comparing the costs and consequences of alternative approaches and courses being considered for solving priority health problems or satisfying population health needs. Costs are either one-off, or may be ongoing. Benefits are most often received over time. To build this effect of time into the economic analysis it is necessary to calculate a payback period. This is the time it takes for the benefits of a project/programme to repay its costs. Many companies look for payback on projects over a specified period of time e.g. three years. Therefore costs (and benefits in cost-benefit analysis) should be discounted to net present value (3,12,13).

Methods of economic evaluation

There are three basic types of economic evaluation with various potential to support decision making process toward rational use of national income and resources: cost-benefit analysis, cost-effectiveness analysis and cost-utility analysis. In addition, there are some variations as well: cost-minimization analysis, cost-consequence analysis, and cost-of-illness analysis with lower potential and importance.

Cost-benefit analysis

Cost-benefit analysis (CBA) is a powerful, widely used technique and tool in a wide range of public decision-making settings intended to improve the quality of public policy decisions (defined according to the change in social well-being or social welfare that they bring about), i.e. for deciding whether to make a change or accept to finance a certain project/ programme. In this analysis the economic and social costs of medical care and the benefits of reduced loss of net earnings due to preventing premature death or disability are

considered. The general rule for the allocation of funds in a cost-benefit analysis is that the ratio of marginal benefit (the benefit of preventing an additional case) to marginal cost (the cost of preventing an additional case) should be equal to or greater than 1. To use the tool, firstly work out how much the change will cost to make or project or a programme will cost to be implemented. Then calculate the benefit coming out from it. Where costs or benefits are paid or received over time, work out the time it will take for the benefits to repay the costs. CBA can be carried out using only financial costs and financial benefits. It is possible to include intangible items within the analysis and to estimate a value for these, which inevitably brings an element of subjectivity into the process. In the cost-benefit approach the objective is to identify and value costs and outcomes in monetary terms (1,3,12-15).

CBA places a monetary value on all elements - on the input side as well as on the output side. It means that all the consequences of each alternative intervention/ project/ programme have to be measured in terms of money and/or converted to money where costs are not directly observable (value of life). Forming 'benefits', means that the consequences are measured on the same unit as 'costs'. This means that all benefits and costs of the intervention/ project/ programme should be measured in terms of their equivalent money value. Furthermore, the benefits and costs of the intervention or project or programme have to be expressed in terms of equivalent money value (dollars, euros, pounds) of a particular time. This is not just due to the differences in the value of money at different times because of inflation. A dollar available five years from now is not as good as a dollar available now. This is because a dollar available now can be invested and earn interest for five years and would be worth more than a dollar in five years. If the interest rate is 'r' then a dollar/ euro invested for 't' years will grow to be $(1+r)^t$. Therefore the amount of money that would have to be deposited now so that it would grow to be one dollar 't' years in the future is $(1+r)^{-t}$. This called the discounted value or present value of a dollar available 't' years in the future (12,13,15).

The net benefit of the intervention/project/programme is just the sum of the present value of the benefits less the present value of the costs. When the inflation has been considered, a worthwhile intervention/project/programme is one for which the discounted value of the benefits exceeds the discounted value of the costs; i.e., the net benefits are positive. This is equivalent to the benefit/cost ratio being greater than one and the internal rate of return being greater than the cost of capital (12,13).

If the discounted present value of the benefits exceeds the discounted present value of the costs then the project is worthwhile. This is equivalent to the condition that the net benefit must be positive. Another equivalent condition is that the ratio of the present value of the benefits to the present value of the costs must be greater than one (12-14).

If there are more than one mutually exclusive projects or programmes that have positive net present value then there has to be further analysis. From the set of mutually exclusive projects/programmes the one that should be selected is the one with the highest net present value. An example of CBA was to use this approach in evaluating new community care packages for mentally ill patients. The obvious lesson to be learnt from this study is that it excluded a variety of dimensions of clinical outcome which could not be valued in monetary terms. If, almost inevitably, CBA is incomplete it is less useful than cost-utility analysis (13,14).

Some costs (intangible costs) or benefits in CBA, i.e. reducing or avoiding pain, grief, suffering, loss of leisure time etc., are not easy to express in monetary value.

'Willingness to pay' approach (i.e. how much patients would be willing to pay for a given health benefit, such as avoidance of pain or disability) is often used to attribute such benefits (10).

CBA sometimes evaluate the benefit of saving human lives. There is considerable antipathy in the general public to the idea of placing a monetary value on human life. The controversy is defused when it is recognized that the benefit of such projects/ programmes is in reducing the risk of death. There are many cases in which people voluntarily accept increased risks in return for higher pay, such as working in the petroleum industry, smelting-works or mining, or for time savings in higher speed in automobile travel. These choices can be used to estimate the personal cost people place on increased risk and thus the value to them of reduced risk. This computation is equivalent to placing an economic value on the expected number of lives saved (12,13).

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is a subset of cost-benefit analysis in which a policy outcome is taken as given and the analysis seeks to identify the least-cost means for achieving the goal (taking into account any ancillary benefits of alternative actions/interventions/treatment). This form of analysis seeks to determine the costs and effectiveness of an activity or to compare similar alternative activities to determine the relative degree to which they will obtain the desired objectives or outcomes. The preferred action or alternative is one that requires the least cost to produce a given level of effectiveness, or provides the greatest effectiveness for a given level of cost. In the health care field, outcomes are measured in terms of health status. In analyses related to health and health care CEA is the net gain in health or in reducing the burden of disease from a specific intervention in relation to its cost. The aim is to determine the least expensive way of achieving the goal, by comparing alternative methods of intervention in order to make a choice. The most cost effective method is the one that achieves the same goal using the least resources or that achieves the most effect per monetary unit of cost (dollar, euro, pound). Cost-effectiveness means choosing the interventions that give most value for money, and giving lower priority to those that contribute little to improving people's health. A low cost per DALY gained indicates a high degree of cost-effectiveness, and therefore an intervention that should be of high priority, given limited resources. CEA is but one tool in the improvement of the health system, but it is a tool that is more under than over-utilized (1-4,10,12,15,16).

An example of this is the study compared dialysis and transplantation for patients with end stage renal failure, to identify the cost per life DALY gained from the alternative (14). Comparison of life years gained for patients with end stage renal disease in the United States found that renal transplantation was less expensive ($3,600 per year of life gained) compared to home dialysis ($4,200 per life year gained) and hospital dialysis ($116,000 per life year gained). Moreover, transplantation provides a higher quality of life. This was perhaps the first example of "cost-utility" analysis, where life years gained, were weighted according to quality of life. This can be expressed as the cost-effectiveness per QALY (4,14).

The 1993 World Bank report, point out various interventions based on public health and clinical services in developing countries compared in terms of DALYs versus costs of the interventions (WB '93, Ted). Such analyses help to construct a basket of essential services on the basis of comparative cost-effectiveness (6,7). Highly cost-

effective interventions include vitamin A supplementation, measles control, and directly observed chemotherapy for TB. A high cost but highly effective intervention presented as chemotherapy for leukaemia in children under age 15. This intervention is justified as benefits are high. The same chemotherapy in a 75 year old would present a low DALY value. CEA studies examine issues such as day surgery versus inpatient surgery, operations versus medications (e.g., for peptic ulcers and coronary heart disease), public versus individual dental prevention (e.g., fluoride versus dental hygienist care), and community versus institutional care. Treatment of psychiatric patients in a large mental hospital, in the psychiatric ward of a general hospital, and in a day treatment centre show day treatment centre care to be least costly, but some measure of the severity of illness and care needs to be added to this assessment. Planning mental health services and facilities with reduced hospitalization requires adequate resources for mental health care in the community, to prevent chronic mental patients from becoming part of the homeless population as has happened in many large cities (4,6).

In the health service, the cost per healthy life-year saved in a rich country could vary between USD 250 for screening and treating newborns with sickle-cell anaemia, and USD 5 million for control of radioactivity emission. In a poor country in Africa, the cost for a life-year saved through short-course drug therapy for malaria can be as low as USD 3, while other interventions like anti-retroviral treatment may cost many hundreds of dollars per healthy year saved, i.e. to get the same result. Combining calculations of cost with measures of effectiveness of interventions and using them to guide policy decisions is a very recent development (2).

Sometimes the least costly method is the least effective. For example, a study showed that prevention of pregnancy by the withdrawal method is least costly but is far less effective than use of the birth control pill. Abortion as a method of birth control may be less costly than use of the pill, but, in addition to the ethical issues, it produces complications and contributes to excess morbidity and mortality in subsequent pregnancies, both for the mother and the newborn (4).

Cost-effectiveness analysis takes into account both the cost and effectiveness of interventions, as a measure of value for cost, but does not answer the question of whether or when the intervention should be done. CEA is not limited to only one specific outcome effect. An intervention-specific group of effects may be used, too. In general, the various medical outcome effects of a treatment cannot be summed up like cost figures. This aggregation necessitates complicated procedures and (potentially problematic) evaluations of the multiple outcome effects of interventions (3,12,15).

Cost-utility analysis

Cost-utility analysis (CUA) is a form of economic evaluation in which the outcomes of alternative procedures, interventions or programs are expressed in terms of a single "utility-based" unit. A widely used utility-based measure is the quality adjusted life year (QALY). Converting various utilities (health states) into QALYs is done by multiplying the utility value by the years spent in certain health state. The aim is to choose the intervention or program that supplies a unit of effect, in terms of QALYs, at lowest cost or to achieve the most QALYs per monetery unit (dollar, euro, pound) of cost. (3,10,12).

The health state utility is a cardinal number, usually between 0 and 1.0, associated with a particular health state. For example, 10 years in a health state with a utility value of 0.5 would result in 5 QALYs (i.e. equivalent to 5 years of perfect health).

This approach gives the cost of producing a given change in outcome (e.g. a QALY). In evaluating alternative treatment for patients with end stage renal failure, it is necessary to extend the cost-effectiveness analysis and to take account of the differences in the quality of survival on dialysis as compared to renal transplantation. The objective of cost-utility analysis is to value, however crudely, improved health in terms of the cost of producing a QALY or well year (13,17,18).

Cost-minimization analysis

Cost-minimization (CMA) analysis is used to compare net costs of alternative healthcare interventions with exactly the same effects (quality and quantity of the effect are the same). The costs of given healthcare interventions are analysed and compared. The cost-minimization approach assumes that the outcomes of the alternatives being evaluated are equivalent and compares their costs. The aim is to choose the intervention that supplies an effect at lowest cost. An example of this is the study of day case versus hospital stay treatment for hernias which involved the randomized control trial of the alternatives. The effects of these alternatives, in terms of number and type of complications, length of recovery period, and recurrence, were very similar and the CMA analysis assumes the results were equivalent (12,14).

EXERCISE

Task 1

After introductory lecture students carefully read the theoretical background of this module and recommended readings. Afterwards they:

- discuss differences among four types of economic evaluation,
- the value of four types of economic evaluation, and
- their limitations in decision-making process.

Task 2

For accomplishing this task students first make three groups in order to analyze use of following three basic types of economic evaluation in papers available in PubMed Central:

- Group 1 – cost-benefit analysis,
- Group 2 – cost-effectiveness analysis, and
- Group 3 – cost-utility analysis. cost-minimization analysis.

All three groups perform following steps:

- go to PubMed Central database, an archive of biomedical and life-sciences journal literature,
- search for selected type of economic evaluation,
- limit search to last 10 years,
- analyze papers you have found by medical discipline, country, etc.,
- make a report (a short review paper),
- present the report to other two groups of students.

REFERENCES

1. Auld MC, Donaldson C, Mitton C, Shackley P. Health economics and public health. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.877-98.
2. Lindstrand A, Bergstrom S, Rosling H, Rubenson B, Stenson B, Tylleskar T. Global health - An introductory textbook. Copenhagen: Narayana Press; 2008.
3. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
4. Tulchinsky TH, Varavikova EA. Chapter 11: Measuring costs: The economics of health. In: Tulchinsky TH, Varavikova EA. The New Public Health: An Introduction for the 21st Century. San Diego: Academic Press; 2000. p.549-89.
5. Donev D, Bozikov J. Financing of health care. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.302-23. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
6. The World Bank. The World Bank development report 1993. Investing in Health. Available from URL: http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&siteName=WDS&entityID=000009265_3970716142319. Accessed: August 14, 2009.
7. Bobadilla JL, Cowley P, Musgrove P, Saxenian H. Design, content and financing of an essential national package of health services. Bull WHO. 1994;72:653-62.
8. World Health Organisation. Macroeconomics and health: Global public health goods for health. The report of the commission on macroeconomics and health, working group 2. Geneva: World Health Organisation; 2001.
9. World Health Organisation. Health21: the health for all policy framework for the WHO European Region. (European Health for All series; no. 6.) Copenhagen: World Health Organization, Regional Office for Europe, 1999.
10. Last J, editor. A dictionary of epidemiology. Fourth edition. New York, NY: Oxford University Press, Inc.; 2001.
11. Mooney G, Jan S, Wiseman V. Measuring health needs. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.1766-71.
12. Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 1993.
13. Watkins T. An introduction to cost benefit analysis. San Jose, CA: San Jose State University Department of Economics [homepage on the Internet]. Available from URL: http://www.sjsu.edu/faculty/watkins/cba.htm. Accessed: April 29, 2008.
14. Maynard A. Cost-benefit analysis of drug treatment. In: Abshagen U, Munnich FE, editors. Cost of illness and benefits of drug treatment. International Satellite Symposium to the IV. World Conference on Clinical Pharmacology and Therapeutics, Mannheim, July 1989. Proceedings. W. Muenchen: Zuckschwerdt Verlag; 1990. p.3-11.
15. Kopp RJ, Krupnick AJ, Toman M. Cost-benefit analysis and regulatory reform: An assessment of the science and the art. Discussion paper 97-19. Washington, DC: Resources for the future; 1997. Available from URL: www.rff.org/Documents/RFF-DP-97-19.pdf. Accessed: August 9, 2009.

16. Jamison D. Cost-effectiveness analysis: concepts and applications. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.903-19.
17. Slesnick D. Empirical approaches to the measurement of welfare. J Econom Literature. 1998;36:2108-65.
18. Norton BG. Economists' preferences and the preferences of economists. Environ Values. 1994;3:311-32.

RECOMMENDED READINGS

1. Auld MC, Donaldson C, Mitton C, Shackley P. Health economics and public health. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.877-98.
2. Donev D, Bozikov J. Financing of health care. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.302-23. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
3. Jamison D. Cost-effectiveness analysis: concepts and applications. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.903-19.
4. Mooney G, Jan S, Wiseman V. Measuring health needs. In: Detels R, McEwen J, Beaglehole R, Tanaka H, editors. Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.1766-71.
5. Watkins T. An introduction to cost benefit analysis. San Jose, CA: San Jose State University Department of Economics [homepage on the Internet]. Available from URL: http://www.sjsu.edu/faculty/watkins/cba.htm. Accessed: April 29, 2008.
6. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**ECONOMIC EVALUATION IN HEALTHCARE: PRACTICAL APPROACH**
Module: 4.3.2	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Pia Vračko**, MD, PhD Candidate, Research Assistant, Teaching Assistant National Institute of Public Health of the Republic of Slovenia, Slovenia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Pia Vračko** National Institute of Public Health of the Republic of Slovenia, Trubarjeva ulica 2, 1000 Ljubljana, Slovenia E-mail: pia.vracko@ivz-rs.si
Keywords	Health economics, economic evaluation, cost-minimization, cost-effectiveness, cost-utility, cost-benefit
Learning objectives	After completing this module students and public health professionals should be able to: • understand economic evaluation terminology; • describe methodology of cost analysis and find data for it; • describe methodology and understand different uses of cost-minimization, cost-effectiveness, cost-utility, and cost-benefit analyses; • understand the importance of certain issues that are important in economic evaluation in healthcare; • independently interpret the economic evaluation studies.
Abstract	Economic evaluation of healthcare interventions is a useful tool in resource allocation planning. It can enumerate the potential costs and value the anticipated consequences of alternative programs, and analyse risks and uncertainties for their successful outcomes. In this module, four different types of economic evaluation are presented: cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis.
Teaching methods	An introductory lecture gives the students insight into characteristics of economic evaluation techniques. The theoretical knowledge is illustrated by case studies. In continuation they should distinguish among four types of economic evaluation, their value, and limitations in decision-making process. In group work, students first perform all four different types of economic evaluation in a simple example, and than present self-selected cases of four types of economic evaluation.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice test and calculation of various measures of economic evaluation in practice.

ECONOMIC EVALUATION IN HEALTHCARE: PRACTICAL APPROACH

Pia Vračko, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

What is an economic evaluation?

Economic evaluation may be defined as the comparative analysis of alternative courses of action in terms of both their costs and consequences (1). Based upon welfare economics, a subfield within economics that has explored the question of how to maximize consumer welfare within a limited budget, it tries to assess the social desirability of a program relative to some other alternative. Economic evaluation can enumerate the potential costs and value the anticipated consequences of a proposed intervention, program, policy or regulatory initiative, and reflect trade-offs in alternatives. The additional expense of the new intervention means that resources have to be redirected from elsewhere and an economic evaluation assesses whether or not the additional benefits generated by the new intervention are greater than the loss in benefits from the reduction in other programmes - that is, is the reallocation efficient? Additionally, an economic evaluation can analyse risk and uncertainty for a successful outcome of a project (2).

The ultimate goal of economic evaluation in healthcare is to maximize net health benefits for all persons in a target population given a range of health care interventions and known resource constraints (3). With the healthcare sector accounting for a sizeable proportion of national expenditures, the issue of efficiency has become a central objective of policymakers within most health systems.

What does economic evaluation involve?

The basics of economic evaluation involve identifying, measuring, valuing, and comparing the costs and consequences of alternatives being considered (1).

The measurement of costs is always in monetary units. The outputs, on the other hand, can come in different forms, and correspondingly there are different types of economic evaluation that are described in next section.

In addition to costs and consequences, a reliable economic evaluation should involve:

1. Clear definition of the viewpoint from which the analysis is performed; an item may be a cost (or consequence) from one point of view, but not from another. For example, patients' travel costs are a cost from the patient's point of view and from the society's point of view, but not a cost from the Ministry of Health's point of view. Possible points of view include those of society, the Ministry of Health, other government ministries, the government in general, the patient, the employer, and the agency providing the program (1). Optimally, the analyst should adopt the societal point of view, which is the broadest one and is always relevant;

2. Time structure of intervention; usually, costs and consequences do not occur at the same time, therefore costs (and benefits in cost-benefit analysis) should be discounted to net present value;
3. Analysis of risk and uncertainty (of successful outcome); uncertainty in economic evaluation can arise because of methodological disagreement among analysis, the data requirements of the study, the need to extrapolate results over time, or the desire to generalize the results of the study to another setting (4);
4. Distributional considerations (distribution of wealth in society, equity), which markets ignore and the healthcare sector considers an essential ingredient.

For in-depth information on theoretical background of economic evaluation, please refer to module Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems (5).

Cost analysis

Cost analysis is used in all four types of economic evaluation. In healthcare field the costs are disaggregated into three categories: direct, indirect and intangible. Direct means directly related to healthcare industry (the stuff, the hospitals and the patients), indirect refers to inputs and outputs outside healthcare industry, and intangible refers to the pain and suffering that are caused or alleviated by a healthcare intervention. For example, physician and nursing expenses are direct costs, and loss of income due to illness is referred to as indirect cost.

According to Drummond et al. (1), steps in cost analysis are the following:

1. Identification of resource categories, for example:
    a. health care resource use
        - hospital resources (investigations, therapy, bed days, out-patient attendances, overheads)
        - community care resources (general practitioner visits, nurse visits, ambulance or hospital car)
    b. patient and family resource use (patient's time, time of relatives, out-of-pocket expenses for transport)
    c. resource use in other sectors (social workers visits, home help visits).
2. Measurement of resource use.
   The most accurate data collection is by clinical trial, as case report forms are completed for patients enrolled in the trial. Normally these record data on clinical events, but they can be modified to include resource use, such as number and type of investigations, date of hospital admission and discharge. In the absence of a trial the two major sources of data on resource use are routine statistics kept at the hospital or by other agencies, and patient's case charts.
3. Valuation of resource items.
   This is dependent on the availability of local financial data. In some settings, there may be data on hospital billings or charges, usually kept by agencies. In other settings, detailed costing studies would be necessary.

Types of economic evaluation

There are four types of economic evaluation (1,6). One type only uses costs; it is called a cost-minimization analysis. The other three types differ according to what kind of consequence they incorporate along with costs.

The most obvious consequence is what the healthcare industry immediately works with, such as diagnostic test outcome or an operation successfully completed. These outputs are called "effects", and expressed in natural units (such as percentage detection).

A broader measure of effects relies on "utilities" (i.e., estimates of the satisfaction of the effects), and the output unit is called a "quality-adjusted life-year", (QALY). The QALY is a measure of the value of health outcomes. Since health is a function of length of life and quality of life, the QALY was developed as an attempt to combine the value of these attributes into a single index number. The QALY calculation is simple: the change in utility value induced by the treatment is multiplied by the duration of the treatment effect to provide the number of QALYs gained. QALYs can then be incorporated with medical costs to arrive at a final common denominator of cost/QALY. This parameter can be used to compare the cost-effectiveness of any treatment (7).

Lastly, the output can be expressed in the same monetary unit as the costs, in which case the consequences are called "benefits".

"Effects" are used in a cost-effectiveness analysis, "QALYs" are used in cost-utility analysis, and "benefits" are used in cost-benefit analysis (8).

Cost-minimization analysis

Cost-minimization analysis (CMA) is used to compare net costs of healthcare interventions with exactly the same effects (quality and quantity of the effect are the same). The costs of given healthcare interventions are analysed and compared. The aim is to choose the intervention that supplies an effect at lowest cost. Under a CMA, treatment 1 would be more cost-efficient (that is, less costly) than treatment 2 if (Equation 1):

$$C_1 < C_2 \qquad \textbf{Equation 1.}$$

C_1 = *cost of treatment 1*
C_2 = *cost of treatment 2*

If we want to calculate the difference between costs of treatment 1 and costs of treatment 2 we use the simple procedure (Equation 2):

$$\Delta C = C_1 - C_2 \qquad \textbf{Equation 2.}$$

ΔC= *cost difference*
C_1 = *cost of treatment 1*
C_2 = *cost of treatment 2*

To summarize, a CMA identifies, measures and compares only input costs of alternative interventions what is, due to its simplicity, an advantage of this type of studies. But, it assumes outcomes to be equivalent. It requires clinical evidence that differences in health effects between alternatives are minimal or not important. The major problem is that the assumptions are in most cases difficult to justify prior to any experimental study. This method, therefore, has only limited application what is a disadvantage of this type of studies. A common example of a CMA is the comparison of drugs that are of the same chemical entity, the same dose, and have the same pharmaceutical properties as each other (innovative or original drug versus generic drug, or generic made by one pharmaceutical company compared with a generic made by another pharmaceutical company). In this case only the cost of the drug itself needs to be compared because outcomes should be the same.

Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is used to compare net costs of healthcare interventions per unit of the effect. A single effect common to the given alternatives, but achieved to different degrees, is considered and expressed in natural clinical units, such as years of life saved, premature deaths avoided or improvements in functional status (units of blood pressure or cholesterol).

The aim is to choose the intervention that supplies a unit of effect at lowest cost. or one can try to achieve the most effect per monetary unit (i.e. euro, dollar etc.) of cost. Under a CEA, treatment 1 would be more cost-effective that treatment 2 if (Equation 3):

$$\frac{C_1}{E_1} < \frac{C_2}{E_2} \qquad \textbf{Equation 3.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2

Comparing the costs and health outcomes of 2 treatments results in 1 of 9 pairings (Figure 1). Since costs and benefits are measured in non-comparable units, their ratio (Equation 2) provides a yardstick with which to assess relative (productive) efficiency (9). It does not, however, enable us to evaluate the relative efficiency of interventions which provide more benefit at greater cost or less benefit at lower cost (10).

For decision a cost-effectiveness grid can be used (6,11). In the case the health outcomes are equivalent (Figure 1, centre column), the choice should be based on cost; when costs are equivalent (Figure 1, centre row), the choice should be based on outcome. When one strategy has better outcomes and lower costs (Figure 1, upper right and lower left boxes), the choice is clear. The decision is difficult only when the strategy that is more expensive also produces better outcomes (Figure 1, upper left and lower right boxes).

	HEALTH OUTCOME		
COST	$E_1 > E_2$	$E_1 = E_2$	$E_1 < E_2$
$C_1 > C_2$	Calculate incremental cost-effectiveness ratio	2 less expensive Choose 2	2 dominates 1 Choose 2
$C_1 = C_2$	1 better effect Choose 1	Makes no difference	2 better effect Choose 2
$C_1 < C_2$	1 dominates 2 Choose 1	1 less expensive Choose 1	Calculate incremental cost-effectiveness ratio

Figure 1. Cost-Effectiveness Comparison of Treatments 1 and 2 and Possible Decisions (6,11).
LEGEND: C=costs, E=effect, 1=treatment 1, 2=treatment 2.

If an intervention is both more expensive and more effective than an alternative, then the criterion for efficiency becomes the incremental cost effectiveness ratio, iCER. It is defined as the ratio of the change in costs of two programmes to the change in effects of the programmes (Equation 4):

$$iCER = \frac{C_1 - C_2}{E_1 - E_2} \qquad \textbf{Equation 4.}$$

iCER = incremental cost-effectiveness ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2

iCER represents the additional cost of an intervention divided by the additional health outcome it achieves; in fact, iCER provides information on how much each additional therapeutic success will cost. It is most useful when a choice must be made among several therapeutic interventions. The lower the ratio, the greater is the health improvement for a given resource expenditure. A ratio increases most rapidly as its denominator approaches 0 (this is, when the 2 interventions provide nearly equal health effects).

A major limitation of cost-effectiveness analysis is its inability to compare interventions with differing natural effects (1). For example, interventions aimed at increasing life years gained cannot be directly compared with those which improve

physical functioning. Cost-effectiveness analysis therefore cannot directly address resource allocation among alternative, unrelated health interventions (10).

Cost-utility analysis

Cost-utility analysis (CUA) is used to compare net costs of healthcare interventions with different effects, when prices do not want to be used explicitly. Therefore, all effects need to be converted to a common unit, a quality adjusted life year (a QALY). The aim is to choose the intervention that supplies a unit of effect, in terms of QALYs, at lowest cost. In CUA, the following criterion is used to replace Equation 2 (Equation 5):

$$\frac{C_1}{QALY_1} < \frac{C_2}{QALY_2}$$ **Equation 5.**

C_1 = cost of treatment 1
C_2 = cost of treatment 2
$QALY_1$ = QALYs achieved by treatment 1
$QALY_2$ = QALYs achieved by treatment 2

In cost-utility analysis, relative efficiency is assessed using an incremental cost-utility ratio (iCUR). It is defined as the ratio of the change in costs of two programmes to the change in utility of the programmes (Equation 6).

$$iCUR = \frac{C_1 - C_2}{QALY_1 - QALY_2}$$ **Equation 6.**

iCUR = incremental cost-utility ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
$QALY_1$ = QALYs achieved by treatment 1
$QALY_2$ = QALYs achieved by treatment 2

In CUA, the use of a single measure of health benefits enables the comparison of the efficiency of resource allocation among alternative, unrelated health interventions. The optimal decision rule involves ranking the incremental cost-utility ratios of different interventions and selecting those with the lowest ratios (best value) until the budget is depleted (12,13). The lower the incremental ratio, the higher the priority in terms of maximizing health benefits derived from a given level of expenditure (13).

The value of CUA is limited by understanding and meaningful interpretation of QALYs. Individuals are not used to buying a QALY and are therefore unfamiliar with the process of trying to obtain a QALY at a lowest cost. Deriving meaningful estimates of QALYs is therefore at the heart of the CUA evaluation.

Cost-benefit analysis

Cost-benefit analysis (CBA) places a monetary value on the consequences of each alternative health interventions. Forming 'benefits' means that the consequences are measured using the same unit as 'costs'. One can tell whether the benefits are greater than costs and thus know whether the expenditure is worthwhile. The basic cost-benefit criterion is (Equation 7):

$$B_1 > C_1 \qquad \textbf{Equation 7.}$$

B_1 = benefits of treatment 1
C_1 = cost of treatment 1

Benefits can also be expressed as a product of effect (E) and price (P). In this case, the Equation 6 can be rewritten as (Equation 8):

$$P_1 \times E_1 > C_1 \qquad \textbf{Equation 8.}$$

C_1 = cost of treatment 1
E_1 = units of health effects achieved by treatment 1
P_1 = attributable price of effects achieved by treatment 1

When there is a financial budget constraint, using funds for one purpose precludes their use for another. In this case, the CBA criterion has an alternative formulation (Equation 9):

$$\frac{C_1}{P_1 \times E_1} < \frac{C_2}{P_2 \times E_2} \qquad \textbf{Equation 9.}$$

C_1 = cost of treatment 1
C_2 = cost of treatment 2
E_1 = units of health effects achieved by treatment 1
E_2 = units of health effects achieved by treatment 2
P1 = attributable price of effects achieved by treatment 1
P2 = attributable price of effects achieved by treatment 2

This criterion enables to choose the intervention that supplies a unit of benefit at lowest cost (7).

To calculate the difference in cost-benefit between the two (or more) programmes being compared in the evaluation, incremental cost-benefit ratio (iCBR) may be used. It is defined as the ratio of the change in costs of two programmes to the change in benefits of the programmes (Equation 10):

$$iCBR = \frac{C_1 - C_2}{B_1 - B_2}$$ **Equation 10.**

iCBR = incremental cost-benefit ratio
C_1 = cost of treatment 1
C_2 = cost of treatment 2
B_1 = benefits of treatment 1
B_2 = benefits of treatment 2

The result of iCBR shows additional euros earned with each additional euro invested in new treatment.

Another useful criterion in CBA is the so called net benefits (NB) criterion (Equation 11):

$$NB = (B_1 - B_2) - (C_1 - C_2)$$ **Equation 11.**

NB = net benefits
C_1 = cost of treatment 1
C_2 = cost of treatment 2
B_1 = benefits of treatment 1
B_2 = benefits of treatment 2

The result of this criterion represents the excess of monetary benefits over costs.

Welfare economics shows that under certain conditions any net excess of monetary benefits over costs represents the gain in welfare by society (14). Cost-benefit analysis therefore makes it possible to determine, firstly, whether an individual intervention offers an overall net welfare gain and, secondly, how the welfare gain from that intervention compares with those from alternative interventions. Increased use of interventions with the greatest net gain will increase efficiency. By valuing all costs and benefits in the same units, cost-benefit analysis compares diverse interventions using the net benefit criterion. Cost-benefit analysis thus simultaneously addresses issues of productive efficiency and issues of resource allocation among alternative, unrelated health interventions.

There has been a general unease with the pricing of the effects in CBA. Commonly, the monetary value is attributed to outcomes through 'willingness to pay' approach (i.e. how much patients would be willing to pay for a given health benefit, such as avoidance of pain or disability, or, in public health, how much a health agency would be willing to pay for prevention/treatment of a disease in a population) (15).

Types of economic evaluation in practice

When relevant effects of the given health care interventions are observed to be similar, CMA is used to compare net costs.

When a single effect is common to the alternative interventions, but achieved to different degrees, and expressed in natural clinical units, CEA is used to compare alternative interventions in terms of net cost per unit of health effect obtained.

Many times, economic analysts may want to assess interventions with outcomes that are either common or not common to the alternative interventions. In these situations, a specific measure of value must be applied to the relevant outcomes to allow for relevant comparisons with a common denominator. In these cases, the CUA, with QALYs as a common unit, and CBA with common monetary unit, may be applied.

CASE STUDY 1: COSTING ALTERNATIVE RADIOTHERAPY TREATMENTS

Introduction

Cost analysis is the first and inevitable step in every economic evaluation. A reliable cost analysis may be complex and sometimes challenging due to many different data sources. Many times, cost data are hard to get. Therefore, it is meaningful to represent a comprehensive cost analysis (with suggested data sources) in this case study. Drummond et al. (1) based it on an actual study, undertaken in the UK (16). For more detailed directions on costing, please refer to the recommended readings.

Description of a case

A clinical trial is being carried out comparing two forms of radiotherapy for patients with head and neck cancer and carcinoma of the bronchus:

- patients receiving conventional therapy are treated once per day, 5 days per week, for about 6 weeks.; they would normally travel on a daily basis to a hospital-based radiotherapy centre to receive care;
- patients receiving continuous hyperfractionated accelerated radiotherapy (CHART) are treated three times on each of 12 consecutive days, including the week-end; because of the intensity and frequency of treatment, patients would normally stay in hospital during therapy, either in a regular hospital ward or in a hostel owned by the hospital.

The different treatment regimens obviously give rise to different costs. However, in addition, there may be differences in the period following treatment for the following reasons:

1. the higher intensity of the CHART regimen might give rise to more side-effects, and hence a greater need for community care after hospital discharge;
2. the CHART regimen might give better tumour control, thereby slowing down the progression of the disease;
3. CHART might reduce the extent of late radiation changes, and a lower incidence of necrosis may also reduce the need for salvage surgery.

The clinical trial will provide an opportunity to gather data on the use of resources by patients in the two treatment groups. You are asked to

1. identify which categories of resource you feel it would be important to assess;
2. indicate how you might measure the use of these resources in physical units;
3. say how you might value the resource consumption in money terms.

Identification of resource categories

Resource use can be considered under the following broad headings (Table 1):

Table 1. Identification of resource categories for economic evaluation.

Resource items	
Hospital resources	• radiotherapy, • bed days, • out-patient attendances, • overheads
Community care resources	• general practitioner visits, • nurse visits, • ambulance or hospital car
Patient and family resource use	• patient's time, • time of relatives, • out-of-pocket expenses for transport (e.g. car, train, or taxi)
Resource use in other sectors	• social workers visits, • home help visits

Measurement of resource use

The fact that a clinical trial is taking place greatly increases the opportunity for accurate data collection as case report forms are completed for patients enrolled in the trial. Normally these record data on clinical events, but they can be modified to include resource use, such as number and type of investigations, date of hospital admission and discharge. Also, the fact that patients are enrolled in a trial provides the opportunity to interview them about the resource use in community care, time taken to travel to hospital, and personal expenditure. They can also be given diary cards to record expenditure or time spent by relatives in home nursing.

In the absence of a trial the two major sources of data on resource use are routine statistics kept at the hospital or by other agencies, and patients' case notes (charts). The quality of these records varies by agency and data are usually more comprehensive at the mail place (clinic) where the patient is being treated. In addition, there are no routine records for patient and family resource use.

Turning to the specific resource items identified above, we might expect to record quantities used as follows (Table 2):

Table 2. Measurement of resource use in economic evaluation.

Resource items	Description
Hospital care	
Radiotherapy	The number of treatment sessions could be recorded, possibly differentiating by length of session and time of day (e.g. normal working hours, after hours, weekends)
Bed days	The number of bed days could be recorded, differentiating by type of hospital ward
Out-patient attendances	The number of attendances could be recorded
Overheads	These would probably be related to the number of bed days or other suitable resource item
Community care	
General practitioner visits	The number could be ascertained, either by asking patients, or by consulting the general practitioners. It may make sense to differentiate between home visits and visits to the practitioner's office
Nurse visits	The number could be recorded as for general practitioner visits above. The purpose of the nurse visit and type of nurse (e.g. general nurse, specialist cancer nurse) would be recorded
Ambulance or hospital car	The number and length of trips could be recorded. Length of trip could be ascertained from the patient's place of residence
Patient and family resources	
Patient's time	The time taken in seeking and receiving care could be estimated by asking the patient. Time off work could be estimated separately
Relatives' time	Relatives could spend time in home nursing and in accompanying patients to hospital. It could be estimated as for patients' time above
Out-of-pocket expenses	Some may be estimated directly in money terms, others may be estimated by asking patients
Resources in other sectors	
Social worker and home help visits	These would be estimated in a similar way to nurse visits above.

Valuation of resource items

It is extremely difficult to give general advice on this because it is so dependent on the availability of local financial data. In some settings, like the USA, there may be data on hospital billings and charges. In other settings, detailed costing studies would be necessary. When using charge data, it is important to:

1. investigate the relationship between charges and costs;
2. record physical quantities as well as charges, so as to facilitate generalization of study results to other settings.

Turning to the specific resource items measured above, we might expect to value them as follows (Table 3).

Table 3. Valuation of resource items in economic analysis.

Resource items	Valuation
Hospital care	
Radiotherapy treatment sessions	In some settings there may be charge data, or average cost figures, for radiotherapy sessions. However, even if these exist, which is unlikely in many locations, they may not differentiate by type of session (for example, normal hours, out-of-hours, or weekend). This distinction is critical to understanding the relative cost of conventional radiotherapy and CHART. Therefore, it is likely that micro-costing would be required. In micro-costing the approach would be to derive the cost of a treatment session from its component parts, namely consultant (medical) time, radiographer time, medical physics time, consumables, equipment, buildings, and departmental overheads. Some survey work may be required, plus data from the hospital finance department on staff salaries, overtime allowances, and equipment prices. Costing of equipment and buildings will require assumptions to be made about useful life and resale value. It would be necessary to apportion these costs to individual treatment sessions. Judgements would also need to be made about which components of hospital overheads (for example, cleaning, building maintenance, or administration) are most appropriately allocated to departments and the allocation basis (for example, square metres, cubic metres, number of stuff, and so on). Some elements of overhead may be better allocated on the basis of in-patient days or number of patients.
Bed days	It may be possible to use the average daily costs (or per diems) for different types of wards, including hostel wards. However, these may be considered too imprecise, in which case micro-costing might be undertaken. This would derive a daily cost for a particular category of ward by considering nurse stuffing levels, medical (consultant) input, and overheads. Because hostel wards may not feature in the standard hospital accounts, micro-costing may be required for these, for example, they may be slightly off site or rely partly on staffing by volunteers. An opportunity cost for volunteer time may have to be inputted. In costing hospital beds it may be decided to make an allowance for the fact that there is usually less than 100% occupancy.
Out-patient attendances	There may be an average cost or charge available for an out-patient visit, although this may not differentiate between oncology and other clinical specialties. Depending on the quantitative importance of this item, micro-costing may be undertaken.

Table 3. Cont.

Resource items	Valuation
Overheads	As mentioned above, these could be allocated to the radiotherapy treatments, to out-patient attendances, or to hospital bed days, depending on the overhead item.
Community care	
General practitioner visits	There may be data available on physician fees for various types of visits (e.g. general assessment, home visit, etc.). Alternatively, there may be nationally available data on the average costs of various general practitioner services. Failing this, micro-costing may be required. This would calculate the cost of practitioners' time (per minute or per hour) and add the cost of travel for home visits. Drug costs would also need to be considered
Nurse visits	The agencies providing the nurses may have data on the average cost of a visit. This may even distinguish between various types of visit. Failing this, micro-costing would have to be employed, taking into account nursing salaries, length of visits, travel time, and nurses time spent in general administration. There may also be some consumables to be accounted for in the cost of nurse visits.
Ambulance and hospital car	Estimates may be available for the average cost per mile travelled. This could be combined with data on the distances involved to generate total costs.
Patient and family resources	
Patients' time	If the time was taken from work-time, the gross salary (including employment benefits) could be used. Different assumptions could be made about the opportunity cost of leisure time
Relatives' time	In general the valuation of this raises the same issues as the valuation of patients' time. The valuation of time spent in informal nursing care is complicated because the relative may also be able to carry out other tasks at the same time.
Out-of-pocket expenses	In general, the financial expenditures made (for example, bus fares) would suffice. However, for some items, such as use of one's private car, the expenditures (say) on fuel would underestimate the true cost. Here, monitoring organizations can often provide data on the cost (per mile or kilometre) of running a car.

Finally, a few rare events, such as hospital admission for particular types of surgery, may be handled separately. Depending on how quantitatively important they seem, case-mix group costs or disease-specific *per diems* may suffice. Alternatively, micro-costing may be undertaken.

CASE STUDY 2: CALCULATION OF ECONOMETRIC MEASURES IN PRACTICE

Data set

The four main methods of an economic evaluation in health care have been introduced in this module. The study questions that follow require the student to select and assemble the appropriate figures from a set of categories of costs and outcomes to conduct each of the four kinds of evaluation.

The evaluation being considered in Table 4 is based upon fictive data on provision of intensive care that involves increased current capital expenditures in order to increase a patient's future survival chances. The costs and consequences for patients are listed in Table 4.

Table 4. Evaluation of intensive care treatment.

Cost or consequence	Before intensive care	With intensive care	Incremental effect*
Cost per additional survivor (to hospital discharge)	3,400€ C_1 *(hospital)*	12,200€ C_2 *(hospital)*	8,800€
Cost per additional survivor (to death)	72,500€ C_1 *(society)*	80,100€ C_2*(society)*	7,600€
Survival rate (to hospital discharge)	65.4% E_{1a}	80.2% E_{2a}	14.8%
Survival time (per live birth):			
a. Life-years	18.8 E_{1b}	27.7 E_{2b}	8.9
b. QALYs	17.4 $QALY_1$	36.0 $QALY_2$	8.6
Earnings	92,200€ B_1	124,200€ B_2	32,000€

LEGEND: *Incremental effect represents the difference, in cost or effect, between the two programmes being compared in economic evaluation.

Cost-minimization analysis

A CMA comparison of intensive care (before versus with intensive care) from: (a) the hospital's perspective and (b) society's perspective is presented in Equation 11 and Equation 12. For calculation Equation 2 should be used.

a) Hospital perspective (Equation 12):

$$\Delta C_{(\text{hospital})} = 12{,}200€ - 3{,}400€ = 8{,}800€ \qquad \textbf{Equation 12.}$$

$\Delta C_{(hospital)}$= *cost difference from hospital's perspective*

From the hospital perspective, cost per additional survivor before intensive care is 3.400€ and with intensive care is 12,200€. The intensive care increases the cost per additional survivor for 8,800€.

b) Societal perspective (Equation 13):

$$\Delta C_{(society)} = 80{,}100€ - 72{,}500€ = 7{,}600€ \qquad \textbf{Equation 13.}$$

$\Delta C_{(society)}$= cost difference from society's perspective

From the societal perspective, cost per additional survivor before intensive care is 72,500€ and with intensive care is 80,100€. The intensive care increases the cost per additional survivor for 7,600€.

Cost-effectiveness analysis

A CEA comparison of intensive care from society's perspective reveals that intensive care provides a better health outcome, but it is also more expensive. Consequently, the incremental cost-effectiveness ratio should be calculated (Equation 14 and Equation 15 for survival rate and survival time, respectfully). For calculation Equation 4 should be used.

a) Survival rate (Equation 13):

$$iCER_{(society)} = \frac{72{,}500€ - 80{,}100€}{65.4 - 80.2} = \frac{-7{,}600€}{-14.8} = 513.5€ \qquad \textbf{Equation 14.}$$

$iCER_{(society)}$ = incremental cost-effectiveness ratio from society's perspective

From the societal point of view, an increase in survival rate per 1% would cost additional 513.5€

b) Survival time (Life-years) (Equation 14):

$$iCER_{(society)} = \frac{72{,}500€ - 80{,}100€}{18.8 - 27.7} = \frac{-7{,}600€}{-8.9} = 853.9€ \qquad \textbf{Equation 15.}$$

$iCER_{(society)}$ = incremental cost-effectiveness ratio from society's perspective

From the societal point of view, a prolongation of survival time for 1 life-year would cost additional 853.9€.

Cost-utility analysis

A CUA comparison of intensive care from society's perspective is presented in Equation 16. To form CUA comparison, only QALYs can be used as a consequence. For calculation Equation 6 should be used:

$$iCUR_{(society)} = \frac{72{,}500€ - 80{,}100€}{17.4 - 36.0} = \frac{-7{,}600€}{-18.6} = 7{,}581.4€ \qquad \textbf{Equation 16.}$$

$iCUR_{(society)}$ = incremental cost-utility ratio from society's perspective

From the societal point of view, one QALY per survivor gained would cost additional 7,581.4€.

Cost-benefit analysis

For CBA comparison of intensive care from society's perspective, the iCBR and Net benefit should be calculated. iCBR is presented in Equation 17 and net benefit is presented in Equation 18. To form CBA comparison, only earnings/benefits in monetary terms can be used as a consequence. For calculation Equation 10 and Equation 11 should be used.

a.) iCBR (Equation 17):

$$iCBR_{(society)} = \frac{72{,}500€ - 80{,}100€}{92{,}200€ - 124{,}200€} = \frac{-7{,}600€}{-32{,}000€} = 0.2375 \qquad \textbf{Equation 17.}$$

$iCBR_{(society)}$ = incremental cost-benefit ratio from society's perspective

iCBR shows that for each 0.2375 euro spent on intensive care treatment, the society would earn 1 euro per additional survivor.

b.) net benefit (Equation 18):

$$NB_{(society)} = (92{,}200€ - 124{,}200€) - (72{,}500€ - 80{,}100€) = \\ = 32{,}000€ - 7{,}600€ = 24{,}400€ \qquad \textbf{Equation 18.}$$

$NB_{(society)}$ = net benefits from society's perspective

The net benefits to the society from the intensive care treatment would be 24,400€ per additional survivor.

EXERCISE

Task 1

In table 5 data for undertaking an economic analysis are presented.

Table 5. Evaluation of treatment B versus treatment A.

Cost or consequence	Treatment A	Treatment B
Cost per additional survivor (to hospital discharge)	2,200€	15,100€
Cost per additional survivor (to death)	60,500€	90,700€
Survival rate (to hospital discharge)	56.4%	91.1%
Survival time (per live birth):		
a. Life-years	22.2	30.8
b. QALYs	16.1	40.6
Earnings	88,300€	134,900€

Undertake and comment:

1. a CMA comparison of treatment B versus treatment A from: (a) the hospital's perspective and (b) society's perspective.
2. a CEA comparison of intensive care from society's perspective.
3. a CUA comparison of intensive care from society's perspective.
4. a CBA comparison of intensive care from society's perspective.

Task 2

In four groups, perform an internet search of bibliographic data-bases (e.g. Medline, Pubmed) to find cases of four types of economic evaluation analysis, each group respectively. The groups should critically discuss the strengths and weaknesses of the selected cases and than present them to other students.

Task 3

The four groups have the roles of decision-makers. They are considering the approval of cases selected in Task 2. What additional information would they want to consider along with economic evaluation? Discuss in groups and present the findings to other students.

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of healthcare programme. Oxford: Oxford University Press; 2005.
2. Zerbe RO, Bellas AS. A primer for benefit-cost analysis. Northampton: Edward Elgar Publishing; 2006.
3. Schulman KA, Seils D. Clinical Economics. In: Interactive Textbook on Clinical Symptom Research. Available from URL: http://symptomresearch.nih.gov/chapter_12/index.htm. Accessed: March 18, 2009.

4. Briggs AH. Handling uncertainty in economic evaluation. In: Drummond MF, McGuire A, eds. Economic evaluation in health care: merging theory with practice. Oxford: Oxford University Press; 2001: p.172-214.
5. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004: p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: March 18, 2009.
6. Rascati KL. Essentials of pharmacoeconomics. Philadelphia: Wolters Kluver/Lippincot, Williams&Wilkins; 2009.
7. Prieto L, Sacristan JA. Problems and solutions in calculating quality-adjusted life-years (QALYs). Health Qual Life Outcomes. 2003;1: 80.
8. Brent RJ. Cost-benefit analysis and health care evaluations. Northampton: Edward Elgar Publishing; 2003.
9. Phelps C, Mushlin A. On the near equivalence of cost-effectiveness and cost-benefit analyses. Int J Technol Assessment Health Care. 1991;7:12-21.
10. Birch S, Gafni A. Cost effectiveness/utility analyses. Do current decision rules lead us to where we want to be? J Health Econ. 1992;11:279-96.
11. Economic Analyses. Merck Manuals Online Medical Library for Healthcare Professionals. Merck & Co., Inc., Whitehouse Station, NJ, USA, 2005. Available from URL: http://www.merck.com/mmpe/sec22/ch328/ch328i.html#CIHHGBJG Accessed: March 18, 2009.
12. Weinstein M, Zeckhauser R. Critical ratios and efficient allocation. J Pub Economics. 1973;2:147-158.
13. Weinstein M, Stason W. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296:716-721.
14. Robinson R. Cost-benefit analysis. Br Med J. 1993;307:859-862.
15. Byford S, Raftery J. Perspectives in economics evaluation. Br Med J. 1998;316:1529.
16. Coyle D, Drummond MF. Costs of conventional radical radiotherapy versus continuous hyperfractionated accelerated radiotherapy (CHART) in the treatment of patients with head and neck cancer and carcinoma of the bronchus. Centre for Health Economics, University of York, York; 1997.

RECOMMENDED READINGS

1. Brent RJ. Cost-benefit analysis and health care evaluations. Northampton: Edward Elgar Publishing; 2003.
2. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
3. Rascati KL. Essentials of pharmacoeconomics. Philadelphia: Wolters Kluver/Lippincot, Williams&Wilkins; 2009.
4. Wenzel H, Hysa B. Economic appraisal as a basis for decision-making in health systems. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2004: p.376-403. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: March 18, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**SWOT ANALYSIS**
Module: 4.4.1	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Andrej Plesničar**, MD, PhD, Teaching Assistant Faculty of Health Studies, University of Ljubljana, Slovenia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia
Address for correspondence	**Andrej Plesničar** Faculty of Health Studies, University of Ljubljana, Slovenia Zdravstvena pot 5, Ljubljana, SI-1000 Slovenia E-mail: andrej.plesnicar@vsz.uni-lj.si
Keywords	Strategic planning, SWOT analysis
Learning objectives	After completing this module students should: • understand the role of strategic planning in public health; • be familiar with SWOT analysis as one of situation assessment tools in strategic planning; • be able to perform a SWOT analysis.
Abstract	SWOT analysis is a strategic planning tool used to evaluate different types of positive and negative factors which may affect desired future outcomes of a project, programme, process, etc. As a case study, a SWOT analysis of the 2nd PH-SEE Summer School for Public Health Professionals is presented. The study was performed by the Slovenian participants of the course at the end of the course.
Teaching methods	An introductory lecture gives the students first insight in characteristics of SWOT analysis. The theoretical knowledge is illustrated by a case study. In continuation they first carefully read the recommended readings. Afterwards they discuss the characteristics of SWOT analysis with other students. Afterwards, in small groups (2-3 students) they perform a SWOT analysis of a given example.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	SWOT analysis of a situation of a selected process.

SWOT ANALYSIS

Andrej Plesničar, Lijana Zaletel-Kragelj

THEORETICAL BACKGROUND

Basic considerations

SWOT analysis definition

SWOT analysis is a strategic planning tool used to evaluate different types of positive and negative factors which may affect desired future outcomes of a project, programme, process, etc. SWOT stands for (1-3):

- S – strengths,
- W – weaknesses,
- O – opportunities, and
- T – threats.

Before describing the characteristics of SWOT analysis, it is important to place this analysis in wider frame of strategic planning.

Strategic planning definition and basic characteristics

There exist several more or less similar definitions of strategic planning, two of them being:

- according to Geyer (2), strategic planning is a process, by which an organisation analyses whether it is effective in its goals and objectives, and it establishes whether the organisation needs to change its direction to fulfil its purpose. Strategic planning helps to respond to the external environment in the most effective way.
- according to Lerner (3), strategic planning is a complex and ongoing process of organizational change. It is based on thorough analysis of foreseen or predicted trends and scenarios of the possible alternative futures, as well as the analysis of internal and external data. It is a qualitative, idea driven process. It integrates "soft" data, not always supported quantitatively, such as experiences, intuition, and ideas, involves the organization in the ongoing dialogue, and aims to provide a clear organizational vision and focus. It is an ongoing, continuous learning process, an organizational dialogue, which extends beyond attaining a set of predetermined goals. It aims to change the way an organization thinks and operates, and create a learning organization;

SWOT analysis in the frame of strategic planning

Strategic planning is a several steps process. According to Geyer (2), these steps are:

1. Getting organised,
2. Analysis of the situation,
3. Development of a strategy,
4. Drafting a plan,
5. Approval of the plan, and

6. Implementation of the plan.

One of the most important steps is analysis of the situation (2). SWOT analysis is one of many tools which could be used in this step of strategic planning, but SWOT analysis is by far the most popular (3).

Description of SWOT analysis

Factors, assessed in SWOT analysis could be internal or external. This tool involves specifying the objective of the project and identifying four groups of factors:

- positive (favourable) internal factors or strengths,
- negative (unfavourable) internal factors or weaknesses,
- positive (favourable) external factors or opportunities, and
- negative (unfavourable) external factors or threats.

These four groups of factors could be presented in a SWOT analysis matrix (Figure 1).

	Positive factors	Negative factors
Internal factors	STRENGTHS	WEAKNESSES
External factors	OPPORTUNITIES	THREATS

Figure 1. SWOT analysis matrix.

SWOT analysis has been widely employed in industry and management as a decision making aid and as an introduction to planning in various applications. It's been also used to uncover new outlooks and identify problems that would impede progress. (1).

Doing SWOT analysis

SWOT analysis is a simple and quick method. The first step would be to make a simple worksheet by drawing a cross, creating four cells: one each for Strengths, Weaknesses, Opportunities and Threats. The next step is to list specific items related to the problems of the programme into the appropriate cells in the worksheet. Alternatively, a more detailed SWOT analysis can be made with the use of the template of a more elaborate worksheet scheme (Figure 2).

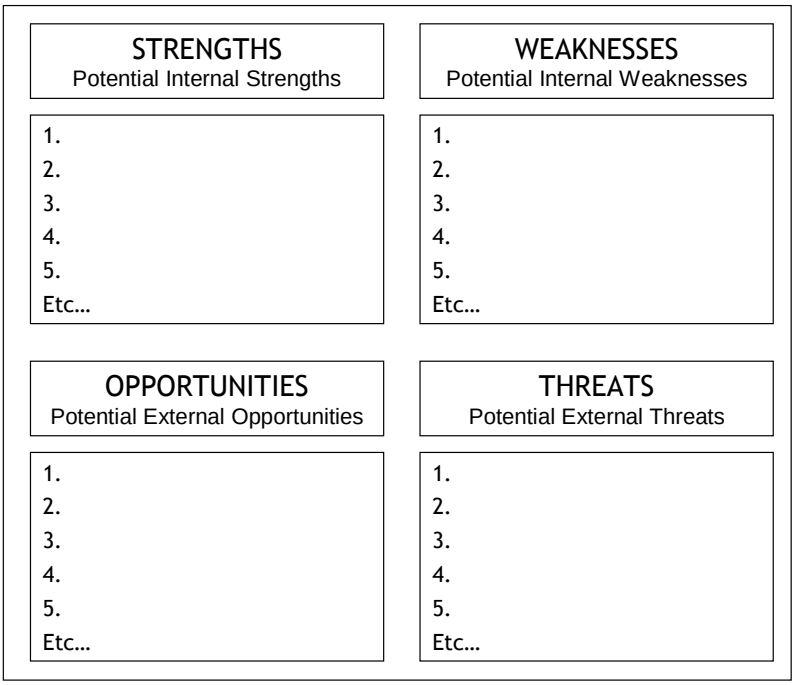


Figure 2. SWOT analysis worksheet template.

It is advised that no more than ten exact points should be listed under each heading (1); all of them should be carefully evaluated in the following sequence:

1. Do the internal analysis.
   Internal analysis examines the advantages and drawbacks of the programme. This can be achieved with the analysis of **strengths** and **weaknesses** of the programme during its course and the impact they have on participants' perception.
2. Do the external analysis.
   External analysis examines the main points in the analysis of the actual professional circumstances and environment of the participants. It also identifies those points that represent **opportunities**, and those that represent **threats** or obstacles to their future achievements. Thus, classify the answers or the data collected as external **opportunities** or **threats**.
3. Enter the collected information into the worksheet (Figure 2).
4. Apply the collected information to devise a strategy that uses the **strengths** and **opportunities** to reduce the **weaknesses** and **threats**, and to achieve the objectives of the education programme in the participants' future professional performance.

SWOT analysis in public health

Strategic planning and particularly SWOT analysis could be used in public health, since public health problems require strategic planning. Among others, strategic planning could be useful methodology in health promotion, especially in mobilizing community partnerships to identify and solve health problems of a population, and informing, educating people about health issues. In fact, SWOT analysis is used in public health issues. One of examples is SWOT analysis of Halton's Healthy Living Project (4). However, SWOT analysis could be used in other fields of public health as well (5,6).

In short, an analysis of "internal" (Strengths and Weaknesses) and "external" (Opportunities and Threats) factors affecting any public health programme may help in making a plan for the future. Irrelevant and outdated programmes could thus be replaced with innovative and competent ones.

