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
CONTENTS

Preface V
List of Authors VII

Chapter 1: METHODS OF STUDYING POPULATION HEALTH 1

1.1 Basic Concepts
1.1.1 Measurement of Health and Disease: An Introduction 3
   Tatjana Pekmezović
1.1.2 Probability – Basic Concepts 13
   Jadranka Božikov

1.2 Description of Health Phenomena and Their Quantification
1.2.1 Organizing and Describing Data 23
   Lijana Zaletel-Kragelj
1.2.2 Frequency Measures: Prevalence and Incidence 63
   Lijana Zaletel-Kragelj
1.2.3 Age Standardization Procedure: Direct Method 93
   Jadranka Božikov, Lijana Zaletel-Kragelj, Doris Bardehle
1.2.4 Measures of Location: Measures of Central Tendency and Dispersion 115
   Gena Grancharova, Silviya Aleksandrova
1.2.5 Measures of Location: Quantiles 147
   Gena Grancharova, Silviya Aleksandrova
1.2.6 Frequency Measures: Estimating Risk 161
   Lijana Zaletel-Kragelj, Jadranka Božikov

1.3 Analysis of Health Phenomena
1.3.1 Measures of Association and Potential Impact 179
   Lijana Zaletel-Kragelj
1.3.2 Cluster Analysis 215
   Anca Vitcu
1.3.3 Total Risk Assessment 245
   Mariana Dyakova, Emilia Karaslová, Hristo Mateev

1.4 Quantitative Study Designs
1.4.1 Introduction to Epidemiological Studies 267
   Enver Roshi, Genc Burazeri
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4.2</td>
<td>Features of Epidemiological Studies</td>
<td>Lijana Zaletel-Kragelj, Ivan Eržen, Doncho Donev</td>
<td>275</td>
</tr>
<tr>
<td>1.4.3</td>
<