SWOT analysis could be also a method for assessing the quality and relevance of public health education programmes. It can be used by groups or individuals as an external assessment of such a programme, or may represent introspection into negative and positive concerns.

CASE STUDY: SWOT ANALYSIS AS AN EVALUATION TOOL OF PUBLIC HEALTH EDUCATION PROGRAMMES - PARTICIPANTS' ASSESSMENT OF 2nd PH-SEE SUMMER SCHOOL FOR PUBLIC HEALTH PROFESSIONALS

Introduction

Parliamentary democracy and free market economy were reintroduced to the countries of South Eastern Europe (SEE) in the last decade of 20th century after an interval of more than 50 years. In the majority of these countries democratically elected governments assumed responsibility in the late eighties and early nineties, and the region has been referred to as being in political and social transition. It is important to note that epidemiological transition had in some of the former socialist countries of SEE occurred long before the beginning of these changes. However, the problems associated with economic restructuring have been in these countries often compounded by the inability of inefficient governments and judiciaries to deal with them properly. Decreasing gross domestic product, extensive environmental damage, rising international debt, massive unemployment resulting from closures of large and ineffective socialist style companies, and growing corruption have been in some cases exacerbated by political instability and war. The extent of social and political reforms and the ability to create a functioning market economy are therefore not uniform in SEE and the population of the region frequently has to cope with the aftermath of economic deprivation, alienation and social inequality. Not least of all, the health expenditures remained on almost basic levels in some of these countries and have contributed to a decade of disturbing trends in public health indicators. As a consequence, the awareness of the need to tackle these problems has gradually emerged among the health professionals and people of SEE itself and Europe as a whole (7,8).

The concept of public health has simultaneously evolved into a broader philosophy that is concerned not only with the provision of preventive and therapeutic health services, but also with other components important for the functioning of health care systems. These include the dilemmas associated with medical personnel and facilities, environment, life styles, economic support and just distribution of health services. The philosophy of New Public Health therefore links the traditional Public Health issues of epidemiology and hygiene with more recent issues of health promotion, social responsibility, economics and management (9). As the New Public Health is concerned with the totality of the health care systems, it corresponds well to the present day circumstances and health problems in SEE. It is thus imperative for the Public Health experts in SEE to continuously upgrade and broaden their knowledge, and to receive extensive parallel education in related fields of biomedical, social and anthropological sciences, economics, management, and advances in technology (5,10,11).

Recent developments in SEE countries have led the international community to establish a comprehensive, long-term conflict prevention strategy for this region. The activity of Stability Pact for SEE, a partnership of more than 40 countries and formed following the European Union initiative in Cologne on June 10th, 1999, has in this role replaced earlier haphazard reactive intervention policies in SEE. It supports the SEE countries in their "efforts to foster peace, democracy, respect for human rights and economic prosperity in order to achieve stability in the whole region" through the implementation and enhancement of regional co-operation, and serves as an instrument to co-ordinate and facilitate the projects of all its partners (12). With regard to the need of preparing public health professionals in SEE to upgrade their knowledge in order to follow and cope with the developments mentioned before, a PH-SEE (Public Health Collaboration in South Eastern Europe) initiative, including Programmes for Training and Research in Public Health, was developed within the framework of Stability Pact for SEE. The 2nd PH-SEE Summer School for Public Health Professionals (2nd PH-SEE-SSPHP) that took place on July 21st-27th, 2002, in Ljubljana, Slovenia, represents a part of the PH-SEE curriculum with clearly stated purposes and objectives (13). Accordingly, the participants were asked to collect and present their assessments and opinions of the 2nd PH-SEE-SSPHP immediately after the lectures and exercises ended.

Lecturers, participants and the set-up of the 2nd PH-SEE-SSPHP

The faculty of the 2nd PH-SEE-SSPHP included twelve invited lecturers. All of them were well known experts in Public Health and some of the related fields such as statistics, management, physics and information technologies in their home countries. Three of the lecturers came from Croatia, four from Germany, and five of them came from Slovenia. An additional lecture was given on an ad hoc basis by a participant from Macedonia. In addition, three members of the Organising Committee were constantly available on the spot to help the participants with equipment or advice if needed.

The majority of the participants of the 2nd PH-SEE-SSPHP came from the SEE countries. The largest number was from Slovenia, while others came from Albania,

Bosnia and Herzegovina, Bulgaria, Croatia, Romania, Macedonia, Serbia (including participants from Voivodina) and Montenegro (Table 1). One of the participants came from Germany, and there were no participants from Republic of Moldova and Kosovo. Altogether, out of 32 participants 26 were females and six were males. All had university degrees in medicine and/or experience in dealing with issues of Public Health. Among the Slovenian participants, nine out of 14 worked at regional Institutes of Public Health. Only five out of 18 participants from other countries held a similar position, with the rest being employed in various government departments and universities of their respective countries. All participants were proficient in English.

Table 1. Regional distribution of the participants of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, Slovenia, July 21st-27th, 2002.

Country	Number of participants	
Albania	3	
Bosnia and Herzegovina	1	
Bulgaria	2	
Croatia	1	
Germany	1	
Macedonia	1	
Romania	2	
Slovenia	14	
Yugoslavia	7	
Voivodina		3
Serbia		2
Montenegro		2

The 2nd PH-SEE-SSPHP took place in the premises of the National Health Insurance Institute (NHII) in Ljubljana, Slovenia. The venue of the 2nd PH-SEE-SSPHP was in the basement of the NHII building in the centre of the city in a well equipped and comfortable lecture room with effective air conditioning system. All the participants had an access to personal computers (PC), usually two per PC, and all the PCs were provided with relevant software programmes and internet links. The language of the 2nd PH-SEE-SSPHP was English, while some of the participants conversed among themselves in other languages as well. Catering was provided by the NHII for all of the participants.

Evaluation of public health education programmes and SWOT analysis

Involvement of students in medical education programmes planning and evaluation on all levels has by now become an accepted practice in some parts of the world. This type of evaluation of Public Health education programmes thus forms an integral part of the programmes' continuous professional and scientific quality assessment and assurance (7,1). It should be as accurate and comprehensive as possible, and it is advisable a certain sequence of steps in the analysis and assessment of a programme is followed (Table 2). It should not be taken as a test resulting in either "pass or fail".

Table 2. A sequence of suggested steps in the analysis of a Public Health education programme.

Step		Description
1.	Assessment of the Public Health education programme; overall assessment:	• assessment of the plans of the programme. • assessment of the procedures of the programme: – planned goals. – exercise objectives
2.	Assessment of the "host" facilities of the Public Health education programme:	• faculty of the programme. • premises available for the programme: – provision of resources (equipment, etc…). – familiarity with the circumstances and conditions
3.	Assessment of the participants of the Public Health education programme:	• capability to carry out tasks as instructed (exercises, etc…). • adherence to the agreed standards

On the last day the lecturers and organisers asked the participants to evaluate the 2nd PH-SEE-SSPHP and present their assessments and opinions in the final discussion of the course. The participants were divided into two groups: the first was made from those from Slovenia, and the second from those from other countries. The participants were asked to discuss the advantages and drawbacks of the 2nd PH-SEE-SSPHP and for this purpose the Slovenian participants chose to use the SWOT analysis in their assessment, as mentioned previously in the text. The participants had 15 minutes to prepare their evaluation immediately after the last lecture at the 2nd PH-SEE-SSPHP ended.

Results of the SWOT analysis of the 2nd PH-SEE-SSPHP

The results of the evaluation are necessarily fragmentary and incomplete despite the fact that SWOT analysis was used. The factors affecting the 2nd PH-SEE-SSPHP are listed in each cell of the SWOT analysis template from top to bottom, according to their importance as perceived by the Slovenian participants (Figure 3).

It is the Slovenian participants' impression that the "host" qualities of the 2nd PH-SEE-SSPHP easily translate into majority of its Strengths. Highly qualified lecturers, well organised seminars with relevant practical exercises, study and exercise materials of good quality, presentations of carefully selected skills and knowledge relevant to the interests of the participants all attest to the success of the faculty of the Summer School, its International Programme Committee and its Organising Committee in their efforts to offer the participants relevant up-to-date information. For example, practical exercises enabled the participants to prepare presentations of comparisons of health indicators to place their national pattern in the context of other countries with similar economic and social circumstances. Well

organised course of the 2nd PH-SEE-SSPHP with exact timing, well equipped premises and excellent catering made the attendance easier and more comfortable for all the participants and added to its success. Although the largest part of participants came from Slovenia, the participants' regional distribution was still regarded to be good. Finally, in Slovenian participants' view, all the participants were highly motivated to attend the 2nd PH-SEE-SSPHP.

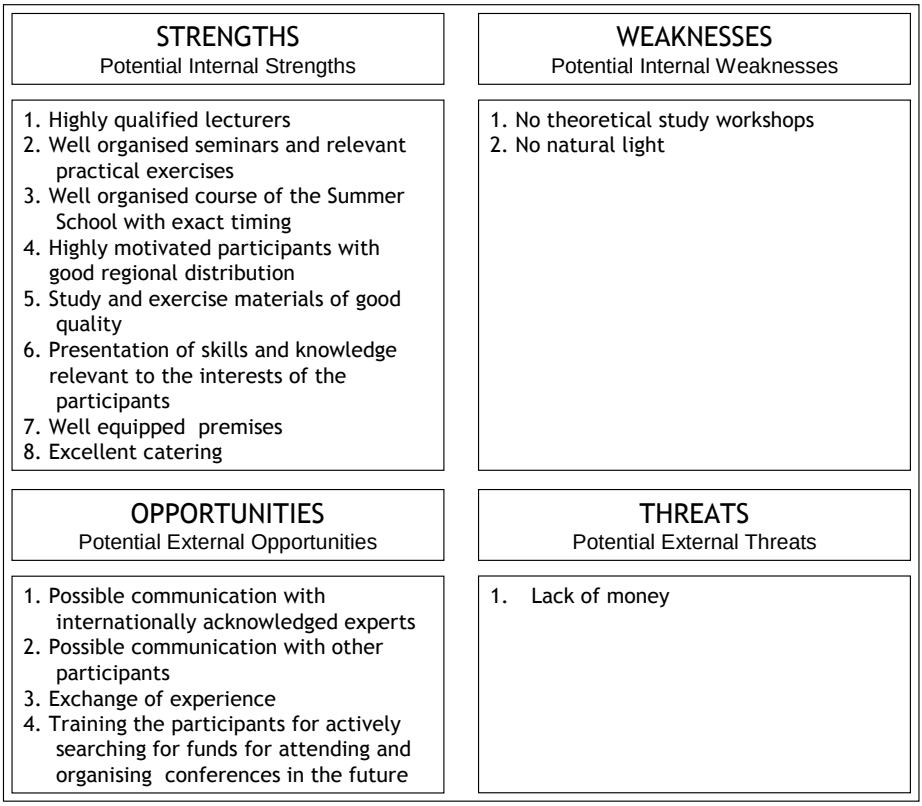


Figure 3. Results of the SWOT analysis of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, Slovenia, July 21st-27th, 2002, performed by Slovenian participants..

The absence of theoretical study workshops was regarded by some of the Slovenian participants as one of the internal weaknesses of the 2nd PH-SEE-SSPHP. The problem was not discussed extensively during the SWOT analysis and the missing theoretical workshops were not clearly defined. In one participant's opinion it was suggested that relevant thematic roundtables should be organised on similar meetings in the future. The lack of natural light was regarded only as a minor drawback of the 2nd PH-SEE-SSPHP.

External opportunities also seem to be closely associated to the "host" qualities of the 2nd PH-SEE-SSPHP. Formal and informal discussions between the experts of the faculty and participants led to the exchange of valuable experience during the course of the 2nd PH-SEE-SSPHP that should be further encouraged in the future. Training the participants for actively searching for funds for attending and organising conferences in the future was regarded as important by the majority of Slovenian participants, but for many of them it seemed overly complicated with little chance of success and not relevant in the context of their professional duties and priorities.

Among the threats the Slovenian participants emphasised the "lack of money". Although none of them had any particular financial problems with regard to the participation at the 2nd PH-SEE-SSPHP, it was regarded as a sort of an "overall" threat to the feasibility of similar Public Health education meetings in the future.

Discussion

Evaluation of the 2nd PH-SEE-SSPHP with SWOT analysis, as performed by the Slovenian participants, stressed its significance in the context of present day economic, social and political circumstances in the SEE countries. It confirmed the importance of continuing education in Public Health in its broader context and provided the participants with some important new practical information and skills. It is clear from the SWOT analysis that the 2nd PH-SEE-SSPHP with its Strengths and Opportunities forms a valuable part of the PH-SEE initiative and its Programmes for Training and Research in Public Health.

It can be seen from the SWOT analysis of the 2nd PH-SEE-SSPHP that no major modifications are needed for the future meetings of similar nature. The 2nd PH-SEE-SSPHP helped to increase the sense of self-esteem of the participants, reinforced their knowledge, and has probably motivated them to better performance. Many of them may in the future assume responsibilities as health planners, policy analysts and administrators in Public Health departments of their governments, and as researchers and teachers in educational institutions, communities, organised medical care groups and voluntary health organisations (9,15). Their expertise may thus facilitate the course of anticipated reforms needed to upgrade quality in health care systems that must cope with changes in the health needs of societies in SEE countries (6-8,13).

SWOT analysis was effectively used in the evaluation of the 2nd PH-SEE-SSPHP. However, as the circumstances change with the passage of time due to national and regional influences, an updated analysis should be repeated frequently (16). Committees and other groups could use a SWOT analysis as a way of introspection into positive and negative concerns when deciding about new Public Health education programmes in SEE and elsewhere. SWOT analysis can help decision makers and individuals expand their vision and facilitate necessary adjustments. In any case, decision making should build on Strengths, reduce Weaknesses, seize Opportunities and react to Threats (17).

In conclusion, it is safe to say that PH-SEE initiative with its Programmes for Training and Research in Public Health programmes makes an important contribution in achieving self-sustaining regional stability in SEE. The 2nd PH-SEE-SSPHP represents one of the steps ahead to achieve such stability, with all of the participants

having a clearer vision where the region, its Public Health experts and health care systems should be heading.

EXERCISE

Task 1

Carefully read the part on theoretical background of this module. Critically discuss the characteristics of SWOT analysis with your colleagues.

Task 2

Carefully read the following two documents:

- CINDI World Health Organization, Regional Office for Europe. A strategy to prevent chronic disease in Europe. A focus on public health action. The CINDI vision. Copenhagen: WHO, Regional Office for Europe; 2004. Available from: URL: http://www.euro.who.int/document/E83057.pdf (Accessed: August 10, 2007).
- Maucec Zakotnik J, Fras Z, Zaletel Kragelj L. The WHO Countrywide Integrated Non-communicable Diseases (CINDI) Programme in Slovenia. In: Donev D, Pavleković G, Zaletel-Kragelj L (editors). Health promotion and disease prevention. A handbook for teachers, researches, health professionals and decision makers. Lage: Hans Jacobs Publishing Company, 2007. p.204-219. http://www.snz.hr/ph-see/Documents/Publications/FPH-SEE_Book_on_HP&DP.pdf

Task 3

Make groups of three to four students and make SWOT analysis of CINDI programme in Slovenia (or some other CINDI country).

Task 4

Discuss results of your SWOT analysis with results of other student groups.

REFERENCES

1. Johnson G, Scholes K, Sexty RW. Exploring strategic management. Scarborough, Ontario: Prentice Hall; 1989.
2. Geyer Y. Strategic planning. Handbook series for community-based organisations. Pretoria: IDASA, 2006. Available from URL: www.idasa.org.za/gbOutputFiles.asp?WriteContent=Y&RID=1703. Accessed: March 18, 2009.
3. Lerner AL. A Strategic Planning Primer for Higher Education. Northridge: College of Business Administration and Economics, California State University; 1999. Available from URL: http://www.sonoma.edu/aa/planning/Strategic_Planning_Primer.pdf. Accessed: March 18, 2009.

4. Samuels T, Perry C. A SWOT analysis of Halton's Healthy Living Project: prospects for sustainability. Executive Summary. Chester: The University of Chester; 2005. Available from URL: http://www.chester.ac.uk/cphr/reports/025.pdf. Accessed: March 18, 2009.
5. Miller MG. Environmental metabolomics: a SWOT analysis (strengths, weaknesses, opportunities, and threats). J Proteom Res. 2007;6:540-5.
6. Dageid W, Duckert F. The process of evaluating a capacity-building support initiative for HIV positive South Africans. Eval Program Plann. 2007;30:381-91.
7. Chenet L, McKee M, Fulop N, Bojan F, Brand H, Hort A, Kalbarczyk P. Changing life expectancy in Central Europe: is there a single reason? J Public Health Med. 1996;18:329-36.
8. Plesničar AF. A glimpse on the cancer problem in Latin America from a Central European perspective: II. Challenge Newsletter. 2000:2:14-22.
9. Tulchinsky TH, Varavikova EA. The new public health. An introduction for the 21st century. San Diego, CA: Academic Press; 2000.
10. Kohler L. Public health renaissance and the role of schools of public health. Europ J Public Health. 1991;1:2-9.
11. Kahn K, Tolmann SM. Planning professional education at schools of public health. Am J Public Health. 1992;82:1653-7.
12. Stability Pact [homepage on the Internet]. Available from URL: http://www.stabilitypact.org. Accessed: Aug 25, 2002.
13. Laaser U, Bardehle D. Health monitoring and reporting systems. Proccedings of the 2nd PH-SEE Summer School for Public Health Professionals, Ljubljana, July 21st-27th. Ljubljana: University of Ljubljana, Faculty of Medicine; 2002.
14. World Federation for Medical Education. World summit of medical education. Edinburgh, London: World Federation for Medical Education; 1993.
15. Roemer M. Genuine professional doctor of public health the world needs. J Nurs Scholarship. 1999;31:43-4.
16. Glass NM: Pro-active management: How to improve your management performance? East Brunswick, NJ: Nichols Publishing; 1991.
17. Balamuralikrishna R, Dugger JC. Swot analysis: a management tool for initiating new programs in vocational schools. J Vocat Tech Training. 1995;12:108-15.

RECOMMENDED READINGS

1. Geyer Y. Strategic planning. Handbook series for community-based organisations. Pretoria: IDASA; 2006. Available from URL: www.idasa.org.za/gbOutputFiles.asp?WriteContent=Y&RID=1703. Accessed: March 18, 2009.
2. Lerner AL. A Strategic planning primer for higher education. Northridge: College of Business Administration and Economics, California State University; 1999. Available from URL: http://www.sonoma.edu/aa/planning/Strategic_Planning_Primer.pdf. Accessed: March 18, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**RAPID ASSESSMENT AND RESPONSE - RAR**
Module: 4.4.2	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Enida Imamović**, MD, DDM Public Health Institute of Federation B&H, Bosnia and Herzegovina **Dragana Nikšić**, MD, PhD Medical Faculty, University of Sarajevo, Bosnia and Herzegovina
Address for Correspondence	**Enida Imamović** Public Health Institute of Federation B&H Titova 9, 71000 Sarajevo, Bosnia and Herzegovina Tel: + 387 33 26 42 00 Fax: + 387 33 22 05 48 E-mail: imamovic@bih.net.ba; e.imamovic@zzjzfbih.ba
Keywords	Rapid assessment, response
Learning objectives	After completing this module students and public health professionals should: • understand the specificities of the rapid assessment and response (RAR) methodology • improve their knowledge of the methods used in „Rapid Assessment“ • know to identify the problem which could be examined by applying RAR in the area where the problem occurred • know to recognise the key community members who can contribute to a successful research, the adoption of intervention measures and their implementation.
Abstract	The method called “Rapid Assessment and Response - RAR” is applicable if we wish to collect the relevant data concerning the area of public health which cannot be provided by the official statistics. This refers to the necessity for a rapid assessment of the real situation in the field with the aim to adopt well-reasoned intervention measures. RAR can be focused on a particular disease (e.g. HIV/AIDS), a particular population group (e.g. the Roma people), or on a particular risk behaviour (e.g. smoking).
Teaching methods	After the introductory lecture, students first carefully read the recommended literature, and then they all together identify a problem which could be examined by applying RAR in the area where it occurred. After that, students work in small groups; every group has the task to develop one research method for the problem they have chosen to examine. Then, the groups hold presentations which they will discuss as well as the possible interventions. The result of this team work is drafting up a RAR protocol for a chosen priority.
Specific recommendations for teachers	• work under teacher supervision: 30%, individual students’ work proportion: 70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: 1 computer on 5 students, access to the Internet and bibliographic data-bases; • training materials: recommended readings and RAR frame Protocol; • target audience: master degree students according to Bologna scheme.
Assessment of Students	Multiple choice questionnaire (MCQ); group presentations: RAR protocol.

RAPID ASSESSMENT AND RESPONSE - RAR

Enida Imamović, Dragana Nikšić

THEORETICAL BACKGROUND

Introduction

Rapid Assessment and Response or RAR represents a relatively new methodological approach to research whose main features are a rapid preparation, rapid implementation and adoption of intervention measures.

With the aim to explain the application of this method, the World Health Organisation has prepared a whole range of guides. For that purpose, a guide titled Rapid Assessment and Response to Injecting Drug Use was issued in May 2008 (1). Following the publication of this guide, others, such as Rapid Assessment and Response to sexual behaviour (2), and a guide aimed to asses risk behaviour among young persons that are particularly prone to risk (3-5).

This method enables us to assess the existing risk situations within a community - situations for which there are no relevant data, e.g. injecting drug use or HIV/AIDS (6-11). The data that have been collected are used for the adoption of well-reasoned intervention measures necessary for the reduction of negative consequences of the existing risk situations harmful for one's health.

RAR is a part of a significant preventive strategy which provides information, skills and interventions linked with the decrease of risk behaviour harmful to health among individuals or groups that have not been effectively provided with the official health care services.

Radical changes of social and cultural norms, among which we include the socio-political and socio-economic changes, as well the still-existing consequences of the war in South-Eastern Europe have contributed to the increase of risk behaviour among the population, especially among the young, about which, unfortunately, we did not have any relevant data. Therefore, in 2001/2002, as part of the project called "HIV/AIDS prevention among the young people of South Eastern European countries", UNICEF enabled the data collection in South Eastern Europe by applying the RAR methodology. Research was carried out in five countries simultaneously: in Albania, Bosnia and Herzegovina, Macedonia, Croatia and the Federal Republic of Yugoslavia with the aim to gain a rapid insight into the risk behaviour of young people and adopt well-reasoned intervention measures (12).

Method

It is possible to carry out a research applying Rapid Assessment and Response method (RAR) both in wider and narrower geographic areas, that is, on the local community level, in certain towns and countries, or even on a larger scale – on the regional level.

Advantages of the RAR methodology are reflected in a rapid assessment of problems, which enables a rapid intervention, an active participation of the local community, use of existing information and methods which are equally important and in the possibility to analyse data on several levels (13-16).

RAR can be focused on:

- a disease or state (HIV/AIDS, sexually transmitted diseases – STD, TB, pregnancy, etc.),
- a specific population group (e.g. a vulnerable group of young people aged 10-24, refugees, the Roma people, prisoners, victims of violence, etc.),
- places where the problem occurs (institutions designated to provide accommodation and upbringing of minors, prisons or a part of the local community at risk),
- risk behaviour (sexual habits, consummation of alcohol, tobacco smoking, drugs, etc).

RAR encompasses the use of several data collection methods. Through research, new data that cannot be attained in another way are collected (primary data).
All the existing data (secondary data) attained from official institutions, governmental and non-governmental organisations, international organisations and data that were collected from another relevant research are used as well. In order to collect primary data, the observation, mapping, poll, focus group and interview methods are used (1-5).

Advisory body

Through the formation of advisory bodies, all the representatives of a community can, in one way or the other, contribute to the success of the research itself or to the adoption of intervention measures and their later implementation are engaged in the research from the very beginning.

Representatives of all sectors can contribute to the rapid assessment and alleviate the direct implementation of project measures thereby contributing to the sustainability of ideas and practices which the project advocates. Those are the representatives of the government and the non-government sectors, media, health professionals, teachers, parents, religious communities similar (17).

Research team

It is very important to engage all the administrative levels of the public sector in the organisational structure of the research, as well as the representatives of the non-governmental sector, frequent members of the examined groups that are frequently "closed" when it comes to providing information for the official sector and therefore for interventions as well. In the first place, the core research team has the task to define the general and specific research goals, key assessment areas, working methods and the time frame.

Rapid assessment methods

By applying rapid assessment we collect quantitative and qualitative data:

- quantitative data provide us with an insight into the distribution of the phenomenon and grant statistical confirmation. They are vital for the planning and implementation of the intervention,
- by collecting qualitative data, we gain a deeper insight into the complexity of the examined phenomenon (e.g. risk behaviour) and into their influence on the examined group (young people aged 15-24). These data are necessary for the specification of goals and intervention design (1-5,16,17).

Collection of existing or secondary data

These data do not have to be directly related to research goals but they are related to the research framework. The sources of secondary data are different and those are: demographic, vital and health statistical data; all of them are relevant, published or unpublished data (project reports of governmental and non-governmental organisations, scientific periodicals, books, and some other publications) as well as the data derived from different media (media files). It is important not to waste precious time on collecting secondary data that do not have a significant value in the end.

Collection of primary data

We collect primary data from all the sources available, such as:

- "ordinary" population,
- members of groups falling within the scope of the research,
- members of risk groups and core risk groups including the former members,
- key-informants (street vendors, tenants, parents, social workers, teachers, policemen, drug dealers, procurers, etc.).

Observations

Applying observation as a research method, by simply observing, listening and recording we gain insight into first-hand information which usually remains "hidden" but is directly related to behaviour, meaning, relations and contexts of risk behaviour.

Mapping

The mapping method is applied simultaneously with the observation method. It is used to illustrate the aspects of the environment in which the persons live, that is, where they practice certain behaviour. The application of the mapping method assists us to identify the areas where the risk behaviour is "more prominent", which also provides the opportunity for targeted interventions to take place. However, one should bear in mind that the research is completely anonymous and that by applying the mapping method we do not wish to compromise the confidentiality of the data collected.

Polls

As a research instrument for collecting quantitative data, we use polls consisting of obligatory questions which, for the sake of standardisation, have to be identical for all the countries in which RAR is implemented. Obligatory questions are posed to all the respondents to which the research refers (e.g. young people aged 14-24), that means, to all the people who participate in the poll, the interview or the focus group.

Apart from the obligatory questions, in order to provide a better insight into the specificity of a situation on their own territory, every country has to define a set of additional questions.

Prior to field work, there is an obligatory pre-testing, after which a definite version of the questionnaire is created.

Focus groups

The focus group method is used in research in order to gain a deeper insight into understanding the group and the social norms related to the behaviour of respondents and to get replies to questions from different respondent groups which cannot be asked in a poll.

When working with focus groups, we use questions that have been especially prepared for respective groups. This method is used when working with respondents from different groups and organisations that can contribute to a better understanding of problems and answer the questions which cannot be answered in another way (e.g. students, parents, teachers, pedagogues, health services, drug takers, nongovernmental organizations representatives, the governmental sector, etc).

Interview

The interview method attempts to get replies to questions which are too personal to answer in a poll or a focus group. Interviews are used to get information from individuals belonging to risk population groups, health workers, teachers, politicians and other relevant representatives of the community.

Data collection techniques

Since we are talking about a research that, apart from the quantity, aspires to collect the qualitative data as well, the "snow ball" strategy is used, that is, data are collected until the saturation moment has been reached. Once we start getting the same data by applying a certain method, or working with a certain group of respondents, which can be determined on the basis of constant data input and analysis, the collection of data by applying that method is to be interrupted. The methods that follow pay attention to questions that had previously not been answered clearly enough.

Ethical issues

Throughout the research, great emphasis is laid on ethical issues. Regardless of the method applied for data collection, the goal of the research is repeatedly emphasised as well as the anonymity and confidentiality of collected data. Respondents are asked for permission for their replies to be recorded, the emphasis being on the principle of confidentiality.

With the aim to collect data on the spot, all the working team members have accreditation cards and a "letter of presentation" which serves to identify them and to explain the general goals of the RAR research. The research being completely anonymous, the confidentiality of data is by no means an issue, and even the detailed maps and observations are used for research exclusively.

Assessment areas

Data are being collected and analysed in terms of the assessment areas under observation.

- context (the entire environment),
- risk and protective behaviour,
- health and social consequences,
- existing interventions.

Within the context assessment, we collect data that influence the existing and situations that bear potential risk, negative health consequences and opportunities for the development/adoption of adequate intervention measures.

The assessment of risk and protective behaviour is carried out on the grounds of data on the scope and nature of risks on the one hand, and types of protection (prevention) on the other, which increase, or decrease the risk that certain phenomena bear among the population.

Data on negative health and social consequences of risk behaviour are collected within the framework of assessment of social and health consequences, and within the framework of intervention assessment. When it comes to the examined phenomena and state the necessity for future interventions the existing interventions data are collected and assessed.

Data analysis

The input and processing of quantitative data is an ongoing process in a chosen programme. The participation in the RAR being completely voluntary, every participant of this research, i.e. respondent can at any time refuse to participate or refuse to reply to any of the questions posed. Questions that were not answered remain "empty" throughout the data input process. Data are analysed in a descriptive way, which implies the use of frequencies, averages and percentages.

The results attained by qualitative methods are first entered into journals of the team members on the field, and later into especially designed activity tables; every individual method-case that has been applied is entered into a separate activity table.

During the meeting of all the local team members, the key responses from all the activity tables are entered into summary tables or mega-tables for all the four areas under observation: context, risk and protective behaviour, health and social consequences and existing interventions.

Triangulation

Data received on the grounds of the same question posed by using a different method are triangulated, which ensures a high data validity.

Definition of triangulation: comparing or crosschecking data that were collected using different methods (e.g. focus group, existing information, polls) and from different data sources (e.g. injecting drug users, health service providers, politicians).

Data that have not been confirmed by means of triangulation cannot be taken into account in the final analysis. Likewise, questionnaires filled in by persons whose age does not correspond to recommended age groups are not taken into account.

Induction

A majority of research questions is defined prior to RAR, while others can be posed during the data collection process. Induction is an approach according to which a hypothesis is stated and then data are collected – data that will confirm, deny or modify that hypothesis. This can be repeated several times with a rapid assessment.

Rapid assessment results and their dissemination

Results are presented in a form of report whereby their rapid dissemination to all interested participants such as politicians, media, donors, health and non-health professionals, the NG

sector, groups of respondents, etc. is of greatest importance. The manners in which the results of the rapid assessment are disseminated are versatile – through media, meetings with decision-makers, health service providers, non-governmental organisations, donors, etc.

This phase is particularly important for the adoption and implementation of intervention measures. Suggestions for intervention have to contain several alternatives; they have to be acceptable to decision-makers, donors and field executors (16-18).

Intervention

Community's response to a phenomenon which is being examined has to be rapid first of all; this implies a rapid adoption and implementation of intervention measures.

The level of intervention can be individual, which means that we can intervene on the level of a certain group, or it can be structural, in terms of a change of legislation. Individual intervention is less efficient because an individual is perceived as a "victim". Increasing the knowledge of individuals about the problem does not have a significant effect because other requirements for the over-all rehabilitation (social, professional, economic requirements) have not been met.

The group and the structural approaches to intervention have an all-encompassing effect because they meet the requirements for sustainable changes in a community (16-19).

Conclusions

Research based on the RAR methodology can be carried out in a wider or narrower geographic area, that is, on the local level (towns, countries) or on the sub-regional and regional levels.

The RAR method is applicable if we wish to collect the relevant data concerning the area of public health which cannot be provided by the official statistics. This refers to the necessity for a rapid assessment of the real situation on the field with the aim to adopt well-reasoned intervention measures applying to a wide spectrum of public and health issues in certain population groups, such as HIV/AIDS and STDs; consummation of alcohol, tobacco and drugs; sexual and reproductive health, violence, injuries and accidents; nutrition and mental health, etc.

The main features of this method are a rapid preparation of the research and a rapid implementation on the field, which enables a rapid adoption of intervention measures (Box 1). The engagement of relevant representatives of the community in the research from the start, contributes to the overall success of the research, the adoption of intervention measures and their later implementation on the field.

Box 1. Rapid assessment and response (RAR) features.

Rapid Assessment and Response – RAR

- *a rapid preparation and implementation,*
- *collect new data only when the existing data are insufficient and unreliable,*
- *crosschecking data derived from different sources (triangulation),*
- *complexity of approach; social, cultural, religious, political and economic context,*
- *engagement and support by the decision-maker,*
- *engagement of the community in all phases of research,*
- *a rapid and all-encompassing intervention.*

EXERCISE

Task 1

Read the theory part of the module carefully and improve knowledge about the RAR methodology and its application by reading the recommended literature and browsing the Internet. Criticise and discuss the application of this method with colleagues.

Box 2. Guide for preparation of the rapid assessment and response (RAR) protocol.

RAR PROTOCOL	
Title	*e.g. RAR to risk behaviour among young people*
Problem formulation	*A situational analysis which points at the problem and its gravity, and the necessity for a rapid assessment of the situation.*
Research topic	*Define the target population (e.g. young people aged 15-24), the disease (e.g. HIV/AIDS) or state (e.g. pregnancy), region (municipality, state)*
The overall goal	*Rapid assessment of risk behaviour among the young and adoption of well-reasoned intervention measures with the aim to improve their health*
Specific goals	*Decreasing risk behaviour among the young (smoking, consummation of alcohol, drugs…)*
Methodology	*Secondary data (specify)* *Primary data (specify the methods)* *Timeframe (set)* *Key participants (specify)*
Data base, data input and data processing	*Define the data base (which), data input (when and who), data processing (when and who)*
Data analysis	*Simulate the analysis in terms of:* • *Context: socio-economic and political environment* • *Risk and protective behaviour* • *Health and social consequences* • *Existing interventions*
Results	*Specify the anticipated results*
Dissemination	*Simulate as to whom (the key participants) and how (methods-media, meeting, round table…) present the results*
Resources	*Simulate the necessary human and material resources*
RAR team	*Advisory body (specify the members according to their occupation, profession and institution they work for)* *Research team members (their number in terms of their profession and role: team leader and deputy, members of the working group, poll-takers – field work, etc.*
Timeframe	*Set in terms of working phases (organigramme)*
Budget	*Simulate the framework budget*

Task 2

Find out if this method had already been applied in the country in which the students live or in the neighbouring countries and in which area had the research been carried out, and then read the final reports.

Task 3

Make a suggestion – choose a priority for the national RAR research (taking into account that it has to be a problem for which relevant, updated data cannot be attained) which calls for a rapid assessment and intervention.

Task 4

Final paper (group work): Create a RAR research protocol (Box 2). for a chosen priority problem

REFERENCES

1. World Health Organization, Regional Office for Europe. The RAR guide on injecting drug use: Social change and mental health. Copenhagen: World Health Organization, Regional Office for Europe; 1998. Available from URL: htpp:// www.documents hiv aids IDU rapid ass resp guide pdf. Accessed: March 20, 2009.
2. UNAIDS and World Health Organization. The RAR guide on substance use and sexual risky behaviour. Geneva: World Health Organization, UNAIDS; 1998.
3. World Health Organization /SAB, UNICEF, UNAIDS AND UNDCP. The Rapid Assessment and Response guide on psychoactive substance use and especially vulnerable young people (EVYP-RAR). Draft for field testing. July 1998. Geneva: World Health Organization /SAB, UNICEF, UNAIDS AND UNDCP, 1998.
4. World Health Organization, Regional Office for Europe. SEX-RAR Guide: The rapid assessment and response guide on psychoactive substance use and sexual risk behaviour. Copenhagen: World Health Organization; 2002. Available at URL: http://www.who.int/substance_abuse/publications/epidemiology/en/index.html. Accessed: March 20, 2009.
5. World Health Organization, Regional office for Europe. Basic principles for effective prevention of HIV infection among injecting drug users. Copenhagen: World Health Organization, Regional office for Europe; 1998.
6. UNICEF. A decade of transition. MONEE Project, SEE/CIS/Baltic.New York: UNICEF; 2000.
7. UNAIDS. Global strategy Framework on young people and HIV/AID. New York: UNAIDS; 2000.
8. UNICEF/Area Balkans. HIV/AIDS in South Eastern Europe. Sarajevo: UNICEF/Area Balkans; 2001.
9. Needle RH, Trotter RT II et all. Rapid Assessment of the HIV/AIDS crisis in racial and ethnic minority communities: An approach for timely community interventions. Am J Pub Health. 2003;93:6970-79.
10. Rhodes T, Stimson GV, Fitch C, Ball A, Renton A. Rapid assessment, injecting drug use and public health. Lancet. 1999;354:65-8.

11. UNIDCP. Rapid Assessment of illicit drug abuse in the Republic of Croatia.. Zagreb: National Commission for Drug Abuse Suppression of the Government of the Republic of Croatia; 1999.
12. Wong E**.** Rapid Assessment and Response on HIV/AIDS among Especialy Vulnerable Young People in South Eastern Europe. Geneva: UNICEF and Canadian International Development Agency; 2002.
13. Trotter RT II, Needle RH, Goosby E, Bates C, Singer M. A methodological model for rapid assessment, response, and evaluation: the RARE program in public health. Field Methods. 2001;13:137-159.
14. McNall M, Foster-Fishman PG**.** Methods of rapid evaluation, assessment, and appraisal. Am J Evaluat. 2007;28:151-68.
15. Brown D R, Hernandez A, Saint-Jean G, et all. A participatory action research pilot study of urban health disparities using rapid assessment response and evaluation. Am J Public Health. 2008;98:28 38.
16. Pope C. Qualitative Research: Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ. 1995;311:42-45.
17. Needle RH, Trotter RT II, Goosby E, Bates C, Von Zinkernagel D. Methodologically sound rapid assessment and response: providing timely data for policy development on drug use interventions. Int J Drug Policy. 2000;11:19-23.
18. Stimson GV, Fitch C., Jarlais DD, Poznyak V, Perlis T, Oppenheimer E, Rhodes T. The Who Phase II Drug injection collaborative rapid assessment and response studies of Injection Drug Use: Knowledge gain, capacity building, and intervention development in a multisite study. Am J Public Health. 2006;96:288-95.
19. Stanton B, Kim N, Galbraith J, Parrot M. Design Issues Addressed in Published Evaluations of Adolescents HIV-Risk Reduction Interventions. J Adolesc Health 1996. 18:387-396.

RECOMMENDED READINGS

1. World Health Organization, Regional Office for Europe. The RAR guide on injecting drug use: Social change and mental health. copenhagen: World Health Organization, Regional Office for Europe; 1998. Available from URL: htpp:// www.documents hiv aids IDU rapid ass resp guide pdf. Accessed: March 20, 2009.
2. World Health Organization, Regional Office for Europe. SEX-RAR Guide: The rapid assessment and response guide on psychoactive substance use and sexual risk behaviour. Copenhagen: World Health Organization; 2002. Available at URL: http://www.who.int/substance_abuse/publications/epidemiology/en/index.html. Accessed: March 20, 2009.
3. World Health Organization /SAB, UNICEF, UNAIDS AND UNDCP. The Rapid Assessment and Response guide on psychoactive substance use and especially vulnerable young people (EVYP-RAR). Draft for field testing. July 1998. Geneva: World Health Organization /SAB, UNICEF, UNAIDS AND UNDCP, 1998.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**MEASUREMENT, MONITORING AND EVALUATION OF PUBLIC HEALTH SYSTEMS: ASSESSMENT OF ESSENTIAL PUBLIC HEALTH FUNCTIONS**
Module: 4.4.3	**ECTS (suggested): 0.50**
Author(s), degrees, institution(s)	**Fimka Tozija**, MD, PhD, Associate Professor National Institute for Health Protection, Skopje, Republic of Macedonia **Dragan Gjorgjev**, MD, PhD, Professor National Institute for Health Protection, Skopje, Republic of Macedonia **Dance Gudeva Nikovska**, MD, MPH, Teaching Assistant Ministry of Health, Republic of Macedonia
Address for correspondence	**Dragan Gjorgjev** National Institute for Health Protection 50 Divizija No 6 1000 Skopje Republic of Macedonia
Keywords	Public health, essential public health functions, measurement, instrument
Learning objectives	After completing the modules students should: • become familiar with governance of public health services • become familiar with the 10 essential functions of the public health • become familiar with the instrument • develop proposals for monitoring and evaluation systems in their countries
Abstract	Although public health (PH) theoretically differ from medical services; and have to be designed to produce so called "public goods", in reality, the public becomes aware of the need for these services when a problem develops and the health system pay high prices in illness, disability/death or large economic costs. Ministries of Health have the steering role to provide health for their respective population; in order to accomplish this mandate, health authorities need to continuously define most striking problems of its population, define which groups are most affected, where are the problems most frequent and whether those problems generate inequities. CDC PAHO methodology and instrument evaluate functioning of public health sector across 11 essential public health functions through set of 49 indicators, measures and sub-measures. Results provide recommendations for improvement of the essential PH functions, reduction of barriers and costs associated with them, identification of critical issues and dilemmas, as well as development of strategies to address them. Repeated measurements over time facilitate consistency quantification between measurement and identification of the "grey zones" in PH system, thus facilitating design of targeted interventions for institutional capacity strengthening.
Teaching methods	Qualitative method, SWOT analysis, instrument application, individual work, group work
Specific recommendations for teachers	To be familiar with the instrument, the health system and the needs of the group.
Assessment of students	Formative assessment - during the course, individual and group work Summative assessment - final SWOT analysis applying the instrument

MEASUREMENT, MONITORING AND EVALUATION OF PUBLIC HEALTH SYSTEMS: ASSESSMENT OF ESSENTIAL PUBLIC HEALTH FUNCTIONS

Fimka Tozija, Dragan Gjorgjev, Dance Gudeva Nikovska

THEORETICAL BACKGROUND

About public health sector assessment

Public health systems in theory are different from medical services, their principal goal being to reduce exposure to disease (e.g. food safety, health education and promotion) and are ultimately imperceptible to the public. Although these services have to be designed to produce so called "public goods" (1) in reality, the public becomes aware of the need for these services when a problem develops (e.g. epidemics or natural disaster). When public health systems fail to provide appropriate response, people and the health system pay high prices in illness, disability or death, incurring large economic costs.

Good governance in health sector

Lately, the terms "governance" and "good governance" are interchangeably exploited in literature analyzing development processes of different countries. According to United Nations (UN) Economic and Social Commission for Asia and Pacific bad governance is to a large extent considered as a *basic reason for all evil in a country* (2), while big donor agencies and international financial institutions fundamentally link their assistance and credits based on reforms undertaken by states to provide for good governance.

The concept of "governance" is as old as human civilization. The simplest definition describes it as "a process of decision making and a process with which those decisions are (or are not) implemented" (2). As a process of decision making, analysis of good governance focuses on formal and informal actors involved in decision making processes, implementation of the decisions, as well as formal and informal actors responsible for their implementation.

UN Institute defines eight principal characteristics of good governance (2). Good governance is:

- participatory,
- consensus oriented,
- respects the principles of accountability,
- respects the principles of transparency,
- respects the principles of efficacy,
- respects the principles of efficiency,
- respects the principles of equity, and
- respects the law and regulations

Aims (stewardship role of health authorities)

Country health authorities, i.e. Ministries of Health (hereinafter referred as MoH) have the steering role to provide health for their respective population; in order to accomplish this mandate, health authorities need to continuously define most striking problems of its population, define which groups are most affected, where are the problems most frequent and whether those problems generate inequities. In addition, MoH has to assess whether health status of the population changes over time and how is this affected by health programs and interventions.

Current trends of decentralization or privatization of the health sector in developing countries compel national health authorities with complex array of responsibilities which can be summarized as (3):

1. definition of criteria for resources allocation that have to be channeled to decentralized public health institutions, with definition of criteria as to measure the needs, performances and impact of interventions;
2. harmonization of the action plans and management models of decentralized agencies, directly responsible for health services provision in their respective areas;
3. definition of the content of essential public health functions that remain central responsibility of the state and, based on complementary criteria, allocation of competencies and resources to different levels of public administration;
4. technical cooperation with decentralized service providers, in order to guarantee steady transfer of authority from central to local levels, as well as development of appropriate institutional capacity for uninterrupted and complete execution of functions;
5. definition of mechanisms for redistribution of existing and capital costs as to compensate for likely inequities that the process of decentralization might generate.

Methodology of the assessment process

The methodology of EPHF (from Essential Public Health Functions) analysis is explorative in nature. According to Kaufmann and Kray, cited by Gupta (1), the approach is analogous to a "polls of experts", emphasizing that "for many dimensions of governance, subjective data is the only data that is potentially informative". Survey performed with this instrument is qualitative and conclusions have to be cautiously interpreted due to limitations in its design and methodology.

Essentially, the methodology is designed to obtain self-evaluation of the public health system by people who know the system well from within (4), rather then providing for "objective" information typically asked in surveys[28].

[28] For example, in order to assess the aspects of central-local government agencies collaboration, it does not ask an objective question such as how many meetings were organized last year, but rather asks questions such as whether local agencies were provided public health data, followed by more specific questions to appraise whether this data were

Results obtained with this kind of survey should facilitate provision of recommendations for improvement of the essential public health functions, reduction of barriers and costs associated with them, identification of critical issues and dilemmas, as well as development of strategies to address them.

Application of this methodology in USA and South American countries have shown that despite the fact that scoring criteria are not completely defined, the instrument is an excellent tool for identification of strengths and weaknesses of public health systems. Repeated measurements over time facilitate consistency quantification between measurement and identification of the "grey zones" in public health system, thus assisting in design of targeted interventions for institutional capacity strengthening (5).

Measurement instrument

Basic instrument for this kind of assessment is standardized questionnaire designed by US Center for Disease Control and Pan American Health Organization[29]. The instrument is currently not available for download from World Bank Institute website.

The questionnaire is divided into 11 segments labeled *Essential Public Health Functions* (EPHF - Table 1). Each function is assessed through set of predefined indicators (total of 49 indicators), specific for each segment. Every segment begins with description of services covered by the respective function; segments are divided into sub-segments with questions specified for the defined indicators.

The instrument is simple, user friendly Excel file[30]. Separate sheet is assigned to each individual EPHF, in addition to the summary sheet for all 11 functions. Graphical presentation of sub-segments is automatically displayed, as scores for each question are entered into designated cells.

Individual answers are recorded as YES (response is assigned value of 1) or NO (response is assigned value of 0). Upon completion of each segment, summary value is obtained for individual indicators (result between 0 and 1, depending on proportion of positive answers for that indicator). Mean value is calculated from the results obtained for each indicator by individual respondents, in order to assess the respective EPHF (Example of the excel sheet for assessment of one EPHF is presented in Figure 1).

The instrument can be modified with an option to assign 0.5 for answers that could not be answered simply YES or NO; examples include answers such as "*partial implementation, existing, but not enforced law/regulation, in phase of preparation* etc.".

made user-friendly, any attempts to evaluate utilization of data, modifications to expand their potential use etc.

[29] In USA, the instrument has been used by several countries for identification of most striking problems in public health services, and consequently used for analysis of the health systems in Latin American countries. Several publications named *Health in Americas* are published, presenting analyses performed in 42 countries.

[30] The instrument is available on-line only if one enrolls in the World Bank Institute e-learning program. (we can only provide Macedonian version of the instrument.

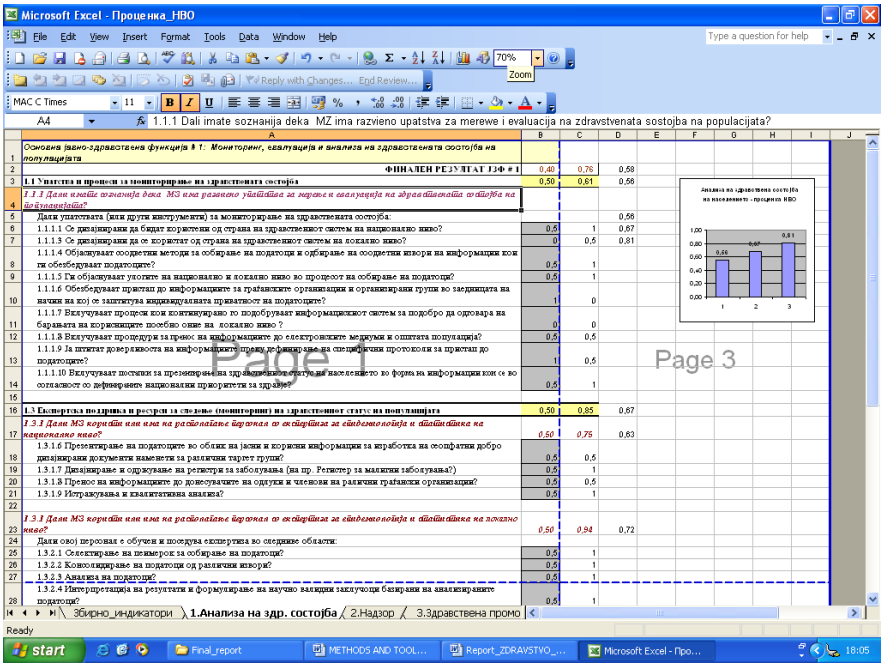


Figure 1. Example of the Excel file for assessment of one EPHF.

Table 1. Eleven Essential Public Health Functions and specific indicators for each area.

Essential Public Health Function	Indicator
Essential PHF#1: Monitoring, evaluation and analysis of health status	1.1. Guidelines and processes for monitoring health status 1.2. Evaluation of the quality of information 1.3. Expert support and resources for monitoring of health status 1.4. Technical support for monitoring and evaluation of the health status 1.5. Technical assistance at the sub-national/local level in monitoring, evaluation and analysis of the health status
Essential PHF#2: Public health surveillance, research and control of risks and threats to public health	2.1. Surveillance systems to identify threats to public health 2.2. Capacities and expertise for public health surveillance 2.3. Capacity of public health laboratories 2.4. Preparedness for an effective response to control threats to the environment and public health 2.5. Technical assistance and support to sub-national/local level in surveillance, research and control of risks and threats to public health

Table 1. Cont.

Essential Public Health Function	Indicator
Essential PHF#3: Health promotion	3.1. Support for health promotion activities, development of standards and interventions for promotion of healthy lifestyles and environment 3.2. Building sectorial and extrasectorial partnerships for health promotion 3.3. National planning and coordination of strategies for information, education and social communication for health promotion 3.4. Reorientation of health services towards health promotion 3.5. Technical assistance and support to sub-national/local level for strengthening health promotion activities
Essential PHF #4: Social participation in health	4.1. Civil society empowerment for decision making in public health 4.2. Strengthening social participation in public health sector 4.3. Technical assistance and support to sub-national/local level for strengthening social participation in public health sector
Essential PHF# 5: Development of policies and institutional capacity for planning and management in public health	5.1. Definition of national and local health goals 5.2. Development, monitoring and evaluation of public health policies 5.3. Development of institutional capacity for management of public health systems 5.4. Management of international cooperation in public health 5.5. technical assistance and support to sub-national/local level in development of policies, planning and management in public health
Essential PHF # 6: Strengthening of institutional capacity for regulation and enforcement in public health	6.1. Periodic monitoring, evaluation and revision of regulatory framework 6.2. Enforcement of law and regulations 6.3. Knowledge, skills and mechanisms for revision, improvement and enforcement of laws 6.4. Assistance and technical support to sub-national/local level and enforcement of laws and regulations
Essential PHF#7: Evaluation and promotion of equitable access to necessary health services	7.1. Monitoring and evaluation of access to necessary health services 7.2. Knowledge, skills and mechanisms for improvement in access to necessary health services for the entire population 7.3. Advocacy and activities for improvement of access to necessary health services 7.4. Assistance and technical support to sub-national/local level for promotion of equitable access to necessary health services
Essential PHF #8: Human resources development and training in public health	8.1. Profile of public health professionals 8.2. Quality improvement of the workforce 8.3. Continuous education and undergraduate public health studies 8.4. Improvement of the workforce for provision of culturally appropriate services 8.5. Technical assistance and support to sub-national/local level in human resources development

Table 1. Cont.

Essential Public Health Function	Indicator
Essential PHF#9: Ensuring the quality of personal and population-based health services	9.1. Definition of standards and evaluation of the quality of personal and population-based health services 9.2. Improvement of patient satisfaction with health services 9.3. Systems for technological management and assessment of health system technology that support decision making in public health sector 9.4. Technical assistance and support to sub-national/local level for improvement of the quality of health services
Essential PHF#10: Research in public health	10.1. Development of agenda for public health research 10.2. Development of capacities for institutional research 10.3. Technical assistance and support to sub-national/local level for public health research
Essential PHF#11: Reducing the impact of emergencies and disasters on health	11.1 Emergency preparedness and disaster management in health 11.2. Development of standards and guidelines that support emergency preparedness and disaster management in health 11.3. Coordination and partnership with other agencies and/or institutions in emergencies and disasters 11.4. Technical assistance and support to the local level to reduce the impact of emergencies and disasters on health

Description of the essential public health functions in the instrument

Each EPHF in the instrument has its own functional identity and covers specific processes that generate explicit results and outcomes. At the same time, they share matching resources and complement each other.

The first EPHF (EPHF#1 - *Monitoring, evaluation and health status analysis),* along with EPHF#5 - *Development of policies and institutional capacity for planning and management in public health,* EPHF#8 - *Human resources development and public health training* and EPHF#10 - *Research in public health* are all examples of systematic functions that support each other or are complementary, covering areas with shared responsibilities and are common for all public health activities.

The third essential public health function, EPHF#3 - *Health promotion* and EPHF#4 - *Social participation in health,* demand collaboration with other public health functions and should have the ability to change operational conditions in the overall public health system and public health through involvement of direct users of health services. EPHF#6 *Strengthening institutional capacities for regulation and enforcement in public health* is instrumental and fundamental, thus providing for appropriate function in the framework of the whole public health system. Similar to aforementioned descriptions, the second EPHF#2 - *Public health surveillance, research and control of risks and threats to public health,* EPHF#7 - *Evaluation and promotion of equitable access to necessary health services* and EPHF#9 - *Ensuring*

the quality of personal and population-based health services are directly linked with the ultimate goal of public health system. Finally, EPHF#11 - *Reducing the impact of emergencies and disasters on health* is an example of the most specific function with defined areas of action and requires support and coordination of all previously described functions.

Application of the instrument

Assessment of public health system functioning in a given country could be performed in two different ways:

1. by a representative sample of individual respondents from relevant agencies. In this case, the instrument is administered individually and each score represents individual opinion of the person to each of the 11 EPHF. Mean value is then calculated for the respective indicator, as explained above, or
2. at a meeting or workshop, convening professionals from relevant agencies (health personnel, academicians, civil sector representatives, other relevant specialists). In such case, there is a need to engage external facilitator who will help build and record consensus responses in the working groups. Usually, the facilitator should read loud the definition, standards, measures and sub-measures of each EPHF to the group assigned to discuss the EPHF. At the end of the workshop, results of the measurement are usually shared with participants, providing examples of the types of analysis that can be done when measuring EPHF, oriented toward the identification of intervention areas in order to improve the institutional capacity of health authorities in exercising EPHF that relate to it. Number of members of each group is arbitrary; however, experience shows that it ranges from 15-30 participants representatives of relevant Governmental and Non-Governmental institutions.

Scoring/results of the measurement

The score for each indicator which represents part of the measurement for each EPHF is based on the scores obtained from so-called *First Tier Statements.* Individual questions for each EPHF (measures and sub-measures pertaining to that statement), could be answered YES (response is assigned value of 1), NO (response assigned with 0) or partial response (0.5[31]).

The calculation of the final score of every first tier statement is essentially the average of the "Yes" and "Partial" responses to the measures and sub-measures. The score assigned to the indicator is the average of results obtained for each of the measures within the indicator and the average of the results of all the indicators in a function determines the score for the performance of that particular EPHF.

[31] Examples include replies e.g. partially, in phase of preparation, existing, but not enforced laws etc. It is recommended to present comments on such responses, as to provide rationale and additional explanation. Partial responses, including comments provided by different groups of respondents could be subject to separate qualitative analysis, or mixed-methods research study design.

Interpretation of results

Conventional scale is proposed for overall interpretation of obtained results of EPHF analysis. It is presented in Box 1.

Box 1. Conventional scale for overall interpretation of results of essential public health functions analysis.

76-100% (0.76-1.0)	*Quartile of **optimal** performance*
51-75% (0.51-0.75)	*Quartile of **above average** performance*
26-50% (0.26 - 0.5)	*Quartile of **below average** performance*
0-25% (0.0 - 0.25)	*Quartile of **minimum** performance*

CASE STUDY: DEVELOPMENT AND APPLICATION OF AN INSTRUMENT FOR MEASUREMENT, MONITORING AND EVALUATION OF 10 ESSENTIAL PUBLIC HEALTH FUNCTIONS IN REPUBLIC OF MACEDONIA

Introduction

The main goal of the analysis of good governance in public health sector in Republic of Macedonia was to develop and apply an instrument for assessment of 10 EPHF that are responsibility of health authorities, i.e., Ministry of Health and its institutions, in order to identify "grey zones" and provide recommendations and directions for improvement.

Methods

The analysis has been performed in the period Apr-Nov, 2007 using the standardized questionnaire developed by US CDC and PAHO (3,4) translated and adapted to be used in Macedonian context. The instrument incorporates 11 areas of EPHF and outlines 49 indicators that are evaluated through answers on questions for measures and sub measures in the defined area. It was agreed to exclude EPHF#11 *Reducing the impact of disasters and emergencies to health*, in order to provide more detailed analysis of the other 10 EPHF.

Our survey was performed throughout three stages:

1. In the first phase, the instrument was translated into Macedonian language, pre-tested on representative sample of government officials and necessary adaptations made to be used in Macedonian context.
2. The second phase included interviews with 3 groups of respondents – central government officials, representatives of government institutions at local level and representatives of non-governmental sector.
3. In the third phase, results were summarized and final report prepared, including recommendations for remedial activities.

Results

Comparative analysis of all 10 EPHF, applying conventional interpretation of results, shows that none of the EPHF prove optimal results, a score shared by all three groups of respondents. Highest score is recorded for EPHF#1 *Monitoring, evaluation and health situation analysis,* while lowest score is documented for EPHF#3 *Health promotion,* EPHF#*8 Human Resource Development and Training in Public Health* and EPHF#10 *Research in Public Health.*

Assessment of the specific indicators for each EPHF (Box 2) has identified existence of "grey zones", which were used to perform subsequent analysis of strengths, weaknesses, opportunities, and threats (SWOT analysis) of governance in public health sector. In Figures 2-4 and Table 2 examples of the results of EPHF analysis are presented.

Box 2. Description of specific indicators for assessment of essential public health functions #2**.**

Indicator #1	*Guidelines and processes for monitoring of health status of population*
Indicator #2	*Evaluation of the quality of information*
Indicator #3	*Expert support and resources for monitoring of the health status of population*
Indicator #4	*Technical support for monitoring and evaluation of the health status of population*
Indicator #5	*Technical support at local levels of public health systems in monitoring, evaluation and analysis of the health status of population*

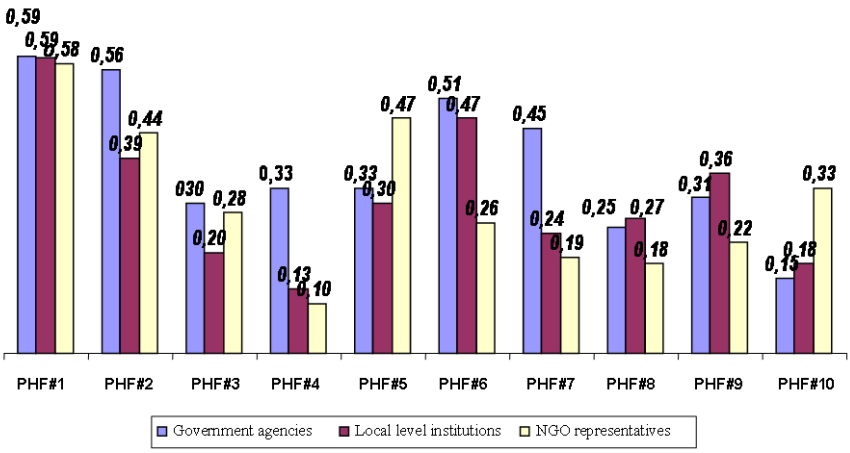


Figure 2. Example assessment of 10 essential public health functions by 3 groups of respondents. LEGEND: PHF = public health function; NGO = non-governmental organizations.

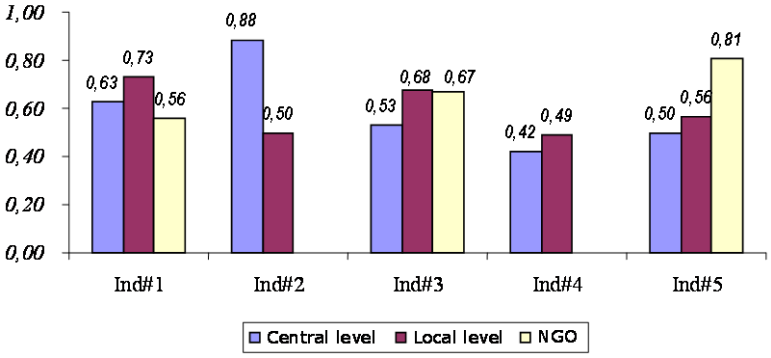


Figure 3. Results for assessment of 5 indicators for essential public health function #1. LEGEND: ind = indicator.

1. Strengths point to existence of guidelines and processes for monitoring of the population health status (Figure 3); existence of epidemiology surveillance network and control over communicable diseases; activities for health promotion; development of norms and interventions for healthy behavior and environment; initiatives foe strengthening civil society involvement into decision making processes; partially defined health priorities at national and local level; random monitoring, evaluation and modification of regulatory framework; advocacy and initiatives for improvement of access to essential health services; respectable portrayal of public health professionals; activities for improvement of patient satisfaction and activities for development of institutional capacities for public health research.
2. Identified weaknesses refer to weak technical support for monitoring and evaluation of the population health status (Figure 3); insufficient capacity and expertise for public health surveillance; absence of national planning and coordination of strategies for health promotion; lack of technical assistance and support to local level institutions for strengthening of civil sector involvement; lack of institutional capacity development for management of public health institutions; insufficient technical assistance and support for institutions at local level in development and enforcement of lows and regulations; absence of knowledge, skills and mechanisms for improvement of the access to essential health services; unsatisfactory capacity building for public health professionals to provide culturally sensitive services; lack of defined standards and evaluation for improvement of the quality of health services and absence of agenda for public health research.
3. Opportunities are foreseen into adopted Strategy for development of integrated health information system; accreditation of the system for laboratories quality assurance; existing strategies (alcohol, illicit drugs, tobacco, food and nutrition, adolescent health, convention for narcotic drugs); programs which

involve citizen associations; opening of Center for Public Health and organization of training for managers of health institutions; activities for EU integration and involvement into local self-governance; constitutionally guaranteed rights for health protection of each citizen and widely accessible network of health institutions; Health Strategy 2008-2015; strengthening of human capacities in local communities; existing Strategic plan at the Ministry of Health and establishment of Department for Injuries Control within National Public Health Institute.

4. Possible threats bring in dysfunctional health promotion bodies at central level; lack of coordination mechanisms with the civil society (Figure 4); delay of the decentralization and EU integration processes, in addition to absence of continuity of health reforms; privatization process and risk imposed with for the access to high quality health services; inadequate human resources policy and high unemployment rate of both medical and non-medical personnel; decision making that is not evidence based.

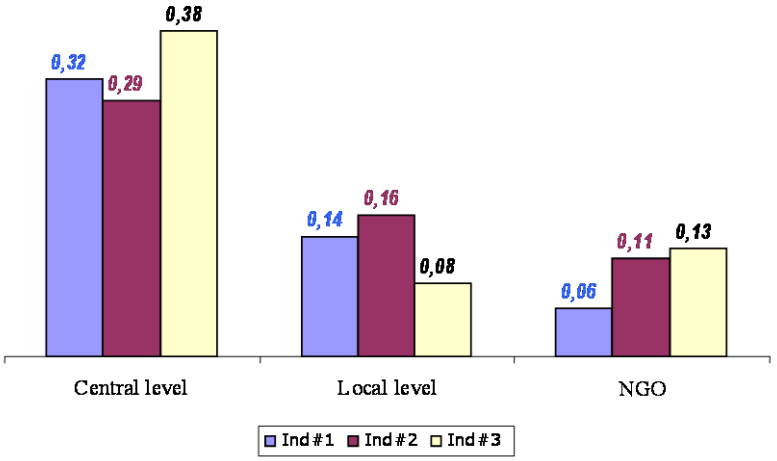


Figure 4. Results for assessment of 3 indicators for essential public health function #4. LEGEND: ind = indicator; NGO = non-governmental organizations.

Table 2. Ten Essential Public Health Functions indicators values in Republic of Macedonia.

Indicator	Value
Essential PHF#1:	
1.1. Guidelines and processes for monitoring health status	0.64
1.2. Evaluation of the quality of information	0.46
1.3. Expert support and resources for monitoring of health status	0.63
1.4. Technical support for monitoring and evaluation of the health status	0.30
1.5. Technical assistance at the sub-national/local level in monitoring, evaluation and analysis of the health status	0.62

Table 2. Cont.

Indicator	Value
Essential PHF#2:	
2.1. Surveillance systems to identify threats to public health	0.62
2.2. Capacities and expertise for public health surveillance	0.29
2.3. Capacity of public health laboratories	0.56
2.4. Preparedness for an effective response to control threats to the environment and public health	0.40
2.5. Technical assistance and support to sub-national/local level in surveillance, research and control of risks and threats to public health	0.45
Essential PHF#3	
3.1. Support for Health Promotion Activities, development of norms and interventions to promote healthy behaviors and environments	0.41
3.2. Building sectoral and extrasectoral partnerships for health promotion	0.19
3.3. National planning and coordination of information, education and communication strategies for health promotion	0.19
3.4. Reorientation of the health services towards health promotion	0.37
3.5. Technical assistance and support to subnational levels to strengthen health promotion activities	0.15
Essential PHF#4	
4.1. Empowering civil society for decision- making in health	0.17
4.2. Strengthening of social participation in health	0.19
4.3. Technical assistance and support to subnational level to strengthen social participation in health	0.19
Essential PHF#5	
5.1. Definition of national and subnational health objectives	0.44
5.2. Development, monitoring and evaluation of public health policies	0.38
5.3. Development of institutional capacity for the management of public health systems	0.07
5.4. Management of international cooperation in public health	0.33
5.5. Technical assistance and support to the subnational levels for policy development, planning and management in public health	0.31
Essential PHF#6	
6.1. Periodic monitoring, evaluation and revision of regulatory framework	0.59
6.2. Enforcement of laws and regulations	0.38
6.3. Knowledge, skills and mechanisms for reviewing, improving and enforcing the regulations	0.44
6.4. Support and technical assistance to the subnational levels of public health in developing and enforcing laws and regulations	0.26
Essential PHF#7	
7.1. Monitoring and evaluation of access to essential health services	0.32
7.2. Knowledge, skills and mechanisms to improve access to essential health services	0.25
7.3. Advocacy and action to improve access to necessary health services	0.47
7.4. Support and technical assistance to the subnational levels of public health to promote equitable access to necessary health services	0.35

Table 2. Cont.

Indicator	Value
Essential PHF#8	
8.1. Description of the public health workforce profile	0.36
8.2. Improving the quality of the workforce	0.17
8.3. Continuing education and graduate training in public health	0.18
8.4. Improving workforce to ensure culturally appropriate delivery of services	0.08
8.5. Technical assistance and support to the subnational levels in human resources development	0.14
Essential PHF#9	
9.1. Definition of standards and evaluation of quality of population-based and personal health services	0.30
9.2. Improving user satisfaction with health services	0.25
9.3. Systems for technology management and health technology assessment that support decision-making in public health	0.24
9.4. Technical assistance and support to the subnational levels to ensure quality improvement in personal and population-based health services	0.31
Essential PHF#10	
10.1. Development of a public health research agenda	0.43
10.2. Development of institutional research capacity	0.47
10.3. Technical assistance and support to the subnational levels for research in public health	0.38

Discussion

Recommendations include:

1. improvement of technical support for monitoring and evaluation of the health status of the population;
2. strengthening public health surveillance capacities;
3. timely and efficient response to unanticipated public health risks, particularly at local level;
4. establishment of intra- and intersectoral coordinated system for health promotion at national and local level, in addition to technical assistance and support to local institutions;
5. establishment of efficient mechanisms and technical support for active involvement of the civil sector in decision making and policy development processes;
6. institution of efficient system for monitoring of implementation of public health policies with technical support to local levels;
7. creation of capable system for monitoring, evaluation and modification of implementation of regulations, with special emphasis to technical support to local level institutions;

8. strengthening the institutional and human resources system for monitoring and evaluation of the access to essential health services and appropriate technical support to local level;
9. promotion and amplification of capacity building of health professionals to provide high quality health services;
10. re-definition of quality standards for health services as to comply with EU regulations, with particular emphasis on patient satisfaction and their active involvement;
11. launch of Strategy and Action Plan for public health research with active involvement of civil society and local self-government, as well as call attention to attain financial resources for strengthening technical and human capacities for research.

EXERCISE

Task 1

Students will break into 4 groups (may break in sub-groups for parallel activity). Each group will:

- analyze 1 EPHF with data from different country,
- perform SWOT analysis for the respective EPHF,
- group presentations (outlined in a handout).

Task 2

Students will break into 4 groups. Each group will:

- analyze all EPHF with data from the same country,
- perform SWOT analysis of all EPHF and comment results,
- group presentations – overcoming the identified gaps for each EPHF.

REFERENCES

1. Das Gupta M, Rani M. India's Public Health System. How Well Does It Function at the National Level? World Bank Policy Research Working Paper 3447, November 2004. Available from URL: http://econ.worldbank.org/external/default/main?pagePK=64165259&theSitePK=469382&piPK=64165421&menuPK=64166093&entityID=000012009_20041130093207. Accessed March 24, 2009.
2. What is good governance? UN Economic and Social Commission for Asia and Pacific. Available from URL: http://www.unescap.org/pdd/prs/ProjectActivities/Ongoing/gg/governance.as. Accessed March 24, 2009.
3. World Health Organization, Pan American Health Organization. Public Health in the Americas. Conceptual Renewal, Performance Assessment, and Bases for Action. Washington, D.C.: Pan American Health Organization, 2002. Available from URL: www.lachealthsys.org/documents/fesp/Public_Health_in_the_Americas_-_Whole_Book.pdf. Accessed March 24, 2009.

4. World Health Organization, Pan American Health Organization. Essential Public Health Functions. In: World Health Organization, Pan American Health Organization. Public Health in the Americas. Conceptual Renewal, Performance Assessment, and Bases for Action. Washington, D.C.: Pan American Health Organization, 2002. p.59-70. Available from URL: www.lachealthsys.org/documents/fesp/Public_Health_in_the_Americas_-_Whole_Book.pdf. Accessed March 24, 2009.
5. World Health Organization, Pan American Health Organization. Country Health Profiles [homepage on the Internet]. Available from URL: http://www.paho.org/english/hia_1998ed.htm. Accessed March 24, 2009.

RECOMMENDED READINGS:

1. Breslow L. From disease prevention to health promotion. JAMA. 1999;281:1030-33.
2. Breslow L. Perspectives: The Third Revolution in Health. Ann Rev Public Health. 2004;25:XIII-XVIII.
3. Chen L, Evans T, Anand S, Boufford JI, Brown H, Chowdhury M, et al. Human resources for health: overcoming the crisis. Lancet**.** 2004;364:1984-90.
4. Chrvala CA, Bulger RJ, editors; Committee on Leading Health Indicators for Healthy People 2010. Leading Health Indicators for Healthy People 2010: Final Report. Washington, DC: Institute of Medicine, Division of Health Promotion and Disease Prevention; 1999.
5. Etzioni A. The Spirit of community. Rights, responsibilities, and the communitaria agenda. New York: Crown Publishers, Inc.; 1993.
6. Fuenzalida-Puelma HL, Connor SS. The right to health care in the right to health in the Americas. Washington, D.C.: WHO, Pan American Health Organization, PAHO Scientific Publication No. 509; 1989.
7. Lippeveld T, Sauerborn R, Bodart C. Design and implementation of health information systems**.** Geneva: World Health Organization; 2002.
8. Martineau T, Martínez J. Rethinking human resources: an agenda for the millennium. Health Policy Planning. 1998;13:345-58.
9. McNabb SJN**,** Chungong S**,** Ryan M**,** Tadesse W**,** Nsubuga P**,** Alemu W**,** Carande-Kulis V and Rodier G**.** Conceptual framework of public health surveillance and action and its application in health sector reform. BMC Public Health. 2002;2:2. Available from URL: http://www.biomedcentral.com/1471-2458. Accessed: June 10, 2009.
10. Narasimhan V, Brown H, Pablos-Mendez A, Adams O, Dussault G, Elzinga G, et al. Responding to the global human resources crisis. Lancet. 2004;363:1469-72.
11. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, Chiller TM, Espitia V, Imtiaz R, Sosin D, Stroup DF, Tauxe RV, Vijayaraghavan M, Trostle M. Public health surveillance: A tool for targeting and monitoring intervention. In:. Disease Control Priorities in Developing Countries (2nd Edition),ed. New York: Oxford University Press; 2006. pp. 997-1018. Available from URL: http://www.dcp2.org/pubs/DCP. Accessed March 24, 2009.
12. Oliver A, Mossialos E. Equity in access to health care: outlining the foundations for action. J Epidemiol Commun Health. 2004;58:655-8.
13. Rootman I, Goodstadt M, Hyndman B, McQueen DV, Potvin L, Springett J, Ziglio E. Evaluation in health promotion: principles and perspectives. Copenhagen: WHO, Regional Office for Europe, European Series, No. 92; 2001.
14. World Health Organization, Pan American Health Organization. Health Indicators: Building Blocks for Health Situation Analysis. PAHO Epidemiological Bulletin. 2001;22:1-5. Available from URL: www.paho.org/english/dd/ais/eb_v22n4.pdf. Accessed March 24, 2009.

15. World Health Organization, Pan American Health Organization. Public Health in the Americas. Conceptual Renewal, Performance Assessment, and Bases for Action. Washington, D.C.: Pan American Health Organization, 2002. Available from URL: www.lachealthsys.org/documents/fesp/Public_Health_in_the_Americas_-_Whole_Book.pdf. Accessed March 24, 2009.
16. World Health Organization. Health Data. Available from URL: http://www.who.int/whosis/en/. Accessed March 24, 2009.
17. World Health Organization. Otawa Charter. Available from URL: www.who.int/hpr/NPH/docs/ottawa_charter_hp.pdf. Accessed: June 10, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**STRATEGIC PLANNING IN HEALTH CARE - GENERAL APPROACH**
Module: 4.4.4	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Doncho Donev**, MD, PhD**,** Professor Institute of Social Medicine, Faculty of Medicine, University "Ss Cyril and Methodius", Skopje, Republic of Macedonia **Neda Milevska-Kostova**, MSc, MCPPM Centre for Regional Policy Research and Cooperation "Studiorum", Skopje, Republic of Macedonia **Adriana Galan,** Public Health and Health Management consultant Institute of Public Health, Bucharest, Romania
Address for correspondence	**Doncho Donev** Institute of Social Medicine, Faculty of Medicine, 50 Divizia 6, MKD-1000 Skopje, Republic of Macedonia E-mail: dmdonev@gmail.com
Keywords	Health policy and strategy; strategic planning; strategic management; strategy implementation
Learning objectives	After completing this module students and public health professionals should be able to understand: • the basic features of strategic planning (SP) in health care, • strategic assessment components and marketing audit, • health policy and strategy concepts at macro and micro level, • generic organization strategies for health care organizations, • the process of health policy formulation, communication, implementation and evaluation.
Abstract	SP is a process of anticipating the future, assessing present conditions and making decisions and actions to achieve particular outcomes set within a policy. There are two types of SP and strategies in health care: at a macro-level - aiming to improve overall health based on a broad comprehensive approach (general strategies) or directed specifically to one or more dimensions of the population health (sectoral strategies); and at a micro level - the process of strategic management for deciding on the vision, mission, goals and objectives of the organization and on the formulation and implementation of strategies that are to be followed in attaining them.
Teaching methods	An introductory lecture gives the students the insight of the basic concepts of strategic planning in health care at macro and micro level. Additional teaching methods for this module are work in small groups, discussion and individual work.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers, LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice test and essay.

STRATEGIC PLANNING IN HEALTH CARE - GENERAL APPROACH

Doncho Donev, Neda Milevska-Kostova, Adriana Galan

THEORETICAL BACKGROUND

Applying the most advanced health planning and programming methods as well as the improvement of the health system organization and performance represent permanent matters of concern for every country.

One of the most famous examples of strategic thinking is the well-known dialogue from Lewis Carroll's "Alice in the Wonderland":

"Cheshire Puss, ... would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where-" said Alice.

"Then it doesn't much matter which way you go," said the Cat.

"- so long as I get somewhere," Alice added as an explanation.

"Oh, you are sure to do that," said the Cat.

This dialogue reflects very well the essence of a good strategy, which is about the results that should be achieved and what the ways are. Planning provides the direction and the sequence of activities in accomplishing the goals and objectives. To make a good plan it is necessary to know what are the results expected. In other words, an organization without a strategy is like a ship without a rudder, going around in circles (1-3).

The Oxford dictionary defines strategy as "planning and directing of the whole operation of a campaign or war; plan, policy". According to this definition, strategy is equivalent with plan and policy. Sometimes the terms are used as synonyms, but it would be a slight difference between them. For instance, World Health Organization's (WHO) definition of "health policy" is: "A formal statement or procedure within institutions (notably government) which defines priorities and the parameters for action in response to health needs, available resources and other political pressures" (2).

Debates in health care today

Health policy is an important and fast-growing area of debate not only within academics and health professionals, but also within politicians, community groups, the media and the public. This is due to the fact that health gets more and more importance in the whole society, and new uncertainties appeared as a result of recent economic and social growth and development; actually the basis for important policy debates. In the developed countries, the health expenditure has risen to such levels representing about 10 per cent of all economic activity and Gross Domestic Product (GDP). Other countries are expected to follow this trend. The health sector has become the largest employer in the society due to the explosion of medical opportunities. However, inequalities in health status and accessibility of health care between countries and within countries are still present and intensifying. The middle of the twentieth century was a period of triumphalism for

medicine, but there are signs that the confidence in the technical solution is close to the end. Technocratic determinants of health care may condition the way in which the operational level is working, but the overall shape of health care has much more to do with political, social, economic and historical determinants than with technology (4,5).

The value of health care

Health care is a process which is not limited to the health sector. In the context of Western medicine, health care is focused on the organised medical care of individuals. This concern is realised through the application of medical technology by health care professionals in an institutional setting. However, health care is not always clearly distinguished from other activities to improve human life and not always valued first and foremost for the production of health. Health care professionals may see their goals not in terms of improving health, as the long-term objective, but rather in terms of managing an organisation, which is sometimes shorter-term day-to-day goal. For some people health care is a right, a human requirement which any decent society will make available for all its members. For others, health care is to be seen as an expenditure and the health sector as producing health care which is the subject of political struggle and capturing votes - a commodity or resource for which people compete.

Beyond technical developments there is also growing recognition of the importance of health for the overall objectives of the societies within the European Community and broader. Health is a key foundation stone of the overall Lisbon strategy of growth, competitiveness and sustainable development. A healthy economy depends on a healthy population. This is doubly important as the European population ages in the coming decades and beside adding years to life it is necessary to add healthy life to years. Development planners have often argued that health care should be viewed as an investment in a healthier and more productive society. An alternative view sees organised health care as a large industry. Health care may be distributed unequally in the society, with favourized groups, defined by class, ethnicity, age or other attributes such as skills, receiving more than others. The issue of market-based approaches to health care versus public sector-based approaches runs through all policy debates today configuring debates about equity, development and sustainability (4,5).

Health policy

Policy should mainly focus on the vision for the future while learning from the past experiences, outlining priorities, setting clear directions and the main objectives within the given societal and economic context, and on the role of reaching consensus, and informing people. Based on the general policy framework, the health plans, programmes and strategies can be further developed, designing the way in which policy objectives will be achieved, establishing short- and medium-term deadlines and setting clear responsibilities for each actor involved. A wide range of policies can contribute to the improvement of population health and healthier life, ranging from employment and social protection strategies to risk factors control, health promotion and protection, and changes in the life style of the population (2,5).

There are four important factors in the process of policy development and implementation in health care and public health:

1. identification of health risks/problems and preventive options,
2. intervention development,
3. policy development, and
4. policy enactment and assurance.

In practice, problem definition, political context, and the policy process are inextricably intertwined, although exerting varying influences at different stages of the policy cycle. Decision-making lies at the heart of the policy process (6).

Health strategy development

Situational analysis represents an important step of the pre-planning phase for a strategy development. It actually evaluates the profile of population's health (can be a "target" population) and the health care system in relation with the internal and external environment. The assessment can be done based on available and reliable health indicators.

The main goal of this step is to identify priority health problems based on valid criteria. Another important goal is to provide data and information necessary to design goals and objectives for the strategy. Data and information collected during this step cover the following domains:

- internal and external environment (review of economic, social and health objectives and policies) - SWOT analysis;
- health status and related determinants (mortality and morbidity rates, disability, burden of disease, life expectancy, lifestyle indicators, trends etc.);
- health system (public/private institutions, accessibility for health care, population coverage with services, patient flow within the health care system, etc.) resources - human, material and financial (7-10).

If there is a functional and valid information system, health indicators constitute a fundamental tool that generates evidence on the population health status and trends. Inequalities in health can be also evaluated, which may - in turn -serve as basis for highlighting the population groups with the highest health needs and identification of critical areas. If existing, health indicators facilitate further monitoring and evaluation of health objectives and goals set up by a strategy or program. The main output of this step is represented by a comprehensive review to inform the strategy, offering a comprehensive picture of the existing situation. Data obtained through the situation analysis also provide a benchmark against which to measure future trends (9).

Problem identification and priority setting process is based on existing health system indicators, on special surveys, and on consensus research. It is actually a process of comparisons and decision-making, based on special methods and techniques for ranking the identified problems according to their importance. Limited resources require priority setting to address competing demands across health system. In order to judge and prioritise the identified problems three main criteria are commonly used:

- *problem's dimension and severity* (incidence/prevalence, premature deaths, potential years of life lost, burden of disease, trends, the size of the population at risk, the impact on medical services, family, society, etc.);

- *intervention capacity* (knowledge on the disease/associated risk factors, prevention possibilities);
- *existing resources for intervention* (existing services, qualified personnel, population accessibility to health services), (9,11).

Strategy formulation is the process leading to the establishment of national health goals. It is crucial for a successful strategy that the goals are formulated through a democratic process, involving a continuous dialogue with target population, as well as with those actors who will have responsibility for its implementation (9,12).

Goals and objectives

A goal

Broadly, a goal is a statement of intended output (13). A goal represents a general aim towards which to strive; a statement of a desired future state, condition, or purpose (14). A goal differs from an objective by being more general and comprehensive, and usually by being long-range rather than short range. Within the health sector, WHO has defined the goal of Health for all by the year 2000, which means the pursuit of the goal that "as a minimum all people in all countries should have at least such a level of health that they are capable of working productively and participating actively in the social life of the country in which they live". Moreover WHO recognizes "that to attain such a level of health every individual should have access to primary health care and through it to all levels of a comprehensive health system". Within the WHO strategy, targets have been defined with indicators to assess progress towards the overall health goal. Member States of WHO have generally endorsed the health goal and the adoption of the targets and the use of the indicators system for assessing progress in health development. Health goals summarize the health outcomes which, in the light of existing knowledge and resources, a country or community might hope to achieve in a defined period of time. A goal should really represent the solution to an identified problem, being realistic at the same time. Goals should be directed toward the vision and principles generally accepted; something the health system wants and expects to accomplish in the future (13,14).

An objective

An objective is a measurable condition or level of achievement at each stage of progression toward a goal (14). Objectives are clearly specifying a relevant timeframe within which they should be met. If goal statements are generally vague, a well-designed objective will be Specific, Measurable, Attainable/Achievable, Realistic and Time-bound (SMART) (see Table 1).

Table 1. Main features of the objective in healthcare planning (9).

SMART Features	Definition/ meaning
Specific	An objective should address a specific target or accomplishment
Measurable	A metric (usually an indicator) should be established to indicate that an objective has been met
Attainable/Achievable	If an objective cannot be achieved, then it's probably a dream
Realistic	Limit objectives to what can realistically be done with available resources
Time-bound	To achieve objectives within a specified time frame

Health Strategy formulation and communication

The aim of health strategy formulation is to determine interconnected medical, public health, and health sector related measures and activities, planned to achieve the goals and objectives. The current level of development of the health system and health care should be the starting point in the formulation of a new strategy. Radical changes are usually easy to implement and rather not recommended. On the contrary, it is very important to preserve and maintain those elements of the health care system that proved to work properly.

Health policy and strategy analysed alternatives can be informed in a number of ways, using some of the already defined policy communication tools:

1. health policy study with policy and strategy options assessment against the existing policy solutions;
2. health policy paper to initiate debate among stakeholders and policy community focused on analysing the policy and strategy alternatives, and the description of the urgency and relevance of the problem would come alongside for argumenting the proposed health policy and strategy solutions;
3. health policy brief in a form of a narrative, rather than as a structured text typical for the policy paper, focused on argumenting pro one and against other alternatives, giving enough arguments to the decision makers for the preferred solution or strategy;
4. oral brief as a from of presentation it goes straight forward into the issue being presented in a very specific way that the audience is directly and on-the-spot confronted with the health policy analysts and decision-makers. It is important the process of strategy formulation to be transparent, and careful consideration needs to be given to the stakeholders interests and what expertise is needed in order to design a feasible strategy. Health strategy needs to include practical and concrete considerations – and needs to provide a time-line that will indicate when steps need to be taken, who needs to be involved, and how their efforts should be coordinated (12,15,16).

Action plan/ implementation plan

The successful implementation of a strategy is determined by some aspects related to the implementation plan, such as: coordination and responsibility, budgets, monitoring and evaluation. More important than the number or the category of actors involved in strategy implementation is the relationship between them, the degree of communication and the clear definition of each actor's roles and responsibilities (2).

The Action Plan sets out the strategic directions and actions for improving the (health) outcomes. The action plan contains besides goals and principles, specific objectives and appropriate actions. It also includes an appendix with a description and assessment of general instruments that can be used, such as administrative system structure, regulations and supervision, monitoring, economics, etc. The plan also includes areas of common interest for health and other sectors, and where better integration or co-operation is needed. An Action Plan is a written outline that defines: 1) What actions need to be done; 2) What resources are necessary to achieve the stated goals and objectives; 3) Who needs to do what; 4) A timeline for accomplishing the goals; and 5) Estimated budgets (9,15).

STRATEGIC PLANNING IN PRACTICE

Strategic planning at a macro-level

General and sectoral global strategies

There is no rule on how extensive should be a public health strategy. International and national practice has shown that public health strategies could be of two main categories:

1. General global strategies,
   These strategies are aiming at improving overall health based on a comprehensive approach, such as:
   - the WHO Strategy "Health for All in the 21st Century" (17);
   - Primary Health Care Strategy (Alma Ata Declaration, 1978) (18);
   - The World Bank Development Report 1993 - Investing in Health (19);
   - Health in All Policies (5);
   - WHO international health guidelines like Guidelines on hand hygiene in health care and Guidelines for drinking-water quality (20,21).

2. Sectoral global strategies,
   These strategies are limited to one or more dimensions of public health and these could address:
   - a specific domain: "Charters, Declarations and other documents from the Global Conferences on Health Promotion: Ottawa 1986, Adelaide 1988, Sundsvall 1991, Jakarta 1997, Mexico 2000 and Bangkok 2005 (22);
   - one or more specific diseases: "WHO Strategy for Elimination of Measles and Congenital Rubella Infection" (23);
   - one or more risk factors: "WHO Global Strategy on Diet, Physical Activity and Health" (24);
   - a specific population category: "Global Strategy on infant and young child feeding" (25);
   - International Guidelines on HIV/AIDS and human rights, 2006 (26).

International, national and sub-national strategies

Another criterion for classification is the geographical coverage. From this point of view, strategies could be developed at international, national, sub-national or local community level (2).

International (broad regional) level

For *European region* international organizations such as the World Health Organization, as well as the European Union (EU), have developed either general or sectoral public health strategies, such as:

- "Health Strategy of the European Community" (27),
- "European Environment and Health Strategy" (28),
- "Community Strategy on Safety and Health at Work" (29),
- "European Strategy for Tobacco Control" (30),

- "European Alcohol Action Plan" (31),
- "European Strategy for the Prevention and Control of Non-communicable Diseases" (32),
- "European strategy for child and adolescent health and development" (33).

European Union recognised that it is time now for health to be put at the centre of EU policy making. Health should be considered as the driver of competitiveness and sustainable development, and positioning health as a driver of economic development is part of this process. With an enlarged EU of 27 Member States there are even clearer health and economic inequalities that must be urgently addressed. The European Community Health Strategy contains three main strands:

1. Improve health information for all levels of society;
2. Create a mechanism for responding rapidly to major health threats; and
3. Address health determinants, notably harmful factors linked to lifestyle.

The Community action programme for 2003-2008 in the field of public health represents the cornerstone of the strategy, being an essential component of the European Community's health strategy (9).

Reviewing the documents elaborated by the European Commission, it is worthwhile to notice that generally, at European level, a strategy in the field of Public Health should rely on some global approaches deriving from the new public health concept (9):

- develop and disseminate a system-oriented approach;
- partnership-oriented approaches;
- tailor interventions to the specific needs of communities;
- integrate a policy orientation into public health practice;
- apply a comprehensive strategy across diverse issues;
- foster cross-disciplinary collaboration and strategy;
- advocate for solutions that address multiple problems; and
- train and support the next generation leaders.

The Health21 targets, set up by the framework policy document of the WHO European Regional Committee - Health21 (17), establish the vision of health for all and represent the framework for developing health policies at national level in the WHO European Region. The Health21 policy has the following main elements:

1. The one constant goal is to achieve full health potential for all;
2. Two main aims for better health guide efforts toward this ultimate goal: a) promoting and protecting people's health throughout the course of their lives; and b) reducing the incidence of and suffering from the main diseases and injuries;
3. Three basic values from the ethical foundation:
    - health is a fundamental human right;
    - equity in health and solidarity in action between countries, between groups of people within countries and between genders; and
    - participation by and accountability of individuals, groups and communities and of institutions, organizations and sectors in health development.

Health21 also suggests four main strategies for action for attaining these targets:

- multisectoral strategies to tackle the determinants of health;
- health-outcome-driven programmes and investments for health development and clinical care;
- integrated family- and community-oriented primary health care; and
- a participatory health development process that involves relevant partners for health, at all levels - home, school and worksite, local community and country - and that promotes joint decision-making implementation and accountability.

European countries are expected to adapt these targets to meet their national conditions, needs and capacities. Many European countries have developed national health strategies based on the principles, vision, aims and targets included in the WHO Health21 framework policy document (9,17).

National level

Strategies developed at national level are usually based on the priority health problems in a country, but they are also aligned to the global or regional strategies, based on the overall strategic determination of the country as well as its political, geographic and economic positioning.

Some national or local strategies are adaptations of international or regional sectoral strategies at the specific country/local conditions, such as “Healthy Cities” or “Health Promoting Schools Programmes”. Strategic planning and programming in health care at the national level means translation of the national health policy into long-term and mid-term action plans for health development with goals and objectives to be accomplished in certain periods of time. National health policy and strategy for health development should lead toward organized health system based on primary health care and integration of well defined national health priorities and programmes which provide appropriate health technology. The process of health planning and programming is continuous. National health strategy should identify starting points for health development; explain paths and ways for involvement of other sectors of society influencing health, as well as to provide analysis of priority problems and constraints, and how to overcome them (2,15).

Strategies of EU countries aim at improving health of everyone and the health of the worse off in particular being more centred on: health determinants, reducing social inequalities, target groups, and the main "killers". Strategies of the South Eastern European countries are more focused on improving the public health infrastructure, health care services or the conformity with international standards (2).

National health policy also provides the foundation for the national health strategy as a framework for planning and strengthening public health activities, programmes and services. It guides in working with the community, non-government agencies, local government councils and other government departments. The Public Health Strategy sets the platform for the Governments' action on health. It identifies the priority areas and aims to ensure that health services are directed toward those areas that will ensure the highest health benefits for the population. National health strategy consists of selected medical, public health, professional, financial and development interventions, together with

estimation of necessary manpower and technology for achievement of health development goals (9,15).

Sub-national level

Formulation of appropriate health policy and strategy at national level lead toward formulation of priority programmes for health development and detailed programming at regional, municipal and the local community level. Local strategies can be developed at sub-country level in order to endorse the national strategy at local community level. Local health strategies identify specific aims and steps towards implementation. These include key linkages with other strategies and plans, in addition to the national action plan (2,15).

Strategic planning at a micro-level

The level of health care (services) organization (HCO) or micro-level strategic planning is defined as anticipating the future, assessing present conditions, and making decisions concerning organizational direction, programs, and resource deployment. This process results in answers to questions of what to do and when, where, how, and for what purposes it is to be done. The process of planning consists of a series of activities that include assessing present information about the organization and its environment; making assumptions about the future; evaluating present objectives and/or developing new ones; and formulating organization strategies and operational programs that, when implemented, will accomplish goals and objectives (1,13).

Planning is often characterized by type and time frame, the individuals who plan, and the various approaches used in planning. Type of planning refers to level (strategic/operational) and scope (broad/narrow).

Strategic versus operational planning

Strategic planning in HCOs is performed at the senior management level (such as management board) with input from other organization members, including the professional staff holding leadership positions. It is broad, all encompassing in scope, and particularly concerned with environmental and contextual assessment.

Conversely, *operational planning* in HCOs is more narrow and limited than strategic planning, is performed at lower levels of the organization, and usually is the concern of the mid-level management of the organization. Operational planning is subservient to, derived from, and must be in harmony with strategic planning - in other words, operational planning is action plan for achieving strategic planning goals. Included are establishing sub-objectives along with operational programs, policies, and procedures in major differentiated units of the organization that may encompass groups of departments or individual departments (1,3,15).

Time frames of planning vary. Strategic planning is long range and encompasses multiple years (usually more than 5 years). Operational planning is usually short term, with a time frame of one year or less. The governing body and senior management are responsible for strategic planning in HCOs. The governing body is responsible for setting the HCO's direction and mission, goals and objectives. Senior management has significant input in formulating organization objectives and is charged with developing and implementing organization strategies to accomplish them. The governing body's role does not include originating strategy. It is responsible, however, for ensuring that proposals

brought before it by senior management are properly prepared and are consistent with the HCO's mission and culture and its responsibilities to stakeholders (1).

Middle-level managers also plan. Generally, they are concerned with short-term operational planning and design and implementation of programs, policies, and procedures in their area of responsibility. This most often occurs at the level encompassing multiple departments. Finally, department managers and first-line supervisors also plan, usually in relation to specific operations or activities such as estimating work load, scheduling work activity, and allocating resources (1,3).

Organization strategies and operational programs

The first step to planning is to understand why it is needed. Having strategies allows an organization to find the best approaches to implementing its mission, build on its assets, recognize its weaknesses, and anticipate any problems that may arise. Strategic planning provides guidance and direction for the staff and a tool for the board to hold the organization accountable to fulfilling its mission. It creates a framework within which priorities can be set and a coherent focus given to program implementation.

Health organization strategies are broad, general programs designed by an HCO to accomplish organization goals and objectives. They are long-term major patterns of activity requiring a substantial commitment of resources. Traditionally, "strategy" is the term reserved to describe the means (way) of accomplishing organization goals and objectives (1,34).

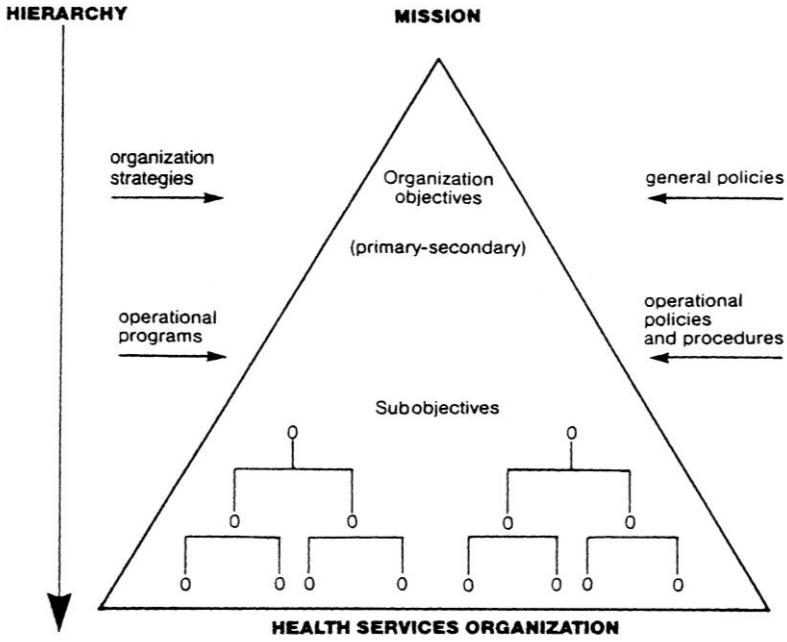


Figure 1. Operationalizing an health care organization's mission (3)

The concept of strategic management and strategic planning at the health care organization level

Strategic management is the process by which a firm or HCO manages the formulation and implementation of its strategy. A *strategy* is the central, integrated, externally oriented concept and coordinated means by which an organization will tend to achieve its goals and objectives. Strategies typically take one of two forms: business strategy or corporate strategy.

1. Business strategy.
   The objective of a business strategy is to spell out how the firm or HCO plans to compete. This plan integrates choices regarding arenas (where the HCO will be active), vehicles (how it will get there), differentiators (how it will compete and be successful), staging (the speed and sequence of its moves), and economic logic (how it obtains its returns).
2. Corporate strategy.
   The objective of corporate strategy is to spell out which businesses/services a firm or HCO will compete in, how ownership by the corporate parent adds value to the business/services, and how this particular diversification approach helps each business/services compete in its respective health care markets (3,13,15,35).

Strategic planning is overt, anticipatory, and long term. Its perspective embraces the whole HCO rather than single departments or units. It is both externally and internally oriented and involves assessment of the HCO vis-à-vis its environments. Strategic planning is the process concerned with:

- formulating HCO primary and secondary objectives, and
- developing organization strategies and the derived operational plans to achieve them.

It is an ends-means process in which objectives are ends to be achieved and strategies are means - the ways to accomplish them.

Organization strategies were described as the HCO's long-range, major patterns of activity requiring a substantial commitment of resources. Those strategies are unified, comprehensive plans (means) that capitalize on the HCO's strengths, take advantage of external opportunities, seek to reduce and/or overcome threats, and mitigate weaknesses. Organization strategies are the means to achieve strategic goals and objectives. Strategic planning requires creative, innovative thinking in providing a good rationale to the process of selecting goals, setting constraints and firm planning for allocating resources (3,13).

Health care organization (HCO) strategic planning

In order to effectively deploy a strategic plan, the organization must align the goals and objectives across all levels and to all functional units. Six stages can be identified in the development of the relationship between macro-systems and micro-systems. There are five critical themes associated with these stages:

- trust making,
- mitigation of constraints and barriers among departments and units,

- creation of a common vocabulary,
- raising awareness of micro-system,
- create cultural and behavioural change, which include the actualization of staff-generated ideas and trust making.

The micro-system, i.e. the small, functional, front-line units, is where the strategic plans become operationalised (15,36).

The components of strategic planning are:

1. Formulating objectives.
2. Strategic assessment/situation analysis.
3. Strategy alternatives and choice.
4. Strategy formulation.
5. Strategy implementation.
   Subsequent to choosing one or more strategies, program implementation occurs through operational planning, resource allocation, and conversion.
6. Strategic control and evaluation.
   Finally, HCO managers control by monitoring and evaluating HCO outputs to determine if the strategies chosen and implemented result in objectives being accomplished (1).

Formulating objectives

Objectives are yardsticks for tracking an organization's performance and progress. Objectives represent a managerial commitment to achieving specific performance targets within a specific time frame. All HCOs have objectives; some are primary, others secondary. Occasionally, objectives are re-prioritized and new ones added. As the ends, targets, and desired outputs, they are the focus of organization activity. Objectives influence selection criteria for strategy choice. If specific strategies, when implemented, are likely to accomplish objectives, several of the choice criteria have been met. The governing body formulates HCO objectives. In doing so, it is affected by other forces: organizational culture, influence of stakeholders, and values/ethics of the choice makers (1,3,13).

1. Organizational culture.
   Organizational culture is the ingrained pattern of shared beliefs, values, and assumptions acquired by organization members over time. Culture is the legacy - what the HCO is and what it stands for - that permeates the HCO. It is known to and, it is hoped, shared by all. Culture shapes acceptable behaviour of members and depicts the desired nature of relationships between the HCO and its stakeholders, including customers and employees. Objectives must be consistent with mission, culture, and the derived relationship with others as reflected by the culture. If the culture is incompatible with mission and mission-oriented objectives, it must be changed. If objectives are inconsistent with mission and mission-oriented culture, the objectives must be changed (1,13).
2. Stakeholders.
   Stakeholders are those constituents with a vested interest in the affairs, actions, and objectives of the HCO. They are individuals, groups, or organizations affected by the HCO who may seek to influence it and its objectives. There are three types of stakeholders: internal stakeholders, such as employees; interface stakeholders, such

as patients and members of the professional staff organization; and external stakeholders, such as third-party payers, government, and the community at large. Each stakeholder group and subgroup may have interests and demands that conflict with those of others and may seek to influence the HCO's priorities and objectives. As new objectives are established by the governing body or the emphasis on existing objectives is changed, these stakeholders will seek to influence the outcome. It is the governing body's responsibility to balance stakeholder demands and ensure that they are compatible with the HCO's mission. Balancing requires maintaining ethical values and social responsibility and preventing inappropriate stakeholder demands from predominating (1,13).

3. Values and ethics.
   Establishing new objectives or modifying present objectives requires choice - a decision by one person or a group. Almost always it is the governing body that makes this choice. Just as culture and stakeholders influence objectives, so, too, do the values and ethics of those who make the choice (1).

Strategic assessment/situation analysis

Strategic assessment is the heart of strategic planning. Essentially, it involves gathering and evaluating information about the past and present and making assumptions about the future. The two major elements are: external environment analysis and internal capability analysis. The marketing audit facilitates both. The "yields" or results of assessment include organizational strengths and weaknesses, and identify external environmental opportunities and threats (SWOT analysis), and risks, issues, and deficiencies confronting the HCO. As observed by others in the business sector, principal assessment activities in strategy formulation include:

1. Identifying opportunities and threats.
   Identifying opportunities and threats in the company's environment and attaching some estimate of risk to the discernible alternatives. Before choice can be made, the company's strengths and weaknesses should be appraised together with the resources on hand and available. Its actual or potential capacity to take advantage of perceived market needs or to cope with attendant risks should be estimated as objectively as possible (1,15).
2. External environmental analysis.
   Environmental scanning to identify external threats and opportunities is critical to strategy formulation and choice. Threats in the environment are events that may adversely affect the HCO. Examples include competition and alternate forms of delivery; change in third-party reimbursement; change in target market demographics and health status; new technologies; changes in accreditation, regulation, and licensure; and status of the economy.

   Opportunities are favourable or advantageous circumstances in the external environment that may benefit the HCO. They include changing demographics and service patterns; decline in primary care physicians in a rural area, which enables a hospital to open a family practice centre. In part, opportunity evaluation includes a market assessment of present services and gaps (heart disease, trauma, perinatology) in service and clientele served (1,15).

Sector analysis is one way in which systematic environmental analysis can be performed and threats and opportunities identified. Among the important sectors are:

- economic (recession, business cycle, capital availability, unemployment), demographics (changes in service area or target market numbers, age, income, location, health status, public health, rise of communicable diseases),
- cultural/ sociological (respect or disrespect for authority, attitudes of employees about work, sexual behaviours, personal ethics),
- political (health care as an actual or perceived right, public policy and federal fiscal responsibility for health care, regulation),
- competition (barriers to entry or exit, new providers, financial incentives, new delivery arrangements),
- technology and support (health personnel, medical education and research, cost and pace of technological development),
- stakeholders (their power and influence, their relative importance to the HCO, and their demands). Competitive position is also part of external environmental assessment. While not focusing on global sectors such as the state of the economy or public policy, it focuses on the HCO's position relative to competitors.

Components evaluated include:

- nature of competition (is it docile/cooperative or aggressive/uncooperative in nature?),
- threat of potential entrants (are other competitors likely to expand products and services in the service area, or are there barriers to entry?),
- threat of substitute products (new technologies such as magnetic resonance imaging, computed tomography),
- power of suppliers and buyers (how much HCO as a supplier of health services is dominant in a market with barriers to entry, are its services specialized, and are there substitute products or services).

External environmental analysis gives an answer to the question "What HCO need to do?" (1,15).

The degree to which external environmental analysis - composed of sector analysis and competitive position - identifies threats and opportunities significantly influences the scope and quality of organization strategies the HCO considers and ultimately chooses. Absence of such analysis amplifies the HCO's risk.

3. Marketing audit

Marketing audit is a systematic evaluation of target markets and their needs. To examine target markets means using many of the environmental components of sector analysis. Consequently, information pertaining to service area, target market, competition, and appropriateness of exchange facilitators is an integral part of strategic planning. Marketing audit gives an answer to the question "Where HCO need to go?" (1,13,15)

4. Internal capability analysis

Internal capability analysis concerns assessing the HCO's strengths and weaknesses, and drawing inferences about comparative advantage or, as some have called it, distinctive competence. Organization strengths include referral patterns, reputation for quality, cost efficiency, qualifications and stature of clinical staff and other professionals, resource (financial) availability and sustainability, the range and types of products and services provided, a cohesive culture, and proactive management. Weaknesses may include shortage of capital, outdated physical plant and equipment, hostile labour environment, poor reputation, aging or decreasing numbers of physicians and health professionals, or reactive management.

Awareness of an HCO's strengths in all functional areas permits conclusions about comparative advantage. Some of these conclusions specify its particular skills, where it excels compared to other HCOs, and how it can be differentiated. Is the HCO the lowest cost producer? Does it have the best reputation? Is it on the cutting edge of technology application? Comparative advantage is a barrier to entry that other HCOs must overcome in order to compete. The systematic appraisal and evaluation of an HCO through internal capability analysis and external environmental analysis identifies internal strengths and weaknesses as well as external threats and opportunities facing the HCO and comparative advantage. SWOT analysis enables managers to identify risks facing the HCO, issues confronting it, organizational deficiencies, and gaps in objectives. Gaps in objectives are the difference or discrepancy between actual and desired results. Internal capability analysis gives an answer to the question "What HCO can/ is able to do?" (1,11,13).

Strategy alternatives and choice

Strategy choice includes formulation of organization strategies for consideration and choice of one or more strategies relative to criteria. Both are done within a particular context. At any given time, an array of organization strategies can be identified, evaluated, chosen, and concurrently implemented by HCOs to achieve objectives. Seldom does an HCO implement only one strategy; it usually implements a combination. Strategies are extensively discussed in the literature, which uses several typologies based on different criteria, i.e. the level of the strategic management, the type of the development of the HCO (extensive, intensive), the direction and extend of development (limited growth, spread growth, liquidation), the way of realization (revolutionary, evolutionary) etc. The most common alternative strategies for HCO are: specialization/niche, vertical and horizontal integration, diversification, retrenchment/divestiture and inter-organizational linkages (1,13,15).

1. Specialization/niche.
   Specialization by product/service and/or market is an organization strategy in which HCOs focus on or emphasize selected products or services, often based on disease or acuity of illness. Some hospitals, for example, are known for their specialization in oncology, organ transplantation, and cardiac surgery; nursing facilities specialized in chronic care; hospices specialized in care of the dying. A niche strategy means focusing on a service area, such as the inner city, or target market, such as ambulatory outpatients. Both strategies are usually implemented in tandem - as in the case of a paediatric hospital that specializes by type of care

(neonatal) and has a niche in a specific target market (children) - and may involve differentiation based on low-cost, high-technology leadership.

2. Vertical integration.
   Vertical integration as an organization strategy occurs when an HCO operates at more than one point on a chain of production and/or distribution. Vertical integration means a "broad range of patient care and support services operated in a functionally unified manner" adding upstream or downstream services. The concept of vertical integration may also be based on the patient's acuity of illness, which can range from acute to chronic. Examples of forward integration for an acute care hospital would be ambulatory care, satellite family practice clinics, and wellness promotion. Backward integration includes long-term care and rehabilitation.

   Integration may also occur in non-service areas - that is additional non-specific activities, such as factors of production that involve make/buy decisions (contracted housekeeping or data processing services for the HCO, or those that supply or manufacture generic pharmaceuticals, prosthetic devices, and intravenous solutions used by a hospital), as well as when a hospital offer fitness services, parking, book selling or educational programs and activities (1,15).
3. Horizontal integration.
   Horizontal integration is an organization strategy in which an HCO expands its core products or services at the same point in the production process and in the same part of the industry. This is usually done to round out product/service lines and to enter new markets. Horizontal integration may be achieved through internal development, acquisition, or merger. An acute care hospital that adds coronary bypass surgery to its existing surgical services or that builds a suburban acute care facility is horizontally integrating (1,37).
4. Diversification.
   Diversification is an organization strategy in which HCOs add new products/services and/or enter new markets. Diversification is usually defined relative to the organization's traditional main line of business and/or core services, and whether the activity is related or unrelated. The strategy of diversification is oriented toward horizontal development spreading the borders of the current core services in conditions of broadening the market. For example, for acute care hospitals, diversification includes adding new non-inpatient-care products/services, such as occupational medicine, women's medicine, and wellness programs, and/or non-acute-care activities, such as rehabilitation, palliative care and substance abuse treatment (1,15).
5. Retrenchment/divestiture.
   A strategy of retrenchment, or downsizing, involves reducing the scope or intensity of products and services, partial withdrawal from a market area, or decreasing capacity in terms of facilities, equipment, or staff. A divestiture is eliminating products or services, withdrawing from a market area, or closing facilities. In highly competitive markets in which the HCO has no comparative advantage or in instances in which demand has decreased, an HCO may implement a strategy of retrenchment - in extreme cases, divestiture. The more commonly implemented strategy is retrenchment. This reduces losses, permits reallocation of resources to more promising services, and, in extreme cases, enables the HCO to survive.

Declining birth rates caused some hospitals to downsize obstetrics; high levels of uncompensated care led others to close (divest) trauma centres; low inpatient occupancy rates caused others to reduce the number of acute care beds (retrenchment) while simultaneously increasing beds in long-term or rehabilitative care (diversifying) (1).

6. Interorganizational linkages.
   Joint ventures, mergers and consolidations, and multi-organizational systems have become prevalent arrangements in health services delivery. For the most part these inter-organizational relationships are of recent origin and represent an organization strategy that is usually coupled with one or more of those previously described: joint ventures can be coupled with vertical and horizontal integration and/or diversification strategies; multi-organizational system arrangements can be coupled with specialization, integration, and even retrenchment strategies (1,37).
7. Context of strategy choice.
   The range of alternative organization strategies considered and those eventually selected is greatly influenced by the context in which choice is made. It includes the type of organization, strategic decision style, managerial philosophy, organizational culture and choice-maker values, portfolio analysis, organizational life cycle, and competitive position. The process of developing alternative strategies, assessment of the alternatives, selection of the most appropriate strategy and writing of the strategic plan at the same time initiate the process of strategy formulation (1,15)

Strategy formulation

Strategy formulation is the process of deciding what to do. Some strategies result from rational and methodical planning processes based on analyses of both internal resources and capabilities and the external environment. In fact, formulation of the strategy starts during the process of situation analysis/ strategy assessment and determining the mission and goals of the HCO. Strategic alternatives are the main possible variants of the future performance of the HCO and they are more concrete/specific than the views to the mission and goals of HCO. Different aspects of strategy are referred to as intended, deliberate, realized, emergent, and unrealized. An intended strategy is the initial plan, whereas the realized strategy is what actually is put in place and succeeds. Thus, parts of the realized strategy can be credited to deliberate choices and actions (i.e., intended strategies that are realized), and parts are due to unplanned ones (i.e., realized strategies that were not deliberate but nevertheless emerged). Finally, some aspect of the initial strategic plan is not realized at all, and drops by the wayside (1,15,35).

Good strategy formulation means refining the elements of the strategy.

Strategy implementation

If strategy formulation is the determination of what the firm is going to do, strategy implementation is how the firm goes about doing it. Strategy implementation is the process of performing all the activities necessary to do what has been planned. Due to the fact that they are interlinked, the two processes are iterative and interdependent from the standpoint that implementation should provide information that is used to periodically modify business and corporate strategy. For successful implementation of the strategy it is necessary to develop clear functional strategies which need to be translated into additional

concrete specific plans and supporting programmes. These functional strategies and supportive programmes relate to the organizational structure and design, marketing, financial resources, human resources etc. In addition, the success in the strategy implementation depends on the appropriate reaction and readiness for change of all those functional spheres and organizational culture, professional qualifications, leadership, motives and incentives, ethics and readiness for changes. Human resources factor is predominant and interrelated with all other factors for successful implementation of the strategy (15,35).

Strategic learning institutions discover their own paths and solutions rather than blindly follow a detailed strategic mandate from administration. Answers to their most critical implementation and adaptive questions will not flow down ready-made from above, but will be tailored to meet the requirements of their own particular situation. Strategic learning organizations have certain attributes in common in developing their own answers: They continuously experiment rather than seek final solutions. They favour improvisation over forecasts, formulate new actions rather than defend past ones, nurture change rather than permanence, encourage creative conflict rather than tranquillity, encourage questioning rather than compliance, expose contradictions rather than hide them. Most importantly, strategic learning organizations realize that successful strategic change is best undertaken as a process of learning. HCOs can no longer afford the illusion of traditional strategic planning, with its emphasis on bureaucratic controls from the top to the bottom. The initial step in discovering ways to improve the capability of HCOs is to adapt continuously while fulfilling their mission. Healthcare leaders must create a shared vision of where an institution is heading rather than what the final destination will be, nurture a spirit of experimentation and discovery rather than close supervision and unbending control, and recognize that plans have to be continuously changed and adjusted. To learn means to face the unknown: to recognize that we do not possess all the answers; to concede that we do not always know what to do; to admit that past actions and solutions may no longer be appropriate, in fact may have been the incubators of today's problems; to question basic assumptions long held about running the institution; and to face the reality that HCO is vulnerable to the political dynamics prevalent in all organizations. Hospitals and other HCOs must seek to develop and maintain a continuing state of readiness in which everyone in the organization, from front-line clinician to senior management, is poised to act in anticipation of and in response to unforeseen changes in the environment and to learn from their own experiences in confronting the future. The strategic plan needs to be a living document, reviewed annually and integrated into the agenda of each board meeting. Strategies can be changed when senior management perceives a threat to the organization, new opportunities for the organization, or a better way to achieve the goals (19,38).

Strategic control and evaluation

Strategic control as a function of the strategic management starts at the moment when strategic goals and objectives are determined and the strategic plan prepared. The main purpose of the control is to recognize over time some deviations from the direction of the selected strategy in order to undertake appropriate corrective actions toward further realization of the goals and objectives of the HCO strategy (15).

Strategic control is realized into *strategic assessment* and *diagnose* from which a conclusion appear about the necessity for change of the strategy. Precondition and

mechanism for more effective strategic control is the conditional (situational) planning for creating alternative strategic plans for the changing circumstances and situation. Basically, the strategic control need to provide answers to the questions: "Do we perform the things right?" and "Do we perform the right things?". Accurate and reliable information to realize strategic changes in the HCO are a precondition for effective strategic control. Such information is expected to indicate the fulfilment of the determined strategic standards (assessment of the implementation) and what are the reasons for recognised deviations (qualitative analysis). The process of strategic control is connected with the strategic evaluation and conclusion which gives an answer to the question "To which extend the strategic goals and objectives are accomplished (15,39).

Each year, the board of the HCO should review both the operational plan and how its objectives have been fulfilled, along with the strategic plan. The board may want to consider establishing a strategic audit committee that meets every two to three years to determine the effectiveness and relevance of the strategy and implementing plan (15,38).

EXERCISE

Task 1

After introductory lecture students, in small groups, should discuss various aspects of the approach and methodology of strategic planning at macro and micro levels.

Task 2

Prepare an oral briefing for a panel of 10 experts in biostatistics, econometrics, and environmental economics. All have extensive knowledge of the health effects of leaded gasoline. Some of them are colleagues with whom you have worked previously. One member of the panel is a respected scientist, who is currently advisor to the Minister. The panel has to take position with respect to the International Lead and Zinc Union, which claims "statistical correlations do not prove causation." The panel is designing a research proposal to investigate the health effects of lead. Your briefing is supposed to help them.

The class is divided into groups of 5, and one student is appointed as group coordinator. The presentation is given in a form of PowerPoint slides, using all its available tools to maximize the effect on the panel of experts. The teacher appoints one person from each group to act as member of the panel, which will compare and evaluate the oral briefings and presentations.

REFERENCES

1. Rakich JS, Longest BB, Daar K. Managing health services organizations. Baltimore: Health Professions Press, Inc.; 1992.
2. Scintee SG. Best practices of public health strategies: Comparison and critical appraisal. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe Public Health Cooperation. Lage: Hans Jacobs Publishing Company; 2005. p.109-24. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
3. Thompson AA, Strickland AJ. Strategic management concept and cases. Tenth edition. Boston, MA: Irwin McGraw-Hill; 1998.

4. Barker K. The Health care policy process. London: SAGE Publications Inc; 1996.
5. Stahl T, Wismar M, Ollila E, Lathinen E, Leppo K. Eds. Health in all policies - Prospects and potentials. Copenhagen: Ministry of Social Affairs and Health-Finland, WHO European Observatory on Health Systems and Policies; 2006. Available from URL: http://www.euro.who.int/document/e89260.pdf Accessed: August 14, 2009.
6. Davis P. Problems, politics, and processes: public health sciences and policy in developed countries In: Detels R, McEwen J, Beaglehole R, Tanaka H. editors Oxford textbook of public health. Fourth edition. New York: Oxford University Press Inc.; 2004. p.937-50.
7. Tchamov K. Planning and programming of health care. In: Bjegović V, Donev D, editor. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.472-90. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
8. Kovacic L, Jaksic Z. Management cycle: From planning to evaluation. In: Kovacic L, Zaletel-Kragelj L, editors. Management in health care practice. A handbook for teachers, Researchers and health professionals. Lage: Hans Jacobs Publishing Company; 2008. p.110-24. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
9. Galan Adriana. Components of a public health strategy. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company; 2005. p.178-89. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
10. Bjegovic V, Wiskow C..Framework for a common regional public health strategy in South Eastern Europe. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company; 2005. p.585-644. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
11. Micovic P. Health management (In Serbian). Belgrade: SZR "Seka"; 2008.
12. Donev D. National public health strategy in Macedonia. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company; 2005. p.381-408. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
13. Anthony NR, Young WD. Management control in nonprofit organizations. Boston:: Irwin; 1994.
14. World Health Organization. European observatory on health systems and policies. Glossary [homepage on the Internet]. Available from URL: http://www.euro.who.int/observatory/Glossary/TopPage?phrase=G Accessed: June 16, 2009.
15. Borisov V. Strategic healthcare management: Philosophy and practice (In Bulgarian). Sofia: Filvest; 2006:.
16. Dunn W.N. Public policy analysis. London: Pearson Prentice Hall; 2004.
17. World Health Organization. Health 21 - The health for all policy framework for the WHO European Region. European health for all series No. 6. Copenhagen: WHO, Regional Office for Europe; 1999. Available from URL: www.euro.who.int/document/health21/wa540ga199heeng.pdf Accessed: August 14, 2009.
18. World Health Organization. Declaration of Alma-Ata International conference on primary health care, Alma-Ata, USSR, 6-12 September, 1978 [homepage on the Internet]. Available from URL: http://www.euro.who.int/AboutWHO/Policy/20010827_1. Accessed: August 15, 2009.
19. The World Bank. The World Bank development report 1993. Investing in Health. Oxford: Oxford University Press;1993. Available from URL: http://www-

wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&siteName=WDS&entityID=000009265_3970716142319. Accessed: August 14, 2009.
20. World Health Organization. WHO Guidelines on hand hygiene in health care. WHO patient safety – A world alliance for safety health care. Geneva: Who; 2009. Available from URL: http://whqlibdoc.who.int/publications/2009/9789241597906_eng.pdf Assessed: August 15, 2009.
21. World Health Organization. Guidelines for drinking-water quality, third edition, incorporating first and second addenda. WHO Water Sanitation and Health (WSH). Geneva: WHO; 2008. Available from URL: http://www.who.int/water_sanitation_health/dwq/gdwq3rev/en/ Assessed: August 15, 2009.
22. World Health Organization. Global conferences on health promotion [homepage on the Internet]. Available from URL: http://www.who.int/healthpromotion/conferences/en/ Accessed: August 14, 2009.
23. World Health Organization. Eliminating measles and rubella and preventing congenital rubella infection. Copenhagen: WHO, Regional Office for Europe; 2005. Available from URL: http://www.euro.who.int/document/E87772.pdf Accessed: August 14, 2009.
24. World Health Organization. Global strategy on diet, physical activity and health. Geneva: WHO; 2003. Available from URL: http://www.who.int/dietphysicalactivity/strategy/eb11344/strategy_english_web.pdf. Accessed: August 14, 2009.
25. World Health Organization. Global strategy on infant and young child feeding. Geneva: WHO; 2003. Available from URL: http://www.who.int/nutrition/publications/infantfeeding/9241562218/en/index.html Accessed: August 14, 2009.
26. UNAIDS. International guidelines on HIV/AIDS and human rights. Geneva: UNAIDS; 2006. Available from URL: http://www2.ohchr.org/english/issues/hiv/docs/consolidated_guidelines.pdf Accessed: August 14, 2009.
27. Commission of the European Communities. White paper - together for health: A strategic approach for the EU 2008-2013. Brussels: Commission of the European Communities; 2007. Available from URL: http://ec.europa.eu/health/ph_overview/Documents/strategy_wp_en.pdf Accessed: August 14, 2009.
28. European Commission. European environment and health strategy. Brussels: Commission of the European Communities; 2003.Available from URL: http://ec.europa.eu/environment/health/strategy.htm Accessed: August 14, 2009.
29. European Commission. Improving quality and productivity at work: Community strategy 2007-2012 on health and safety at work. Brussels: Commission of the European Communities; 2007. Available from URL: http://eur-lex.europa.eu/Notice.do?checktexts=checkbox&val=443914%3Acs&pos=1&page=1&lang=en&pgs=10&nbl=1&list=443914%3Acs%2C&hwords=&action=GO&visu=%23texte Accessed: August 14, 2009.
30. World Health Organization. European strategy for tobacco control. Copenhagen: WHO, Regional Office for Europe; 2002. Available from URL: http://www.euro.who.int/tobaccofree/Policy/20030826_3 Accessed: August 14, 2009.
31. World Health Organization. European alcohol action plan 2000-2005. Copenhagen: WHO, Regional Office for Europe; 2000. Available from URL: http://www.euro.who.int/document/E67946.pdf. Accessed: August 14, 2009.
32. World Health Organization. European strategy for the prevention and control of non-communicable diseases. Copenhagen: WHO, Regional Office for Europe; 2006. Available

from URL: http://www.euro.who.int/Document/RC56/edoc08.pdf Accessed: August 14, 2009.
33. World Health Organization. European strategy for child and adolescent health and development. WHO, Regional Office for Europe; 2005. Available from URL: http://www.euro.who.int/document/E87710.pdf. Accessed: August 14, 2009.
34. O'Sullivan MJ. Strategic learning in healthcare organizations. Hosp Top. 1999;77:13-21.
35. Carpenter MA, Sanders WG. Strategic management concept and cases - A dynamic perspective. Second Edition. New Jersey: Pearson Education, Inc.; 2007.
36. Kosnik LK, Espinosa JA. Microsystems in health care: Part 7. Jt Comm J Qual Saf. 2003;29:452-9.
37. Angrisani DP, Goldman RL. Predicting successful hospital mergers and acquisitions - A financial and marketing analytical tool. New York: The Haworth Press, Inc.; 1997.
38. Perrone M, Johnston J. Presenting: Strategic planning: Choosing the right method for your nonprofit organization. Washington DC: BoardSource; 2007.
39. Kim KK, Michelman JE. An examination of factors for the strategic use of information systems in the healthcare industry. MIS Quarterly 1990;14:201-15.

RECOMMENDED READINGS

1. Rakich JS, Longest BB, Daar K. Managing Health Services Organizations. London: Health Professions Press, Inc.; 1992.
2. Scintee Silvia Gabriela. Best practices of public health strategies: Comparison and critical appraisal. In: Galan A, Scintee G, editors. Public Health Strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company, 2005. p.109-24. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
3. Tchamov K. Planning and programming of health care. In: Bjegović V, Donev D, editors. Health systems and their evidence based development. A handbook for teachers, researchers, health professionals and decision makers. Lage: Hans Jacobs Publishing Company; 2004. p.472-90. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
4. Kovacic L, Jaksic Z. Management cycle: From planning to evaluation. In: Kovacic L, Zaletel-Kragelj L, editors. Management in health care practice. A handbook for teachers, Researchers and health professionals. Lage: Hans Jacobs Publishing Company; 2008. p.110-24. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
5. Galan Adriana. Components of a Public Health Strategy. In: Galan A, Scintee G, editors. Public Health Strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company; 2005. p.178-89. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
6. Bjegovic V, Wiskow C. Framework for a common regional public health strategy in South Eastern Europe. In: Galan A, Scintee G, editors. Public Health Strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company, 2005; p.585-644. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**PRIORITY SETTING IN HEALTH CARE**
Module: 4.4.5	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Vladimir Lazarevik**, MD, MPH, Teaching Assistant Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia **Doncho Donev**, MD, PhD, Professor Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia **Dance Gudeva Nikovska**, MD, MPH, Teaching Assistant Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia **Blaško Kasapinov**, MD, Teaching Assistant Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University, Skopje, Republic of Macedonia
Address for correspondence	**Vladimir Lazarevik** Institute of Social Medicine, Faculty of Medicine Ss Cyril and Methodius University-Skopje, 50 Divizija No 6, 1000 Skopje, Republic of Macedonia E-mail: vlazarevik@gmail.com
Keywords	Priority setting, resource allocation, rationing, evidence-based decision-making
Learning objectives	After completing this module students should be able to understand: • basic concepts and approaches in priority setting (PS) in health care (HC); • components and levels of decision-making process; • ethical principles and the concepts of equity; • economical approach of PS; • stakeholders to be involved into the decision-making process; • main phases of the PS process.
Abstract	PS is one of the most important issues in HC policy. There are no simple and easy solutions in distributing HC resources and setting priorities in HC, so it is crucial that countries develop transparent approaches to facilitate the decision-making process. PS requires explicit debate about the principles and criteria that are used to make decisions about allocating HC resources. Despite numerous attempts, there is no commonly accepted methodology. The PS process can be approached at macro- or micro-level. Based on the international knowledge and our practical experience we have developed cycle for priority setting that can be easily adopted in many settings.
Teaching methods	An introductory lecture gives the students insight into basic concepts and approaches of priority setting in HC. In continuation, students discuss different aspects and components of the process for priority setting in HC. In group work, they apply some models for priority setting.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers, LCD projection, access to the Internet and bibliographic data-bases (PubMed Central); • training materials: recommended readings or other related readings;
Assessment of students	Multiple choice test and essay

PRIORITY SETTING IN HEALTH CARE

Vladimir Lazarevik, Doncho Donev,
Dance Gudeva Nikovska, Blaško Kasapinov

THEORETICAL BACKGROUND

About the concept of priority setting

All health care systems face problems of justice and efficiency related to setting priorities for allocating a limited pool of resources for health care of the population. There are no simple and easy solutions in distributing health care resources and setting priorities in health care (1).

Very frequently the institutions or authorities who make these decisions have disagreements over the principles guiding such decision making process. In all these settings distrust, speculation and impartiality both among the public and end users very frequently surrounds the priority setting process (2,3). This particularly is the case in young democratic societies with young institutions and huge political interferences over the priority setting processes.

Many countries and institutions has tried to set up committees, forums, advisor boards or other official bodies to articulate clear principles that should govern the priority setting process and how can health funding best contribute to health gains in ways that are affordable, fair, and consistent with national values and aspirations. Despite numerous attempts, unfortunately there is no commonly accepted methodology that could be taken from one country or institutions and simply implemented elsewhere. There are various alternatives and options, all depending on the specific settings, culture and traditions where such decisions are made (4-6).

Priority setting (also known as resource allocation or rationing) means the distribution of limited resources among competing programs or people for healthcare rationing by different actors involved in the decision making. Despite priority setting's increasing prominence in policy and academic discussion, it is still unclear what the level of understanding and acceptance of priority setting is at different levels of health care (7,8).

In general, priority setting was found to involve four steps (9):

1. Identification of health care needs,
2. Allocation of resources,
3. Communication of decisions to stakeholders, and
4. Management of feedback from them.

Levels of decision-making and priority setting

Priority setting in health care takes place at all levels within the health system. It is useful in the approach to these problems to divide the decision making process into two main categories: the decision making process that is organized at the *macro-level* (national, regional-district and local community level), and decision making process that takes place at *micro-level* or at the level of healthcare organizations and the level of the individual patient (10,11). Similar categorization of the priority setting process as a complex interaction is pointing out that it occurs at the national level (or macro-level), the regional or district level (or mezzo-level) and the patient level (or micro-

level) (12). Decision-makers at different levels of the health care system may follow different equity concepts. Depending on the features of the health care system, different decision-makers are involved in setting health care priorities. In most public health care systems, there are political decision-makers at the national level and at regional administrative levels. In addition, decisions are made by physicians at the level of the individual patient (13).

Williams defines priority setting as who gets what at whose expense (14). The "what" can be either organs from donors, laboratory procedures or, most commonly, money for services and human resources.

Decisions at the clinical level are crucial for the implementation of equity concepts. At this level interventions actually reach people and policy strategies can make a difference. Faced with the clinical reality, health care providers are likely to follow their individual ethical principles, rather than centrally defined policies (13).

Macro-level decision-making and priority setting

Macro-level decisions are the proposed solutions for resource allocations that affect health care needs of a population at an aggregate policy level, from national to the local community level. It acknowledges a need to spend funds wisely confronting many challenges such as innovative technologies, pharmaceutical advances and demographic trends toward older populations, which will strain already stressed health resources, including plant and budgets (6). It also has implications for the ways states prioritize health spending as a proportion of the gross national spending.

National policymakers or decision-makers at higher levels (such as donor agencies) and lower levels (local governments, even individual clinicians) may profoundly influence resource allocation (13).

Many authors therefore assume implicitly that society's concept of fairness is translated into equity concepts by a benevolent decision-maker. Attempts to elicit equity concepts from the general public have been made, applying concepts of fairness into operational priority setting principles (13).

The examples of such decisions may include:

- distribution of health care facilities at specified geographical region,
- purchasing and distribution of medical equipment on a national scale, or
- introduction of new mandatory vaccine program such was recently case with including human papilloma virus (HPV) vaccine in the regular program for immunization in Macedonia, decision that has sparkled fierce debate in among professionals and the local media (15),
- outbreak of new pandemic influence A H1N1 flu virus popularly called Swine Flu. The decision makers in these countries are forced to make bold decisions in order to secure prevention for their populations purchasing vaccines, anti viral therapy (Tamiflu), and other alternatives.

All these decisions directly influence the health budgets.

Key issues relevant to macro-level priority setting included: application of evidence, values, incentives, physician and public involvement (16).

There is a strong understanding of the importance of evidence-based practice for priority setting at a macro level, but there are significant barriers and challenges in terms of capacity, resources, infrastructure and organizational culture (16).

Priority setting, particularly at a macro level in which decision makers are required to compare disparate patient groups and health outcomes, is necessarily based on values and beliefs, which actually influence decisions and priorities. Values and guiding principles need to be more apparent and serve as a "checkpoint" in the priority setting process to ensure that there is congruence between the decisions and the values (13,16).

The application of evidence in decision-making is a subject of considerable attention in the policy literature as an important component of macro-level priority setting (16).

Micro-level decision-making and priority setting

On the other hand the micro-level decisions take place at the clinical or organizational level and examples of such decisions are type of treatment for the patients, duration of hospital stay, referral of a patient for further examination or treatment abroad, or excluding patient for a given treatment (17). Clinician decision is based on severity of disease and the expected outcome of the intervention. However, these decisions would very much depend of the specific clinical setting, availability of medical technologies, therapies and other accessible alternatives for treatment.

Hospitals and regional health authorities must set priorities in the face of resource constraints. Decision-makers seek practical ways to set priorities fairly in strategic planning, but find limited guidance from the literature. Hospitals and regional health authorities elsewhere are facing significant resource allocation challenges. Priorities must be set among competing opportunities because demand for health care exceeds available resources (18).

The priority setting process at the level of health care organization should be supported by leadership development and change management strategies to strengthen institutional capacity for priority setting decision-making (18).

Criteria used for priority setting

Many criteria for priority setting have been proposed and debated. The most common criteria from a review of the literature, and criteria used in the BOD study are summarized in Table 1. This list is not complete, only some of the most commonly debated criteria are included (12).

Criterion is the societal wish to **maximize general population health**. This has indeed been the basis of many national disease programs in the past century. A second set of criteria relates to the **distribution of health** in the population. Societies may give high priority to interventions that target vulnerable population groups such as the poor, the severely ill, or children or women of reproductive age, because they are more deserving of health care than others. Also, societies may give high priority to the economically productive people to stimulate economic growth, or low priority to people who require health care as a result from irresponsible behaviour (e.g. smoking). A third set of criteria responds to **specific societal preferences**, e.g. for acute care in life threatening situations, or for curative over preventive services. A fourth set of criteria relates to the **budgetary and practical constraints** that policy makers face when implementing interventions, including costs and availability of trained health workers, and may take these into account when choosing between interventions. A fifth set of criteria are political criteria (19).

Table 1. Some of the criteria that are most widely used and debated for priority setting (Table is adopted from Kapiriri L, Norheim OF) (12).

Criteria	Comments
Medical	
• cost effectiveness of intervention, • expected outcome of treatment, • costs of treatment, • effectiveness of treatment, • severity of the condition, • quality of evidence on effectiveness, • urgency of need for care	• most of the medical criteria are well supported by previous studies, • there are differences in the degree of support for different criteria (for example, between cost-effectiveness and severity of illness), • in some studies, respondents valued equality of access to care above outcome of treatment
Non-medical	
• age • gender • race • social status • responsibilities • mental (or learning) capabilities • physical capabilities • area of residence • time on waiting list • political views • community view • number of people benefiting from an intervention • genetic background and sexual orientation • first-time or second-time transplants • likely work status after transplants • patient's lifestyle responsible for cause of disease	• there is lot of controversies over the non-medical criteria, • most of the literature agrees that it is important to include some non-medical criteria, however, studies have failed to come to consensus about which criteria are important, • criteria vary according to who is asked, • some researchers have proposed that unless such criteria have a direct influence on the outcome of treatment it should not be considered.

Creating a priority setting framework

Promoting health and confronting disease challenges requires action across a range of activities in the health system. This includes improvements in the policymaking and stewardship role of governments, better access to human resources, drugs, medical equipment, and consumables, and a greater engagement of both, the providers of services and the public (13).

There is no universal best method of priority setting that works in all circumstances and all countries. The application of simple "rules' is not sufficient as it fails to take into account the wide range of factors and constraints that may influence the priority setting process. In addition, different societies will have

different ideas about what is "optimal." Optimization requires setting of objectives, which may vary between countries and regions. However, two key objectives have received worldwide attention: maximization of health and reduction of inequalities in health (13).

A priority setting framework should be made up of goals, rationalising principles, units of focus, and protocols for decision-making. The basic problem confronting national policymakers is to select an optimal portfolio of programs within the nation's means, as reflected in a fixed budget constraint.

In addition to the groupings of the priority setting process in health care on macro and micro levels, the complexity of the issue should also be approached looking over two different but important perspectives: *ethical* and *economical* (20).

Ethical principles and equity concepts

In pluralistic societies and democratic institutions it is reasonable to expect that people who make the priority setting decisions would disagree over the ethical principles which guide these processes. There are people who believe that highest priority should be given to those people who are worst off, while others would restrain from such approach. There are also supporters to approaches of biggest aggregate benefits to all members in the group. Daniels and Sabin argue that it is easier to establish process that is considered fair for decision making, than to agree on its core principles. In a number of papers these authors have developed the concept called "*accountability for reasonableness*" (18,21-24). This concept has been primarily developed to serve as tool for priority setting and rationing, but its use has been extended to other areas of health policy such as health inequalities as well as human approach in health (25,26). As authors suggested the ethics of the concept how to hold decision makers accountable relies on several key pillars:

- the process should be fully transparent about the grounds for decisions;
- the decisions must be based on reasons that all actors see as relevant; and
- there should be established process how one could challenge the decision in the light of new evidence and arguments.

Many countries with publicly accountable health care systems such as those of Canada, UK, Sweden and others have applied similar ethical framework of accountability in setting their priorities taking into account public attitudes, community expectations and human ethics (27).

It is not easy to determine the reasons for sometimes treating people differently. In considering equitable distribution of healthcare resources, care must be taken to ensure that any differences between individuals or groups that are used to justify different treatment are morally relevant differences. Thus, differences in race, sex or income are not seen as morally relevant. More controversial issues include whether factors such as personal responsibility for health and the presence of dependents are morally relevant in decisions about priority setting in health care (24).

Economic approach and evaluation in priority setting

The supporters of the economical approach as one could expect would propose economic methods and analyses in setting priorities. Despite many problems and

critics to the economical methods they are most widely used with focus on the positive and best results and outcomes. Economists use tools such as:

- cost-benefit analysis (CBA) is an economic method to estimate the cost of certain project in relation to the benefits resulting of the project implementation. CBA makes the decision making process more rational, consisted and clearly communicated;
- cost effectiveness analysis (CEA) compares the cost of two different methods reaching the same goal, but differently to the CBA does not measure the benefits in its monetary value.
- a quality adjusted live year (QALY) is specially developed index that integrates economic values with quality of the life of the individuals. QALY calculates the value of certain healthcare intervention and provide possibility to make direct comparison between different interventions. The general idea is that a high priority health care activity is one where the cost per QALY is as low as it can be (24,28-30).

Some philosophers' objects QALY as relevant measure on the grounds that it has preference towards age and younger people have higher benefits (28,31).

There is a weakness of current approaches to priority setting using cost effectiveness techniques and a broader approach is recommended to resource allocation and purchasing using cost benefit and stakeholder analysis (13).

A review of the literature on priority setting in healthcare adopts an economic perspective on the problem of choosing the optimal portfolio of programmes that can be afforded from a limited national healthcare budget. The traditional economic approach, proposes maximizing health gain (however measured) subject to a budget constraint, which implies ranking programs according to their cost-effectiveness ratio. However, this traditional approach is subject to three important difficulties: limitations in economic evaluation methodology, incorporating equity principles, and practical constraints. These suggest a need for a fundamental rethink of the role of cost-effectiveness analysis in priority setting (13).

Methodological concerns include identifying whose perspective to adopt, the generalization of results to multiple settings, the treatment of uncertainty and timing, and the treatment of interactions between programs. Most equity considerations can be captured in two broad headings:

- equity related to some concept of need; and
- equity related to access to services.

In principle equity concerns can be incorporated into an economic approach to priority setting with relative ease. However we find that many contributions to the debate on equity concepts are theoretical and remote from practical implementation issues. The traditional cost-effectiveness approach generally ignores the numerous practical constraints arising from the political, institutional, and environmental context in which priority setting takes place. These include the influence of interest groups, the transaction costs associated with policy changes, and the interactions between the provision and financing of health services (13).

More recently, one of the economic approaches described in the literature that has been tested and successfully used in practice is program budgeting and marginal

analysis (PBMA) (13,32). PBMA is a pragmatic, economic framework that identifies how resources are being spent before looking into potential changes in service provision, at the margin, to maximize benefit and minimize cost. PBMA can address questions of efficiency and equity, and it can be utilized either within or across programs of care (micro and macro PBMA).

The framework can be operationalized by answering five questions about the use of resources:

- What resources are available in total?
- In what ways are these resources currently spent?
- What are the main candidates for more resources and what would be their effectiveness?
- Could any areas of care be provided to the same level of effectiveness but with fewer resources, thereby releasing resources to fund candidates for more resources?
- Should some areas of care, despite being effective, have fewer resources because another program is more cost-effective?

All these methods have its advantages and disadvantages but the decision what to use is very much content specific. Regardless of the methods used always there are two key economic principles that guide the process of priority setting. The first is the principle of *opportunity costs.* This is economic concept which answers the question what is the second best alternative in financing certain project. In other words decision makers should ask themselves prior making the final decision what do they need to give up in order to proceed with the planned investment.

The second economic principle is that of *margin* which refers to shifting or changing the resources mix. In this scenario the decision makers should answer if more resources are available than planned, how should be spent. Or, if there are fewer resources than expected, how to take resources from sectors producing least benefit and reorienting in the sectors that will yield bigger benefits (13,32).

Stakeholders and actors in priority setting

Faced with continuous ageing of the population, together with increasing health care costs, it is inevitable that health budgets, elsewhere, will be subject to increasing pressure and allocating finite health resources will be more difficult and hotly contested. But who should resolve these problems? Is there a place for citizen involvement, or are decisions best left to others? How much of the decision-making should be left to professionals? Which different players in the health area, including health professionals and members of the public, should be included in the decision-making process related to priority setting? There are some essential questions related to the actors in the priority setting process: Is it appropriate priority setting in health care to be made just by policy analysts and decision-makers? Should the public be involved in priority setting? Or should physicians be gatekeepers of medical resources? Or to focus on the more general question: who should make the decisions (at the different levels) (33).

Stakeholders should be engaged in the priority setting process. Although the organisational executive would ultimately be accountable for making the priority

setting decisions, stakeholders could be engaged particularly as key informants through expert and broader stakeholder consultation. This consultation should include both internal stakeholders (e.g., staff, patient advisory groups) and external stakeholders (e.g., institutional partners, community groups, government officials) (18,33).

Politicians and government officials as policy-makers

When confronted with such complex problems, policy-makers tend to use intuitive or heuristic approaches to simplify complexity, and in the process, important information may be lost, and priority setting is ad-hoc. Or worse, they act out of political self-interest and prioritize interventions according to their own objectives. In other words, policy makers may not always well placed to make informed well-thought choices involving trade-offs of societal values (19).

Evidence-based decision-making may also be at variance with political ambitions. Some examples, without details, illustrate the point i.e. the argument for the retention of small hospitals in rural communities may be more persuasive when political patronage is secured. Although there is evidence that clinical outcomes can be at risk when low patient volumes are combined with inadequate facilities and limited staff, many communities value their hospital as a pivotal community resource. Its importance does not depend entirely on the results of health interventions but also on a contribution to a sense of community pride, connectedness and integrity, issues that are often central to political interests. As a result political support may be biased towards maintaining a hospital service even if the evidence points the opposite (6).

Managers and decision-makers

There are specific responsibilities of the Board and senior management in relation to the priority setting process which should be clarified explicitly and upfront. Decision-makers identified some confusion about these responsibilities given that clinical service priority setting involved an overlap of the strategic responsibility of the Board with the operational responsibility of Senior Management. Board members and Senior Managers sometimes are delineating their respective roles and responsibilities in the priority setting process even could be identified a number of elements that are critical to the design of the priority setting process itself (18).

Clinicians and General Practitioners

There is a perceived misalignment of incentives between physicians (who are not directly accountable to the health region and paid on a fee-for-service basis directly from the government) and the health region, which may be compelled to limit access and constrain spending. This mismatch of incentives translates into a lack of constancy of purpose, poor cooperation and contradictory targets. In fact, inadequate physician involvement in macro-level priority setting has been identified as a reason for perceived inconsistent prioritization of services. Challenges for direct physician involvement include time, but perhaps more fundamentally, the philosophical dilemma physicians face in making macro-level tradeoffs (i.e., rationing decisions) that may create explicit "winners" and "losers." (16).

Physicians were seen to provide valuable input into decision-making both as experts and as evidence drivers, as they bring specific expertise about their particular service areas. The role of physicians, however, should not be expected to include the broader mandate of priority setting across major service areas (16).

Patients and the public/citizenery

Another ongoing challenge in priority setting is in involving the public to inform resource allocation. Generally, there was scepticism regarding the public's level of understanding of funding issues and the range of issues facing decision-makers (16).

Determining how health care resources should be allocated - often termed rationing or priority setting - has traditionally been carried out by health care personnel, usually doctors but increasingly managers. Because priority setting entails assigning value to different outcomes, however, the consumers of health care have a legitimate claim to be involved in the process of allocating resources (34).

Public should have greater involvement in setting health care priorities, but progress on this front has been slow. Merits of the public involvement are compelling and in keeping with a democratic society's desire for an informed, fully engaged citizenry. There are at least 4 reasons for believing that public engagement in setting health care priorities has value (8):

- for the public funds and uses the health care system, citizens are the most important stakeholders of the health care system. thus, legitimacy and fairness demand that they be at the priority setting table,
- greater involvement of the public in policy-making is in keeping with the principles of democracy,
- empowering people to provide input in decisions that affect their lives encourages support for those decisions, which in turn improves the public's trust and confidence in the health care system, and
- public involvement provides a crucial perspective about the values and priorities of the community, which should lead to higher quality, or at least greater acceptance of, priority setting decisions.

Priority setting process

Using the available literature and practical experience in health policy making we were able to develop priority setting cycle. This priority setting cycle is consisted of six main phases:

1. Creating the priority setting forum;
2. Principles or rules of the process;
3. Economical considerations;
4. Questions each decision making process should follow;
5. Defining the ethical issues and dilemmas;
6. Elaborating the decisions publicly.

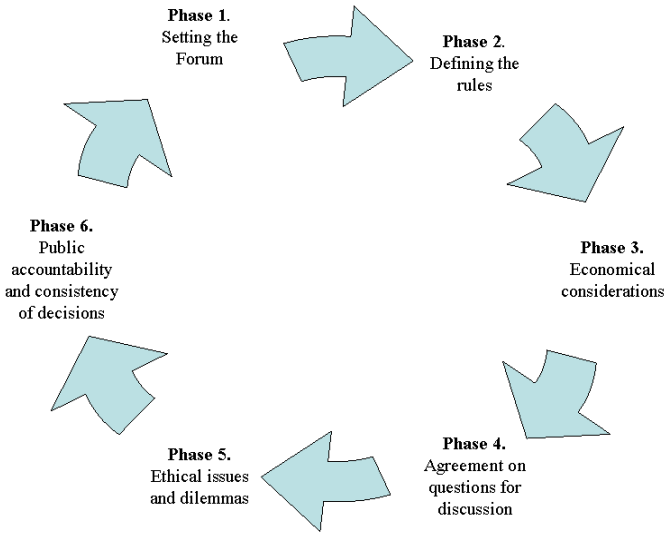


Figure 1. Priority setting cycle.

The process can be applied at the macro and at the micro levels.

Phase I - Creating the priority setting forum

The priority setting process should start by setting up relevant body that would be able to deal with the issues of priority setting. Such body can be established on national, regional or even at the local community level, as an advisory board to the national or local community authorities and decision makers dealing with health care priority setting. It is not important how the body would be called. There are number of examples. It can be named as priority forum; advisory board, steering committee for priority setting, etc. For the purpose of this module we would name our priority setting body as Priority Forum (hereafter referred just as Forum).

More important aspect of the Forum is the composition of the members and its ability to deal with the priority setting issues, pressures and the processes.

In order to make the work manageable it is recommended Forum to be consisted of smaller number of members that can vary from 7 (seven) to 11 (eleven). The appointment of the members of the Forum by the Minister of Health and the local community authorities should follow certain pattern, defined criteria for selection and the duration of the work in the Forum. The members vote to elect chair and vice chair of the Forum.

The work of the Forum is transparent, and people can be invited to observe how the decision making process takes place.

The forum will act through meetings, focus group and interviews with senior managers and clinicians to critically reflect on current and historical macro-level priority setting practices and to identify key components for inclusion in a macro priority setting model with multiple decision criteria (16).

The structure of the Forum should be multidisciplinary and its role should be clearly and explicitly defined in advance of priority setting with shared accountability for priority setting across the clinical and administrative leadership. Engaging the medical leadership in a decision-making role was identified as key to developing a successful priority setting process. The engagement of other non-medical clinical leaders (e.g., nursing leadership) was also thought to be important for ensuring the legitimacy of the priority setting process (18).

Phase II - Rules and principles of the priority setting process

The second phase is the agreement among the members of the Forum on the principles of work. The experience shows that in the democratic societies one could expect disagreements over principles that would guide the process. Daniels suggests that it would be able more easily to agree on what members of the group would consist as fair process of decision making. Here we refer to the concept of accountability for reasonableness that can be applied and adapted with various modifications specific for different settings (21-23).

Forum would promote the importance and benefits of the priority setting in health care, and by ongoing informal and formal training would build understanding among senior managers and clinicians of economic principles for priority setting activity (16).

Priority setting criteria should be clearly defined and understood by decision-makers and stakeholders. Data/information should be collected to support their application in the priority setting process. Initially identified and proposed criteria by the Forum could be further refined through stakeholder engagement and tested with decision-makers to ensure a common interpretation of each criterion and consistency in their implementation (18).

Phase III - Economical considerations

The third phase needs to take into account the economical consideration in setting the priorities. Different authors suggest various approaches, but the most applicable are different types of cost effective analysis using QALY. There have been serious critiques to these models mainly due to certain ethical consideration mentioned earlier. However, economical methods enable the decision makers to establish monetary value or referent costs the authority or organization is able to afford in investment in new therapy, program or intervention (4). Certainly the monetary value per QALY gained very much depends on the financial abilities of the organization which makes the decisions. Establishing guiding cost per QALY gained would enable decision makers to have consistency in treating similar cases equally in similar situations.

Important interactions occur between methods of financing health care and priority setting processes, even though in principle, these two processes should be independent of each other. Forum should illustrate the impact that particular forms of financing (collective insurance, private insurance, complementary insurance, and direct user charges) may have on provider and patient behaviour, which can influence who gains access to health care. They may also influence the size of the revenue base available for funding health care.

Phase IV - Agreement and questions for discussion

The fourth phase is the definition of set of questions each priority setting process should follow. As an example we would use set of questions that can be used even without having available economic data. However, economic methods can also be applied or the experience from elsewhere can be utilized. For instance using the methodology and documents available from the organizations specifically developed for setting priorities such as National Institute of Clinical Excellence from UK with appropriate financial modification of the value in certain programs can be successfully implemented in different settings.

The following set of questions would enable the group to utilize the available resources in order to respond to the concerns (4):

- What is the evidence that the suggested intervention or program is effective?
- How good is the evidence?
- What is the value of the effect?
- Is it worth paying the costs in relation to the value of the effect and ability to afford the investment?

These questions represent just one example who (4) the process can be guided. However, each particular group can define its own set of questions that all members of the group can agree to be relevant and acceptable.

Development and implementation of this approach to priority setting at the macro level will lead to improvement the efficiency in resource allocation, as well as more consistent priority setting at micro level by physicians in everyday clinical work with individual patients.

Phase V - Ethical issues and dilemmas

The fifth phase refers to the ethical dilemmas than may arise as a result of the decision making process in setting the priorities, as well as the consistency of similar decisions in various settings. Some authors suggest that this priority making decision always balances over two opposite ends.

The first position is called maximizing the minimum or whether the patients treated with certain therapy, treatment or intervention "X" are in worst starting position than patients treated with therapy "Y"? The supporters to this concept are named maximiners. The basic logic behind this concept is that priority should be given to those individuals or groups who are in the worst situation, in order to level up their health condition compared to the rest of the populations.

The second extreme position in setting priorities is which of the two interventions will result in biggest net health benefit for the population as a whole. The group of supporters who are in favour in this option are named maximizers. The logic behind this concept is based on the egalitarian principles where all members of the group are entitled to certain benefits, but the priority is given to the investment or therapy that will bring biggest net benefit for the group. Daniels suggest that the dilemma that is open to health policy makers can be summarized as follows: Should we give equal opportunities to those who may have even the smallest benefit from the intervention or investment, or should we use the resources to help those who have the biggest chances for success? In neither case

there are simple solutions, and most appropriate decisions would depend on the culture and traditions of the organization or country where such decisions are made.

Phase VI - Public accountability and consistency of decisions

The final phase of the priority setting cycle is the accountability to the public and elaboration of the decisions. In order to prevent any disputes and conflicts that may arise as a result of the decision making process, all members of the group should agree on the process that has led to the final decision. The group should document the case and make it easy available to the public and interested parties. Transparency should be the key component of the priority setting process that is fair. An effective communication strategy should be developed to ensure a transparent priority setting process. The purpose of the communication strategy should be to ensure that stakeholders know and understand the scope and necessity of priority setting decision-making, the degrees of freedom within which priority setting would take place (including explicit identification of any "sacred cows" that would be immune from priority setting), and the particularities of the priority setting process (who will do what, how the process will work, and why). In addition, the rationales for priority setting decisions should be communicated to stakeholders and should clearly demonstrate how these decisions are defensible in light of the priority setting criteria and available data/information.

Conclusions

The purpose of this module is not to promote particular concept, or to argue what is the best way of approaching priority setting in health care. Our intention was to summarize most frequently used and practically applicable approaches, and to present extract of ideas that would help health policy decision makers how to deal with priority issues in their settings where often there is lack of adequate human resource capacities, constant political pressures, frequent changes of the leadership as well as time constrains to set up priorities without devoting enough time and resources for its proper analysis.

Not all countries or health institutions within the countries are able always to train, recruit, or have high quality technical staff that would be able to apply relevant economic analysis and priority setting methods in their settings. We have pointed out that the most important aspects are the ability of the people within the organization, community or health authority to be engaged and participate in the process; and for the decision makers to assure that the process is transparent, consistent and open for critique and further improvements.

CASE STUDY: INTRODUCTION OF HPV IMMUNIZATION IN MACEDONIA

Introduction

The HPV vaccination episode in Macedonia illustrates real case scenario in political decision making surrounded with lot of controversies, mistrust and not clearly defined purpose. Although official justification of *introduction of HPV immunization in Macedonia* was for public health purposes, the way how to decision was made open the

room for many speculations on possible corruption and financial benefits for the importers of the vaccine:

The process

The main characteristics of the process of introduction of HPV vaccination in Macedonia was:

- the Government of the Republic of Macedonia in 2008 initiated discussion on possible introduction of HPV vaccination for girls in Macedonia;
- in order to secure funding for the project, the Government has asked the minister of health to talk to the representatives of the World Bank and to secure funding from the budget of the existing Health Sector Management Program financed by the loan from the World Bank (available from URL: www.moh-hsmp.gov.mk). The discussions were not successful. The Bank was not convinced that reallocation of the resources for the purpose of supporting purchasing of HPV vaccines from the already planned activities with the frame of the project would be more beneficial for the Macedonian population. In their written communication to the minister of health they have pointed out that the available evidence on the effectiveness of the HPV vaccine is not conclusive (35,36) and that there are many other priorities Government could do for the same funds. In addition, eventual acceptance of the Bank to reallocate the funds needed to follow serious administrative procedures and amendments in the Loan Agreement, procedure that may take up to six months to be completed;
- the outcome of these talks had no effect on the initiative of the government. The government made its final decision to introduce HPV vaccination for girls aged 9-15 years as voluntary program for year 2008 (15,37). The government secured 4 million denars (around 650,000 Euros) for purchasing 9,000 vaccines to vaccinate 3,000 girls in the specified age group. Whole process and the final decision was more ad-hoc intervention, than clear and transparent decision making process;
- in order to promote the new program the Government and the Ministry of health started huge promotion of the vaccination against HPV viruses in all printed and electronic media. This campaign created huge interest for the vaccination against HPV among the population, but also a lot of controversies and dilemmas among the professionals and media;
- only authorised public health institutes were eligible to distribute and apply the vaccination. The first problem that institutes experienced was the small number of the available vaccines, compared to the big interest among the population. Moreover, there were parents of older girls (aged above 15) who also requested vaccines to be available for their daughters;
- under the public pressure, the government extended the eligible age for voluntary vaccination for girls from initial 9-15 years, to 9-26 years of age. It was also officially announced that the HPV vaccination will be included in the mandatory calendar for immunization of the population as of year 2009;
- the vaccination process was organized on "first come first serve" basis. Those families and girls who were informed, interested and able to obtain the free HPV vaccines were vaccinated for the first dose. All others were registered in the

waiting lists for the vaccination as of next year. Again, whole process was not followed by proper priority setting cycle and media and general public characterized the action as populist and primarily delivered for political reasons;

- in April 2009, few months after official announcing of the latest decision for extending the eligible age among girls for HPV vaccination, the Government was faced with economic and financial crisis. The Ministry of health despite the initial promises and official statements was forced to cut the planned budget and to withdraw from already announced plans for introduction of mandatory and free HPV vaccination as of 2009. This embarrassing situation was widely communicated with all media (38). The biggest problem appeared for those families and girls who have already taken the first dose of the vaccination in 2008, and were waiting for the booster dozes. In addition, many families felt that they were left down by the Government since the process was not fair. Initial cohort of vaccinated girls was privileged, compared to the rest of the eligible female population for HPV vaccination. There was no consistency in the decisions of the Government,
- only after three days of official withdrawing, the Government again under the pressure of the public and interest groups announced that the HPV vaccination will proceed as initially planned (39). In order to vaccinate the eligible population of 15,000 girls, the Government needed to allocate 3.5 million Euros. Official justification was that agreement was made with the importers of the HPV vaccine to postpone the payment for the next year.

The temporary epilogue

There are still many people in Macedonia, both among the professionals and the wider public, who are not convinced how the whole process will proceed and how it would be financed.

EXERCISE

A scenario

Health Authority has limited resources to finance new public health intervention in the community. There are two options for interventions. The first is complete renovation of the toilets in the primary schools. The second is purchasing modern ECG equipment and laboratory for the primary health care clinic. Both interventions have its advantages. If toilets are renovated it is expected that increased hygiene in the schools will reduce the morbidity rates from diarrhoeal diseases among the students by 45% within the year. On the other hand, if the equipment for better diagnosis of the cardiovascular diseases is purchased, the citizens who are at risk of cardiac diseases may be prevented of sudden heart attack what would reduce the mortality by 15% within a year.

The community has in total 20,000 citizens, out of which 5,000 are students in the primary schools. Over 30% of the citizens are aged over 50 years and are considered to be at risk of heart attack.

Imagine that your group is in charge of the distribution of the health care resources and priority setting in the local community (municipality, village). You are asked to

provide proposal and justify your decision to the Council of the local community who needs to make final decision.

Note: You are confronted with ethical dilemma. This imaginary scenario deals with two subgroups of one population and the key difference is their age. Some philosophers argue that it is wrong to treat years at different age with equal value, while others disagree. You may wish to propose solution to this problem (40-42).

Task 1

Students will break into 4 groups. Each group should have the role of Priority Forum and work on opposite interventions:

- propose to the Local community council what intervention according to your expertise and knowledge should have priority and why?
- provide justification for the methodology used;
- what were your ethical considerations;
- group presentations (outlined in a handout).

Task 2

Students will break into 4 groups (in sub-groups for parallel activity). Each group will:

- perform opportunity cost analysis answering the question What is the second best alternative in financing of the project/intervention/therapy?
- what is the available evidence? (students are encouraged to use Internet search and relevant literature over the class if available to support its claims).
- group presentations (outlined in a handout).

Task 3

Final task is a group discussion combined with one-on-one interviews. Students may wish to use the HPV case study as basis for their discussions. Other examples are welcome. Questions guiding the group discussion and one-on-one interviews:

1. What is your perception of what priority setting entails in the Health Care System (HCS)?
2. What is your overall reflection on the HCS's current priority setting practices and those in the past?
3. What have been the major driving forces behind priority setting in introduction of HPV vaccine? What do you think the major driving forces should be?
4. Is it clear what are the values and guiding principles of the priority setting process explained in the case study? To what extent are these values considered when setting priorities?
5. To what extent does the HCS have an organizational culture conducive to using evidence in priority setting activities?
6. What capacities currently exist within the HCS to build a macro-level priority setting model? (E.g., organizational structure, information sources, links to the university, etc.)
7. How could priority setting practices/processes in the HCS be improved?

8. What is your vision for priority setting models and practices in the future?
9. What further information or training do you think the organization needs to get to your ultimate priority setting model?

REFERENCES

1. Holm S. Goodbye to the simple solutions: the second phase of priority setting in health care. Br Med J. 1998;317:1000-2.
2. Ham C, Pickard S. Tragic choices in health care: the case of Child B. London: King's Fund; 1998.
3. Ham C, Melver S. Contested decisions: priority setting in the NHS. London: King's Fund; 2000.
4. Hope T, Reynolds J, Gruffiths S. Integrating cost effectiveness with other values. In: Rodhes R, Battin MP, Silvers A, editors. Medicine and Social Justice. Oxford: Oxford University Press; 2002. p.144-55.
5. Ham C. Priorities setting in health care: learning from international experience. Health Policy. 1997;42:49-66.
6. Durie M. A framework for priority-setting in health care. Wellington: Massey University; 2004. Available from URL: http://temata.massey.ac.nz/massey/fms//Te%20Mata%20O%20Te%20Tau/Publications%20-%20Mason/M%20Durie%20A%20Framework%20for%20Priority%20setting%20in%20Health%20care.pdf Accessed: Apr 15, 2009.
7. Omar F, Tinghog G, Tinghog P, Carlsson P. Attitudes towards priority-setting and rationing in healthcare - an exploratory survey of Swedish medical students. Scand J Public Health.2009;37:122-130.
8. Canadian Priority Setting Research Network. Public engagement in setting priorities in health care [homepage on the Internet]. Available from URL: http://www.canadianprioritysetting.ca/html/documents/recent_highlights/09-public_engagement_setting_priorities.pdf. Accessed: August 21, 2009.
9. Menon D, Stafinski T, Martin D. Priority-setting for healthcare: Who, how, and is it fair? Health Policy. 2007;84:220-233.
10. Klein, R. Dimensions of rationing: Who should do what? Br Med J. 1993;307:309-311.
11. Beroe K. Priority setting in health care: on the relation between reasonable choices on the micro-level and the macro-level. Theor Med Bioeth. 2008;29:87-102.
12. Kapiriri L, Norheim OF. Criteria for priority-setting in health care in Uganda: exploration of stakeholders' values. Bulletin of the World Health Organization. 2004;82:172-179.
13. Hauck K, Smith PC, Goddard M. The economics of priority setting for health care: A literature review. Washington, DC: The International Bank for Reconstruction and Development/The World Bank; 2004. Available from URL: http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/Chapter3Final.pdf. Accessed: August 18, 2009.
14. Williams A. Priority setting in public and private health care systems: a guide through the ideological jungle. J Health Econom. 1988;7:173-83.
15. Government of Republic of Macedonia. Program for mandatory immunization for 2008 (in Macedonian). Official Gazette. 2008:56.
16. Mitton C, Patten S, Donaldson C, Waldner H. Priority-setting in health authorities: Moving beyond the barriers the Calgary Experience. Healthcare Quarterly. 2005;8:49-55. Available from URL: http://www.longwoods.com/product.php?productid=17154 Accessed: May 17, 2009.
17. Klein R, Day P, Redmayne S. Managing scarcity-priority setting and rationing in the National Health Service. Buckingham: Open University Press; 1996.
18. Gibson JL, MartinDK, Singer PA. Setting priorities in health care organizations: criteria, processes, and parameters of success. BMC Health Serv Res. 2004;4:1-8. Available from

URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=518972&blobtype=pdf. Accessed: August 24, 2009.
19. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc. 2006;4:1-9. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1560167&blobtype=pdf. Accessed: August 24, 2009.
20. Smith N, Mitton C, Peacock S. Qualitative methodologies in health-care priority setting research. Available from URL: http://www3.interscience.wiley.com/journal/121452189/abstract?CRETRY=1&SRETRY=0 Accessed: May 17, 2009.
21. Daniels N, Sabin J. The ethics of accountability in managed care reform. Health Affairs. 1998;17:50-64.
22. Daniels N. Accountability for reasonableness. Br Med J. 2000;321:1301-2.
23. Daniels N, Sabin J. Accountability for reasonableness: an update. Br Med J. 2008;337:a1850.
24. Jain R. Dilemas in setting priorities. Available from URL. http://www.expresshealthcaremgmt.com/20040731/management01.shtml Accessed: May 16, 2009.
25. Daniels N. Just Health: Meeting health needs fairly. New York: Cambridge University Press; 2008.
26. Gruskin S, Daniles N. Justice and human rights: priority setting and fair deliberative process. Am J Public Health. 2008;98:1573-7.
27. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight Countries**.** Int J Equity Health 2008, **7:4**:1-13. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=2248188&blobtype=pdf. Accessed: August 12, 2009.
28. Santric-Milicevic M. Setting community health priorities. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company, 2005:208-26. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
29. The World Bank. The World Bank development report 1993. Investing in Health. Available from URL: http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&siteName=WDS&entityID=000009265_3970716142319. Accessed: August 14, 2009.
30. Bobadilla JL, Cowley P, Musgrove P, Saxenian H. Design, content and financing of an essential national package of health services. Bullet WHO. 1994;72:653-62. Available from URL: http://whqlibdoc.who.int/bulletin/1994/Vol72-No4/bulletin_1994_72(4)_653-662.pdf. Accessed: August 24, 2009.
31. Bognar, G. Age-weighting. Econom Phil. 2008;24:167-89.
32. Mitton C, Donaldson C. Health care priority setting: principles, practice and challenges. Cost Eff Resour Alloc. 2004;2:1-8. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=411060&blobtype=pdf. Accessed: August 24, 2009.
33. Shrimpton B, McKie J, Hurworth R, Bell C, Richardson J. A focus group study of health care priority setting at the individual patient, program and health system levels. Clayton, Victoria: Monash University, Faculty of Business and Economics, Centre for Health Economics; 2007. Available from URL: http://www.buseco.monash.edu.au/centres/che/pubs/rp16.pdf. Accessed: Apr 15, 2009.
34. Dicker A, Armstrong D. Patients' views of priority setting in health care: an interview survey in one practice. Br Med J. 1995;311:1137-9. Available from URL: http://www.bmj.com/cgi/content/full/311/7013/1137 Accessed: May 17, 2009.

35. Haug CJ. Human Papilomavirus Vaccination-reasons for caution. N Eng J Med. 2008;359:861-2.
36. Kim JJ, Goldie SJ. Health and economic implication of the HPV vaccination in the United States. N Eng J Med. 2008;359:821-32.
37. The government of the Republic of Macedonia. Start of the vaccination against HPV viruses (in Mecedonian) [homepage on the Internet]. Available from URL: http://www.vlada.mk/?q=node/1201. Accessed: August 21, 2009.
38. Stojanovska A. There is no money for HPV immunization (in Macedonian). Vest (daily newspaper). April 22, 2009;2654.
39. Balaban L. HPV vaccination will proceed (in Mecedonian). Dnevnik (daily newspaper), April 24, 2009;3961.
40. Hope T, Hicks N, Reynolds DJM, Crisp R, Griffiths S. Rationing and the Health Authority. Br Med J. 1998;317:1067-69. Available from URL: http://www.bmj.com/cgi/reprint/317/7165/1067. Accessed: August 24, 2009.
41. Daniels N. Just health care. Studies in philosophy and health policy. Cambridge: Cambridge University Press; 1985.
42. Lockwood, M. Quality of life and resource allocation. In: Bell JM, Mendus S, editors. Philosophy and medical welfare. Cambridge: Cambridge University Press; 1988.

RECOMMENDED READINGS

1. Daniels N. Accountability for reasonableness. Br Med J. 2000;321:1301-2.
2. Daniels N. Just health: Meeting health needs fairly. New York: Cambridge University Press; 2008.
3. Hauck K, Smith PC, Goddard M. The economics of priority setting for health care: A literature review. Washington, DC: The International Bank for Reconstruction and Development/The World Bank; 2004. Available from URL: http://siteresources.worldbank.org/HEALTHNUTRITIONANDPOPULATION/Resources/281627-1095698140167/Chapter3Final.pdf. Accessed: August 18, 2009.
4. Kapiriri L, Norheim OF. Criteria for priority-setting in health care in Uganda: exploration of stakeholders' values. Bulletin of the World Health Organization. 2004;82:172-179.
5. Mitton C, Patten S, Donaldson C, Waldner H. Priority-setting in health authorities: Moving beyond the barriers the Calgary Experience. Healthcare Quarterly. 2005;8:49-55. Available from URL: http://www.longwoods.com/product.php?productid=17154 Accessed: May 17, 2009.
6. Rodhes R, Battin MP, Silvers A, editors. Medicine and Social Justice. Oxford: Oxford University Press; 2002.
7. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight Countries. Int J Equity Health 2008, **7:**4:1-13. Available from URL: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=2248188&blobtype=pdf. Accessed: August 12, 2009.
8. Santric-Milicevic M. Setting community health priorities. In: Galan A, Scintee G, editors. Public health strategies. Stability Pact for South Eastern Europe public health cooperation. Lage: Hans Jacobs Publishing Company, 2005:208-26. Available from URL: http://www.snz.hr/ph-see/publications.htm. Accessed: August 14, 2009.
9. The World Bank. The World Bank development report 1993. Investing in Health. Available from URL: http://www-wds.worldbank.org/external/default/main?pagePK=64193027&piPK=64187937&theSitePK=523679&menuPK=64187510&searchMenuPK=64187283&siteName=WDS&entityID=000009265_3970716142319. Accessed: August 14, 2009.
10. Williams A. Priority setting in public and private health care systems: a guide through the ideological jungle. J Health Econom. 1988;7:173-83.

Chapter 5

SUPPORTIVE METHODS/ TOOLS/TECHNOLOGIES

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**DATABASES AND THEIR ORGANIZATION**
Module: 5.1.1	**ECTS (suggested): 0.5**
Author(s), degrees, institution(s)	**Josipa Kern**, MD, PhD, Professor Andrija Štampar School of Public Health, Medical School, University of Zagreb, Croatia **Slavica Sović**, MD, PhD candidate Andrija Štampar School of Public Health, Medical School, University of Zagreb, Croatia
Address for correspondence	**Josipa Kern** Andrija Štampar School of Public Health, Rockefellerova 4, 10000 Zagreb, Croatia E-mail: josipa.kern@snz.hr
Keywords	Metadata – data description; Database, Data communication
Learning objectives	After completing this module students should: • understand concept of data organisation • be aware of importance of using information and communication technology (ICT) in medical/public health practice and research • be familiar with preparing and collecting data relevant for health care practice and research • understand the basic principle of databases and be able to create a simple database and use it
Abstract	Medical/health data can be of very complex structure because of a number of entities and corresponding attributes, usually changing in time. Database is a basic concept of data organization enabling complex and flexible relationships between data. Database can ensure quality of data by checking them at input. In using, database enables to combine different entities and corresponding attributes as well as their values in order to make very complex analysis. In this module: • the basic terms in data organization are described (entity, attribute, attribute value, data, information), • methods of data description (metadata or codex of attributes) are explained, • relationship between data are compared, and • an easy-to-use database management software is presented.
Teaching methods	Teaching methods will include lectures, and individual practical work. An introductory lecture gives the basic theoretical knowledge on data organization and information production. After the introductory lecture students will work individually in computer laboratory on solving problems of medical practice.
Specific recommendations for teachers	• work in small groups; • work under teacher supervision/individual students' work proportion 50%/50%; • facilities: a computer laboratory; • equipment: PCs (1 PC per student); • software: MS Office (Access predominantly) • target audience: graduate and postgraduate students according to Bologna scheme.
Assessment of students	Assessment is based on multiple choice questionnaires and a case study.

DATABASES AND THEIR ORGANIZATION

Josipa Kern, Slavica Sović

THEORETICAL BACKGROUND

Basic terms (entity, attribute, attribute value)

Reality includes objects - real and abstract, living and inanimate. The objects are described by variety of characteristics.

Objects can be:

1. Living beings (people, animals, plants)
2. Things (drugs, medical facilities, medical equipment)
3. Phenomena (disease, surgery, intervention, service)

Characteristics of objects are:

1. Identification number, name, age, etc. (for person)
2. Height, weight, blood pressure, health status, blood count, lung function, etc. (for patients)
3. Type, form, name, dose (for medication)
4. Name, last name, specialty (for health workers)
5. Type, name, number of beds (for hospitals)
6. Type, scope, who has implemented it (for intervention), etc.

Characteristics can be perceived and measured: described by text, measured by using specific instrument, estimated at a specific or own scale.

Results of observing, estimating and measuring are:

1. Descriptions - free text (history, discharge summary, notes, remarks, and suggestions), etc.
2. Categories, classes or grades – names or codes (for gender: male and female, or M and F; for marital status: married, unmarried, indented, and widow or 1, 2, 3 and 4)
3. Numbers (age: 24 year old; number of erythrocytes: 3.5 million; blood pressure: 155/80 mm Hg; body temperature: 38.2^{o}C, etc.)

Such objects are usually called *entities*, their characteristics are *attributes*, and the results of observations and measurements are *attribute values* (or *data*). Schematic, relations among the above-mentioned terms are shown in Figure 1:

Different kind of entities can be mutually connected. For example, a person has two eyes, two hands, two kidneys, more teeth. Or, a family can have few children. These relationships can be shown as “one to one”, or 1:1, “one to more”, or 1: n, “more to more”, or n: m. Figure 2 shows that a father has several children, and each child has only one father. Therefore, the relationship should be 1: n. Figure 3 shows that patient is treated by several doctors and each doctor treats several patients. Therefore, the relationship should be n: m. Figure 4 shows a simple relationship between objects in which, one object of one type belongs to only one object of another type. Therefore, the relationship should be 1: 1.

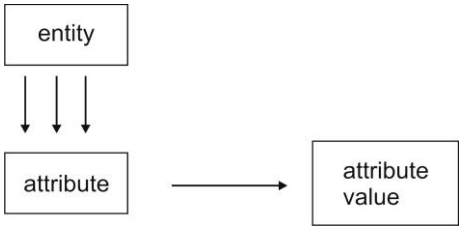


Figure 1. Relationship between entities, attributes, and attribute values.

Figure 2. Entities: *father* and *child;* Relationship*: 1: n*

Figure 3. Entities: *patient* and *doctor*; Relationship: *n:m*

Figure 4. Entities: *maxilla* and *mandible*; Relationship: *1:1*.

There are two types of attribute values: *qualitative* and *quantitative.* Quantitative attribute values are expressed by number, most often in certain units, or as results of counting. Examples of quantitative attribute values are: 180 (mm Hg for *systolic blood pressure*) or 5 (children for attribute *number of children in the family*). Qualitative attribute values are expressed by words (text), categories, usually predetermined. Examples of the qualitative attribute values are: male (for the attribute *gender*), Janko (for the attribute *name*), immovable (for the attribute *mobility of patients*). Usually, for qualitative attribute value we use *codes*, one or more characters that replace the value of attribute. For example, a character / letter M may mean male, and F female gender. M and F are codes.

Attribute enabling to identify an entity is called the *primary key*. It is apparent that the last name, or last name combined with first name, or even with the name of the father cannot be the primary key. There is more than one person with the same value of mentioned attribute/attributes: more than one person with same name and same name of her/his father. Therefore in practice we use numbers being able to

identify a person. In Croatia, so called PIN (personal identification number) is consisted of thirteen digits.

Attributes, not being primary keys should be called *secondary keys*. Secondary key does not identify a single entity. It identifies a group of entities that have the same value of particular attribute.

Data and Information

Data are facts telling something about an entity. As such, attribute values are data. The better knowledge of the entities can be got by data processing, by getting *information* (Examples 1 and 2).

Example 1.

*Height and weight of a patient were measured (height=1.9 m, weight=130 kg). Willing to tell something of the adiposity of the patient, the body mass index (BMI) should be calculated: BMI= weight/(height*height). For this specific case BMI=36. Considering the common limits: normal (BMI <30) and adipose (BMI =>30) persons, this patient with BMI= 36, is adipose. In this case* ***BMI is information****.*

Example 2.

To hypertensive patient, with prescribed antihypertensive therapy, blood pressure was measured in several occasions: 190/110, 170/100, 150/100, 170/95, 165/100 (mm Hg). Monitoring this set of data (and taking into account the medical knowledge about normal /abnormal blood pressure), we get information that the ***blood pressure is not normalised*** *because: therapy is not adequate, or patient does not take the medicine according to physician's instructions. In this case* ***blood pressure is not normalized is information****.*

Figure 5 is indicating link between the concepts of data and information. So, the information could be considered as raw material that should be processed to achieve finished product- information. Information is needed for making decisions. Information enables evidence based decision making (Examples 3 and 4).

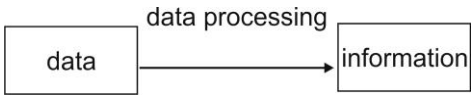


Figure 5. Concept of "data-information".

Example 3.

Obesity of patient is basis for making decision about diet, which can be the first step in solving health problems (e.g. for patients with high blood glucose level).

Example 4. *Continuously increased blood pressure is the basis for checking if the patient actually has followed the instructions. If she/he has followed it, then therapy should be modified (medicine or dose), and if she/he has not followed it, it should be found out why it is so (perhaps because of side effects), and therefore make the appropriate decision about changing therapy or informing patients about the potential consequences of his/her acts.*

Some other examples of information are:

1. Elevated blood pressure - obtained by organizing data on blood pressure in the categories "'above normal", or "normal" with respect to predetermined values (e.g. 140/90).
2. Average age of population group - calculated from the individual ages.
3. Diagnosis - obtained by summarising symptoms, examination results, laboratory findings, etc.

Data and information are not fixed categories. Data (in one situation), may be information in the other. For example, diagnosis is an information (in the case of one patient), but may also be data (in the case of group of patients) from which the prevalence of this diagnosis (diseases) can be estimated.

Formally, each data processing results in obtaining information. But it should be noted that information becomes information in the true sense of this word only if it encourages a decision, an action, and an intervention. Otherwise, such information is not required; it overloads the system, and remains a "dead letter on paper".

Hierarchical Levels in Data Organization

Usually, data should be written in a digital form: a character or sequence of characters. Various sets of characters could be used: decimal numbers (0, 1,...,9), binary (0, 1) or the numbers from another system (octal, hexadecimal, etc.). Sometimes letters of an alphabet could be used. Which letters are included in a set of characters depends on language. In free texts form, special characters such as dot, comma, etc. can not be avoided. In some situations characters like + (plus), - (minus) or similar should be in use.

In reality, the most frequently used set of characters is the standard set of characters. It is consist of decimal numbers, alphabet and special characters. Letters are usually considered as those of English alphabet, whereas special characters (š, č, ć, đ, ž, Š, Č, Ć, Đ, Ž) required for Croatian and similar languages are considered as special characters.

The standard set of characters can be found on keyboard, an input device.

Some types of data are written in analogue form. Examples of analogue data are classical X-ray images, ECG curves printed on paper, etc.

All data, digital or analogue, should be integral parts of patient record. They must be in some way connected to each other and available for discussion of information. Collected at a medical department, outpatient clinic or in a field research, (for treatment, prevention or management) data are mostly mixed, digital and analogue, although the most prevalent type of data is digital one. Here will be discussed organization of digital data.

Recording of the data requires one or more characters. For example, age of person (year) requires up to three characters. If person is 45 (years), two characters are required (4 and 5), but if he/she is 103 (years) three characters are required (1, 0, and 3). Number of characters (for a specific attribute and its values) defines the *field.*

Considering an entity with several attributes, several fields are needed. All of these fields together make a *record* of the entity (in medicine and health care: medical record, health record, personal health record etc.)

All entities of the same type, e.g. all patients (and their corresponding records) make a *file*.

Different types of entities (and their corresponding records) related in some way (by relationships: 1:1, 1:n, or n:m) make a *database*.

The database is the highest level in the organization of data. It can be described in different ways. For example, it can be said that the database is set of data which are connected to each other with controlled redundancy; data in the database should serve for different purposes.

In the hierarchical series: "Character- Field- Record- File- Database", database is the only level in the organization of data that can not be realized without modern information technology, without computer technology.

CASE STUDY

Introduction

Designing the electronic record

Electronic records as parts of database should be structured in some way. It means that:

1. Attribute names should be clearly denoted
2. Attributes are selected according to purpose
3. It should be clear what could be attribute values for each of particular attributes

Record can have an unstructured part. Sometimes, it will be necessary to have image or description in the form of free text (e.g. opinion of specialists) in the record. If the "opinion of specialists" as the attribute in the record has its specific place, it can be said that the record is still structured, although the processing of data recorded in the form of free text could potentially create a problem. Production of information from such data will not be easy, for example, it will not be possible to apply traditional statistical approach.

Structured records have specific format for recording data and they are often called the formatted records. Formatted record is described by the Codex of attributes.

Codex of attributes

The Codex of attribute (so called metadata) is a list and description of attributes for certain entity. It can be summarized in a table with information as follows:

1. ID number of attribute
2. Name of attribute
3. Units of measurement
4. Name of classification or list of values of attributes and associated codes

5. Number of positions (field size)
6. The number of digits

Figure 6 is presenting an example of Codex of Attribute

Number of the Attribute	Name and Description of the Attribute	Field Size	Number of the Decimal Places
1	Number of the Health Card	4	0
2	Personal Identification Number	13	0
3	Name and Surname	40	0
4	Gender F - Female M - Male	1	0
5	Height(cm)	3	0
6	Weight (kg)	4	1
7	Smoker 1. Yes 2. No 3. Former	1	0
8	Date of Visit	8	0

Figure 6. Example of Codex of Attribute.

Codex of attributes – principles

Principles that should be followed for creation of Codex of attributes:

1. Choose appropriate way for identification of entity, attribute for identification
2. Declare are attributes/data qualitative or quantitative,
3. For quantitative attributes/data choose appropriate units of measurement,
4. For qualitative attributes/data select/define the classification and coding system,
5. Identify groups of attributes and describe each particular attribute,
6. Take care of redundancy (it should be under control),
7. Assess the relevance of attributes (considering the objectives and purpose).

Attribute for identification can serve as the primary key. Choice of measurement unit provides clearness during data collection. Classifications that are defined must be:

1. *Complete* (all possible values must be specified in the classification)
2. *Exclusive* (the same entity can not have more than one value)
3. *Adequate* (classification must match the needs of the profession, it must provide the information enabling to achieve the desired goal)

Database Design

Scheme of database

Creating a database starts with data modelling. Data modelling is a process which starts with defining and analyzing of user requirements, and ends with construction of stable and flexible database.

First step should be to define entities, attributes and their values, relationships among the entities (1:1, 1: n, n: m) as well as attribute that identify the entity (primary key). Next step is the development of conceptual data model.

Conceptual model should reflect the real system, shown in form of diagrams. Well-designed conceptual data model tend to avoid redundancy.

Logical modelling starts from the conceptual data model and user's requirements. Logical models contain records, fields, primary key and relationships.

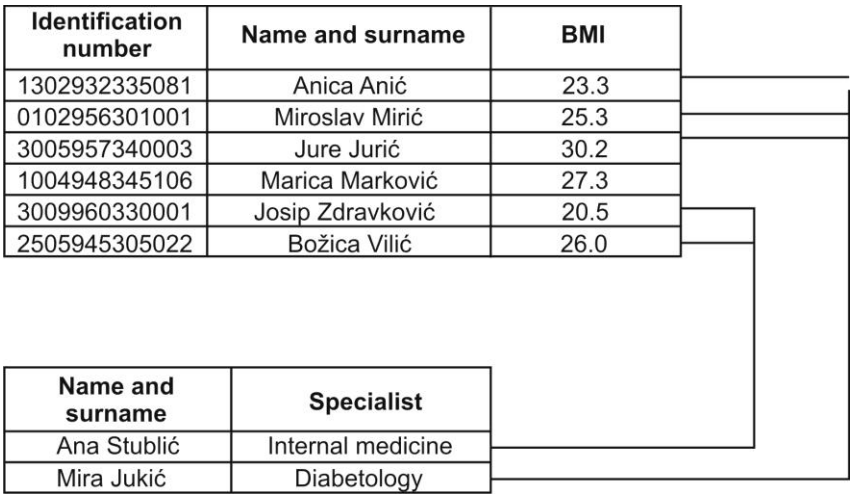
Next step is physical modelling, the realization of logical models of databases on the computer using Data Base Management System, DBMS. An example of conceptual model of database is presented in Figure 7. In this example there are two entities (doctor and patient), the sign "*" denotes that primary key is in the patient's records. List and description of attributes should be added to the scheme.

Figure 7. Example of conceptual model of database.

Today, there is a series of software products that belong to the group of DBMS. They have been built and applied for different platforms, at different levels of computers, on different operating systems. Some of them are designed for personal computers, some for workstations, and some for large computers. Some of them are designed for the DOS operating system, some for Windows, some for UNIX and its version, some for MVS, operating system OS2, and so on. Although there are differences between DBMS, it is enough to become familiar with one of these systems to learn basics of database design and use.

Loading of Database

Loading database includes entering data about entities in the predesigned records. Person who collects data should enter them into database (Figure 8).



Identification number	Name and surname	BMI
1302932335081	Anica Anić	23.3
0102956301001	Miroslav Mirić	25.3
3005957340003	Jure Jurić	30.2
1004948345106	Marica Marković	27.3
3009960330001	Josip Zdravković	20.5
2505945305022	Božica Vilić	26.0

Name and surname	Specialist
Ana Stublić	Internal medicine
Mira Jukić	Diabetology

Figure 8. Result of the implementation of data model, in which there are two types of entities with attribute identification number, name and surname, BMI for entity PATIENT, and name and specialist for entity DOCTOR

Use of database

Simple examples of information that can be obtained from the data entered in the database are:

1. The number of entities with a particular value of attributes (e.g. number of persons with high BMI).
2. The average value (e.g. systolic/ diastolic pressure).
3. A list of selected data for the individual (e.g. name and surname and BMI of the person with identification number 130292335081).
4. Range of values of certain attributes (e.g. height of people younger than 7 years).

The information are obtained by setting question using certain system for working with databases, applications for computing, or statistical software product that accepts data exported from the database (created by DBMS).

Access – an example of database management system

Microsoft Access is an easy-to-use database management system based on Windows operating system, part of Microsoft Office, and applicable on PCs. Access enables creating database, making forms for loading the database, extracting some specific data, and using for reporting. Access provides possibility

of exporting data in a format acceptable to the majority of statistical and other software products.

Through the series of images we will show the use of a software product Microsoft Access - the system for the design of the database and work with the database (DBMS). Steps can be summarized as:

1. Choosing of a folder for storage of the database
2. Giving a name to database that will be a series of tables (relations, types of records)
3. Creation of tables (relations, types of records) for the selected entities
4. Creation of a links between tables (relations, types of entities)
5. Development of a form (or few of them) for data entering into the table
6. Data entering for individual entities
7. Design of a table that contains extracted data from few original tables (selected attributes and/ or selected entities)
8. Creation of reports from the tables, original or derivated.

Choosing of a folder for storage of the database, and giving a name to database

In Figures 9 and 10 creation of the new database is shown.

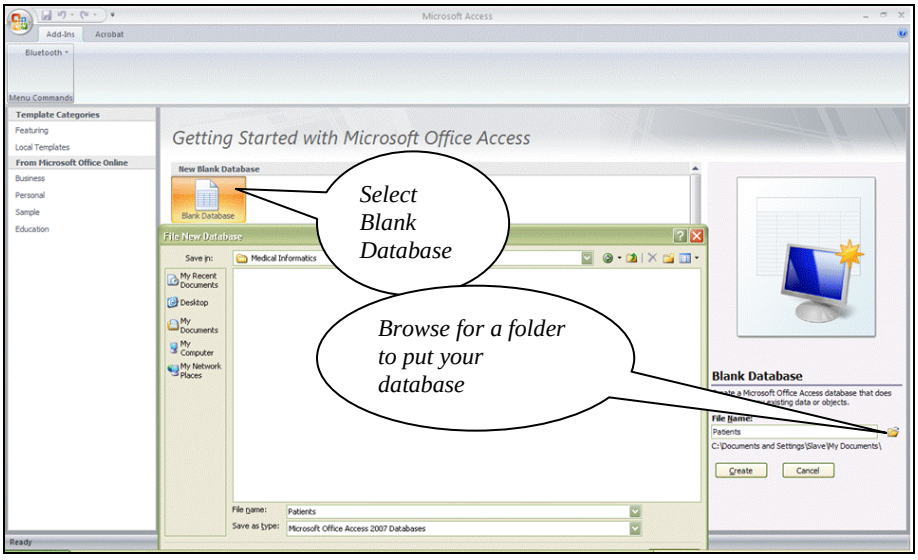


Figure 9. Creation of the new database "*Patients"* in the folder *"Medical Informatics."*

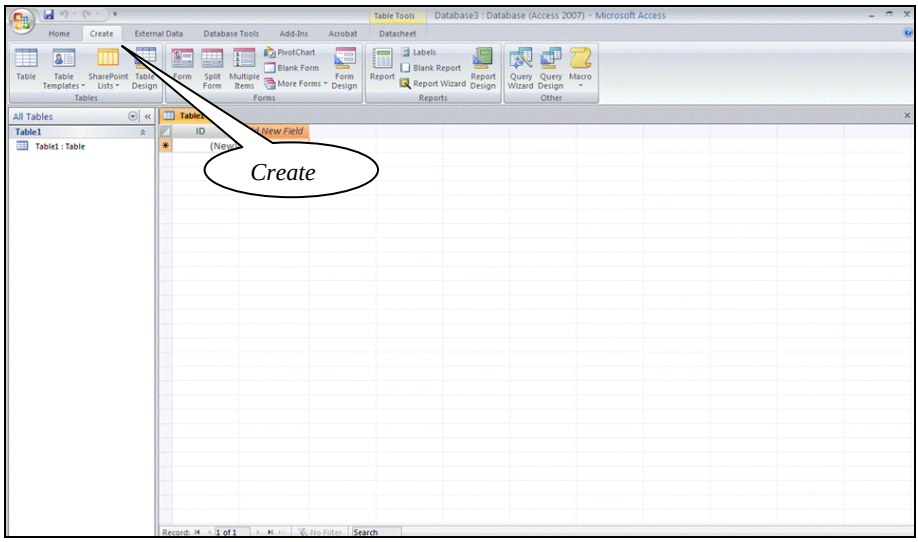


Figure 10. Card *Create* for definition of *Tables*, *Queries*, *Forms*, *Reports*, and definition of elementary table in database.

Creation of tables (relations, types of records) for the selected entities

In Figures 11 to 16 creation of tables for the selected entities is shown.

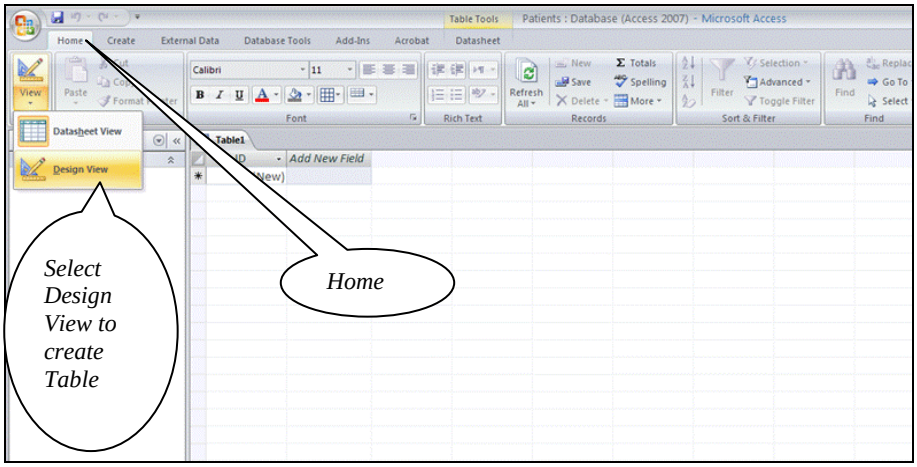


Figure 11. First step is definition (name of attribute, data type and description of attribute) of a *Table.* Go on *Home Card and s*elect *Design View*. Save table as "*Patient."* .

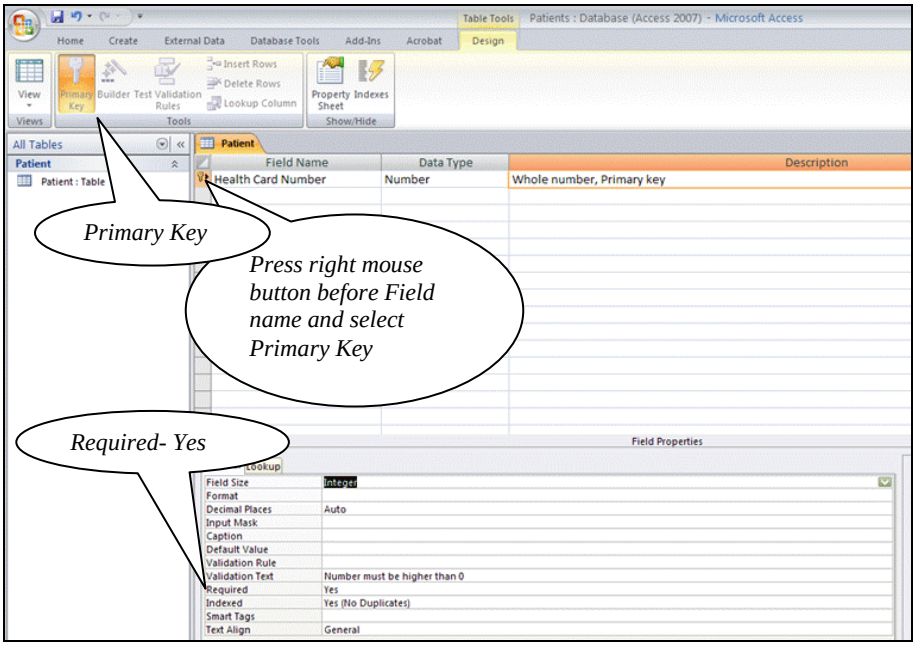


Figure 12. Attribute *Health Card Number* (Data type is number, Field Size Integer) will serve as *Primary Key*. For that reason this attribute is required.

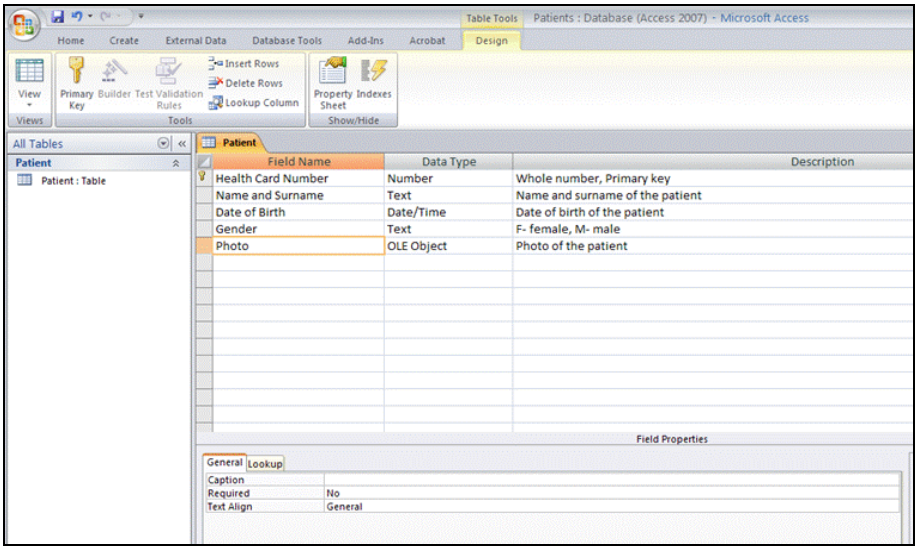


Figure 13. Attributes: *Name and Surname* (Data type is text, Size 30, Required: Yes), *Date of birth* (Data type is date/ time, Required: No), *Gender* (validity rule control recording data in database), *Photo* (Data type OLE Object).

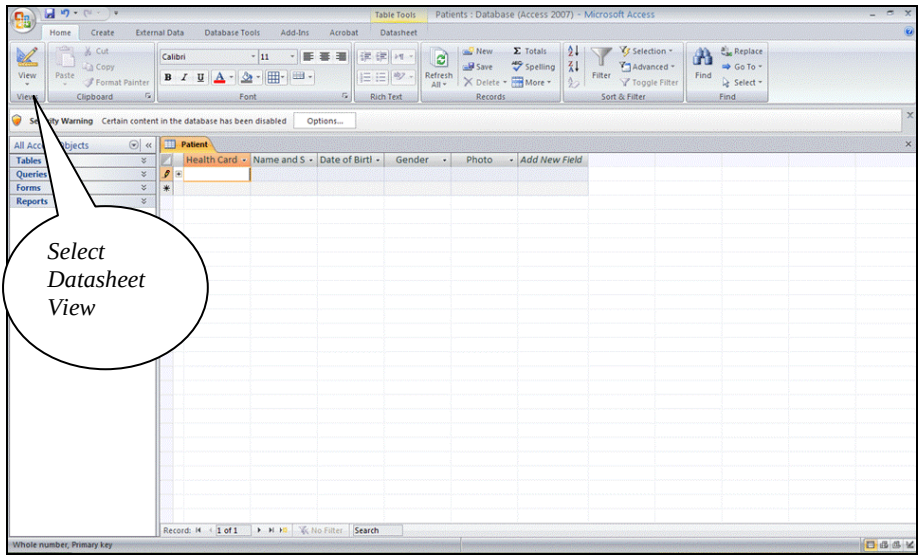


Figure 14. Result of the creation *Table* "*Patient"* in *Datasheet View*.

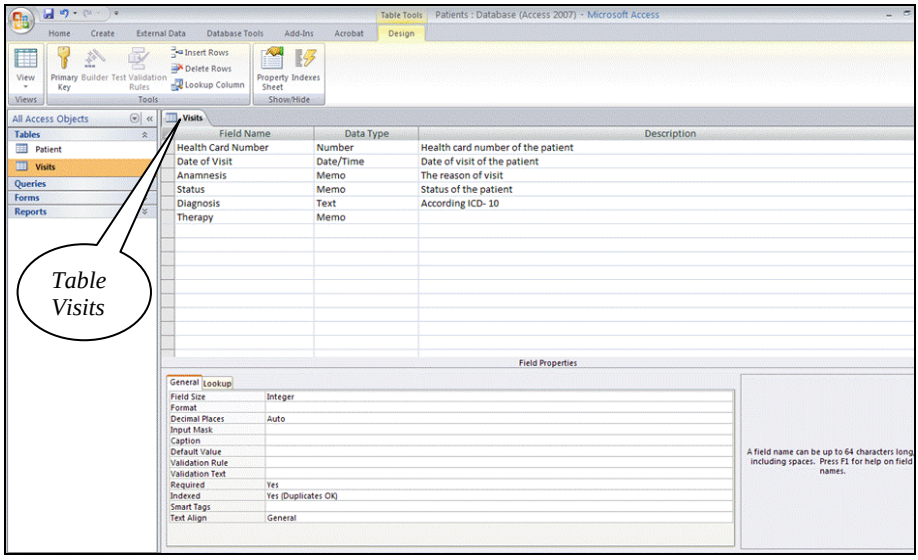


Figure 15. Creation of the new *Table* "*Visits"* in *Design View*. From *Creation* Card, select *Table* in *Design View*. Add attributes. Attribute *Health Card Number* serve as Primary Key and should be added to table "*Visits"* since it will serve in creation of *Relationship*.

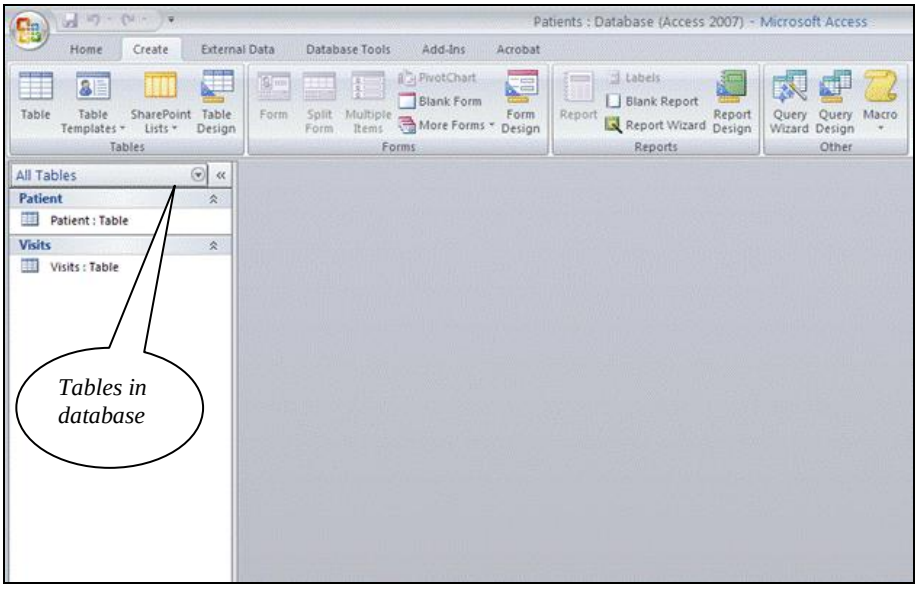


Figure 16. Database "*Patients"* contains two *Tables* ("*Patient*" and "*Visits"*).

Creation of the links between tables

In Figures 17 to 20 creation of the links between tables is shown.

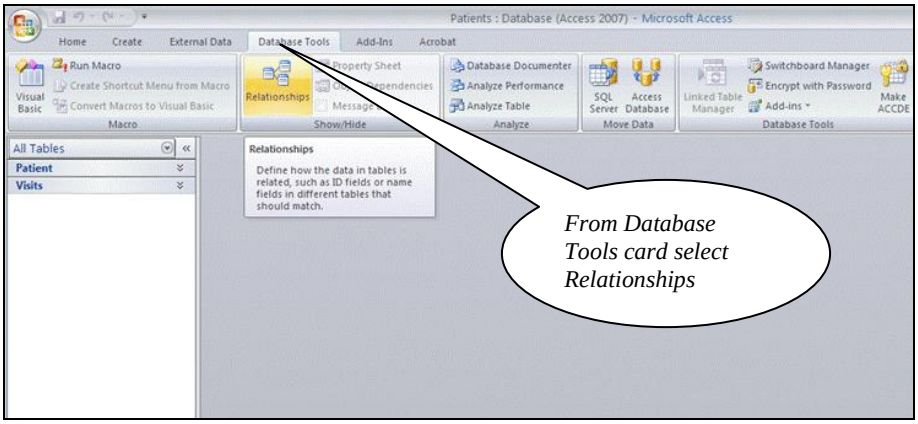


Figure 17. Next step is creation of *Relationship* between tables.

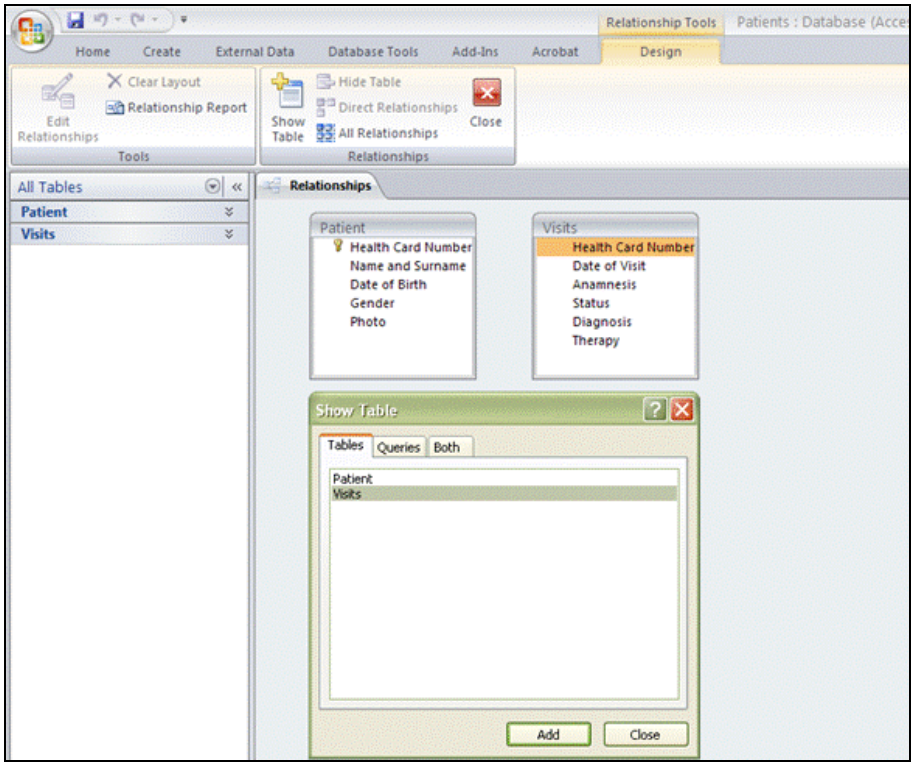


Figure 18. Selection of the *Tables ("Patient" and "Visits")* for the *Relationship.*

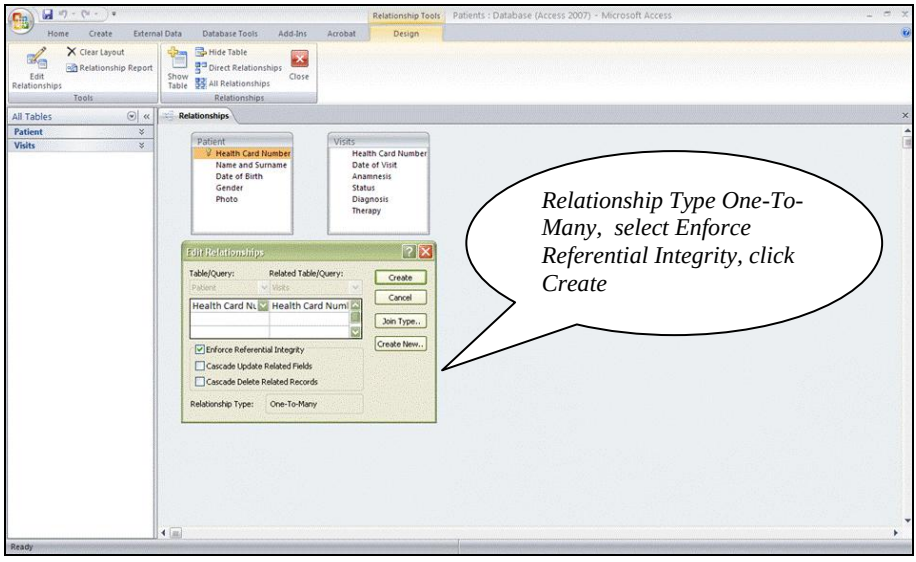


Figure 19. Definition of the Relationship type

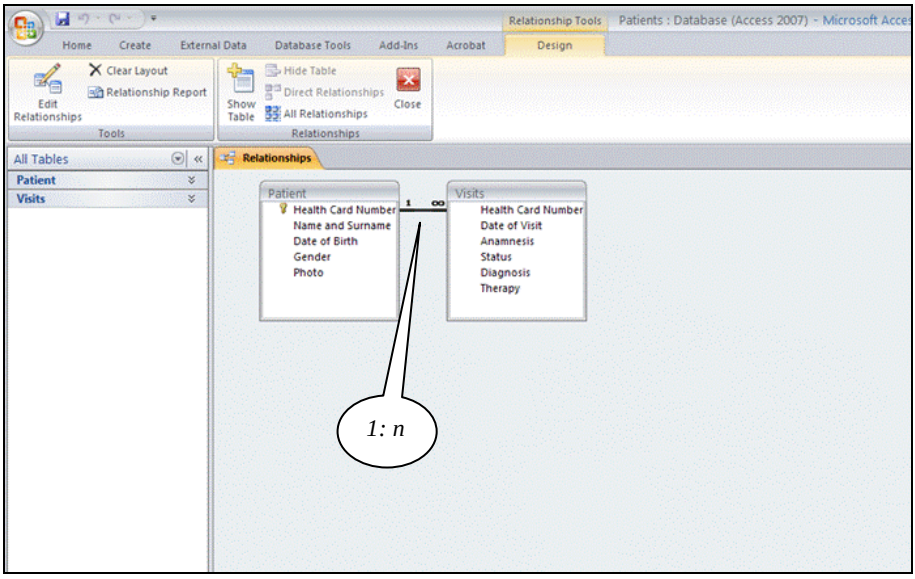


Figure 20. Result of the *Relationship* creation.

Development of a form (or a few of them) for data entering into the table

In Figures 21 to 28 development of a form for data entering into the table is shown.

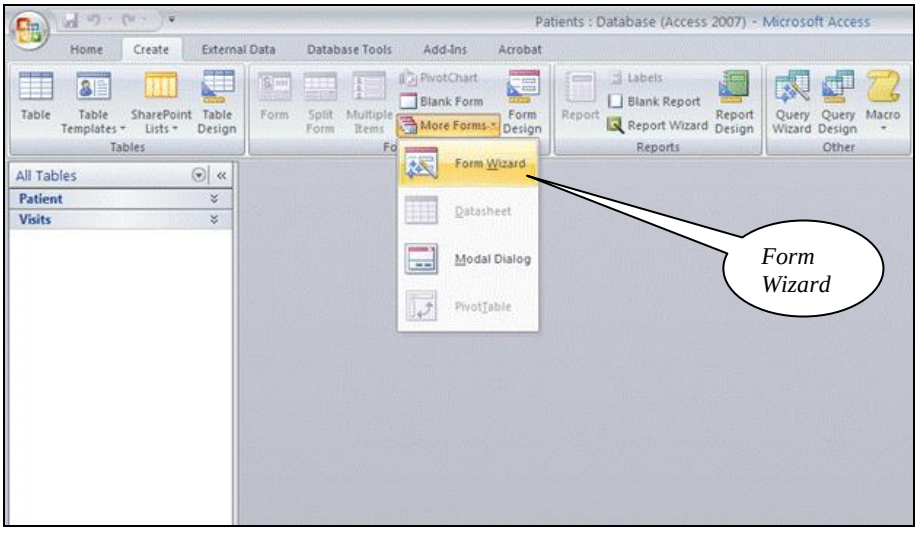


Figure 21. Next step is creation of a *Form* for data entering using *Form Wizard.*

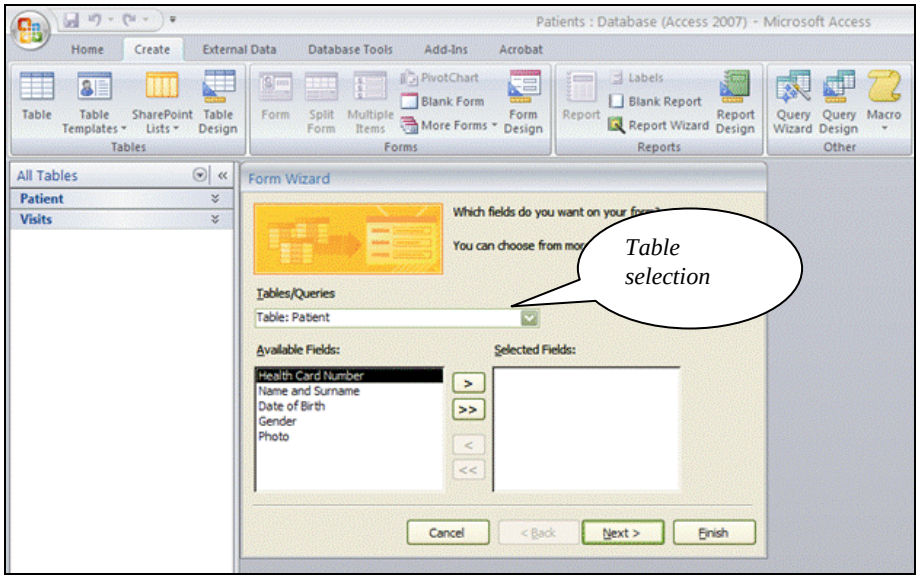


Figure 22. Creation of the *Form* for table *Patient.*

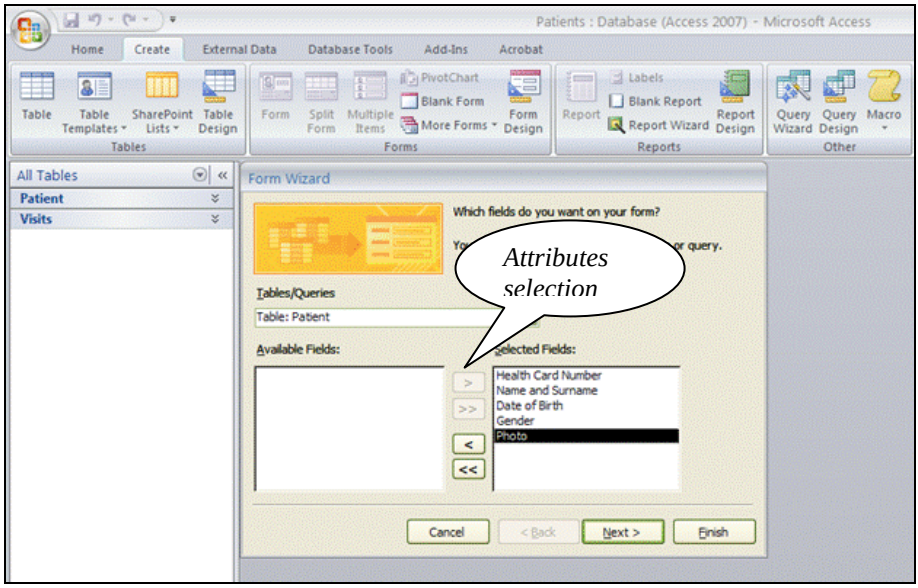


Figure 23. Selection of attributes for the *Form* from table *Patient.*

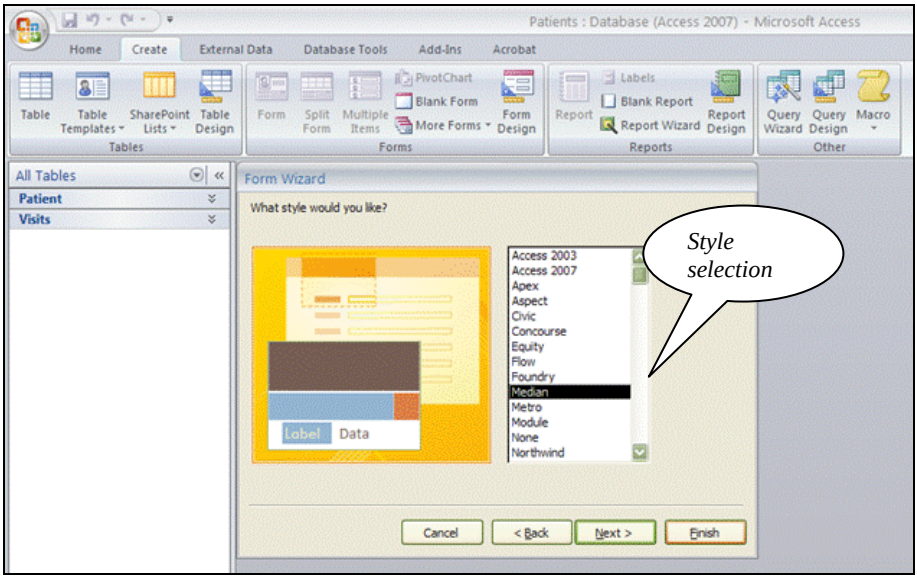


Figure 24. Next step in the *Form* creation: "What style would you like?"

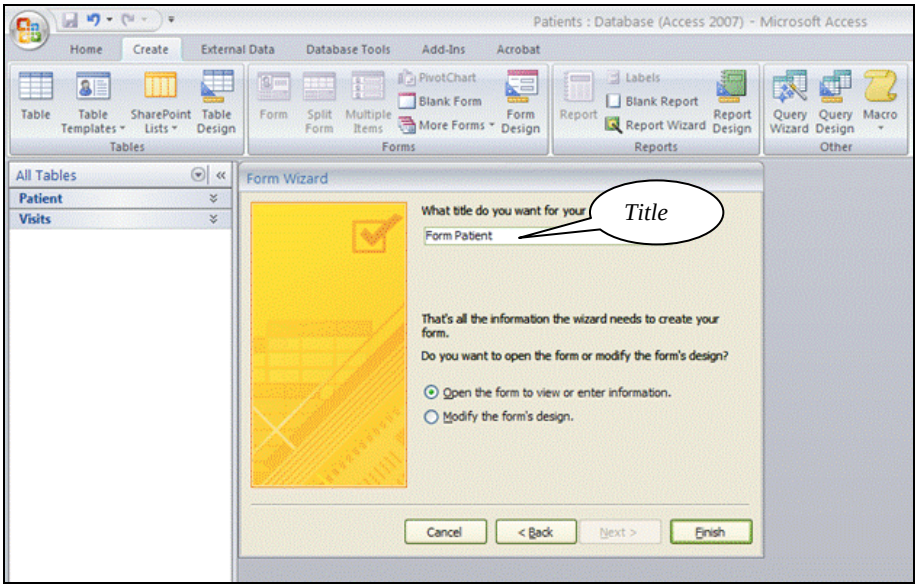


Figure 25. Next step in the *Form* creation: "What title do you want for your form?"

In the same way create a *Form* for table *Visits*.

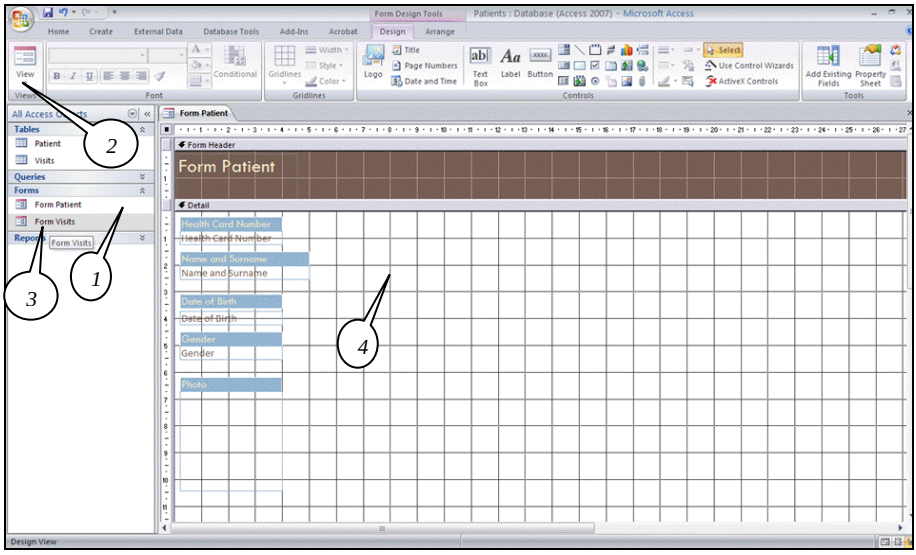


Figure 26. Linking of the *Forms:* insert the *Form Visit* into the *Form Patient*. Open *Form Patient* (1) in *Design View* (2), drag *Form Visits* (3) with left mouse button, and drop it into the *Form Patient* (4); Save changes.

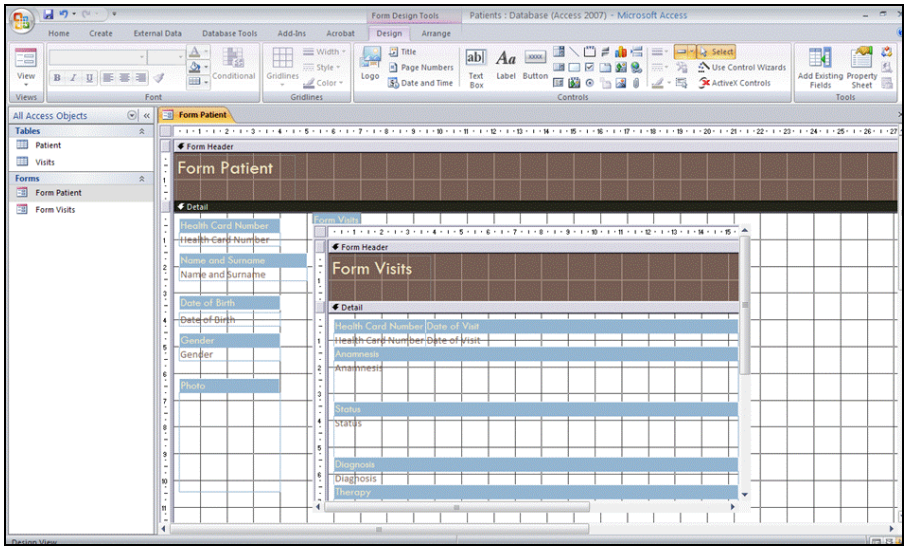


Figure 27. The *Form Visits* is dropped into the *Form Patient.*

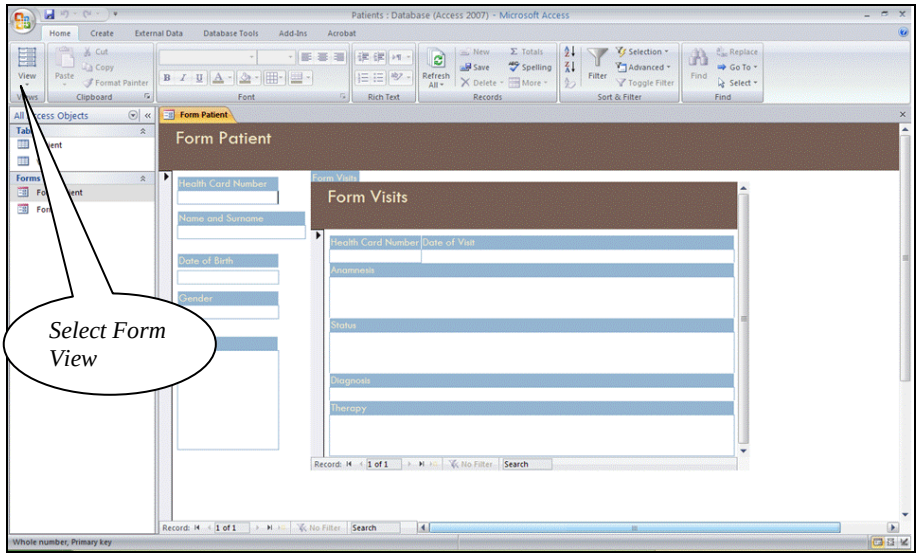


Figure 28. Composition of two forms is the *Form* for entering data in database.

Data entering for individual entities

In Figure 29 data entering for individual entities is shown.

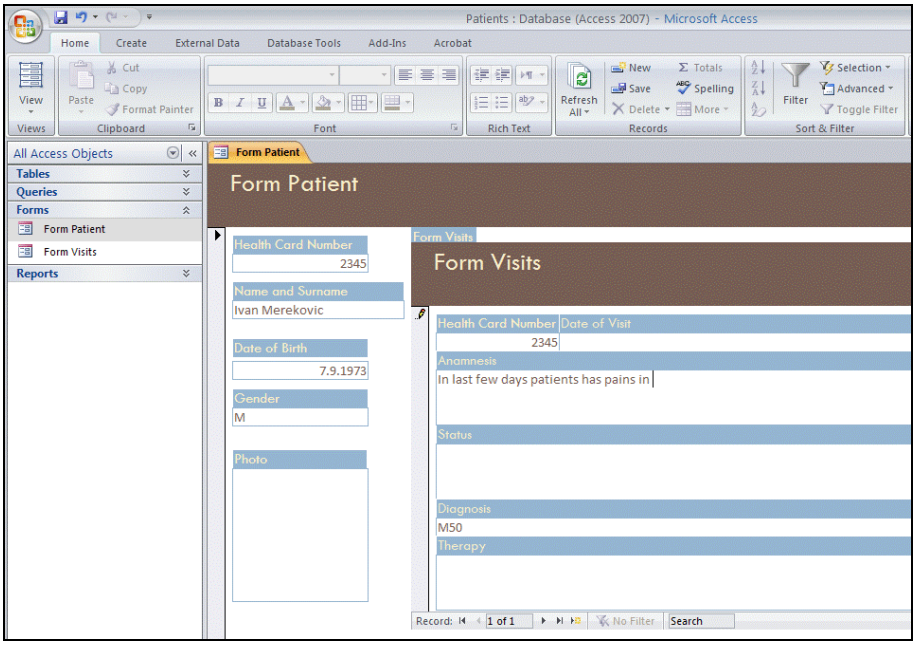


Figure 29. Data entry for individual visit.

Design of a table that contains extracted data from few original tables (selected attributes and/ or selected entities)

In Figures 30 to 38 design of a table that contains extracted data from few original tables is shown.

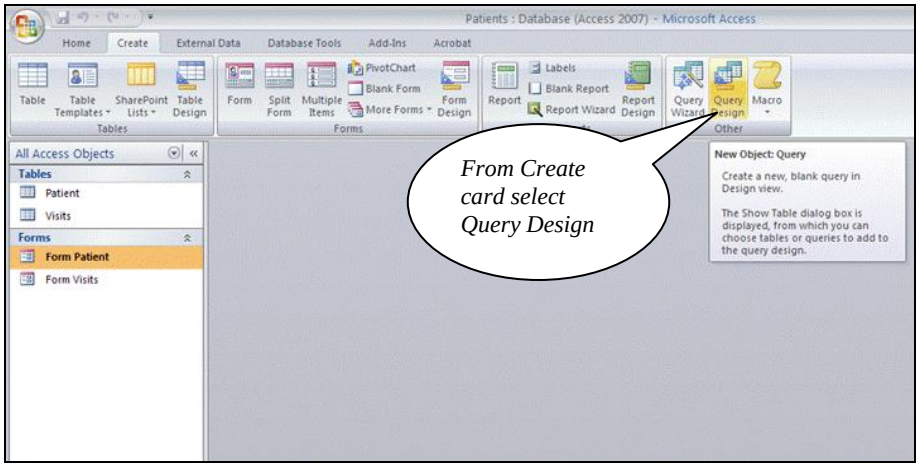


Figure 30. Creation of a table using option *Queries.*

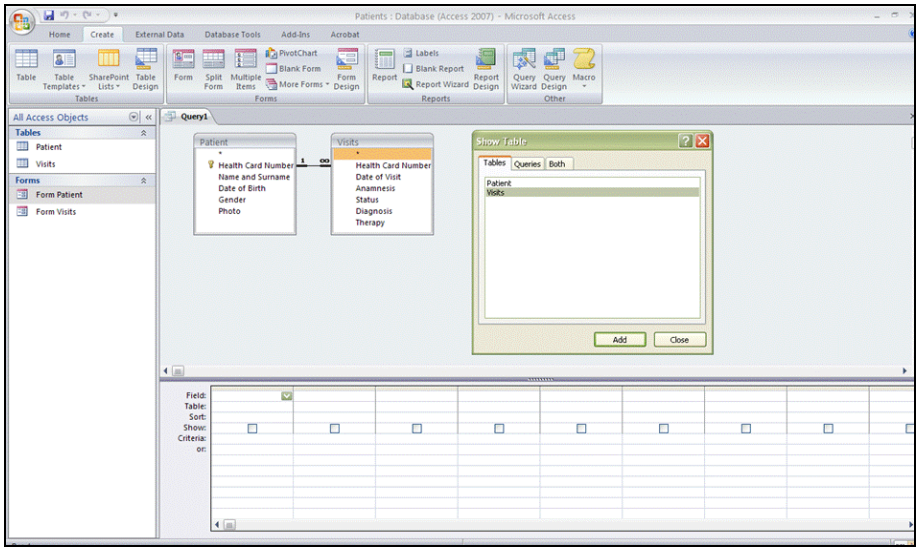


Figure 31. Selection of the tables (*Patient, Visits*) for creation of a new *Table.*

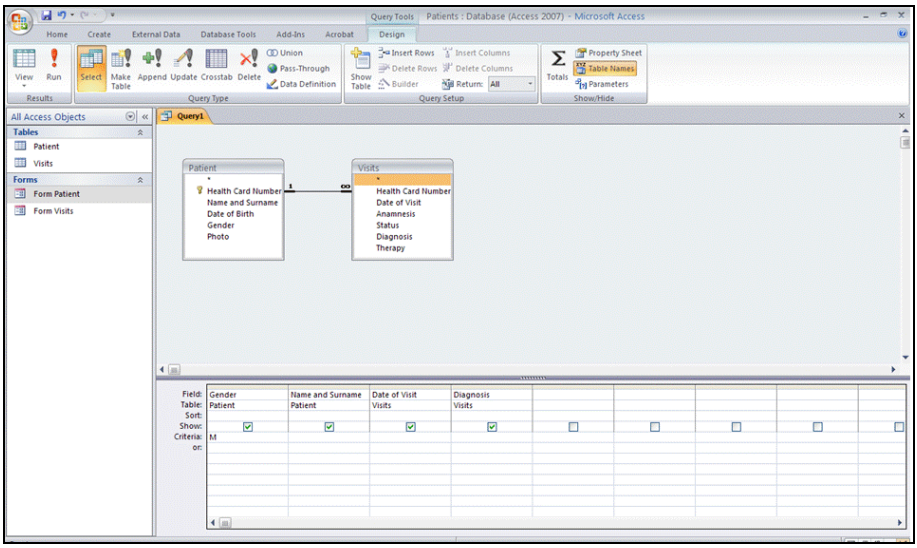


Figure 32. Definition of the criteria (*Gender= M*) and content (*Name and Surname, Date of Visit, Diagnosis*) for the new *Table.*

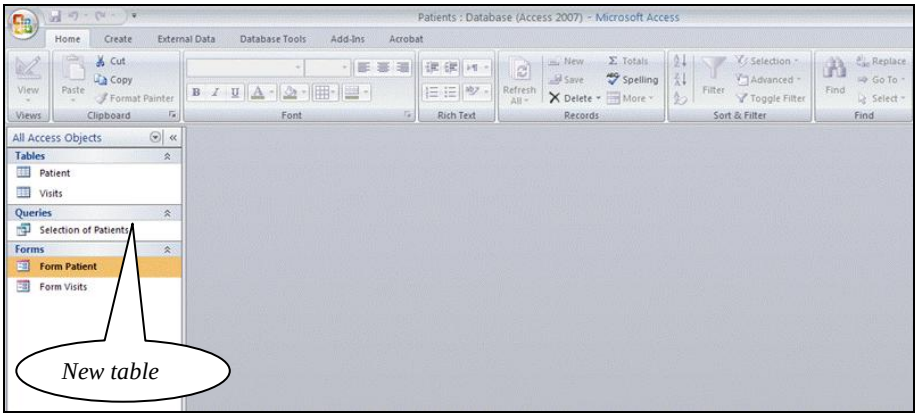


Figure 33. Result of the creation of the new table *Selection of Patients.*

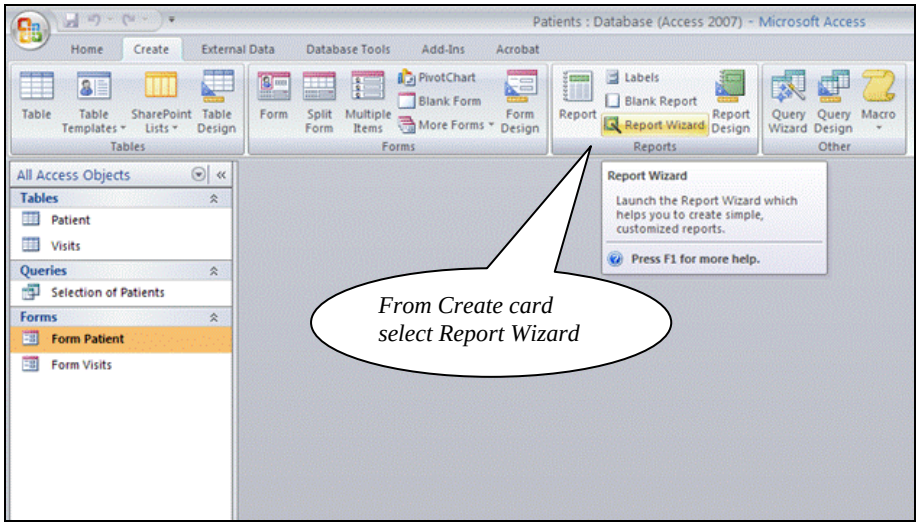


Figure 34. Creation of a *Report* using *Report Wizard*

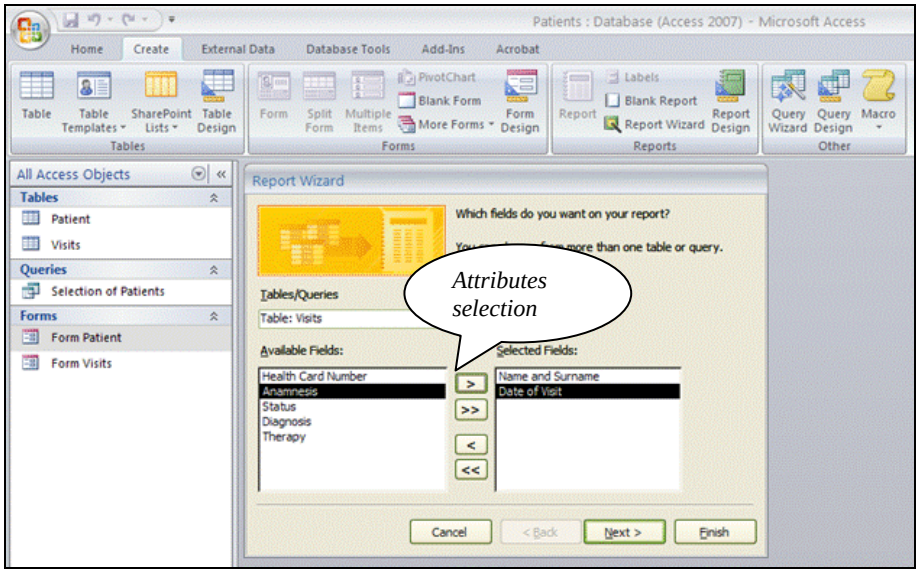


Figure 35. Creation of the *Report*, adding of attributes from table *Visits*

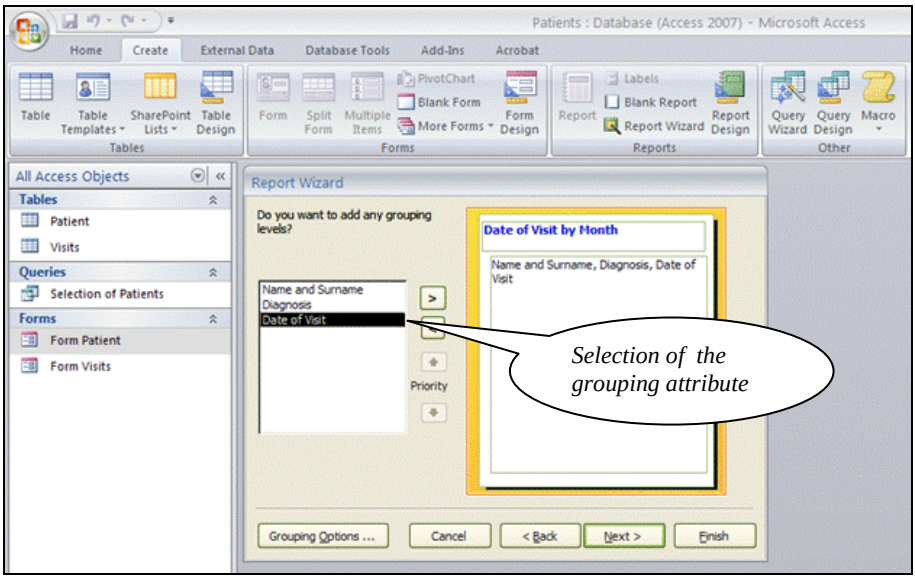


Figure 36. Creation of the *Report*, grouping according to the attribute *Date of Visit.*

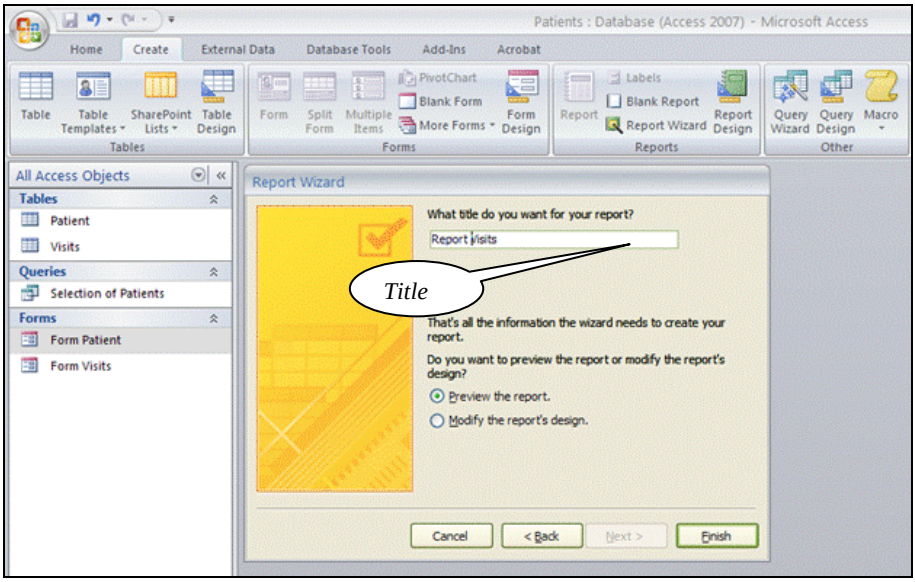


Figure 37. Creation of the *Report*, selection of the *Title.*

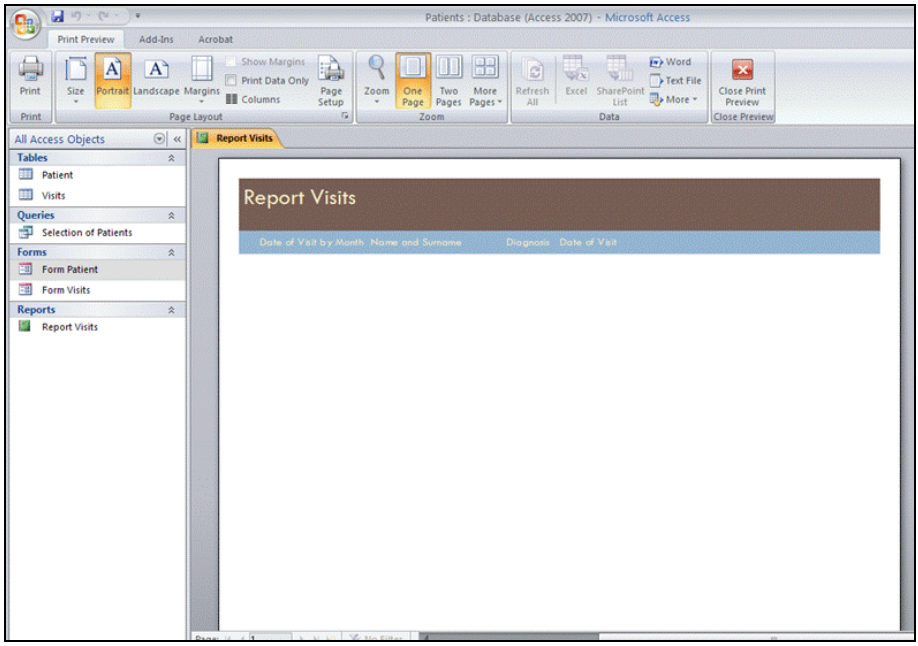


Figure 38. Result of the *Report* creation.

REFERENCES

1. Wempen F. Teach Yourself VISUALLY Microsoft Office Access 2007 (Teach Yourself VISUALLY (Tech)). Singapore: Visual 2007.
2. Baždarić K, Polašek O, Fišter K, Petrovečki M. Organization of Data [in Croatian]. In: Medical Informatics [in Croatian]. Kern J, Petrovečki M, editors. Zagreb: Medicinska naklada 2009.
3. Bointon L, Bourke A, Rourke C. Training for ECDL Syllabus 5 Office 2007: A Practical Course in Windows XP and Office 2007. Dublin: Blackrock Education Centre 2010.
4. Kern J. Methods in Medical Informatics [in Croatian]. Zagreb: Medicinska naklada 2006.
5. Mašić I, Riđanović Z, Pandža H, Mašić Z, editors. Medical Informatics. 2nd edition. Sarajevo: Avicena 2010.
6. Hoyt R, Sutton M, Yoshihashi A. Medical Informatics: Practical Guide for the Healthcare Professional 2007. Lulu.com 2007.

RECOMMENDED READINGS

1. Scott T, Rundall TG, Vogt TM, Hsu J. Implementing and Electronic Medical Record System: Successes, Failures, Lessons. Oxford: Radcliffe Publishing 2007.

2. Shaw N. Going Paperless: Guide to Computerization in Primary Care. Abington: Radcliffe Medical Press Ltd. 2001.
3. Coile RC Jr. The Paperless Hospital: Healthcare in a Digital Age. Chicago: Health Administration Press 2002.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**PUBLIC HEALTH CAPACITY BUILDING: ADULT EDUCATION PRINCIPLES AND METHODS**
Module: 5.2.1	**ECTS (suggested): 0.40**
Author(s), degrees, institution(s)	**Gordana Pavleković,** MD, PhD, Professor Andrija Štampar School of Public Health, Medical School, University of Zagreb, Croatia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Anja Kragelj**, Student of Andragogy, Chair of Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
Address for correspondence	**Gordana Pavleković** Andrija Štampar School of Public Health, Rockefellerova 4, 10000 Zagreb, Croatia E-mail: gpavleko@snz.hr
Keywords	Adult education, andragogy, didactical methods
Learning objectives	After completing this module students should: • know the definition and characteristics of adult education; • be familiar with different teaching methods used in adult education; • be able to use in practice some of group teaching methods of adult education.
Abstract	Like it holds in medicine that children are not diminished adults, and that paediatrics is not only curtailed internal medicine, it holds in educational process that adults are not augmented children. Adult education methods are in fact the same methods we could find in educational process of children, but the application of these methods is rather different. The module is describing principles of adult education, adult education process characteristics and teaching methods appropriate in public health capacity building, and provides some skills in using group education methods.
Teaching methods	An introductory lecture gives the students first insight in characteristics of adult education process, and understanding of phenomena. Afterwards, students first carefully read the recommended readings. Afterwards they form small groups. Every group is given a problem concerning some public health issue to be taught. In a group, students first use the proposed teaching method, and present the use of this method in front of other students. In continuation, then try to find the best teaching method for another public health theme. Every group prepares presentation of method, chosen as optimal. Other students critically discuss their choice.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire for individual assessment (50% of the score); assessment of group work (50% of the score).

PUBLIC HEALTH CAPACITY BUILDING: ADULT EDUCATION PRINCIPLES AND METHODS

Gordana Pavleković, Lijana Zaletel-Kragelj, Anja Kragelj

THEORETICAL BACKGROUND

Introduction

Teaching adults is not the same as teaching children

Like it holds in medicine that children are not diminished adults, and that paediatrics is not only curtailed internal medicine, it holds in educational process that adults are not augmented children in this process. It is true, that adult education methods are the same methods we could find in educational process of children, but the application of these methods is rather different (1). However, these methods are the same by the name but their forms, and educational tools and instruments used in the process, are different.

Basic terms in educational process

To proper understanding of educational process, we need to distinguish between basic phenomena of this process, being education, teaching, learning, studying, and knowledge.

1. Education.
   Education is described/defined as:
   - a process of teaching, training and learning, especially in schools or colleges, to improve knowledge and develop skills (2);
   - a long-term systematic process of gaining of knowledge, and developing of capacity, skills, habits, and personal characteristics (3,4).
2. Teaching.
   Teaching is described/defined as:
   - giving lessons to students in a school, college, university, etc.; to help somebody learn something by giving information about it (2);
   - a process of introduction of students/learners to the process of studying, helping students/learners in this process, and reflective, methodological and methodical (systematic) guiding students/learners in their studying activity (3,4).
3. Learning.
   Learning is described/defined as:
   - gaining knowledge or skill by studying, from experience, from being taught, etc. (2);
   - a direct changing/modifying of an individual with his/her own activity, that is being provoked by intrinsic (internal) needs, or extrinsic (external) incentives/stimuli/motives (3,4);
4. Studying.
   Studying is described/defined as:

- the activity of learning or gaining knowledge, either from books or by examining things in the world (2).

5. Knowledge.
   Knowledge is described/defined as:
   - The information, understanding and skills that an individual gains through education or experience (2);
   - a system of captured facts, and generalization and development of abilities, skills and habits, that an individual captures more or less permanently, and puts to use (3,4).

Andragogy and pedagogy

The process of education of adults has its own scientific branch known under the term andragogy. On the contrary, the scientific branch that deals with education of children is known under the term pedagogy (Figure 1).

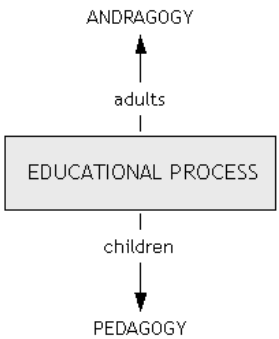


Figure 1. Scientific branches which support educational process of adults and children.

1. Pedagogy.
   The word comes from the Ancient Greek "paidagogeo". This term is a complex term, composed of two basic terms "pais", that means "child", and "ago", that means "to lead". The term "paidagogeo" litterarly means "to lead the child". In Ancient Greece, "paidagogos" was a slave who supervised the education of his master's children. This involved taking them to "didaskaleion", that mean "school", or a "gymnasion", that mean "gym" (5). Thus, this term is related to educational process of children.

   Nowadays, there exist several very similar definitions of pedagogy, three of them being:
   - according to The Advanced Learner's Oxford Dictionary (2), pedagogy is the study of teaching methods;
   - according to Merriam-Webster Online Dictionary (6), pedagogy is the art, science, or profession of teaching;
   - in TheFree Dictionary we could find among others that pedagogy are the principles, practice, or profession of teaching (7).

These contemporary descriptions do not imply that pedagogy is related to children, but the name itself does."

2. Andragogy

Andragogy as a science has no long tradition, but andragogy as a process of education of adults is very old (8). The first use of the term "andragogy" - as far as it is known today - was found with the German high school teacher Alexander Kapp in 1833 (9). This term also, is based on greek terminology, despite this does not originate from Ancient Greece. It was constructed analogically as the term "pedagogy". The first part of this term "andr-", is basing on Greek term that means "adult male" (8,10). Thus the term means "to lead the adult male". Because of "masculinisation", still stirs up disapproval (8). Also the relation to the leisure time activities was disputable. The term "andragogy" has today wider meaning, like the term "gymnasium", and it is not in any case related only to the leisure time activities.

This discipline was in the past more developed in Europe, especially in France, Yugoslavia, and Holland (10). Several former Yugoslavia experts from this field are very well known, among them Serbian expert Dusan Savicević, and Slovene expert Ana Krajnc (8). On the other hand, we need to mention also Malcolm Knowles. He brought in the seventieth the concept of science of andragogy to the United States of America (US), and developed this concept in his way which is well known (11). There exist also several critiques of this model of educational process (12), but this is beyond the scope of this module.

There exist more or less similar definitions of andragogy (8). Krajnc, for example, defined andragogy as the science of education and teaching of adults, while Savičević as relatively independent scientific discipline within more general science of education, that research education and learning of adults from all aspects (8).

Originally, pedagogy and andragogy were seen as two different and separated processes, but current theory sees them on a continuum, with pedagogy on one end and andragogy on the other, rather than two separated processes (Figure 2) (13).

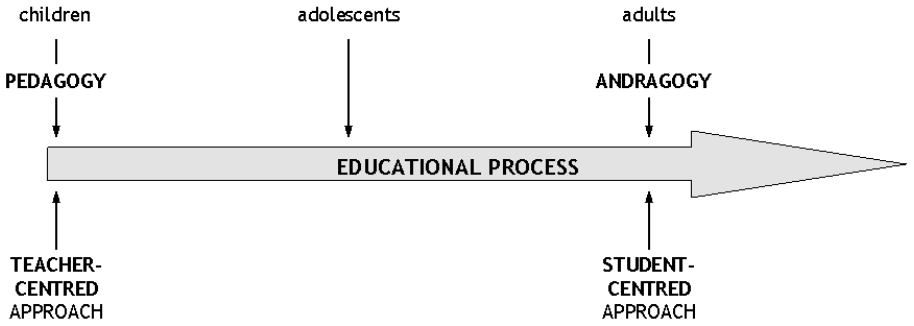


Figure 2. The continuum of educational process, and position of pedagogy and andragogy.

We will discuss some of differences, which distinguish both concepts later on, but in this place would be worthwhile to emphasize maybe the most important distinction between them: andragogic approach is a student-centred approach, while pedagogic approach is teacher-centred (11,13,14) (Figure 2).

Didactics

A term, which is also used very frequently in the context of education, is "didactics". The Croat expert of didactics, Vladimir Poljak, defines in his book didactics as "the branch of pedagogy, which is studying general rules of education" (3). In fact, it is studying the effective applicability of these rules, effective ways and forms of education (15). The word itself has Ancient Greek origin. The word "didaskein" means "to teach" (3).

In andragogy, didactics is defined in a wider way than in pedagogy. In andragogy it is not limited only on institutional or formal education (15).

Capacity building

If we are presenting these methods as important for capacity building, we need to clarify also two terms, being capacity and capacity building (16-19):

- capacity is defined as the ability of individuals, organizations or systems to perform appropriate functions effectively, efficiently and sustainably. Capacity has various dimensions:
  - it is not static but is part of a continuing dynamic process,
  - it does not exist on it own, but is linked with performance,
  - it is an instrument for an individual, team, organisation or system to achieve objectives and
  - capacity always contributes to sustainability.
- capacity building or development is the process by which individuals, groups, organizations, institutions and societies increase their abilities to perform core functions, solve problems, define and achieve objectives, and understand and deal with their development needs in a broad context and in a sustainable manner.

These definitions, adopted by UNDP and UNESCO, emphasize the continuing process of strengthening of abilities to perform core functions, solve problems, define and achieve objectives, and understand and deal with development needs.

Why teaching adults is not the same as teaching children?

If the person is involved in education of adults, it is essential for him/her to be familiar with characteristics of child versus adult person in educational process. Children and adults in educational process differ in many respects. Some of them are summarized in Table 1 (13,20-26).

In short, the adult student in the educational process is usually characterized by following characteristics: he/she is a self-directed, self-motivated manager of personal learning who collaborates as an active participant in the learning process and who takes responsibility for learning (13).

Table 1. Some differences between adults versus children in the educational process (13,20-26).

Feature	Children in educational process	Adults in educational process
Previous experience	Children have no life or other experience when entering educational process. They base their knowledge on the information they get in educational process from the teacher. Adolescents could enter the process with some limited amount of life or other experience.	Adults already have (various amount) of life and other experience when entering educational process. In this process, they add new knowledge to existing one. Additionally, they compare new knowledge to the old one, or they recall an old knowledge and put it into new frame.
The aim	Children have no aim and goals in the educational process. They even usually not think about this perspective of the educational process. The teacher points out the aim of learning.	Adults usually have clear aim and goals of the educational process. If not so, they determine them with the help of a teacher.
Source of motivation	Because children do not have aim and goals, they do not have intrinsic motivation in the educational process. The motivation is almost exclusively extrinsic (marks, fear, punishment, rewards, passed exam, scholarship, etc.)	Because adults have clear aim and goals, they usually have clear intrinsic motivation in the educational process. They also have different kinds of extrinsic motivation (e.g. requirement for competence or licensing, an expected promotion, adaptation to job changes, compliance with company directives, etc.).
Level of motivation	Because motivation is almost exclusively extrinsic, the level is usually lower than in adults.	Because motivation is extrinsic and intrinsic, the level is usually higher than in children. High level of motivation is in some cases conditioned by self-financing of education.
Level of autonomy and self-direction	Usually low.	Usually high.
Level of control over learning process	Usually low.	Usually high.

Skills in principles and methods of educational process and public health

The role of educational process in public health

In public health, education has an important role. It takes place in capacity building as well as in public health profession itself in education of educators of lay people and in some cases the lay people as well (Figure 3).

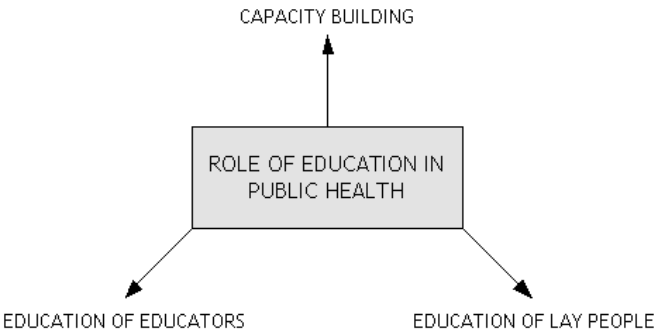


Figure 3. The role of education in public health.

For every profession, high quality and interesting educational process for capacity building is of utmost importance. This is more and more obvious also in public health (27). Public health workforce is becoming more and more important in establishing and/or maintaining public health care systems effective and efficient as much as possible (27). But is this workforce educated enough, especially if we consider the need for evidence-based public health, which needs highly educated and competent workforce in public health research? In US for example, only 44% of the public health workforce has formal public health education (27). European Commission Public Health Executive Agency (PHEA) states that many European Union (EU) member states and candidate countries have insufficient institutional and professional capacity for public health and that the process of reforming relevant services is slow. Compared to the US and other industrialized countries as well as some emerging economies, the relative lack of public health capacity in the EU is striking (28). But not only public health workforce is important. PHEA points out, that certain amount of public health skills is required for all kind of workforce involved in operating of health care systems (for example for medical doctors) (28). Skills should be of that amount that experts from professions involved in operating of health care systems could communicate between themselves in favour of end-users of these systems (for example patients).

Why in public health skills in andragogy are more important than skills in pedagogy?

In public health capacity building, which takes its place at universities, we meet different population groups, but majority of these population groups could be classified according to developmental psychology in early adulthood (29):

- adolescence: from puberty to about age 22-24 years;
- early adulthood: from about age 22-24 to about age 40-45 years;
- middle adulthood: from about age 40-45 years to about age 65 years;
- late adulthood: from about age 65 years to death.

If we take into consideration the developmental process of higher education in Europe (i.e. Bologna process), and Association of Schools of Public Health of the European Region, ASPHER, developmental recommendations for public health profession, we expect (or we already have) following target groups:

- students of the united first- and second-cycle (undergraduate and graduate) study of medicine and dental medicine. These students, which are mostly of age 19-25 years, are no longer children, but also not yet fully adults (23). In fact, they are finishing the adolescence phase, and are increasingly becoming young adults (23). The educational process of this population group is an exceptional art, demanding skills in pedagogy as well as in andragogy. It should pass from teacher-oriented to the student-oriented concept of education;
- students of the first-cycle (undergraduate, bachelor) study of public health or related studies, e.g. Health management study at Rijeka Medical School, Croatia (30), or Bielefeld University, Germany (31). This target population, which is of age 19-22, is very similar to the students of the first half of study of medicine or dental medicine - they are finishing the adolescence phase and are becoming young adults;
- students of the second-cycle (graduate, master) study of public health (or branches of public health, e.g. health management, health promotion, etc.). These students are at least at the phase of younger adults (usually of age 22-24). Since the higher education in Europe is in the process of reorientation from full-time study to part-time study along to be employed, especially starting from master degree study according to Bologna process, older and older population is expected in this cycle. Working adults who will want to succeed in their profession will create a new majority among second-cycle students;
- students of the third-cycle (postgraduate, doctoral) study of public health or related studies are definitely adults of at least 24 years of age, and majority of them are employed;
- students of the professional third-cycle (e.g. postgraduate specializations of doctors in medicine or dental medicine) study of public health or related studies (of at least 25 years of age) are similar to the students of previous group. The principal difference between them is that this study is more professionally, and less scientifically oriented.

We meet almost only with adult population also when we use the educational skills in public health profession itself, irrespective they are used in education of educators, or in education of the lay people. Even if the end target group of public health messages is a group of children or adolescents, public health experts are not usually involved in educational process of these age groups by themselves. They mostly educate educators of these population groups.

This is the reason that we are emphasising andragogical concept in this module, and not the pedagogical one. Adult students have unique needs, especially if they are employed. Before presenting these needs, we need to be acquainted with basic classifications and principles of didactical methods, irrespective they are used in andragogical or pedagogical process.

Teaching methods

Classification of teaching methods

Teaching or didactical methods could be classified in several ways, and these classifications could intertwine, since some methods could be seen from different perspectives (1,3,8,32-34). Ana Krajnc, the famous Slovene expert of andragogy, in her book presents three classifications: according to number of students, according to frontality, and according to formality of the presentation of teaching matter (Figure 4) (1).

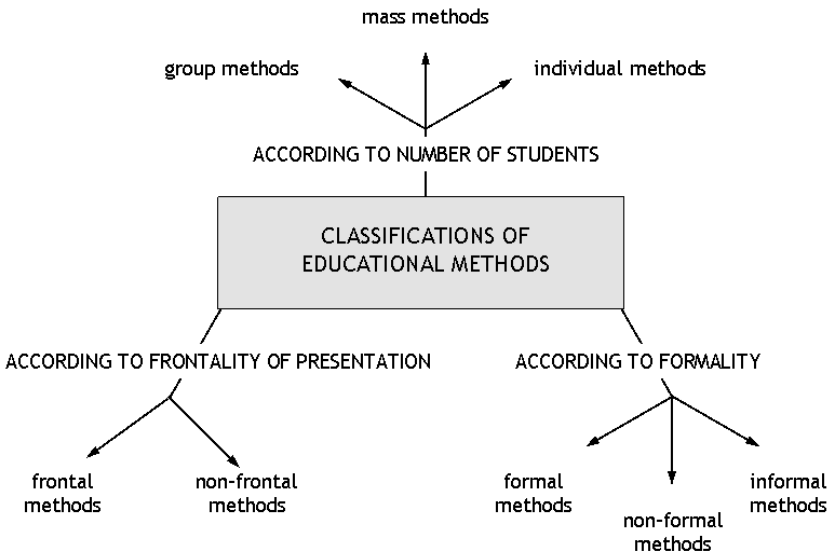


Figure 4. Three classifications of educational methods.

1. Classification according to number of students.
   Number of students certainly influence the choice of teaching methods used in educational process. It is clear that teacher would use a lecture rather than a consultation method when he/she has to teach a mass of 200 students at the same time. According to this criterion, methods are classified in:
   - mass methods of education,
   - group methods of education, and
   - individual methods of education.

   In further considerations we will use this classification, thus it will be discussed in details later.

   Mass, group, and individual methods can pass over one into another, depending on the educational objectives. When the end objective of educational process is adoption of general knowledge, or formation of general image/conception on the issue, then it is enough to use mass education methods. When the end objective is formation of personal viewpoints, and attitudes, group work is more appropriate. Practical exercising in small groups is especially recommended when end objective of educational process is adoption of practical skills (35). In all cases, educational process need to be supplemented by individual forms, at least guided individual study of provided recommended readings.
2. Classification according to frontality of the presentation of teaching matter.
   According to this criterion, methods are classified in:
   - frontal methods of education – in frontal methods, teacher (a transmitter) transmits his/her knowledge to passive audience of students (receivers) (1). The term "frontal" is originating from the position of the teacher - he/she is standing in front of the audience while presenting teaching matter;
   - non-frontal methods of education - in non-frontal methods, students are active and they are at the same time receivers and transmitters of knowledge (1).
3. Classification according to formality.
   Nevertheless the educational process is mostly understood as guided process of acquisition of knowledge, known also as formal, there exist also non-guided way. According to this criterion, methods are classified in:
   - formal methods of education – formal education refers to regular schooling that follows a normal pattern – admission students at about age six or over, promotion from grade to grade on a yearly basis, and use of a curriculum that covers a wide range of knowledge (19). It is directed towards acquiring of formally valuable and recognized qualification, level of education (from primary to the postgraduate), and title;
   - non-formal methods of education – non-formal education refers to educational activities delivered to targeted groups, where there is a possibility to provide attention to individual learners. These activities may include for example courses and workshops that meet specific needs of society and it members (19). It is directed towards knowledge as educational achievement/result, and not as administrative one;

- informal methods of education – informal education refers to learning channels, such as mass media and mass publicity campaigns, where is little or no possibility for attention to the individual (19).

Mass methods of education

Mass methods of education could be first divided in formal and informal. Informal mass education is taking place all around us all the time through mass media, journals, books, internet etc. (1). This part of mass education is out of the scope of this module.

In formal mass education, usually frontal methods are used. The most well known method is lecture (1). Other well known and also frequently used method is demonstration (Figure 5).

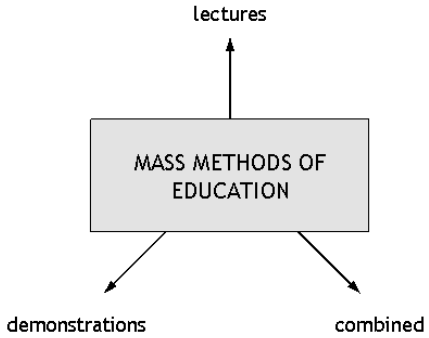


Figure 5. Some mass methods of education.

1. Lectures.
   According to Bergevin et al, the lecture is a well-prepared oral presentation of a subject by a qualified person (36). Lectures are basic teaching method in higher education (37).

   This teaching technique, like others, has its strengths and limitations (36-38). Among strengths we could classify:
   - can facilitate transmitting a message across to an audience of most any size, especially they are useful for mass education and education of large groups;
   - saves time (particularly of teachers);
   - saves money;
   - gives feeling of safety to students;
   - presents factual material in direct, logical manner;
   - stimulates thinking to open discussion, etc.

   Major limitations are common to all frontal methods of teaching:
   - audience is passive (only reception);
   - attention and acceptability of students varies;
   - communication in usually one way, unless teacher stimulates students to ask questions;

- it does not permit assessment of progress of students;
- it does not permit individual tempo of learning;
- experts are not always good teachers, etc.

Lectures to be effective need clear introduction, content limit and summary.

According to Krajnc (1), lectures could take different forms/types, like:

- explication;
- narration, story telling;
- explanation;
- interpretation;
- description
- argumentation, etc.

Usually, two or more of them are combined at one lecture.

Krajnc, Pavleković and Levine give some ideas for better effectiveness of lectures (1,26):

- the lecture should be of maximum of 90 minutes of duration;
- the lecture should be organized, planned ahead;
- the content of the lecture should be logical in order of explanation;
- objective of the lecture should be clear;
- not too much information should be put in one lecture not to transgress the students' threshold of information uptake;
- periodic breaks that relax students by informally discussing the ideas that have been presented should be encountered;
- for keeping students' attention, examples, anecdotes, etc., should be included;
- visuals – charts, slides, and similar should be used, to allow students to see what they have been told;
- it is recommended to allow students for questions;
- it is also recommended to provide opportunities for small group discussion if possible, etc.

In adult learning, lectures could be only the introductory method, preparing students for more active work in small groups (discussion) or even individual (dialogue), which will be (some of them) discussed later on (1). Adults appreciate active involvement.

2. Demonstrations.

By using this method, teacher shows how to perform a certain task (32).

Demonstrations are usually classified into two forms/types (26,36):

- method demonstration – illustrates how to do something in a step-by step procedure;
- result demonstration – shows the results of some activity, practice or procedure through evidence that can be seen, heard, or felt.

Group methods of education

Groups could be larger or smaller. Average large groups are as large as 20-25 students (1,35). This size of groups is regarded as being large and methods used are different than are used in small groups, which are on average large as 6-7 students (35). For using some teaching methods, even smaller groups are necessary. In larger groups different methods, usually of less active approach, are applied than in smaller groups (1,35)

Working with large groups

Methods used in large groups are very similar to that, used in mass adult education, being lectures and demonstrations (1). The only difference is that these methods, used in groups, even as large as 20-25 students, are more personal. Mass lectures, given for a mass of students as large as 150-200 students or even larger, are as effective as lectures given on television. The lecturer is far distant from the audience, almost anonymous. He/she has no personal contact with the audience, and they should be even supplied with loudspeakers (1).

Working with small groups

Teaching methods for small groups are much different than for large groups. They involve students actively. Teacher is in the role of coach and facilitator, giving students instructions how to work and what are the goals of their work (35).

Classical method used in education of medial doctors are laboratory and clinical exercises, but there exist also many other techniques (26,35,39), which could be very useful in public health education process. On this place only some of them could be presented (Figure 6).

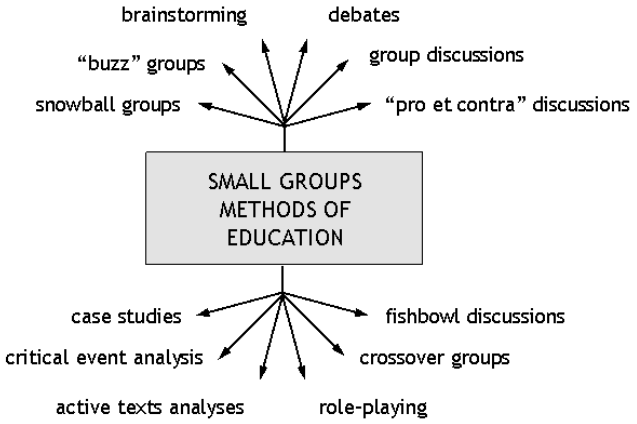


Figure 6. Some small groups methods of education.

1. Text/document analysis (1).
   A key text/document on topic under discussion is given to students to be studied. Each group could study the text from a specific point of view.

This is excellent way to introduce a new topic or new concept for warming-up students before using other more pretentious methods of active learning.

2. Group debate (1,32).
   Group debate is a less pretentious active teaching method. In activity the teacher, as well as students are involved. In debate, each participant gives certain part of the topic under discussion as an answer, complement, or explanation to the previous speaker in the debate, or several of them. Speakers need continually pay attention to other speakers'. In mind, they analyze his/her/their speeches, and search for something not told yet, or not enough explained yet. At the end, the teacher supplements the missing part, if necessary.

   Despite this method is one of the oldest methods of education, it was neglected in pedagogical didactics, since for performing this method, students need to have certain amount of preliminary knowledge and experience, and enough of self-confidence to participate in the debate.

   In andragogy, this method could be very effective in strengthening and consolidating studying material, as well as for gaining new knowledge.
3. Brainstorming (26,32,35,36).
   Brainstorming is a method of free generating, outpouring and sharing of ideas related to the topic under discussion between students. The process has three phases.
   - in the first phase, students are encouraged to employ creative thinking and voicing their ideas. All ideas are accepted at the beginning of the process and no response, regardless of how useless or impractical, is omitted. Original ideas are very welcome;
   - in the second phase, the explanation, and categorization. of ideas takes place;
   - in the third phase, the ideas are analyzed, evaluated, combined, refined etc.

   This method can be excellent to help a group of students think creatively. It is very useful in situations when new ideas related to issue of interest are wanted.

   Despite its usefulness this method has limitations. It is time-consuming and it is limited to the abilities of the participants. Also teacher's professional skills are important.
4. Group discussion (1,26,35,36).
   According to Bergvin et al. (cited by Seaman) (36), group discussion is purposeful conversation and deliberation about a topic of mutual interest among 6-20 participants under the guidance of a trained participant called a "leader".

   Discussion is not any more only presenting (like in lectures or in debates), or mutual supplementing of teaching matter. It is harmonization of contradictory standpoints, opinions, finding, statements, etc. of students about topic under discussion. Each student is confronted with the fact, that he/she must arguing for his/her standpoint, opinion, etc., or justifying it. It could also change it, if other students convince him/her opposite.
5. "Buzz" groups (1,35,39).
   "Buzz groups" technique is used when very concrete task is to be solved in very limited time. Pairs of students (i.e. neighbours in a teaching room) are formed to

discuss a question set by a lecturer. This method increases participation and engagement of students.

This method could be also used in groups of students to discuss/express difficulties they would have been unwilling to reveal to the whole large group.

It could be used in groups in which students know each other very well, but it also could be used with a goal that students become less inhibited. It is easy to implement in any size of student group and in most teaching rooms, even in lecture hall.

The first variant of this method

6. Snowball groups (32,34,39).
Snowball groups could be seen as an extension of "buzz" groups. The second name of this method is "pyramids".

This method involves progressive doubling of small groups. First, a pair of students is addressed to perform an activity. Pairs join up to form fours, then fours to form eights. These groups of eight students report back to the whole group (usually one representative of this group is a reporter) at the plenary session.

The disadvantage is that students could become bored with repeating the same task. For overcoming this problem, a sequence of increasingly more complex tasks could be applied. Another disadvantage is that it is a little time-consuming.

On the other hand advantage is that this method allows students to think for themselves or in small groups before open discussion. It also ensures good participation in plenary discussion.

7. Fishbowls (32,34,39).
Fishbowl method (or aquarium) involves one group observing another, active group. For the purpose of this method, two circles are made. The active group is sitting in the inner circle. It discusses an issue or topic while the outer circle group is observing/listening the discussion. The outer group looks for themes, patterns, soundness of arguments, etc. in the inner group discussion, analyze the inner group's functioning as a group, etc.

The disadvantage is that outer group students could become bored. For overcoming this problem, groups could switch places. Another variant is that participant from outer group that wants to make a contribution can switch the place with one participant from inner group when he/she is silent. This can happen many times during the discussion.

8. Crossover groups (39).
In crossover groups method, students are divided into subgroups that are subsequently split up to form new groups in such a way that maximal crossing over of information is assured. For the purpose of this method, students are labelled by two codes.

If we use for example two sets of codes, being A, B, C and 1, 2, 3, students are labelled as A1, A2, A3, B1, B2, B3, C1, C2, C3, etc. By asking all As, Bs, and Cs to work together, and then all 1s, 2s, and 3s, students could be mixed in such a way that each student meets at least one member from all other primary groups.

This method is excellent for mixing people and information. It is simple to organise. The plenary discussion could be omitted.

9. Case study (26,36).

By case study method, students analyse specific situation that they may be facing in the future. They search for solving the situation that has been presented.

One of the very important uses of this teaching method is in assessment how much students have learned and how comfortable they will be in using their knowledge in solving problems in the future.

10. Role-playing (26,35,36).
    In role-playing, students are provided with a situation in which they are involved, and a role to play. They are asked to adopt a role, and play it. Afterwards they are required to debrief the experience, analysing the scenario, discuss their feelings, etc.

    This method is extremely helpful in gaining skills in interacting with other people, since it is an effective way to get students familiar with realistic situations. On the other hand, the major disadvantage is that shy students could find this method stressful. For overcoming this problem, a role is given to a group of students who work out the details, but the role is actually played only by one of them.

    Role-playing could be performed in small groups, as well as a frontal role-playing in front of the total group.
11. "Pro et contra" discussion (34,35).
    "Pro et contra" discussion is a special type of group discussion. It could be also partially seen as role playing.

    Two groups of students are given opposing positions to defend an opinion, standpoint, etc.
12. Critical event analysis (35).
    This method is basing upon very concrete personal experience of students. Each student presents his/her concrete event, related to the topic under discussion. On the basis of personal story in the group discussion concrete solutions could be gained.

Problem-Based Learning (PBL)

PBL was originally developed at the Faculty of Health Sciences at McMaster University around 1965. PBL is an approach to learning and instruction in which students tackle problems in small groups under the supervision of a tutor. In most of the cases, a problem consists of a description of a set of phenomena or events than can be perceived in reality. These phenomena have to be analysed or explained by the tutorial group in terms of underlying principles, mechanisms or processes. The tools used in order to do that are discussion of the problem and studying relevant resources. Typically, tutorial groups range from 5 to 10 students and one tutor, who may not only be skilled in group process, but expert in content, or may be primarily a facilitator. (40,41)

The core concept in PBL is that the process of problem identification and problem solving can be a powerful way of learning because it is active, goal directed, and germane. Problem identification opens the door to a personal investment in problem-solving.

PBL has the following cognitive effects on student learning:

- activation of prior knowledge – the initial analysis of a problem stimulates the retrieval of knowledge acquired earlier;
- elaboration on prior knowledge through small group discussion, both before or after new knowledge has been acquired – active processing of new information;
- restructuring of knowledge in order to fit the problem presented - construction of an appropriate semantic network.

- learning in context. The problem serves as a scaffold for storing cues that may support retrieval of relevant knowledge when needed for similar problems.
- since students will tend to see problems presented as relevant and since they engage in an open-ended discussion, epistemic curiosity can be expected to emerge.

In Table 2, differences between traditional, teacher-directed learning and problem-based learning are presented.

Table 2. Differences between traditional, teacher-directed learning and problem-based learning.

	For the student	**For the teacher**
Problems with traditional teacher-directed learning	Passivity Perceived lack of relevance of knowledge transmitted Often poor intellectual engagement Downloading of answers for which the students have not understood Little opportunity or incentive to conceive of questions that the teachers has not articulated Often lack of practice in kinds of collaboration which will be valuable in professional roles Relative lack of practice in setting personal learning goals, priorities, and methods	Tendency to use canned didactic talks and become jaded Frequently, lack of stimulation from fresh points of view and novel questions Relative lack of opportunity for cross-disciplinary faculty contact Often little opportunity for role modelling and mentoring, based on personal knowledge of students
Benefits with problem-based learning	Practice in actively identifying the main problems and issues in a case or other complex situation Practice in identifying and prioritizing helpful resources for learning Practice in systematic formulation of problem and hypothesis-setting Increased professionalism, separating vigorous intellectual discourse from personal relationships, as a member of a collaborative learning team Increased stimulus to learn underlying science and other principles as a result of relevant problem-solving	A satisfying sense of grater connection to student learning as a result of being in close touch with student thought processes A stimulus to rethink educational goals and methods Increased cross-disciplinary stimulation For many, a chance to learn or review up-to-date knowledge in fields close to one's own or otherwise interest

Even if there is some overlap, PBL differs from case-based learning. Illustrative cases are legitimate ways of enriching the learning process, but the method and goals differ from those of PBL. PBL cases require the steps of problem-identification, hypothesis formulation, analysis and validation.

PBL can be described as a cyclical process with several steps (Figure 7) (42). Group members must participate in all stages and movement between stages is not necessarily unidirectional.

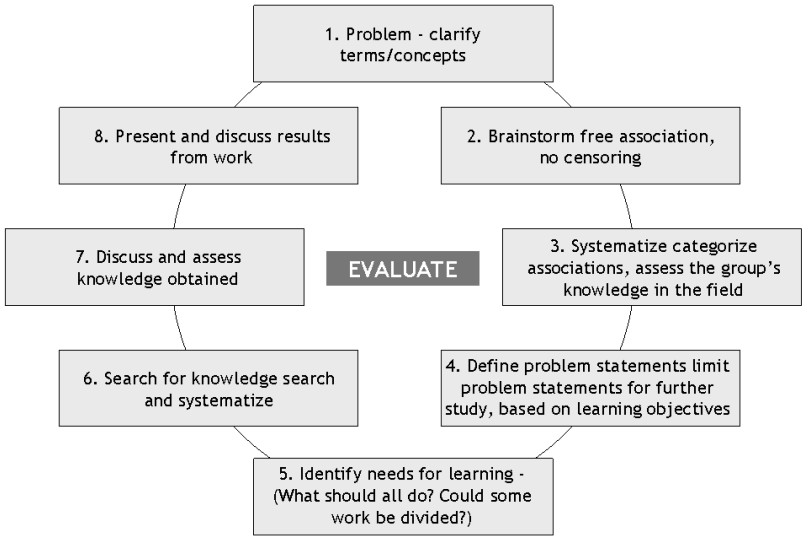


Figure 7. The PBL circle. Adapted from Fosse, 2006 (42).

PBL places greater emphasis and responsibilities on students and tutors/teachers than the traditional teaching methods.

1. Students' responsibilities.
   In PBL, students' responsibilities are being present at the group meetings, being active at tutorials and between tutorials and between tutorials, takes responsibilities for moving through stages of the process, reflection upon and challenging principle ideas and concepts as they emerge in the group process.
2. Tutor's role.
   The tutor's role is one of moderator and facilitator. They are not there to lead the process or provide a definitive answer. They act as "sign post" and steer the process in a productive direction. Initially tutors contribute to identifying learning objectives and formulating problem statements, and ensure the chosen learning objectives are realistic. An effective PBL tutor will ask questions that stimulate students to work on a problem in greater depth. Further, the tutor should give advice on how to collect information. At last, but not least, contribute to solving group conflict.
3. Common responsibilities.
   Additionally, both students and tutor have the common responsibilities. They are responsible for the group functioning well, socially and academically. It is important that feed-back flow in both directions, both from the tutor to the students and from the students to the tutor. The group should agree at the outset some mutual expectations and ground rules. At the start of each subsequent session the group should reflect upon the group process.

Individual methods of education

By the term "individual methods of education", first association is on self-studying, but these methods are much more than only self-studying (1). In addition to self-studying, individual methods of education are also methods of distance education and methods of personal contact (Figure 8).

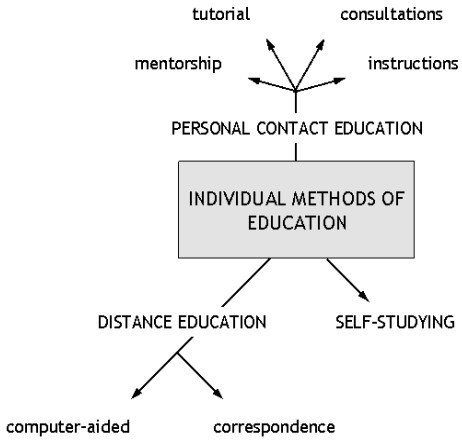


Figure 8. Some individual methods of education.

Individual methods of education enable individualized process, oriented towards individual student (1,32).

Today, with explosion of communication technologies, distance education is very popular, and it is not any more limited on individual distance learning only. With new e-communication technologies, also group distance learning is possible. But this group of methods is out of the scope of this module. Because of its expansion it needs to be presented in a separate module.

Despite extensiveness of communication technologies, an important part of individual methods of education represents methods of personal contact education. In this type of education, teacher is in contact with one student at a time only (1). This part of the module will be dedicated to selected personal contact methods, being consultations, instructions, tutorials, and mentorship.

1. Consultations (1,26,36).
   Consultation is specific teaching method in a form of a mutual discussion, advising, clarification of certain problem, and answering questions. In consultation hours, students could also get information about recommended readings and similar. Consultations are not obligatory method - students seek contact with teacher voluntarily.
2. Instructions (1,26,36).
   Instructions are very specific form of communication between the teacher, and the student, and it refers to narrower segment of educational process.

   According to Dunn and Dunn (36), individualized instructions focus the instructional process on individual skills, abilities, interests, goals, and rate of

learning. One of important features of this method is providing learning activities specifically formed to the needs, interests, and abilities of an individual student.

In comparison to consultations, instructions are one way communication. An instructor explicates the topic while student listens.

3. Tutorials (26,36).
   A tutorial educational method is performed when a single student needs specific help. The focus is usually the specific problems or concerns of the student.
4. Mentorship (1).
   In academic sphere, special method of education called "mentorship" or "mentoring" is applied for the most focused postgraduate study. In this educational relationship, a teacher is called "a mentor", and a student "a mentee". A mentor in this educational process is a trusted teacher who advise and guide his/her protege/protegee, how to learn and gain skills.

   Mentorship is to the certain extent similar to consultations, but between both methods there are huge differences. If consultation is a single educational episode, mentorship in opposite is a continuous process. It is a series of interrelated educational episodes, all directed to attain the educational objective.

   Mentorship is inevitably supplemented by intensive self-study of a mentee.

How effective are different teaching methods?

Before planning which didactical methods out of mass, group or individual methods of education would be used in educational process, and when, it would be useful to know how effective are these methods. This effectiveness is visualized by learning pyramid of the US National Training Laboratories, Bethel, Maine (43) (Figure 9).

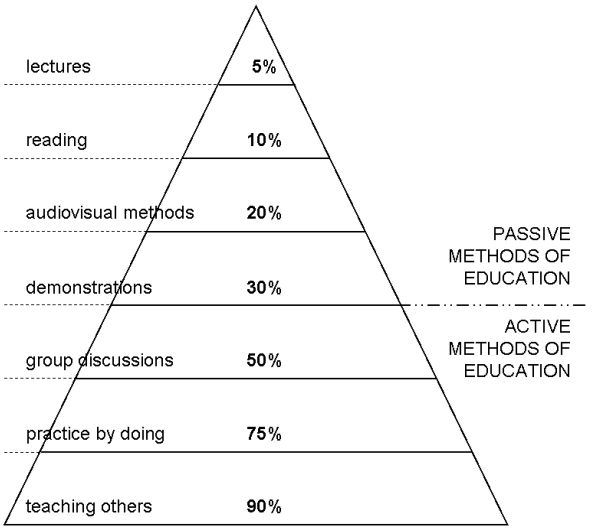


Figure 9. Learning pyramid with average learning retention rates (adapted from National Training Laboratories, Bethel, Maine) (43).

From Figure 9 it is obvious that active methods of education, what mean the methods by which students learn with experience, have much higher retention rate than passive methods, especially those by which abstract learning is required. The far lowest retention rate is achieved at lectures. This traditional educational method is in Croatia and Slovenia still the basic method of education at academic level (37). The situation is very similar in other countries of South Eastern Europe. All this is true also for the public health education in these countries.

In conclusion, it could be worthwhile to summarize basic educational methods. The summary is presented in Table 3.

Table 3. Summary of basic educational methods.

	Content	**Characteristics**	**Carrying out**
Lecture	Knowledge, important information, connection to the theory	Brief transfer of information, limited effects	At the first look easy to carry out
Exercise	Showing and practicing skills and procedure training	"Arteficial circumstances", "out of common sense", often too simple	Organisational difficulties, equipment
Seminar	Description and analysis of case, setting down and solving the problems	Interesting, very often too much abstract, carry out as lectures	Well prepared students and very restrained teachers
Group work	Analysis of case, solving the problems, „difficult" situations	Learning from and support by peers, close interaction with teachers	Well prepared cases and problems as an impulse, clearly defined tasks, expensive
Practical work	Real professional life	Very useful learning, but not always systematic, can easily be superficial	Teachers as „role model" and leader, learning from experience
Reading	„Active" reading or content overview	Need for „critical" reading	First impulse and evaluation are important

Adult education process characteristics

Previously we summarized some differences between adults versus children in the educational process (Table 1), but also the process itself has different characteristics in andragogy versus pedagogy. At the beginning of this module we emphasized maybe the most important distinction between them: andragogic approach is a student-centred approach, while pedagogic approach is teacher-centred (11,13,14) (Figure 2). In andragogical process teacher is not a teacher in a classical sense anymore. In adult education, he/she is more a guide, a "lighthouse", a facilitator, a supervisor, than a "teacher" (1,13,22). Some other differences are presented in Table 4.

Table 4. Some differences of the process of adult education versus children education (22,23).

Feature	Educational process of children	Educational process of adults
The role of the teacher	The teacher's authority is very significant	Teacher is more a guide than a main source of information
Educational programme flexibility	Teaching programmes are based on educational programmes, which are based upon the educational requirements of the society. Teacher has little possibility to be flexible	Educational programmes are more flexible and oriented towards the ability of students, their previous experiences, their needs, etc..
Teaching methods	Mostly frontal methods of teaching are applied.	Active teaching methods are applied more frequently (group work, projects, discussions, etc.).

How to use educational knowledge in public health capacity building?

We are aware that for improving situation in the public health profession the shift from traditional passive educational methods to more active methods is strongly needed. Collins proposes in her paper ten tips in this process for radiologic education, but these tips could be used in public health as well (13):

- connect life experiences and prior learning to new information.
- involve participants in the learning process, serving as a facilitator and not just a supplier of facts.
- create educational programs that are organized with clearly defined elements, clearly showing how the program will help participants reach their goals.
- help learners see a reason for learning something by making it applicable to their work or other responsibilities of value to them.
- acknowledge the experiences that adult participants bring to the learning environment, allowing for opinions to be voiced freely.
- show learners how the learning will benefit them and create a comfortable and appropriately challenging learning environment.
- limit lecturing and provide opportunities for sharing of experiences, questions, and exercises that require participants to practice a skill or apply knowledge.
- accommodate different learning styles by offering a variety of training methods (e.g. group discussion, role-playing, lecturing, case studies, panel/guest expert, games, structured note-taking, individual coaching, demonstration, and variation in media used) and by using visual, auditory, and kinesthetic techniques.
- provide opportunity for feedback from self, peers, and instructor.
- promote group interaction.

We will illustrate now how to use this knowledge in practice using two case studies.

CASE STUDY 1: THE COURSE "ART OF TEACHING IN MEDICINE: CROATIAN MODEL OF TRAINING IN MEDICAL AND PUBLIC HEALTH EDUCATION"

Needs for training of trainers: The story so far

In recent decades, we have been seen dramatic changes in the challenges facing the health profession. The nature and process of training for health workers has therefore undergone considerable change. Today, those involved in medical and public health education must adapt and find more innovative methods for teaching and learning what is now recognised as a much more complex subject embracing both the clinical and social paradigms associated with modern medicine.

Medical School University of Zagreb has recognised the need for changes in the way its medical curriculum is delivered and the need to improve the quality of medical study. Several activities have been implemented at Zagreb to meet those challenges:

1. a new curriculum has been developed and implemented based on the core competences for doctors working in the 21st century identified from the academic and professional societies;
2. the introduction and development of innovative teaching and learning methods;
3. more opportunities for student feedback and evaluation of the curriculum;
4. formal training courses for those staff involved in delivering medical and public health education, particularly young teachers at the beginning of their professional careers, and quality assurance of the training provided.

Since 1990, the Department for Educational Technology at Andrija Stampar School of Public Health, has been responsible for the development and implementation of a programme of initiatives to improve the quality of medical education, Working on the project "The using video and computer technology in continuing medical education for Primary Health Care and Public Health", they had recognised the need for professionalism in medical education. A principle activity is to run a series of workshops, lasting three days, aimed at highly motivated and experienced teachers, to encourage the use of innovative concepts and principles in medical education.

As a result of this programme, the Croatian Association for Medical Education (CAME) was established in 1992, bringing together medical teachers from four existing Medical Schools in Croatia. The activities of CAME and the Department for Educational Technology at Andrija Stampar School of Public Health are mostly aimed at further training of medical and public health teachers. There are two main activities: formal courses covering basic of adult pedagogy and interactive workshops devoted to experienced teaching problems.

Formal course entitled "The Art of Teaching in Medicine" are aimed for younger teachers who are just entering the field of medical and public health education. They are organised twice a year for a group of about 25 participants and the educational methods are mostly learner-centred. The participants are provided by diploma and it is an important prerequisite for teacher's advancement.

The workshops which last 2.5 days are aimed to experienced teachers, they are subject focused and the main educational method is experiential learning. Following subjects have been discussed until now: curriculum development, a lecture - the oldest craft of teachers, problem-based learning, learning from experience, cognitive psychology and medical education, continuing professional development, student's - teacher's relationship, assessment in education, etc.

The both activities are very well evaluated, they are complementary, but the workshops are connected with higher motivation of participants and are considered to be more important strategy for professional development of teachers in biomedicine and public health.

"The Art of Teaching in Medicine": Aims and objectives

The course aims are to:

- encourage and enable medical and public health teachers early in their career to learn and put into practice contemporary approaches in teaching and learning medicine;
- understand the contextual factors influencing this process;
- encourage and motivate teachers to improve the quality of their work, and
- share experiences with others.

It gives a participant:

- a basic understanding of the concepts of study for health professionals, including challenges and dilemmas in teaching and learning;
- a theoretical framework for understanding factors influencing the quality of teaching-learning process;
- a rang of evidence-based strategies, both traditional and innovative methods, to use in medical and public health education;
- a framework for planning, implementation and evaluation in the field of education;
- awareness of ethical issues relating to medical and public health education.

Participants: To whom the course is addressed?

Since 2000, the course was attended by more than 200 participants, mainly staff at Medical School Zagreb. All participants have been young or early career teachers from a range of different medical disciplines (clinical medicine, out-patient care, basic sciences, and public health). Since 2006, the course has been made compulsory requirement for staff involved in medical and public health education within the faculty, and part of the career development process for assistant professors.

Working format and main contents

The Course consists of six learning days with morning and afternoon sessions. A typical daily programme is presented in Table 5.

Table 5. A typical daily programme at the course entitled "The Art of Teaching in Medicine" at Andrija Stampar School of Public Health, Croatia.

Time	Topic
A. 8-9.00am	Self-directed study (readings tasks related to the daily content)
B. 9-10.30am	Introductory topic, short lecture, discussion and exchange of practical experiences
C. 11-13.00pm	Exercise, small group work, tutorials, demonstration
D. 16-17.00pm	Panel discussion (invited guests – students, patients, medical school Board representatives-dean, vice-deans etc.)

In Box 1, all topics of the Course by day are listed.

Box 1. Topics at the course entitled "The Art of Teaching in Medicine" at Andrija Stampar School of Public Health, Croatia, by day.

Day 1:
- *A. Introduction to the course. Expectations. Pre-test (MCQ)**
- *B. Basic Problem-based learning, problem solving*
- *C. Educational strategies and effectiveness of medical and public health education*
- *D. Panel discussion: Europe and the World: Challenges in education for health professionals*

Day 2:
- *A. Self-study, preparation for daily topics*
- *B. Curriculum planning and modules development, developing learning objectives*
- *C. Teaching and learning tools: Module planning and implementation*
- *D. Panel discussion: Student as a partner in teaching-learning process*

Day 3:
- *A. Self-study, preparation for daily topics*
- *B. Basic educational tools in basic, clinical and public health sciences*
- *C. Bedside teaching/Teaching community health*
- *D. Panel discussion: Patient/community as a partner in teaching-learning process*

Day 4:
- *A. Self-study, preparation for daily topics*
- *B. Traditional vs. Innovative methods in teaching medicine: Lecture*
- *C. Lecture delivering (using video for self and peer feed-back)*
- *D. Evaluation*

Day 5:
- *A. Self-study, preparation for daily topic*
- *B. Teaching aids (video and computerized technologies, handouts, textbooks, simulators, skill laboratory, e-learning etc.)*
- *C. Principles and methods in assessment*
- *D. Panel: Assessment and exams as a part of teaching-learning process*

Day 6:
- *A. Final test (the same MSQ)**
- *B. Staff development and academic standards*
- *C. Panel: Strategies and actual policies of high education. Perspectives of Croatian Medical Schools development.*
- *D. Course evaluation and future plans*

Process and impact evaluation

Process evaluation

On the last day of the programme participants are asked to evaluate the course, using quantitative and qualitative methods (written questionnaire and group discussion). Summary of findings are as follows:

- according to the feedback the course fulfilled the expectations of participants (mean grade 4.0±1.2, ranking from 1 the lowest to 5 the maximum grade),
- all participants confirm it has been useful and has benefited their everyday teaching process (mean grade 4.2±0.9),
- those elements of the course scoring most highly in the evaluation were:
  - builds on and expands their previous teaching experiences,
  - directing interests to teaching,
  - stimulation of exchange of experience with colleagues from different departments,
  - peer discussion as a continuous support for their advancement in teaching,
  - satisfaction in communication and sharing doubts with experienced teachers,
  - panel discussions with guests and an overview of present situation in medical education.
- those aspects of the course evaluated less highly and where improvements could be made:
  - insufficient participation of "older teachers" and key decision makers within the faculty,
  - doubts as to whether implementation of skills mastered in the course will transfer to practice.

The course impact

The course impact was analysed using students' assessment of all medical teachers within the Medical School Zagreb. Student feedback was compared with their mean grade, with those who passed the course, and their views as to how many of the teachers implemented innovative teaching and learning (modules) within the medical curriculum. Summary of findings are as follows:

- the mean grade for all faculty (overall assessment by students on scale 1 the lowest to 7 the maximum grade) is 5.1±2.3, for young or early career who had participated in the training programme, this was higher at 5.9±1.2,
- during the last seven years 15% of the trained faculty has introduced innovative modules in their teaching practice, mainly in electives.

Lessons learned

In comparison with other training for trainer's courses in medical education, with similar aims and objectives, this programme has at least two positive characteristics or strengths:

1. Trainers are not experts in pedagogical or didactical knowledge and skills. They are well experienced medical doctors/teachers aware to share their own experiences and reflections with participants.
2. Putting together a mixed group of participants (with different pre-knowledge, field of work, interest's expectations and teaching environment) has a positive influence on their motivation and encouragement.

Beside positives, at least two dilemmas remain:

1. Do we need (formally) certificated medical teachers or voluntary contribution of those who love "the art of teaching medicine", and will personally and with enthusiasm invest in their own development?
2. Impact evaluation (positive assessment of those who finished the course) although positive using students' questionnaires can be biased and their response could be influenced by factors not directly related to the training programme (for example student expectations, teachers personality, student learning styles etc.). Further evaluation and research is needed to expand and clarify the findings from this study and influence the development and implementation of training programmes in medical education.

In summary, the CAME training course in innovative teaching and learning is shown to be important for creating a new culture of teaching and learning in medicine and public health. It has been particularly successful in increasing awareness and motivation of the young or early career members of medical faculty for mastering the fundamental skills in medical education. However, participation of older, more senior or long-serving teachers are somewhat limited, resulting in the current variation in the learning experiences of medical students. The ongoing availability of CPD workshops in innovative teaching and learning in medical education is essential for both the development of specific skills for application of acquired principles in everyday practice as well maintaining the motivation and enthusiasm of those responsible for medical and public health education.

CASE STUDY 2: TEACHING STATISTICAL METHODS IN PUBLIC HEALTH ON POSTGRAUATE PUBLIC HEALTH COURSE LJUBLJANA MEDICAL FACULTY, SLOVENIA

Introduction

Usually, all users of statistical methods with less mathematical background, especially those who use them only occasionally, meet big problems when using these methods. Since these methods represent very important tool in research in public health, it is very important to enable public health students the learning approach based on comprehension of statistical methods not only using them as a "cook book".

For this aim to be attained, learning must be approached with consideration. It is undoubtedly important:

- the teaching matter to be distributed logically and units should be smoothly associated one to another;

- not to use too many equations or mathematical expressions, especially not in professions with less abstract way of thinking (in medicine in particular);
- to enable the students to use the methods in practice, meaning transferring the emphasis from the lectures to practical work on an example (learning with experience);
- to watch carefully all the time if the students are following the explanation and practical use of the teaching matter; the teacher must strive for not concluding a subject until it is accepted by the majority of students;
- to use new pedagogical approaches at lecturing, for example working in smaller groups, learning with discussion etc., whenever possible;
- to use adequate tools, meaning learning with the statistical program of good quality; nevertheless certain calculations shall be carried out manually or partly manually, particularly where this shall be necessary for the procedure to be easier understandable;
- students shall be proposed, if not supplied with, corresponding studying material.

The main purpose of teaching public health students statistical methods with comprehension is for them to be able to have entire process of a public health research, from the definition of a problem as a starting point, through collecting the relevant data and analysing them, finally to interpretation of the results of the analysis under control. This does not necessarily mean that they will need to do the entire process by themselves. Lack of general acquaintance with research tools including statistical methods most often results in a public health worker, not be able to translate the results of an analysis into technical language of his or her profession. This is usually followed by a very complicated interpretation, which can be very hard to understand, especially to decision makers and politicians to whom the results are intended for. In other words, teaching statistical methods with comprehension means to make future public health experts to be able to speak the same language as statistic experts, who are responsible for technical realisation of the methods.

Process of teaching

In teaching statistical methods in public health at postgraduate level of education in Slovenia we try to follow the principles just described as much as possible. It seems that this could be the right way to achieve the aim already described.

In continuation, a short description of the process of teaching statistical methods at the 2 semester (60 ECTS) postgraduate course of public health, which is a compulsory part of the professional third-cycle study of public health postgraduate specialization in public health for doctors of medicine or dental medicine (for others is optional), is given. This method is used now successfully for about 10 years.

Educational objective

End objective of educational process is adoption of practical skills. For achieving this objective, practical exercising in small groups is especially recommended (35).

Target group

Target group are adult postgraduate students of 25-40 years of age, occasionally some students up to 55 years of age are expected. According to developmental psychology classification, this target group is classified in early to middle adulthood (29), thus andragogical principles are to be used.

Educational programme

The experiences of long standing show that the clear elements of teaching matter in statistics, and right sequence of these elements are crucial for success.

We conceived the whole teaching process on the relationship analysis, which is situated centrally. Accordingly, teaching starts with theoretical concept of relationship analysis. This is followed by construction of thinking process from the most simple building blocs of statistics – basic concepts and methods, gradually to more complex ones. When the technique is absorbed, we return to the relationship analysis. The realization of lessons is designed in modular form:

1. Introduction:
    - MODULE 1: Introduction to relationship analysis in public health, and use of statistical methods (objective: understanding basic concept of relationship analysis);
2. Principles of statistical methods:
    - MODULE 2: Statistical concepts and statistical describing of data (objective: understanding of basic statistical concepts, sample distributions and their meaning in statistical inference);
    - MODULE 3: Statistical inference (objective: getting familiar with techniques of statistical inference using worked examples);
3. Using statistical methods in relationship analysis:
    - MODULE 4: Principles of relationship analysis – univariate analysis (objective: transfer of basic principles of statistical methods to simple relationship analysis; students understand, and are able to perform analyses independently);
    - MODULE 5: Principles of relationship analysis - multivariate analysis (objective: extending the methods for simple relationship analysis to more complex ones, getting familiar with principles of multivariate relationship analysis; understanding technique of frequently used statistical methods in public health, i.e. logistic regression; students are able to perform analyses under the guidance of the teacher);
    - MODULE 6: Specific designs (e.g. repeated measurements analysis).

Based on our experiences, students must be given enough time for basic concepts. It is hard to say how much time would this take since groups of public health students are differing considerably. At the beginning of single course, the groups of students mostly are possible to be of very heterogeneous background knowledge. A good teacher shall be able to carry out the whole teaching matter with no regard to the emphasis laid to individual methods.

Teaching process

Teaching method

For successful outcome of teaching process it is essential to use teaching methods with emphasis on the active involvement of students in the process. The approach of learning with experience in small groups based on global worked example using high quality computer programme is used, combined with large group methods, being introductory lecture and demonstration (method demonstration, as well as result demonstration). Occasionally, "buzz" groups method is used.

Size of groups

In a large group, there are 12-18 students.

In each small group there are no more than three students with optimal number of two. Every single small group of students, together with a computer constitutes one working unit. The students are distributed considering following criteria, if possible:

- one student in a group is familiar with working on a computer;
- one student is not a beginner of a postgraduate specialistic study of public health.

This approach enables students to actively share their knowledge and help themselves inside each small group.

Other active involvement

Additionally, discussion inside groups is stimulated, as well as discussion between groups. Also, the comparison of results between working units, whenever possible, is stimulated.

Data material and teaching tools

Data material

For already several years at teaching statistical methods in public health in Slovenia data collection which enables learning such methods in quite a pleasant way has been used. These are the data collecting within the Perinatal Informational System of Slovenia (PISS) (44), which is considered to be one of the permanent collections of medical data of the highest quality with the many years' tradition. It was started in 1987 when collection of data started on a uniform form in all Slovene maternity hospitals.

Data material for teaching is only a small piece out of the whole collection PISS, prepared especially for this purpose. Safeguard of personal data is assured so that all personal identificators have been removed. Moreover, only the data are selected from the whole collection which shall be used for the teaching purpose. Special attention was paid to include such data, which are possible to play the role of the "effect" as well as the data which can play the role of "cause". Also, they are chosen to have diverse statistical characteristics as to be able to represent the majority of statistical methods.

The basic data material for all statistical activities is composed of more than 6000 statistical units, representing the model of a population. For teaching different statistical methods, samples of various sizes are then randomly selected from the

population database. The smallest sample is composed of 30 units and the largest one of 800 units. The students shall select some of such samples from the population while performing practical work by themselves and others shall be prepared by the teacher. First ones are, on principle, different at each working place and shall be, at the end of the corresponding exercise, deleted by the students or by the teacher. Others are the same for all working places, constantly present there and are used for many times. Data material is composed of the following permanent databases:

1. Population database;
2. Database with a sample of 100 units – main database for teaching principles of statistical methods;
3. Database with a sample of 500 units – additional database for teaching principles of statistical methods;
4. Database with a sample of 800 units – main demonstrational database and main database for teaching multivariate statistical methods.

Notwithstanding the number of units in a single database, all databases are including the same statistical variables.

On this data, most of statistical methods could be presented in an obvious way as well as numerous possible traps of their misuse.

Teaching tools

Adequate equipment is also very important for teaching statistical methods in a proper manner, which is carried out mostly by the means of a computer with a statistical program of suitable quality.

At the University of Ljubljana, the statistical program SPSS (Statistical Package for Social Sciences) has been used for this purpose for already a few decades. At the beginning, it was possible to be used only on main university computer. However, for the last about ten years the University has the licence to enable the distribution of the program within its members under favourable conditions for each personal computer. Two main purposes are so achieved: a program of good quality is used within the University, enabling the results to be compared within the university and as this program is widely spread also enabling the results to be compared to other countries.

However, whenever it is necessary for easier understanding of methods, the program can be supplemented by using a calculator and statistical tables.

Teaching material

On general, materials for modules are composed of:

1. Short summary of the subject covered by the unit
2. Exercises for practical work.
3. List of data files for practice.
4. Teaching material necessary for the tasks which are foreseen to be calculated manually.
5. Copies of the results of analysis carried out by the means of the SPSS program on a sample of 800 units.

Studying material

Studying material prepared for students, is composed of:

1. A short manual for working with the SPSS program, so that they were able to start working by themselves with a program from the very beginning.
2. Short summary of the subject, which was not intended as being the only source of statistical theory for the students, but to help directing their flow of thinking and as a help with practical work.
3. Collection of exercises for practical work.
4. Copies of the main results of the SPSS programme, so that the students are able to mark down their comments and explanations of printouts for helping them at studying.
5. Advised textbook for deepening their theoretical knowledge, which could be find in a Central Library of Faculty of Medicine or bought for a reasonable price.

Comments

Using described methods, we consider several principles of andragogy:

- students are actively involved in the educational process,
- teacher is a facilitator and not just a supplier of facts;
- use of life experiences and prior education is stimulated, as well as group interaction;
- feedback from teacher is given in few days;

The effectiveness of educational process is being evaluated through statistical analysis project, which is presented to other students.

EXERCISE

Introduction to Task 1: Micro-teaching exercise

Rationale

Lecture, the most often used method in teaching process, is only one of several ways of passing on factual information, giving an overview and motivating participants in teaching and learning process. Lectures should be carefully prepared, clearly structured and have defined SMART (specific, measurable, achievable, realistic and timing) aim and objectives. To lecture effectively requires that attention be paid to presentation skills and the use of any visual or technical resources.

Learning objectives

Learning objectives are:

1. Practice lecturing an 8-10 minute segment of a planned teaching exercise;
2. Analyse one's own and others' styles in a micro-teaching format.

Task 1

Participants:

- prepare an 8-10 minute lesson on topic of their own choice,

- they present the lecture to their colleagues using any format that they choose[32]. The lecture is videotaped. Both performance/presentation (verbal and non-verbal) is analysed as well as components of the lecture.

It is important to have a constructive feedback from the audience. Colleagues may be able to give an objective critique and provide positive advice on content and delivery. On the other hand, personal reflection may be undertaken by personal viewing a video of own presentation.

Introduction to Tasks 2-4: Working with small groups

Tasks 2-4 are basing on following assumptions:

- we have a large group of 24 students. We split this large group to 4 small groups of 6 students,
- the main teaching matter is health promotion. Accordingly, all tasks are basing on teaching modules, published in the manual for teachers in PH-SEE network "Health promotion and disease prevention. A handbook for teachers, researches, health professionals and decision makers" (45),
- students have already read the theoretical background of this module.

Task 2

Each group is given a theme to be virtually taught from above mentioned manual (Table 6):

[32] Important notes in lecture delivery:

1. Lecture is often arranged in three sections: introduction, body and conclusion.
2. Introduction provides motivation and enthusiasm to inspire the audience to concentrate for the main body of the lecture. It could be useful to provide a statement of the objective(s) of the lecture (e.g. „At the end of this lecture you should be able to....") and indicate the key points which are to be noted on the way (outline).
3. It is crucial to the success of the lecture that you maintain the audience's attention. Asking yourself the following questions will make this easier to accomplish:
   - Where should I stand?
   - How should I speak?
   - What should I do?
   - When should I change style?
   - How to communicate with the audience?
4. Any additional materials (illustrations, handouts, PowerPoint or other multimedia presentation) presented in a lecture must be of good quality.
5. Interruptions and technical faults provide a test of the lecturer's skills. It may be wise to work out in advance a strategy for dealing with some of the following before being faced with them: uninvited questions, participants falling asleep, technical failures, etc.

Table 6. Themes to be virtually taught.

Group	Theme	Chapter in the manual
Group 1	Concepts and Principles in Health Promotion	1.1 Concepts and Principles in Health Promotion
Group 2	Healthy Public Policy	1.2 Healthy Public Policy
Group 3	Media and Health	4.3.2 Media and Health
Group 4	Smoking cessation campaigns	5.6.1 Quit and Win Campaign

Students are obliged to read the Theoretical background parts of proposed modules to be familiar with the contents. They can also read the Case study part, but they are supposed to omit the Exercise part.

Task 3

Every group is proposed to use different teaching method for small groups to perform. In Table 7, the teaching methods to be used are listed.

Table 7. Teaching methods to be used.

Group	Teaching method	Tips
Group 1	Snowball groups	Snowballing is started in this case from single student, than form twos, and at the end make plenary session.
Group 2	Fishbowl	Three students are in the inner circle, and three in the outer.
Group 3	Crossover groups	Students are labeled as A1, A2, B1, B2, C1, and C2. First all As, Bs, and Cs work together, and then all 1s, 2s, and 3s work together
Group 4	»Pro et contra« discussion	From the group of 6 students, two groups of 3 students are formed. One group prepares arguments in favour of public health campaigns, in the other group are the opponents.

Students at the plenary session perform indicated teaching methods in front other 18 students.

Task 4

From the manual (45), each group chose another module and propose a teaching method that is according to their opinion the most suitable. Afterwards they present to other students in short the theme they have chosen and the proposed teaching method. Other students discuss

REFERENCES

1. Krajnc A. Methods, forms, and means in adult education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.101-221.
2. Oxford Advanced Learners Dictionary. Seventh Edition. Oxford: University Press; 2005.
3. Poljak V. Didactics (in Slovene). Ljubljana: Državna založba Slovenije; 1974.
4. Tomić A. Selected didactical topics (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani; 2001.
5. The Online Etymology Dictionary. Available from URL: http://www.etymonline.com/index.php. Accessed: July 24, 2008.
6. Merriam-Webster Online Dictionary. Available from URL: http://www.merriam-webster.com/dictionary/medicine. Accessed: July 24, 2008.
7. TheFreeDictionary. Available from URL: http://www.thefreedictionary.com/. Accessed: July 24, 2008.
8. Ličen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
9. Reischman J. Andragogy. History, Meaning, Context, Function. Available from URL: http://www.uni-bamberg.de/fileadmin/andragogik/08/andragogik/andragogy/index.htm. Accessed: July 24, 2008..
10. Smith MK. "Andragogy", the encyclopaedia of informal education. Last update 1999. Available from URL: http://www.infed.org/lifelonglearning/b-andra.htm. Accessed: July 24, 2008.
11. Knowles MS, Holton EF, Swanson RA. The adult learner: the definitive classic in adult education and human resource development. 6th ed. Amsterdam: Elsevier, 2005.
12. Burge L. Beyond andragogy: Some explorations for distance learning design. J Dist Educat, 1988. Available from URL: http://cade.athabascau.ca/vol3.1/burge.html. Accessed: July 24, 2008.
13. Collins J. Education techniques for lifelong learning. Principles of adult learning. RadioGraphics 2004;24:1483–1489. Available from URL: http://radiographics.rsnajnls.org/cgi/reprint/24/5/1483. Accessed: July 24, 2008.
14. Imel S. Teaching adults: is it different? ERIC Digest No. 82. Available from URL: http://www.ericdigests.org/pre-9211/teaching.htm. Accessed: July 24, 2008.
15. Krajnc A. Andragogical didactics. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.13-38.
16. UNDP. Capacity development resources book. New York: UNDP, 1997. Available from URL: http://mirror.undp.org/magnet/cdrb/Default.htm. Accessed: July 24, 2008.
17. UNDP. Capacity assessment and development, in a systems and strategic management context. Technical advisory paper 3. New York: UNDP; 1998. Available from URL: http://mirror.undp.org/magnet/Docs/cap/CAPTECH3.htm. Accessed: July 24, 2008.
18. UNESCO. UNESCO thesaurus. Available from URL: http://databases.unesco.org/thesaurus/. Accessed: July 24, 2008.
19. International Institute for Educational Planning. Guidebook for planning education in emergencies and reconstruction. Paris: UNESCO; 2006. Available from URL: http://www.unesco.org/iiep/eng/focus/emergency/guidebook.htm. Accessed: July 24, 2008.
20. Krajnc A. Experience of adults and education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost, 1979; p.69-82.
21. Krajnc A. Motivation of adults for education. In: Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.83-100.

22. Ruzgiene L. Strategies and reflections upon adult education. In: Wie AL (editor). Improving adult learning. Nesna: Hogskolen I Nesna; 2003. p.14-22. Available from URL: http://www.hinesna.no/system/files/bilder/57.pdf. Accessed: July 24, 2008.
23. Jakšić Ž. Student-centred teaching. In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching in medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.87-95.
24. Lieb S. Principles of adult learning. Available from URL: http://honolulu.hawaii.edu/intranet/committees/FacDevCom/guidebk/teachtip/adults-2.htm. Accessed: July 24, 2008.
25. O'Brien G. What are the principles of adult learning? Available from URL: http://www.southernhealth.org.au/cpme/articles/adult_learning.htm. Accessed: July 24, 2008.
26. Levine SJ. Effective teaching. A set of papers focusing on the teaching of technical information to adults. In: Michigan State University. Approved trainer of registered technician self study manual. Available from URL: http://www.pested.msu.edu/PesticideApplicatorCertification/RTAT%20Manual/RTATSection7.pdf. Accessed: July 24, 2008.
27. Bryan RL, Kreuter MW, Brownson RC. Integrating Adult Learning Principles Into Training for Public Health Practice. Health Promot Pract OnlineFirst; April 2, 2008 as doi:10.1177/1524839907308117. Available from URL: http://hpp.sagepub.com/cgi/rapidpdf/1524839907308117v1.pdf. Accessed: July 24, 2008.
28. European Commission Public Health Executive Agency. Conference on professionalization and capacity building in public health in South-Eastern and Eastern Europe: the legal and educational framework, Luxembourg, 20/21 February 2008. Draft report. Luxembourg: PHEA, 2008.
29. Zupančič M. The subject and history of developmental psychology. In: Marjanovič-Umek L, Zupančič M, editors. Developmental psychology (in Slovene). Ljubljana: Znanstvenoraziskovalni inštitut Filozofska fakultete; 2004. p.6-27.
30. School of Medicine Rijeka. Study of organization, planning, and management in health care (in Croatian). Available from URL: http://www.medri.hr/studiji/organizacija/o%20studiju.htm. Accessed: July 24, 2008.
31. University of Bielefeld, Faculty of Public Health. Bachelor of Science in Health Communication. Available from URL: http://www.uni-bielefeld.de/gesundhw/english/academic/bhc/index.html. Accessed: July 24, 2008.
32. Loginova N. Adult education methods and methodologies. In: Wie AL, editor. Improving adult learning. Nesna: Hogskolen i Nesna, 2003. p.30-42. Available from URL: http://www.hinesna.no/system/files/bilder/57.pdf. Accessed: July 24, 2008.
33. Piderit GA, Quijano LF. What are the main techniques used in adult education? Available from URL: http://fcis.oise.utoronto.ca/~daniel_schugurensky/faqs/qa14.html. Accessed: July 24, 2008.
34. Puklek Levpušček M, Marentič Požarnik B. Work in small groups for active study (in Slovene) Ljubljana: Filozofska fakulteta, Center za pedagoško izobraževanje; 2005.
35. Vrcić-Keglević M. Learning in small groups, seminars. In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.87-95.
36. Seaman DF. Adult education teaching techniques. College Station, TX: A&M University; 1977. Available from URL: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/35/8b/79.pdf. Accessed: July 24, 2008.

37. Pavleković G. Lecture. In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The Art of Teaching in Medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.145-151.
38. Cantillon P. ABC of learning and teaching in medicine: Teaching large groups. Br Med J. 2003;326:437-440. Available from URL: http://www.bmj.com/cgi/reprint/326/7386/437. Accessed: July 24, 2008.
39. Jaques D. ABC of learning and teaching in medicine: Teaching small groups. Br Med J. 2003;326:492-494. Available from URL: http://www.bmj.com/cgi/reprint/326/7387/492. Accessed: July 24, 2008.
40. Borrows HS, Tamblyn RM. Problem-Based Learning. New York, Springer Publishing Company; 1980.
41. Schmidt HG. Foundations of problem-based learning: some explanatory notes. Medical Education. 1993;27:422-32.
42. Fosse M. PBL. In: ETC-PHHP/EUMPH Summer School: Strategies for Health in Europe. Readings. Galway: ETC-PHHP; 2006.
43. National Training Laboratories. Learning pyramid. Bethel, MA: National Training Laboratories. Available from URL: http://www.tcde.tehama.k12.ca.us/pyramid.pdf. Accessed: July 24, 2008.
44. Institute of Public Health of the Republic of Slovenia. Perinatal Informational system of the Republic of Slovenia. Available from URL: http://www.ivz.si/index.php?akcija=podkategorija&p=44. Accessed: July 24, 2008.
45. Donev D, Pavleković G, Zaletel-Kragelj L, editors. Lage: Hans Jacobs Publishing Company, 2007. p.378-395 Available from URL: http://www.snz.hr/ph-see/Documents/Publications/FPH-SEE_Book_on_HP&DP.pdf. Accessed: July 24, 2008.

RECOMMENDED READINGS

1. Cantillon P. ABC of learning and teaching in medicine: Teaching large groups. Br Med J. 2003;326:437-440. Available from URL: http://www.bmj.com/cgi/reprint/326/7386/437. Accessed: July 24, 2008.
2. Dent JA, Harden RM, editors. A practical guide for medical teachers. London: Elsevier; 2005.
3. Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching in medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005.
4. Jaques D. ABC of learning and teaching in medicine: Teaching small groups. Br Med J. 2003;326:492-494. Available from URL: http://www.bmj.com/cgi/reprint/326/7387/492. Accessed: July 24, 2008.
5. Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979.
6. Wie AL, editor. Improving adult learning. Nesna: Hogskolen i Nesna, 2003. p.30-42. Available from URL: http://www.hinesna.no/system/files/bilder/57.pdf. Accessed: July 24, 2008.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**TELE-EDUCATION AS A NEW METHOD OF MEDICAL AND PUBLIC HEALTH EDUCATION**
Module: 5.2.2	**ECTS (suggested): 0.20**
Author(s), degrees, institution(s)	**Izet Mašić** MD, PhD, Professor Chair for Medical Informatics, Medical faculty, University of Sarajevo, Bosnia and Herzegovina
Address for correspondence	**Izet Mašić** Chair for Medical Informatics, Medical faculty, University of Sarajevo, Bosnia and Herzegovina E-mail: imasic@lol.ba
Keywords	Distance learning, Bosnia and Herzegovina
Learning objectives	After completing this module students should: • know the definition and characteristics of tele-education; • be familiar with advantages and disadvantages of tele-education in comparison to classical methods of education.
Abstract	Increase and development of distance learning (DL) technologies over the past decade has exposed the potential and the efficiency of new technologies. Benefit and use of contemporary information technologies is the area where medical informatics got the most on understanding and importance. Definition of DL as "use of technologies based on health care delivered on distance" covers areas such as electronic health, telehealth (e-health), telematics, telemedicine, tele-education, etc. as web based education. For the need of DL there are various technologies and communication systems from standard telephone lines to the system of transmission digitalized signals with modem, optical fiber, satellite links, wireless technologies, and wed sites, etc. Web based education represents health education on distance, using IC technologies and Internet, as well as continuous education of a health system beneficiaries and use of electronic libraries, data bases or electronic data with data bases of knowledge.
Teaching methods	An introductory lecture gives the students first insight in characteristics of distance education. The theoretical knowledge is illustrated by a case study. After introductory lectures students first carefully read the recommended readings. Afterwards they discuss the characteristics of distance education with other students.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Multiple choice questionnaire.

TELE-EDUCATION AS A NEW METHOD OF MEDICAL AND PUBLIC HEALTH EDUCATION

Izet Mašić

THEORETICAL BACKGROUND

Distance learning - definition and description

Distance learning is conventionally defined as: any educational or learning process or system in which the teacher and instructor are separated geographically or in time from his or her students; or in which students are separated from other students or educational resources (1-3). The most important factor which influences the changes occurring in education has been the installation and development of the Internet and electronic multimedia techniques. Distance learning does not preclude traditional learning processes; frequently it is used in conjunction with in-person classroom or professional training procedures and practices.

Distance learning is used for self-education, tests, services and for the examinations in medicine, i.e. in terms of self-education and individual examination services .The possibility to work in the exercise mode will image files and questions is an attractive way for self-education (4-7). The standard format of the notation files enables to elaborate the results by commercial statistic packets in order to estimate the scale of answers and to find correlation between the obtained results. The method of multi-criterion grading excludes unlimited mutual compensation of the criteria, differentiates the importance of particular courses and introduces the quality criteria. By using computers and teleconferencing technology and through partnerships with local communities, institutions and the private sector, an open, effective, virtual learning community is now in place. Sites are located in college and university campuses, hospitals, schools, libraries, community centres and private companies. Courses are also being delivered to private homes.

For the need of e-health, telemedicine, and tele-education there are various technologies and communication systems from standard telephone lines to the system of transmission digitalized signals with modem, optical fibber, satellite links, wireless technologies, etc. There is no doubt that Internet causes “revolution” in all above, and the latest its possibilities are distribution of virtual medical instruments and medical data in real time and possibility of use in primary health care, even for some diseases with bed prognosis. This revolution how information is stored, transmitted and accessed has extremely important implication for the health sector, especially now when embarking on a global effort to renew the tenets of Health for All based on primary health care and disease prevention, health promotion and costumer education, in the context of service delivery guided by the equity, quality, effectiveness and efficiency. According to Grimson at al in Dublin, “the need to participate in continuing professional development or continuing medical education is considered to be at the very least highly desirable and more likely mandatory. One way by which this can be facilitated in a timely and cost-effective manner” is the use of Information Communication Technologies (ICT) (7-10).

Traditional way of learning and learning from the distance

The latest researches shows that the format of instructions itself has no important influence on the students' achievements if access and availability to information technologies is assured as well as usage of the adequate content of education. In the assessment of the authentic situation the following issues should be addressed:

- results of different tests prepared by lectors has trend to show advantages in comparison with traditional learning methods and there is significant distinction in affirmative attitude to educational materials between distance and traditional learning,
- traditional methods demonstrate better organization and they are clearer in respect to distance learning,
- organization and needs for more efficient influence of distance learning very often improve traditional methods by teachers,
- future research should be focus on critical factors in determining student involvement in development of educational process,
- the variety of teaching and learning options provided by technology allows education to be provided in an appropriate manner to a broader student demographic then ever before.

Facts about distance learning and tele-education

Distance learning enables permanent learning (lifelong learning), students can improve themselves professionally and independently, at their own tempo, at place and time that they choose by themselves, they can choose great deal of subjects which offer different institutions, teachers-individuals; students go through materials for learning by speed of their own and as many times as they want. The place can be chosen – it depends on media which is used for learning material (they can learn at work or from home). Themes access which are not offered by studies in that field – students find and attend the programs which they are interested in, although they are not offered by educational or business institutions in place where they live in or work. Taking part in top-quality and most prestigious programs – student can "attend" at least some studies at the top-quality institutions or studies held by lecturers that are very famous experts without changing their place of living. Choosing this way of learning – active or passive learning, different kinds of interaction: "Classical" written material and writing down their own lecture notes, interactive simulations, discussion with other students (e-mail, tele-conferences).

Practical work with different technologies – they get not just information about that they learn, but additional knowledge and skills about using computer, CD players and video recorders. Independent learning – teachers learn too from students who independently ask for information source.

The meaning of education/learning to distance can be expressed by the definition: that is a form of education which is in process permanently, or most of the time, all or most of the tasks of teaching and learning separately during the time and space between teacher and student.

Pedagogical and organizational improvements have fundamental importance. It is in use both interaction teacher-student and interaction student-student. Phases of synchronized and synchronized learning are combined. Individual and group works are also combined. If all these forms are involved in educational process, they mutually supplement each other, as a last resort. Traditional education as well as contemporary education is supported by informatics technologies in unique system of flexible education. In order to use advantages of flexible education, it is necessary to combine different forms of learning, during the preparation phase and development of every educational course in appropriate way.

Distance learning is not simply a set infrastructure, but rather a concept of learning that incorporates different technologies and learning media. Within the province, different video, audio and computer tele-conferencing systems, along with Computer Based Training, Computer Managed Instructional systems and other media are being integrated technologically, instructionally and organizationally. The tele-education concept crosses all jurisdictions among institutions both within and outside the province, public and private, at any level of education, to anywhere including institutions, workplaces and the home. Tele-education, tele-teaching, tele-training, tele-mentoring, and tele-accreditation have been clearly demonstrated and are now common practice.

Advantages and disadvantages of distance learning

Distance learning compared to traditional way of learning had many its advantages as well as disadvantages. Some of the main advantages of distance learning are:

- the economical factor,
- student has 24 hour access to needed information,
- he/she is given the opportunity to learn the subject in his/hers own time and speed,
- he/she can access learning material independently of place or time,
- he/she is given the opportunity to learn how to work independently,
- using e-mail or chat rooms he/she is able to contact professor or his/hers assistant if there are any questions or confusions regarding lectures; etc. (2,3).

Fundamental advantages of flexible education in terms of classical education are:

- more efficiency;
- increase capacities of educational institutions;
- education can be easily adopted to the needs of education on-the-job;
- costs of educational process are smaller;
- it is possible to distribute the education uniformly, thus the new educational programs are available for fields outside of educational and economic centres;
- it enables the possibility of access to the foreign educational resources to the various institutions;
- superior quality of the knowledge gained.

Many critics consider that using e-mail or chat rooms to obtain contact with the professor is actually the main disadvantage of this system of learning. Question arises whether this way of professor-student communication is helpful to students because face-to-face contact is missing as well as the opportunity of student-professor relationship building.

CASE STUDY: DISTANCE LEARNING IN MEDICAL CURRICULUM AND IMPLEMENTATION OF DISTANCE LEARNING AT MEDICAL FACULTY, UNIVERSITY OF SARAJEVO, BOSNIA AND HERZEGOVINA

Background

By late 1994, the Internet includes 3.2 million computer nodes spread across more than 57,000 institutions in more then 80 countries, with an estimated 30 million users. By the end of the century the Internet linked more then 400 million persons.

The 2002 Eurobarometer survey showed that an average of 78% of EU medical GPs were online, with at the highest level – 98% in Sweden and 97% in United Kingdom. Number of "online patients" grows every day as well. The 2003 Eurobarometer survey on health information sources shows that 23% of Europeans use Internet for health information and that 41% of the European population considers that Internet is a good source of information on health (3).

In spite of fact that Bosnia and Herzegovina (B&H)is last or second before last country in Europe in use of Internet technologies, there is a group of enthusiastic people accompanying Prof. Izet Masic, MD, at Medical Faculty and University of Sarajevo who have been making significant effort to improve poor digital literacy in the University and among medical professionals. E.g. every 60^{th} citizen of B&H uses Internet (50.000 citizens of 3.000.000 citizens in B&H) while in Slovenia every forth Slovenian is familiar and uses Internet technologies for example.

In October 2003, University of Sarajevo began with Distance learning education, opening University Distance Learning Centre. Opening the University Distance Learning Centre, as coordination body and leader in all activities in connection to Distance learning, has provided opportunity for development and growth of this kind of lifelong education.

In relation to above project conducted by the University Tele-information Centre (UTIC) and as continuation of two-year project Possibilities of introduction of Distance learning in Medical curriculum, the Cantonal Ministry approved and supported a new project; Introduction and implementation of Distance learning in medicine. Platform for the course of distance learning is achieved in collaboration with UTIC. University Tele-information Centre, established as part of University of Sarajevo and first ISP in B&H in 1996 (11). It is scientific-organizational unit of the University of Sarajevo for improvement of scientific-research work and through UTIC members of the University can be

gathered in the unique computer-communication structure. Objectives of UTIC are:

- to connect members of the University of Sarajevo with similar institution in the country and abroad due to more efficient use of scientific, research and educational resources,
- use of educational data bases and other information for the needs of the University and its members,
- development an integration of informatics computer technologies in education,
- creation of flexible infrastructure which will enable e-Learning to be accessible to all students and University staff,
- improvement of general digital literacy of academic population,
- development of top quality educational content which could be integrated in the actual European processes of e-Learning revolution.

With their help centre for distance learning, "LUCIS CENTRUM", is created (Figure 1). We hope that this is just a beginning step towards improvements of the B&H education system and that this project will serve as an indicator towards that future.

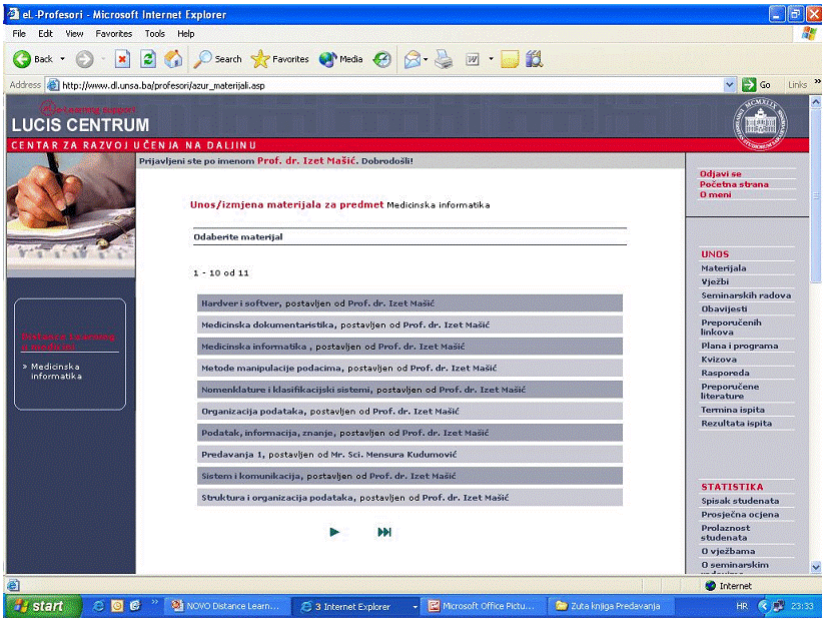


Figure 1. Uploaded materials for subject Medical Informatics at the web page of Lucis Centrum of University of Sarajevo.

On UTIC web site, seven students enrolled from Medical faculty, for the subject Medical Informatics are able to learn from the distance location. So far, teaching staff uploaded eleven lectures at the web site:

1. Hardware and software,
2. Medical documentations,
3. Medical informatics,
4. Methods of data manipulation,
5. Nomenclatures and classification systems,
6. Data organization,
7. Data, information and knowledge,
8. Lectures 1,
9. System and communication,
10. Structure and data organization, and
11. Expert systems.

Beside the materials it is possible to upload and download the following:

- practical works,
- seminar work,
- information,
- recommended links,
- plan and programs,
- quiz,
- schedule,
- recommended readings,
- examination schedule, and
- examination results.

Basically software application has two interfaces: teacher and student interface. Access from any of these is very simple and fast (1-3).

Education content of distance learning

Lecture contents will be presented in our virtual classroom. In our case, learning material from the subject of medical informatics, and later, hopefully from other medical subjects, will be available on web site, www.e-learning.ba (Figure 2).

In this "classroom," learning materials, power point lecture presentations as well as practice exercises with step-by-step instructions, are easily accessible to students. Moreover, on this web site, students will be able to find subject relating literature as well as English version of the presentations (Figure 2). To access this information requires only one click on a download option (Figure 3) as well as one second patience depending on student's internet speed connection. In short, our virtual classroom gives students the opportunity to access needed information, at any time, and in any place without having to be bound to the classroom.

Figure 2. Content of Medical Informatics page at at the web page of Lucis Centrum of University of Sarajevo.

Student is able to browse through the "classroom" using standard navigations (Figure 3):

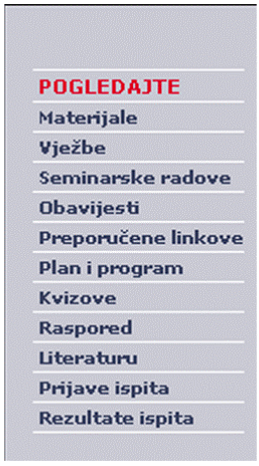


Figure 3. Navigator bar at the web page of Lucis Centrum of University of Sarajevo..

Using these simple navigations, maximal efficiency and fast access to needed material is possible.

As we can see on the Figure 3, all links are in chronological order according to the plan and program of the lecture as well as practice (materials, practical works, seminar work, information, recommended links, plan and programs, quiz, schedule, recommended readings, registration to examination, and examination results).

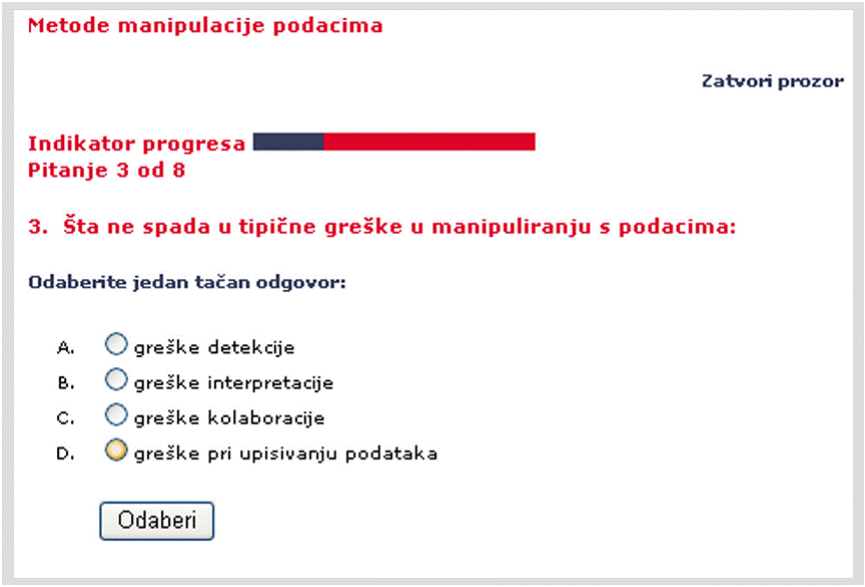


Figure 4. An example of a "Quiz" window at the web page of Lucis Centrum of University of Sarajevo.

Special attention is given to the link "kvizovi" (quiz). In order for the student to check his/hers progress (to test his/hers knowledge of the lecture he/she studied), every lecture is followed by quizzes. Quiz questions are multiple choice questions (Figure 4) and they are based on the lecture content. After every quiz, student receives "feedback" regarding his/hers progress. Results are given in terms percentages (one needs 51% of correct answers to pass the quiz). On this way student has absolutely control over his/her work.

Advantages and disadvantages of distance learning in Bosnia and Herzegovina

In many universities across B&H students' contact with professors are almost impossible (unless one needs to orally take the exam), due to many professors other jobs or responsibilities; students are mainly able to communicate with professor's assistants. Moreover, thru traditional way of teaching, during the lectures, students from their professors obtain mostly the information which they can find in the literature or on the internet. Rarely, there is student-professor

interaction or lecture discussion during the class. From this one can conclude that an ambitious student using tele-education will experience minimum lose.

We live in "Age of Information". These technologies are changing the way we perceive the world, how wee think and communicate with another. Established cultures are being transformed and new cultures are forming. New virtual environment affects the way we build our sense of who we are. Some characteristics of the Internet with which are people in B&H know a little are:

- large volume of users and potential users,
- lack of physical boundaries which allows for the manipulation of time and space,
- information can be accessed in a concurrent fashion using different media,
- concept of redundancy.

In the virtual environment we are applying for information in a way that is expanding our senses and one must to take into account that experience is occurring in the context of the virtual environment. Information without a context has no meaning.

Our experiences in application of tele-education at biomedical universities

The greatest progress was made in the area of tele-education and distance learning in B&H. Distance learning does not preclude traditional learning processes; frequently it is used in conjunction with in-person classroom or professional training procedures and practices. Distance learning is used for self-education, tests, services, and for the examinations in medicine, i.e. in terms of self-education and individual examination services. The possibility to work in the exercise mode with image files and questions is an attractive way for self-education (12-15).

Very first serious initiative was generated by World University Service of Austria (WUS Austria) in B&H. During 2002 and 2003 WUS Austria, through its programs, Distance learning 2002 and Distance learning 2003 year, supported the development of the educational processes at B&H universities. At Medical faculty of University of Sarajevo at Chair for Medical informatics since 2002 is in progress realization of the project named: "Possibilities of introducing of distance learning in medical curriculum", approved by the Federal and the Cantonal ministry of science and education. The purpose of the project is to facilitate improvement of educational process at biomedical faculties, applying contemporary educational methods, methodologies and information technologies in accordance with strategy and goals proclaimed by Bologna declaration. Pilot project was realized during school years 2003-2005, theoretical and practical education of subject Medical informatics are adapted to the new concepts of education using world trends of education from the distance. One group of students was included in the project finalized by electronic exam registration and electronic exam on 20 June 2005, publicly, in the Physiology amphitheatre of the Medical faculty in Sarajevo (Figure 5).

Figure 5. First electronic exam and held on 20 June 2005, publicly, in the Physiology amphitheatre of the Medical faculty in Sarajevo, Bosnia and Herzegovina.

Bologna process, which started across European countries, provides us to promote and introduce modern educational methods of education at biomedical faculties in B&H. Chair of Medical informatics and Chair of Family medicine at Medical Faculty of University of Sarajevo started to use web based education as common way of teaching of medical students. Satisfaction with this method of education within the students is good, but not yet suitable for most of medical disciplines at biomedical faculties in B&H. Web sites of Chair for Medical Informatics and Family Medicine and UTIC are shown in the Figure 6.

Bosnian Society for Medical Informatics (BHSMI) is very proactive in promoting telemedicine and tele-education as part of it. The last event organized by BHSMI was Special Topic Conference named e-Health and e-Education held in December 2005 (Figure 7).

Distance learning in medicine has impact on telemedicine and practicing medicine as well. Basic skills of the use of computers and networks must be a part of all future medical curricula. The impact of technical equipment between patient and the doctor must be understood, and the situation where the diagnosis based on live voice or picture is different from a normal doctor-patient contact. In some areas telemedicine requires unique techniques. Tele-robotically discipline guaranties differ from what surgeons normally learn. Telemedicine, and distance learning as a

prerequisite for it, is lest suited for doctor-to-doctor consultation, and the first contact to a doctor should always be a face-to-face consultation (15,16).

Figure 6. Web sites of Chair for Medical Informatics, the University Tele-information Centre of University of Sarajevo, and Chair for Family Medicine at Medical faculty of the University of Sarajevo.

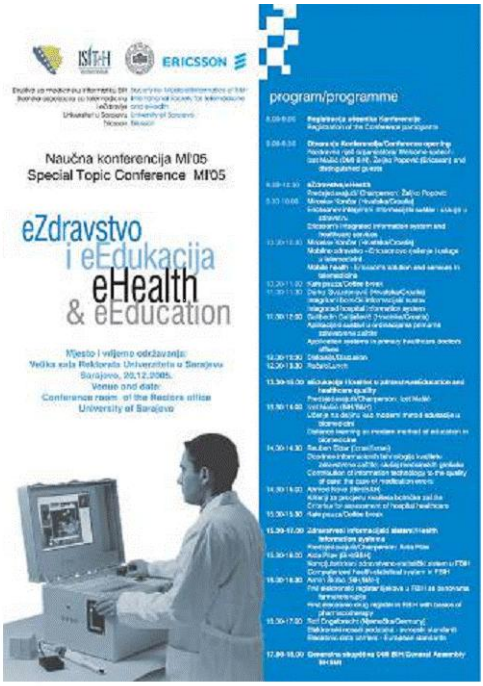


Figure 7. Special Topic Conference – e-Health and e-Education.

Conclusions

Distance learning in medicine has impact on telemedicine and practicing medicine as well. Basic skills of the use of computers and networks must be a part of all future medical curricula. The impact of technical equipment between patient and the doctor must be understood, and the situation where the diagnosis based on live voice or picture is different from a normal doctor-patient contact (10). In some areas telemedicine requires unique techniques. Tele-robotical guaranties differ from what surgeons normally learn. Telemedicine, and distance learning as a prerequisite for it, is lest suited for doctor-to-doctor consultation, and the first contact to a doctor should always be a face-to-face consultation.

Expected outcomes of the project Introduction and Implementation of Distance learning in medicine are:

- development and integration of informatics-computer technologies in medical education,
- creation of flexible infrastructure which will enable access to e-Learning by all students and teaching staff,
- improvement of digital literacy of academic population,
- ensure high educational standards to students and teaching staff, and
- to help medical staff to develop "Lifelong learning way of life".

The health sector is one of the most evident potential beneficiaries of the Internet revolution and World Wide Web resource in the present and in the future, when the tools now available and the system's reliability and efficacy as a whole will be further incremented and improved.

In order to have Distance Learning minimum of requirements should be in place: Minimal infrastructure: equipment and software, from laptop to notebook, projector and scanner to reliable and quality software, appropriate marketing in Bosnia and Herzegovina environment, education of the educators and administrators, Electronic archive (and digital library), support from the officials (moral and financial) and sustainable funding.

Distance learning in medicine is a project started by Prof. Izet Masic, MD, and his team at Chair of Medical Informatics of Faculty of Medicine at University of Sarajevo in year 2002 as first pilot study of introduction of Distance learning education at biomedical faculties in Bosnia and Herzegovina for improving the educational system in our country (17). Distance learning is a trend used in many developed countries spreading fast throughout the rest of the world. We hope that this will be just a starting ground to our attempts of improvement of our education system. This is for the first time in B&H that students are able to apply for exam, have exam and results on the web site as well as his/her picture and all other significant data.

EXERCISE

Task 1

Did you experience e-learning by yourself in the past? Please, do the following task:

- if yes, write down your experiences (positive and negative),
- if not, write down your expectations of e-learning,

- make groups of three to four students and discuss your experiences or expectations,
- make conclusions of your discussion,
- present these conclusions on a plenary session to other groups.

Task 2

Carefully read the theoretical background on this module and theoretical background of the module Public health capacity building: adult education principles and methods of this book. Please, do the following task:

- discuss the advantages and disadvantages of tele-education in comparison to other types of education in a group of three to four students,
- make conclusions of your discussion,
- present these conclusions on a plenary session to other groups.

Task 3

Using the ICT, try to find out if in your country the distance learning at the university level is available to the students. Present results of your "research" to other students.

Task 4

From the Internet download the following free eBook:

Preston DS, Nguyen TH, editors. Virtuality and education:A reader. Oxford: Inter-Disciplinary Press; 2004. Avialable from URL: http://www.inter-disciplinary.net/publishing/id-press/ebooks/virtuality-and-education-a-reader/. Accessed: June 30, 2009.

Make 5 small groups of students and on the basis of discussion papers in this book, please:

- find and discuss case studies of distance learning (tele-education),
- discuss emergent skills in higher education,
- discuss professional challenges and pedagogical opportunities pertaining to virtuality in higher education,
- dymensions of e-learning,
- evaluation of e-learning.

REFERENCES

1. Mašić I, Novo A, Kudumović M, Mašić Z. Education of medical informatics at medical faculties in Bosnia and Herzegovina (in Bosnian). Acta Inform Med. 2005; 13: 3-8.
2. Masic I, Novo A, Kudumovic M, Masic Z. E-learning at Medical Faculty of University of Sarajevo. Acta Inform Med. 2005; 13: 132-5.
3. Masic Z, Novo A, Masic I, Kudumovic M, Toromanovic S, Rama A, Dzananovic A, Bander I, Masic M, Guso E, Balta E. Distance Learning at Biomedical Faculties in Bosnia and Herzegovina. Connecting medical informatics and bio-informatics. In: Engelbrecht R,

Geissbuhler A, Lovis C, Mihalas G, editors. Procedings of MIE 2005. Amstedram: IOS Press; 2005. p.267-72.
4. Engelbrecht R, Ingenerf J, Reiner J. Educational standards - terminologies used. Stud Health Technol Inform. 2004;109:95-113.
5. Masic I, Ramic-Catak A, Kudumovic M, Pasic E. Distance learning in the medical education in B&H In: E-Health & Education Proceedings. Zagreb: E-Health & Education Proceedings; 2002. p.17.
6. Masic I, Bilalovic N, Karcic S, Kudumovic M, Pasic E. Telemedicine and telematic in B&H in the war and post war times. Eur J Med Res. 2002;7(supl.1): 47.
7. Hovenga J. Bricknell L. Current and future trends in teaching and learning. Stud Health Technol Inform. 2004;109:131-142.
8. Mantas J. Future trends in health informatics - theoretical and practical. Stud Health Technol Inform. 2004;109:114-127.
9. Mantas J. Comparative educational systems. Stud Health Technol Inform. 2004;109:8-17.
10. Sosa-Iudicissa M, Oliveri N, Gamboa CA, Roberts J. Internet, telematics and health. Amsterdam: IOS Press, 1997.
11. University tele-informational centre (UTIC) (in Bosnian) [homepage on the Internet]. Available from URL: www.utic.net.ba. Accessed: June 10, 2009.
12. Zimitat C. Designing effective on-line continuing medical education. Med Teach. 2001;23:117-122.
13. Harden RM, Hart IR. Centre for Medical Education, University of Dundee, UK. An international virtual medical school (IVIMEDS): the future for medical education? Med Teach. 2002;24:261-7.
14. Harden RM. A new vision for distance learning and continuing medical education. J Contin Educ Health Prof. 2005;25:43-51.
15. Mašić I, Riđanović Z. Informational technologies in medicine and health care (in Bosnian). Acta Inform Med. 1995;3:5-8.
16. Google.com [homepage on the Internet]. Available from URL: www.google.com. Accessed: June 10, 2009.
17. Medicinski fakultet Univerziteta u Sarajevu. Medical informatics (in Bosnian) [homepage on the Internet]. Available from URL: http://www.unsamedinfo.org/index.php?NAV=opci_podaci. Accessed: June 10, 2009.

RECOMMENDED READINGS

1. Harden RM, Hart IR. Centre for Medical Education, University of Dundee, UK. An international virtual medical school (IVIMEDS): the future for medical education? Med Teach. 2002;24:261-7.
2. Harden RM. A new vision for distance learning and continuing medical education. J Contin Educ Health Prof. 2005;25:43-51.
3. Masic I, Ramic-Catak A, Kudumovic M, Pasic E. Distance learning in the medical education in B&H In: E-Health & Education Proceedings. Zagreb: E-Health & Education Proceedings; 2002. p.17.
4. Mašić I, Riđanović Z. Informacijske tehnolgije u medicini i zdravstvu; Acta Inform Med. 1995;3:5-8.
5. Zimitat C. Designing effective on-line continuing medical education. Med Teach. 2001;23:117-122.

METHODS AND TOOLS IN PUBLIC HEALTH A Handbook for Teachers, Researchers and Health Professionals	
Title	**DESIGNING AND PLANNING EDUCATIONAL PROGRAMMES IN PUBLIC HEALTH**
Module: 5.2.3	**ECTS (suggested): 0.30**
Author(s), degrees, institution(s)	**Gordana Pavleković,** MD, PhD, Professor Andrija Štampar School of Public Health, Medical School, University of Zagreb, Croatia **Lijana Zaletel-Kragelj**, MD, PhD, Associate Professor Chair of Public Health, Faculty of Medicine, University of Ljubljana, Slovenia **Anja Kragelj**, Student of Andragogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia **Nataša Škerget**, Student of Pedagogy at Department of Pedagogy and Andragogy, Faculty of Arts, University of Ljubljana, Slovenia
Address for correspondence	**Gordana Pavleković** Andrija Štampar School of Public Health Rockefellerova 4, 10000 Zagreb, Croatia E-mail: gpavleko@snz.hr
Keywords	Adult education, educational programme, planning
Learning objectives	After completing this module students should: • know the definition and characteristics of educational programme; • be familiar with different models of educational programme planning; • be familiar with educational programme planning process.
Abstract	Programme planning is a delicate phase in educational process. If by education we mean the provision of systematic organized learning experience, programme planning is the core phase. The module is describing principles of educational programme planning and provides some skills in planning process through presenting case studies.
Teaching methods	An introductory lecture gives the students first insight in characteristics of educational programme process and its characteristics. After introductory lectures, students first carefully read the recommended readings. Afterwards they form small groups. Every group is given a problem concerning planning of some public health issue to be taught. In each group, students try to plan a programme and present it to other students. Other students critically discuss their plan.
Specific recommendations for teachers	• work under teacher supervision/individual students' work proportion: 30%/70%; • facilities: a lecture room, a computer room, rooms for small-group work; • equipment: computers (1 computer on 2-3 students), LCD projection, access to the Internet and bibliographic data-bases; • training materials: recommended readings or other related readings; • target audience: master degree students according to Bologna scheme.
Assessment of students	Elaborated example of a plan for educational programme.

DESIGNING AND PLANNING EDUCATIONAL PROGRAMMES IN PUBLIC HEALTH

Gordana Pavleković, Lijana Zaletel-Kragelj, Anja Kragelj, Nataša Škerget

THEORETICAL BACKGROUND

Basic terminology

Before discussing the process of planning educational programmes, it would be worthy to clarify terminology in this field. Two terms are used in relation to planning an educational process, being "educational programme" (or "program") and "curriculum" (1,2). To understand them, we need first to clarify some other terms.

Programme

Several definitions exist on what the term programme means, among which we can find the following:

- a formal set of procedures to conduct an activity (3);
- a set of projects designed to achieve common, long-term goals (4);
- a set of organized but often varied activities (a programme may encompass several different projects, measures and processes) directed towards the achievement of specific objectives (5)
- a generic term taking into account a collection of coordinated projects/sub-programmes, single measures, processes and/or services aimed at achieving a common aim. The key elements of a programme include its inputs, processes, outputs, outcomes, impacts. It is defined and limited in terms of time, scope and budget (6).

On general, a programme is a planned sequence of activities designed to achieve specified objective. It has several steps, being needs assessment, programme planning, programme implementation, and programme evaluation:

- needs assessment - this is a step at which information about community health problems are gathered. On the basis of these information assumptions on how the needs/problems could be addressed, and the objectives/goals that should be reached are set up;
- programme planning - planning step - uses the assumptions set up at the previous step to plan a programme of activities;
- programme - implementation step - refers to the follow-up of the activities in accordance with the plan. Implementation could be expressed in terms of operational or action plans which commonly outline concrete activities, time frames, responsibilities, budgets etc., for the achievement of different objectives of the programme;
- programme evaluation - when running a programme, we want to know how far the programme went and how effective it is in achieving its goals/objectives. We are able to answer these questions by performing the so-called programme evaluation process. A programme evaluation is a periodic review and assessment of a

programme to determine, in light of current circumstances, the adequacy of its objectives and its design, as well as its intended and unintended results (4). This process bases on continuous careful monitoring of the course/implementation of the programme.

Like in other fields, also on the field of education, we meet programmes. In this case we are talking about educational programmes. As mentioned at the very beginning of this module, in tight relation to this term is another term, being "curriculum".

Educational programme (program) versus curriculum

According to Ličen (2), today both terms are used in planning of adult education process but while the term "program" is more frequently used in American literature, the second is more frequently used in British one.

1. Educational programme.
   The term "educational programme" has several meanings, depending on context (2). It could designate several elements, a group of elements, or a whole system of elements:
   - in the case that it designates a plan for educational process, it means a design and documentation for implementation of educational process,
   - in other circumstances it means all educational activities, planned in selected environment (e.g. a school programme, a programme of an educational centre, etc.),
   - in a narrow meaning, it designates a sequence of teaching units, planned to reach selected goals in a given time. In this case, Jakšić is rather talking about "teaching plan" (1).
2. Curriculum.
   Basically, the term "curriculum" originates from Latin word "currere" that means "to run". Accordingly, "curriculum" means "the way the process runs" or "a course" (2,7).

   In British didactics, with the term "curriculum" different learning and educational programmes prepared for participants of educational process are denoted (2). In the past, according to Jarvis (8), the concept of curriculum was typical for elementary education, while in adult education the term "programme" was used. This terminology tried to made distinction between children and youth education, and adult education concept. Additionally, the term "programme" characterized more non-formal way of education, which was used in adult education process (8,9).

   In andragogy today, the term "curriculum" is becoming more and more a general term for all educational programmes, irrespective of being more or less structured, and irrespective of being "opened" (usually in andragogy) or "closed" (usually in pedagogy) (2,8). Educational programmes for adults are not only non-formal. In fact, they are put in a continuum. There exist a pleiade of different programmes, from very non-formal to very structured, from very general to very specific. On the structured part of this continuum are programmes which are very close to traditional curriculum (traditional institutionalized) (8)

   We will give some explanations related to types of educational programmes later on. In fact, we could define a curriculum as organized agreement on

educational process and its content, to fulfil predefined objectives (8), yet in the field of andragogy such a definition should be understood as flexible as possible.

In higher education, which is somewhere in between pedagogy and andragogy (10), this term is usually used to designate the frame document in which professional and political agreements and decisions on educational role of a higher education institution (1).

At the end we could conclude that both terms could be used for the purpose of this module, but we decided to use the term "educational programme" in the meaning of a design of educational process, and documentation for its implementation. The reason is that in public health, as stated in a module "Public health capacity building: adult education principles and methods" in this book, we are facing with both parts of the educational continuum of adults. In capacity building we have very structured programmes, very similar to traditional curricula, while in educating process of educators and lay people, we are facing non-formal education process. Since this module is not intended to get to know students with knowledge how to plan higher education curricula, but rather more general knowledge, in the second part of this module, a planning process of a non-formal education programme will be presented.

From the PH academic sphere point of view, the term "curriculum" is more specific and is more connected to formal education in PH in higher education (capacity building) while the term "educational programme" is more general.

Planning

Planning could be defined as attempting to shape and control events in the future.

We plan because we want to increase chances to achieve the preset goals and objectives as much as possible. "The days are now past when the teacher produced a curriculum like a magician produced a rabbit out of a hat and when the lecturer taught whatever attracted his or her interest. It is now accepted that careful planning is necessary if the programme of teaching and learning is to be successful." (11).

As in many other planning processes, also in educational programme or curriculum planning process general recommended set of steps to be followed exists, being used by different experts slightly differently. In Table 1, ten steps in curriculum planning proposed by Dent and Harden are presented (12).

Table 1. Ten steps in curriculum planning according to Dent and Harden (12).

Step
1. Identifying the need
2. Establishing the learning outcomes
3. Agreeing the content
4. Organization of the content
5. Deciding the educational strategy
6. Deciding the teaching methods
7. Preparing the assessment
8. Communication about curriculum
9. Promoting an appropriate educational environment
10. Managing the curriculum

Educational programmes planning models

In adult education, planning of an educational programme is a key part of this process. There exist several models of planning (2,8,13). According to Caffarella (13) the "program planning models consist of ideas about how programs should be put together and what ingredients are necessary to ensure successful outcomes.". They could be different in shapes and sizes. They could be very simple with steps 1-5 for example, or very complex, using highly developed flow charts or in-depth qualitative descriptions (13). Some models are linear, some are non-linear (8,13):

- in linear models, the planner is expected to start at step one and follow each step in order until the process is completed. This may be helpful to newcomers, but soon loses its appeal since it does not represent the day to day working reality of most programme planners,
- non-linear, dynamic or non-sequential model allows programme planners to address a number of the components simultaneously, and to rearrange components to suit the demands of different situations. In these models of programme planning is seen as a process that consists of a set of interacting and dynamic elements or components.

Among the most well known models are those proposed by Houle (Houle's decision points approach), Knowles (Knowles' andragogical model of programme planning), and Caffarella (Caffarella's interactive planning model) (13-15). In health education also PRECEDE/PROCEED model is well known (16). The detailed discussion on these models is far beyond the scope of this module.

In summary, the educational programme needs to define the learning objectives and learning outcomes, the setting in which it should be performed and the standard to which it should be performed.

Educational programmes planning

Elements of an educational programme

Irrespective if we name a programme of an educational process "curriculum" or "educational programme", it has indispensable elements. If we assume that every educational process and every educational activity could be fully planned, then the most simple structure of an educational programme (according to Tyler) could be used, being (8):

- objectives,
- content,
- methods of teaching and learning, and
- evaluation.

In practice it is not so easy, since the list of indispensable elements depends on how a programme planner understands the concept of education, and the concept of organized educational process.

In summary, Leinster describes the scope of the educational programme planning as follows in Table 2 (17).

Table 2. The scope of educational programme planning (modified Leinster's proposals) (17).

The scope	Description
Content	What knowledge, skills and attitudes should the programme cover? What are the learning objectives and learning outcomes of the programme?
Delivery	How will the learning be delivered? What teaching or learning methods will be used?
Assessment	Will the students' learning be tested?
Structure	How will the content be organised? How will learning and teaching be scheduled?
Resources	What staff, learning materials, equipment and accommodation is needed?
Evaluation	How will the organisers know that the educational programme has been effective in delivering the learning objectives or learning outcomes?

Educational programme planning steps

Different planning models list different but yet similar steps in planning process. In this module we are presenting the general proposal (1,2,8):

1. Analysis of the circumstances and potential participants.
   Since educational programme is to be planned for particular environment its influences should be known and considered at the very beginning of the educational programme planning process. Analysis of these influences is important for assessment of internal (i.e. tradition of the educational institution, prevailing patterns of communication, its mission, financial resources, manpower, equipment, etc.) and external factors (educational market, competing institutions, local community characteristics, etc.). Both could be favourable or unfavourable.

   Second very important task at this first step of the planning process is analysis of the structure of potential participants of the educational programme. The data include data on gender, age, education background, socio-economic situation, geographical distribution, etc. For programme planner it is particularly important to be familiar with the educational background and previous knowledge of their target audience, since programmes should be planned to meet the educational needs and levels of the participants:
   - basic level - for those who have no previous knowledge of the topic,
   - intermediate level - for those who have some experience and basic knowledge in the topic being addressed,
   - advanced level - for those who are experienced, knowledgeable of the topic being addressed.

2. Needs assessment.
   At this step, the needs and interests of the potential programme participants are identified. They should be identified on the basis of data sources (i.e. surveys) rather than on the basis of programme planner's opinions. This kind of surveys usually include questions on educational level of the participants, topics of the most interest to participants, speakers of most interest to participants, participants' preferences on time, date, and location of the programme, reasons that might keep a participant from attending (transportation problems, financial limitations etc.).

There is often a mismatch between what is expected and the competencies gained from the training programme. The relevance or appropriates of educational programmes are the most important in designing and planning.

According to Dunn and co-workers, different approaches could be used to identify the curriculum needs (18).

- consultation with the stakeholders;
- a study of errors in practice;
- critical-incident studies;
- task analysis;
- study of star performers, etc.

3. Educational programme goals, learning objectives, and learning outcomes. After the outline of the programme is developed, learning goals, objectives and outcomes should be identified.

   Learning objectives and learning outcomes are connected to the domains of learning (Table 3).

Table 3. Cognitive processes in Bloom's revised taxonomy for knowledge and descriptors of outcomes (19,20).

Bloom's taxonomy	**Meaning**	**Outcome descriptor**
Remember	retrieving relevant material from long-term memory	recognise, recall
Understand	determining the meaning of instructional massages, including oral, written, and graphic communications	interpret, make examples, classify, summarise, infer, compare, explain
Apply	carrying out or using a procedure in a given situation	execute, implement
Analyse	breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose	differentiate, organise, develop/change attitude
Evaluate	making judgements based on criteria and standards	check, critique
Create	putting elements together to form a novel, coherent whole or make an original product	generate, plan, produce

Educational programme goals are broad statements that define what the programme is expected to achieve. There exist short- and long-term goals. Short-term goals could be easily stated in behavioural terms and could be accomplished in a short time while long-term goals are more general and usually could not be accomplished in a reasonable amount of time.

Learning objectives are the statements that map out the learner's and planner's tasks needed to reach the goal. They state the specific knowledge, attitudes and behaviour changes needed to achieve the goal. In Table 4 some examples of learning objectives are presented.

Table 4. Some examples of learning objectives.

Learning objectives
Upon completion of the programme, the participants will (be able to): • have an understanding of … • define and describe the … • distinguish the different types of … • gain new knowledge of … • gain skills on … • become aware of …

4. Content.
   The primary focus of a curriculum is many times on what is to be taught and when, leaving to the teaching profession decisions as to how this should be done. Many planners of an educational programme believe that the starting point for constructing a curriculum lies in the formulation of content. Many of them think in terms what content students should learn. The fact is that when starting from this point not taking into consideration other curriculum elements there could be problems. It is very important to employ objectives in some way when constructing an educational programme. For formulating objectives analysis of the circumstances and potential participants is of enormous importance. In practice, also, how a topic is to be taught often determines what will be taught.
5. Teaching methods.
   In parallel with determining the content, decision must be taken about how it will be delivered. While there is debate over which learning methods are the most effective and efficient, there are some principles that should be taken into account. It is well known that knowledge is applied most effectively when it is learn in the context in which it is to be applied. It is also accepted that active learning is more effective than passive learning.

   In this step the means through which the changes to be made are planned. Teaching methods like lectures, lectures with discussion, demonstrations, seminars, videos, role playing, case studies, etc. could be used. These methods are extensively discussed in a module "Public health capacity building: adult education principles and methods" in this book.
6. Financial plan and marketing.
   Educational programmes will almost always involve expenses of varying degrees. Regardless the expenses are minimal or large, they should always be planned in advance. Expenses mainly include speaker expenses, meeting/lecture room expenses, meal expenses, educational material expenses, and marketing expenses.

   Sources for covering expenses could be fees or other sources of budgeting. When fees are to be charged they should be large enough to cover the costs but small enough to be marketable to target audience. It is worth to look at the costs of similar programmes and set registration fees accordingly.

   Marketing the programme is also important for successful realization of the educational programme. A brochure or flyer is one of the most useful promotional materials. All information about the programme (e.g. description of the course content, learning objectives, speakers names and qualifications,

level of the programme, prerequisite skills or knowledge, date, time and location, etc.) should be accurately described. In addition to brochures/flyers, also other methods of publicity could be used, being press releases, public announcements, posters, etc.

7. Course of the programme.
   Course of the programme include identification of the speakers and their recruitment, location of the educational programme and logistic.

   The basic criteria for selecting appropriate speakers are their knowledge and experience in the field of content of the educational programme, and their ability to mediate this knowledge to other people. In the process of recruitment a summary of the programme, should be discussed, the target audience described, and the level of the programme clearly specified. Length of the session, teaching aids, and logistic of the programme should be clearly introduced to potential speakers.

   The location for carrying out the educational programme should be carefully chosen. Criteria for selection should include characteristics like adequate premises, clean and safe environment, close public transportation or available parking etc. There are usually many choices of meeting facilities e.g. hotels/motels, congress centres, community centres, schools, etc. It is important that size of the lecture rooms match the size of the expected audience (it is difficult to concentrate in a crowded or half full room). Also the possibility for providing meals and/or refreshments should be taken into consideration.
8. Evaluation.
   The last step to be planned in educational programme planning process is evaluation of teaching. This step has many different meanings:
   - finding out the value of teaching,
   - systematic way of learning and using the lessons learned,
   - complex process of assessment,
   - measurement and judgement,
   - getting the best out of what you do for all involved, collection of information etc.

   In respect to evaluation of teaching work it is frequently emphasized the importance of:
   - students' opinion about the quality of teaching and engagement of teachers,
   - measures outcomes of teaching (final knowledge of students),
   - analysis of curricula and teaching by Faculty administration,
   - occasional visits of departments and discussions with teaching personnel, etc.

   This evaluation of teaching is rarely or never done. Also, in the world documents which very much emphasize the quality improvement in teaching process, the methodology for achieving these goals is not precisely defined, until, perhaps, recently (21,22).

   In students' opinion about the quality of teaching one of the most common methods is student's anonymous questionnaire. However, it is generally believed that students' evaluation about teaching contribute to

quality of teaching, although it is thought that its methodology needs to be still improved (23).

Programme evaluation involves getting feedback from participants as to how the programme met their needs and expectations. It should be done immediately after the programme ends and be carried out by completing an evaluation form. Areas covered in the evaluation form usually include programme content, speaker effectiveness, facilities, handouts and other materials. Nevertheless, the evaluation form should be short and easy to complete. It should also allow the participant to make comments or suggestions.

One of methods that could be used in teaching evaluation is so called SITE (from Structured Interactive Teaching Evaluation) method[33] (24,25).

CASE STUDY: DESIGNING AND PLANNING EDUCATIONAL PROGRAMME "ENVIRONMENT PROTECTION – WHAT COULD WE DO?"

Introduction to the Case Study

This Case Study is a virtual case study, basing on a seminar paper at Ljubljana University Faculty of Arts, Department for Andragogy and Pedagogy, subject General andragogy II (26). Its characteristics are as follows:

- the virtual plan of the educational programme with public health content was constructed upon recommendations for planning adult education programmes of Slovenian Institute for Adult Education Ljubljana University Faculty of Arts (1,8),
- consultations, related to the public health content were given at Ljubljana University Faculty of Medicine, Chair of Public Health.

[33] Structured Interactive Teaching Evaluation (SITE) is an innovative method in evaluation. Starting from the fact that the evaluation of the teaching process is humane and not a mere administrative relationship with students and teachers as the partners, the original method of a structured interactive teaching evaluation was developed at the Department for Educational Technology, Andrija Stampar School of Public Health, Medical School, University of Zagreb. The SITE method has several characteristics:

- the evaluation process does not comprise all students but only randomly selected ones in the number corresponding to the number of all teachers in a particular course,
- participants in the evaluation process (selected students and teachers) are formally appointed by the Dean,
- final interactive meeting through «face-to-face» conversation is facilitated by a person from «outside».

The study on effectiveness and efficiency of the SITE method was done, comparing the results obtained by a traditional anonymous student questionnaire. The SITE method produced better impact on preparedness and active participation of both students and teachers, and influenced their responsibility in the evaluation process. The structured conversation added awareness to the teaching process and coordination to the proposals how to reach quality assurance jointly as equal partners. The SITE method was also efficient in terms of workload, time and cost of application.

For the purpose of this module, the seminar paper was slightly supplemented and changed by its authors. For example, the background of the planning process was set to a slightly changed scenario.

Assumptions, underlying the virtual case study

Underlying assumptions to the new scenario are as follows:

1. Assumption 1.
   The Ministry for Health of the Republic of Slovenia in cooperation with the Ministry for Environment of the Republic of Slovenia in the frame of sustainable development announced a call for application for a grant in maximal amount of 2,500 € for educational programmes related to environmental health. One of tender fields was "Health, environment, and waste disposal".

   The background for this tender field is rapidly growing problem of waste disposal in Slovenia. We are facing with increasing quantity of waste while there are less and less suitable locations for waste deposits. This situation is tightly connected with problems of era of postmodernism.

   In postmodernism, consumption has developed to undreamed extensions. In people it raises values, which are directed towards individuals and into the present time, not taking care of community's interests, or the future. As a result, negative consequences for the mankind are already starting to show. The problem has grown to the point at which we are starting to think over sustainable development. Hopefully, it is not too late.

   According to the World Commission for Environment and Development (WCED), sustainable development is a kind of development, that suits today's needs, but at the same time, it doesn't jeopardize chances of future generations to satisfy their own needs. The strategy of sustainable development embraces three pillars: economic development, social development and protection of environment. We are aware, how important the protection of environment is, therefore we need educational programmes, that would include that topic.
2. Assumption 2.
   The Chair for Public Health of Ljubljana University Medical Faculty decided to prepare a proposal for an educational programme and apply for a grant. To fulfil all necessary standards of good educational programme, it was decided to ask also experts in pedagogy/andragogy to participate.

The process of designing, and planning

Analysis of existing circumstances and potential participants

Analysis of existing circumstances

In the process of brainstorming where is most necessary to start educate people how to preserve our nature healthy for today and tomorrow, we came to the conclusion that personal body care means are becoming a real threat to the nature. Among them are disposable diapers, and manufactured washing agents.

1. Disposable diapers.

When searching for rationale we have found that on the Technical University in Graz, Austria, in 1997 the study about pollution according to diapering system was done (27). The study has shown amazing results: production of disposable diapers contributes at very high degree to the pollution while pollution due to transport and due to waste is comparable.

When using washable diapers, pollution is much smaller and washing contributes most to it. This factor can be decreased by choosing environment friendly washing agents and by dosing them wisely up to 50% less then indicated from the producers. Thus we decided to go this direction.

2. Disposable sanitary towels.
   Additionally, we decided to include in our consideration also a problem, to the certain extent similar to the first one – use of disposable sanitary towels women use at menstruation. Enormous quantities of waste are produced by using this kind of sanitary towels.

   In relation to this problem, we have found out that there exist environment friendly substitute, being a menstrual or menses cup. In Canada in 1995 even a clinical trial was performed to determine whether menses cup is well tolerated by menstruating women (28). The results showed that menses cup may be an acceptable method for some women for coping with menstrual flow. The menstrual cup exists now for decades, but only today its use is maybe put on the agenda. It was developed in the 1930s and went on commercial sale at around the same time as the first tampon. But unfortunately at that time it was not considered proper for women to touch their vaginal area. The manufacturers of the tampon overcame this obstacle by providing disposable applicators that can still be found floating in our toilet bowls today. Additionally, the tampon was commercially more attractive, as customers need to repurchase month after month. In contrast menstrual cup manufacturers do not expect to make repeat sales. Sales of tampons soared and massive advertising campaigns were launched. Disposable feminine hygiene products soon became established necessities in modern western society.

Thus we decided to analyze the circumstances of how could we preserve nature healthy by using selected environment friendly personal body care means (Box 1).

Participants

Who would be the participants of such an educational programme? If we limit to diapers, young parents are obviously the end target group. But not only this population group should and could be target group. Maybe educating young parents first wouldn't be the best choice. To targeting this population group as widely as possible, professionals that are coming in contact with young parents, well educated in this respect, seems to be the first choice.

But who would be the most suitable to make young parents aware of fact, presented above? We came to the conclusion that the best choice would be medical doctors and nurses. They are namely those, who accompany the child from even before birth on. That is why we should start to educate them, because they would be the ones, who will pass their knowledge to young parents.

Additionally, when searching for information on existence of educational programmes already offered in Slovenia about the similar topic, we discovered that

nongovernmental organization (NGO) Association Štorklja (a Stork) already holds lectures for parents on washable diapers in Maribor and Ljubljana. Since we didn't want to compete with this NGO, but rather supplement its efforts, we decided to target previously mentioned population group of medical doctors and nurses.

So, finally we decided to organise this educational programme for health workers, where they will be able to get needed knowledge about alternative hygienic accessories and skills, how to pass that gained knowledge to young parents (Box 1). Additional rationale behind this decision is, that these professionals will be on the frontline not only in educating young parents how to behave friendly to environment, but also they should be prepared to answer numerous questions of end target group, etc.

Box 1. The plan of educational programme "Environment protection – what could we do?" - Section 1: Analysis of the circumstances and potential participants.

1. ANALYSIS OF THE CIRCUMSTANCES AND POTENTIAL PARTICIPANTS

1.1. ANALYSIS OF THE CIRCUMSTANCES

Body care means are becoming a real threat to the nature, since they represent considerable part of environmental pollution. Among them are also disposable diapers, manufactured washing agents, and disposable sanitary towels.

Disposable and eco-diapers

If using disposable diapers in a diapering period, the baby needs approximately 5,000-6,000 disposable diapers, and consequently contributes about 1 tone of waste. It is estimated that disposable diapers today represent about 8-10% of waste in industrialized countries, contributed by only about 1% of population. Disposable diapers seem at the first sight harmless for environment, since they are made of paper. But unfortunately this is far from the truth. To make them ultra-absorbing very dangerous chemical components are used: sodium poly-acrylate (ultra-absorbing gel), toxic organo-chloridic combinations and polyvinyl-chloride (PVC) that makes disposable diapers impermeable for wetness. This hazardous substances cause water pollution and it is suspected that they have toxic and carcinogenic or bio-accumulative effects, in particular because the they decompose very slowly - when deposited disposable diaper needs 500 years for decompose (29). Additionally, they are a source of greenhouse gases since babies' excrements are thrown to waste deposit in the disposable diaper and decompose there. Also many micro organisms can be found in thesa diapers. They slowly trickle trough toward the underground waters and can possibly infect the drinking one.

On the other side, the use of cotton diapers or other natural materials represents a great contribution towards environment protection, and what is even more, each and every family can take a part in it, if they only decide for environment-friendly way of changing their babies. If we use natural materials, the excrements in diapers are being removed into sewage system. Besides that each baby needs only 20-25 cotton diapers that could also be used for other babies. Special cotton diapers are eco-diapers. These are even more environment friendly. Eco-diapers are washable diapers, made of eco-cotton. Eco-cotton is cotton that grew without being sprayed with insecticides and herbicides and without chemical supplements. It had been harvested without any artificial additives. That kind of cotton is the most natural material and it pampers baby's skin as softly as disposable diapers (29).

Environment-friendly washing agents

Use of eco-diapers is related to another environmental problem – the use of washing agents. But pollution when using washable diapers is much smaller and washing contributes most to it. This factor can be decreased by dosing manufactured chemical washing agents wisely up to 50% less then indicated from the producers, or by choosing environment friendly washing agents.

Box 1. Cont.

Environment-friendly washing agents are detergents, made on soap-basis. Into that group we could place all products, that are soap-based, and all natural washing agents, for example washing nuts, soda, vinegar or ashes (29).

Menses cup

Enormous quantities of waste disposal are produced by using disposable sanitary towels, women use at menstruation as well.

There exists environment friendly substitute, known as menstrual or menses cup. The menses or menstrual cup is a reusable menstrual cup around two inches long and made from soft silicone rubber. It is worn internally like a tampon but collects menstrual fluid rather than absorbing. Unlike tampons the menses cup is not a disposable product, so only one is needed (30).

1.2. ANALYSIS OF POTENTIAL PARTICIPANTS

From public health point of view two population groups could potentially be the most suitable participants of this educational programme. The first group, that is also an end target group, is the group of young parents, or young future parents. The second group is consisted of health workers that are in contact with young parents, being medical doctors like obstetricians and paediatricians, midwives, and nurses like nurses in the community health centres, as well as home care nurses.

We decided to organise this educational programme for health workers, where they will be able to get needed knowledge about alternative hygienic accessories and skills, how to pass that gained knowledge to young parents.

Optimally, at least one medical doctor and one nurse from each participating health institution should participate at the seminar.

Needs assessment

The next step is assessment of needs for an educational programme for educating people about harmful consequences of using environment unfriendly body care means and how to contribute to lessen this problem. We concentrated above all to disposable diapers usage problem.

We knew that sources of information on using disposable diapers are in Slovenia very limited, but we have found out, that NGO Association Štorklja is leading a project entitled "Preserving the Water by Using Diapers Friendly for Earth and Water" (29). This project was granted from Regional Center for Environment for Slovenia and United Nations Development Program (UNDP) and Global Environment Facility (GEF) Danube Project (29,31). It is partnered by Association RODA from Croatia (29,32). The project emphasizes awareness-raising, environmental education, policy issues and public participation in decision making related to specific pollution problem of using disposable diapers. According to Association Štorklja, in Slovenia, 99% of the population that uses diapers is using diapers that are one time use only, not realising how disputable they are from ecological point of view.

We could conclude that need for educating young parents of harmful consequences of using disposable diapers in Slovenia is obvious, and the problem should be tackled wisely (Box 2).

Box 2. The plan of educational programme "Environment protection – what could we do?" - Section 2: Needs Assessment.

2. NEEDS ASSESSMENT

In Slovenia, 99% of the population of young parents is using diapers that are one time use only, not realising how disputable they are from ecological point of view. This leads to the estimate that in Slovenia every year about 45.000 children under age 2,5 years are using diapers what makes about 100 millions pieces of disposable diapers or 20.000 tons of waste that need to be deposited (29).

On the other hand, in Slovenia we are facing with problem of lack of suitable locations for waste disposal. People are refusing to have waste deposit in their neighbourhood, especially if the waste is produced in urban settlements, and are deposited in rural areas.

These facts clearly indicate that there should be something done. Waste incineration seems to be one of possible solutions, but not a long-term one, since incinerators could also pollute the environment. Speaking in the medical terms, the best solution of all is to act etiologically, what means reduction in production of waste. For this could be realized, people need to be educated and enlightened.

Regarding problem of disposable diapers, education of young Slovene parents of harmfulness of disposal diapers on one hand, and use of washable diapers instead as an intelligent and rational choice, seems to be reasonable intervention. On Slovene market cotton-made and washable diapers could be purchased, but only a minor percent of young parents use them. The reason is, that one-time-use-only diapers are much more handy than washable ones. At the same time, parents don't realise, how much damage they make to the environment, when they are using disposable diapers instead of washable ones. That is why we should make them conscious of positive effects the use of washable diapers brings.

To be as much effective as possible, the problem should be tackled by educating directly young parents, and by educating health workers for they could educate young parents as well.

Learning objectives

We identified learning objectives at the macro and at the micro level as presented in Box 3.

Box 3. The plan of educational programme "Environment Protection – What could we do?" - Section 3: Learning Objectives.

3. LEARNING OBJECTIVES

3.1 MACRO-LEVEL LEARNING OBJECTIVES

As the purpose of this programme the employees in community health centres and other health institutions should become aware of different ways, how to protect the environment, and of effects the environment has on people's health. Furthermore, we wish them to gain skills, how to pass to other people the knowledge they have got about how to protect the environment from pollution and about instruments to achieve that.

3.2 MICRO-LEVEL LEARNING OBJECTIVES

The participants:

- *will become aware of the consequences of the environmental pollution on general,*
- *will become aware of the health impact of the environmental pollution,*
- *will gain new knowledge about positive effects of the use of environment-friendly hygienic accessories, from the point of view of different professions,*
- *will gain skills, how to pass their knowledge about environment-friendly hygienic accessories to other people, especially to young parents, and how to use them.*

Contents

While modelling components of the contents of proposed educational programme, we tried hard to form such a programme that would help the participants to gain a total knowledge (Box 4).

Box 4. The plan of educational programme "Environment protection – what could we do?" - Section 4: Contents.

4. CONTENTS

Introductory lecture on ecology

An expert from the field of ecology is going to present negative consequences of non-ecological behaviour. Everything will be illustrated.

After the lecture, selected sections of the documentary film ″Unpleasant truth″ will be presented.

Lecture on ecological problem of disposal hygienic accessories, especially disposal diapers

In continuation, an expert from the field of environmental health is going to give a special attention to the problem of environment pollution caused by disposal diapers and disposal sanitary towels. Results of different studies will be presented, as well as suggestions how to cope with this problem.

Is this problem concerning health workers?

Health workers are mainly directed in curing a disease and similar, while prevention is not yet in the first plan, especially when is related to community health rather than to a health of an individual. Yet health workers are and should be those experts people have confidence in, and their advices and ideas still count in the community. They will be acquainted with the concept of empowerment of lay people how to behave healthy to the nature and the community.

Special attention will be given to that, how health workers in community health centres, working with young parents and future parents, and maternities, can do something, what would contribute to acting more ecologically in relation to the ecological problem of disposable diapers.

Lectures of different experts from several fields of activity about positive effects, that come as a result of using cotton diapers

Experts from different spheres are going to tell their points of view on use of washable cotton diapers instead of disposable ones and present their opinions. In such a manner, the participants will be able to realise, that the use of eco-diapers has desirable effects in the field of protecting the environment as well as in the specialized medical fields like dermatology and paediatrics.

The participants will watch A DVD film, entitled ″Environment - and Earth-friendly diapers″, that was prepared in 2006 by nongovernmental (NGO) Association ″Štorklja″.

Representatives of the same NGO will demonstrate cotton-made diapers. Sellers of cotton diapers will exhibit these products at the exposition for participants of the programme could come in touch with them.

Lecture on natural washing agents

There is going to be a lecture on what natural washing agents are, and on positive effects of the use of them. Different experts, who collaborate with the Association ″Štorklja″, will present results of some studies.

Lecture on menses cups

There will be a lecture on what a menses or a menstrual cup is, how it works, and its history. A results of a clinical trial will be presented.

Box 4. Cont.

Representatives of the Association ″Štorklja″ will demonstrate an example of a contemporary menses cup.

Work in groups
In small groups, participants will think about how to transmit gained knowledge to other people, especially young and future parents. Each group will, with a help of a facilitator, prepare an assignment. Every assignment will be about finding new ways, how to promote the environment friendly products like eco-diapers, natural washing agents, and menstrual cups.

Teaching methods

Teaching methods should be appropriate for the target group – young and middle-aged adults with college or university medical/health sciences background. We considered this and andragogical concepts of educational process were applied (Box 5).

Box 5. The plan of educational programme "Environment protection – what could we do?" - Section 5: Teaching Methods.

5. TEACHING METHODS

Following teaching methods will be used:

- *lectures will be performed for introduction to each session,*
- *demonstrations of environment friendly hygienic means will be uses whenever possible,*
- *buzz groups after introductory lectures,*
- *reflections and discussions in large groups,*
- *brainstorming, while working in small groups.,*
- *discussions in small groups,*
- *plenary discussions, etc.*

The educational programme will be organized as a seminar. The seminar is going to take place at week-end.

Budget determination and marketing

From organizational point of view we considered financial plan of the programme as well as marketing (Box 6).

Considering budgeting, we assumed that the programme will be granted with 2,500 €. ″House of Nature″ Company, a company providing environment friendly products, and NGO Association ″Štorklja″, will co-finance the programme. Also, each participant will be charged a minor seminar fee in the sum of 25 € (Box 6). All incomes and expenses were calculated on 60 participants.

Box 6. The plan of educational programme "Environment protection - what could we do?" - Section 6: Financial Plan and Marketing.

6. FINANCIAL PLAN AND MARKETING

6.1 FINANCIAL PLAN

Incomes

Grant	*2,500 €*
House of Nature Company	*1,000 €*
Association Štorklja	*1,000 €*
Seminar fees	*1,500 €*
Total	***6,000 €***

Expenses

Administration costs	*200 €*
Material costs	*200 €*
Lecturers' fees	*900 €*
Accommodation costs for lecturers	*1,200 €*
Rent for the congress room	*250 €*
Use of LCD projector (11€ per hour)	*135 €*
Development of the programme	*500 €*
Beverages	*200 €*
Coffee breaks	*770 €*
Lunch	*1,400 €*
Total	***5,755 €***

6.2 MARKETING

Chair of public health of Ljubljana University Medical Faculty will send the invitation letter to join the seminar to every community health centre, and every maternity. At the same time, it will be recommended at least one medical doctor and one nurse should take part at it.

Course of the programme

First thing regarding the organizational part of the plan was to identify appropriate lecturers. While modelling components of the lectures, we tried hard to form a programme that would help the participants to gain a comprehensive knowledge. That is why we have asked experts from different fields of activity to share their knowledge. We thus planned to invite experts from different fields of public health being ecology, environmental medicine, and social medicine on one side, and experts from different fields of clinical medicine, being paediatrics, obstetrics, gynaecology, and dermatology on the other.

We tried to assess how many participants would attend such a seminar. We wanted to be realistic and set the number of participants to 60. However, there are about 70 community centres in Slovenia at the moment, and 14 maternities. Thus, the number could be even bigger.

We planned to invite also the Association "Štorklja" to participate in education of educators of young parents as well.

Regarding the location of the seminar, we wanted first to locate it on an eco-farm, but unfortunately no such a farm offers a room, where the lectures could take place. That is the reason we have decided the seminar should take place in business centre Ajda (complex of Terme 3000, Prekmurje). This congress centre is very well known and location seemed to be very appropriate, because most of participants live in that area, so

they should not have any problems with transport. That location is in the region, the education is ment for and beside that, Spa and congress centre of Moravske toplice is getting bigger and more significant all the time.

The planned course of the educational programme is presented in Box 7.

Box 7. The plan of educational programme "Environment protection - what could we do?" - Section 7: Financial Plan and Marketing.

7. COURSE OF THE PROGRAMME

7.1 ORGANIZATIONAL ASPECTS

***Number of participants**: 60*

Programme performers:

Coordinators: NS, expert of adult education
AK, expert of adult education

Lecturers: MF, Dipl.Ing.Chem., specialist in ecology
IE, MD PhD, specialist in environmental medicine
LK, MD PhD, specialist in epidemiology
PN, MD, PhD, specialist in paediatrics
ZN, MD, PhD, specialist in gynaecology
AM, MD, PhD, specialist in dermatology

Representatives of the Association "Štorklja"
Representatives of the House of Nature Company

***Duration**: 2 days*

***Location**: Business Centre Ajda, (Terme 3000), Moravske toplice, Slovenia*

7.2 REALIZATION

The seminar will last for 2 days, being Friday and Saturday. Lectures will be held by experts in ecology, environmental medicine, epidemiology, paediatrics, obstetrics and gynaecology, and dermatology. Representatives of House of Nature Company and Association "Štorklja" are also going to take part at it. In continuation the agenda of the lectures is presented.

The lectures will be variegated by picture material and short film.

Constituent element of the seminar will also be the reflexions of the participants – they shall have the chance to express their opinion.

The participants will have the chance, to look at the products that will be presented during the seminar.

On Saturday, the participants are going to work in small groups. Each group will have to brainstorm about, what would be the best way to promote natural products and to convince young parents to use them.

At the end of the seminar the participants shall present their reflexions, and discuss about it.

7.3 SEMINAR AGENDA

FRIDAY

9.00- 9:30	*Registration of participants*
9:30-10:00	*Opening of the seminar and welcome to the participants*
	Coordination: NS, AK

Box 7. Cont.

Session 1	***WE SHOULD BE WORRIED ABOUT OUR ENVIRONMENT***
	Coordination: NS, AK
10:00-10:30	*Introductory lecture on ecology (lecturer MF)*
10:30-11:00	*Selected sections of the documentary film "Unpleasant truth"*
11:00-11:30	*Coffee break*
11:30-11:50	*Ecological problem of disposal diapers (lecturer LK)*
11:50-12:10	*Is this problem concerning health workers? (lecturer IE)*
12:10-12:30	*Reflexions of the participants*
12:30-14:00	*Lunch*
Session 2	***COTTON DIAPERS***
	Coordination: NS, AK
14:00-15:00	*Presentation of eco-diapers, short film – Štorklja, House of Nature*
15:00-15:30	*Washing agents, a disadvantage of cotton diapers(lecturer LK)*
15:30-16:00	*Coffee break*
16:00-16:15	*Ecologist and his point of view (lecturer MF)*
16:15-16:30	*Environmental health expert and his point of view (lecturer IE)*
16:30-16:45	*Paediatrician an his/her point of view (lecturer PN)*
16:45-17:00	*Dermatologist and his/her point of view (lecturer AM)*
17:00-17:45	*Reflexions of the participants (coordination AK, NS)*
SATURDAY	
Session 3	***NATURAL WASHING AGENTS, MENSES CUP***
	Coordination: NS, AK
9.00- 9:30	*Types and use of natural washing agents (lecturer MF)*
9:30- 9:45	*Environmental health expert and his/her point of view (lecturer IE)*
9:45-10:00	*Dermatologist and his/her point of view (lecturer AM)*
10:00-10:30	*Reflexions of the participants (coordination AK, NS)*
10:30-11:00	*Coffee break*
11:00-11:15	*Menses cup and its history (lecturer LK)*
11:15-11:45	*Presentation of menses cup and its use – Štorklja, House of Nature*
11:45-12:00	*Gynaecologist an his/her point of view (lecturer ZN)*
12:00-12:30	*Reflexions of the participants (coordination AK, NS)*
12:30-14:00	*Lunch*
14:00-15:30	*Work in groups (coordination AK, NS)*
15:30-16:00	*Coffee break*
16:00-16:45	*Reports of the groups (coordination AK, NS)*
16:45-17:30	*Reflexions of the participants - all participants and lecturers*
17:30-17:45	*Closure of the seminar*

Evaluation

At the end we considered evaluation of the educational programme through targeting all three parts involved: health workers, users of eco-products, and companies that sell discussed product (Box 8).

Box 8. The plan of educational programme "Environment protection - what could we do?" - Section 8: Evaluation.

8. EVALUATION

The educational programme will be evaluated through reflexions of participants. We will ask the participants kindly to share their opinions about what has been told during the lectures. In the end of the seminar, a plenary discussion shall follow, where they will be able to express their opinion on the whole happening. We will deal out some sheets of paper and the participants should write down their reflexion. Like this, we will get the information about the quality of our programme (its contents as well as its organization).

Six months after the seminar will take place, we will make a research, if people are more interested in the products that have been presented in the seminar. We will ask every community health centre, if they could write a report on promotion activities that refer to eco-products. We are also going to make an opinion poll and ask parents of babies and women:

- *if they have noticed any promotion of that products,*
- *if they maybe use one of them themselves, and*
- *what is their opinion on these products.*

We plan to apply to the companies that sell discussed products to write a report on sales. We will ask them, if they think, making doctors known with these products has helped to increase people's interests in buying them.

Conclusion

Planning, designing, and then implementation of such an educational programme is a challenge, since it is an opposite to all commercials that are just about how practical and how much one product costs. What is sad, nowadays, our society prefers to buy products that don't need any maintenance - it is certainly easier, if we can, after we have used something, just throw it away. We are literally being pushed into that sort of acting if we want to do something for our children and grandchildren.

Acknowledgement

This sample plan was written under mentorship of Prof. Nives Ličen, PhD, University of Ljubljana, Faculty of Arts.

EXERCISE

Groups of two, at maximum three students are formed.

Task 1

After introductory lecture, students carefully read the part on theoretical background of this module and corresponding recommended readings.

Task 2

As a second task they develop a teaching module. After this exercise, participants will be able to prepare and present their own module (learning objective).

In a process of development of a teaching module they should follow the following teaching module structure:

1. Title
2. Author
3. Keywords
4. Rationale (educational needs)
5. Participants/students (to whom)
6. Duration
7. Learning objectives
8. Teaching methods
9. Planning of implementation
10. Assessment of participants/students
11. Module evaluation
12. Recommended readings
13. References

Presentation of a teaching module should be done in 30 minutes and assessed by other colleagues by using a questionnaire sample of which is presented in Table 5.

Table 5. An example of a questionnaire for an assessment of a teaching module.

	Question	**Not at all**				**Very much**
1.	Module appropriate for target participants/students	1	2	3	4	5
2.	Learning objectives clear/well defined	1	2	3	4	5
3.	Learning objectives achievable in planned time	1	2	3	4	5
4.	Teaching methods appropriate	1	2	3	4	5
5.	Implementation plan well presented	1	2	3	4	5
6.	Methods of student assessment appropriate	1	2	3	4	5
7.	Evaluation of the module well defined	1	2	3	4	5
8.	Innovative module	1	2	3	4	5
		Poor				**Excellent**
9.	Overall assessment of the module	1	2	3	4	5
10.	Oral presentation	1	2	3	4	5

Comments and recommendations to author (please describe in few words):

REFERENCES

1. Jakšić Ž. Curriculum and planning of lessons (in Croatian). In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.114-119.
2. Ličen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
3. Last JM. A dictionary of epidemiology. Oxford: Oxford University Press; 2001.
4. World Health Organization, Regional Office for Europe. Handbook for process evaluation in noncommunicable disease prevention. CINDI 2000. Copenhagen: World Health Organization, Regional Office for Europe; 1999. Available from URL: http://www.euro.who.int/document/E66338.pdf. Accessed: August 24, 2009.
5. United Kingdom Evaluation Society. Public Library/Glossary. Glossary of evaluation terms [homepage on the Internet]. Available from URL: http://www.evaluation.org.uk/resources/glossary.aspx. Accessed: August 24, 2009.
6. Swiss Federal Office of Public Health. Evaluation. Glossary of evaluation terms. Berne: Swiss Federal Office of Public Health; 2005. Available from URL: http://www.bag.admin.ch/evaluation/02357/02603/index.html?lang=en&download=M3wBPgDB/8ull6Du36WenojQ1NTTjaXZnqWfVp3Uhmfhnapmmc7Zi6rZnqCkkIR2fXqDbKbXrZ6lhuDZz8mMps2gpKfo. Accessed: August 24, 2009.
7. The online etymology dictionary [homepage on the Internet]. Available from URL: http://www.etymonline.com/index.php. Accessed: August 24, 2009.
8. Svetina M. Educational programmes for adults from planning to carrying out (in Slovene) Ljubljana: Andragoški center Republike Slovenije; 1998.
9. Krajnc A. Methods of adult education. Andragogical didactics (in Slovene). Ljubljana: DE, Delavska enotnost; 1979. p.101-221.
10. Jakšić Ž. Student-centred teaching (in Croatian). In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.87-95.
11. Harden RM. Curriculum planning and development. In: Dent JA, Harden RM, editors. A practical guide for medical education. London: Elsevier; 2005. p.10-8.
12. Dent JA, Harden RM. A practical guide for medical teachers. Edinburgh: Churchill Livingstone; 2001.
13. Caffarella RS. Planning programs for adult learners: a practical guide for educators, trainers, and staff developers. San Francisco, CA: Jossey-Bass Inc.; 1994.
14. Smith MK. Malcolm Knowles, informal adult education, self-direction and anadragogy. The encyclopedia of informal education. London: Infed; 2002. Available from URL: http://www.infed.org/thinkers/et-knowl.htm. Accessed: August 24, 2009.
15. Warren RM. Program planning and development in adult education: Where we are at the beginning of the 21st Century. Muncie, IN: Ball State University. Available from URL: www.hiceducation.org/Edu_Proceedings/Ruth%20M.%20Warren.pdf. Accessed: August 24, 2009.
16. Campbell C. Health education planning models - a review of the literature - part II. Starkville, MS: Mississippi State University; 2001. Available from URL: http://msucares.com/health/health/appa2.htm. Accessed: August 24, 2009.
17. Leinster S. The undergraduate curriculum. In: Dent JA, Harden RM. A practical guide for medical teacher. London: Elsevier; 2005. p.19-27.
18. Dunn WR, Hamilton DD, Harden RM. Techniques of identifying competences needed by doctors. Medical Teacher. 1985;7:15-25.
19. Quality Assurance Agency for Higher Education [homepage on the Internet]. Available from URL: http://www.qaa.ac.uk. Accessed: August 24, 2009.

20. Anderson LW, Krathowohl DR, editors. A taxonomy for teaching, learning and assessing: a revision of bloom's taxonomy for educational objectives. New York, London: Longman, Addison Wesley; 2001.
21. Guilbert JJ. Educational handbook for health personnel. Geneva: WHO; 1987. Available from URL: http://whqlibdoc.who.int/offset/WHO_OFFSET_35.pdf. Accessed: August 22, 2009.
22. Standards and guidelines for quality assurance in the European higher education area. Helsinki: European Association for Quality Assurance in Higher Education; 2005. Available from URL: www.enqa.eu/files/BergenReport210205.pdf. Accessed: August 22, 2009.
23. Metcalfe DH, Matharu M. Students' perception of good and bad teaching: report of a critical incident study. Med Educ. 1995;29:193-7.
24. Pavlekovic G, Bradamante Z. Structured Interactive Teaching Evaluation (SITE): an innovative method in evaluation. In: Association for Medical Education in Europe. Book of abstracts. Cotone Congressi Genoa, Italy, 14-18 September 2006. Dundee: Association for Medical Education in Europe; 2006. p.90. Available from URL: http://moodle.med.bg.ac.yu/moodle/file.php?file=/1/amee_2006_final_abstracts.pdf. Accessed: August 24, 2009.
25. Pavleković G. Evaluation of quality of teaching: why SITE (in Croatian)? In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005. p.299-305.
26. Škerget N, Kragelj A. Educational programme: Environment protection what could be done by ourselves. Subject: General andragogy II. A seminar paper. Ljubljana: University of Ljubljana, Faculty of Arts; 2008.
27. Association Storklja. Preserving the water by using diapers friendly for earth and water [homepage on the Internet]. Available from URL: http://www.storklja.si/t3/index.php?id=1103#1954. Accessed: August 24, 2009.
28. Cheng M, Kung R, Hannah M, Wilansky D, Shime J. Menses cup evaluation study. Fertil Steril 1995;64:661-3.
29. Association Storklja [homepage on the Internet]. Available from URL: http://www.storklja.si/. Accessed: August 24, 2009.
30. Mooncup. Whai is it [homepage on the Internet]. Available from URL: http://www.mooncup.co.uk/menstrual_cup_whatisit.html. Accessed: August 24, 2009.
31. Regional Environmental Center. Danube regional projects: Regional grants [homepage on the Internet]. Available from URL: http://www.rec.org/REC/programs/NGO_support/grants/RegionalDanubeGrants/ResultsRound2.html. Accessed: August 24, 2009.
32. Association Roda [homepage on the Internet]. Available from URL: http://www.roda.hr/. Accessed: August 24, 2009.

RECOMMENDED READINGS

1. Caffarella RS. Planning programs for adult learners: a practical guide for educators, trainers, and staff developers. San Francisco, CA: Jossey-Bass Inc.; 1994.
2. Dent JA, Harden RM, editors. A practical guide for medical education. London: Elsevier; 2005.
3. Jakšić Ž. Curriculum and planning of lessons. In: Jakšić Ž, Pokrajac N, Šmalcelj A, Vrcić-Keglević M, editors. The art of teaching medicine (in Croatian). Zagreb: Medicinska naklada, Hrvatsko društvo za medicinsku edukaciju; 2005.
4. Ličen N. Introduction to adult education (in Slovene). Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za pedagogiko in andragogiko; 2006.
5. Svetina M. Educational programmes for adults from planning to carrying out (in Slovene) Ljubljana: Andragoški center Republike Slovenije; 1998.

INDEX

Prepared by Luka Kovačić and Lijana Zaletel-Kragelj

A

B

C

D

Q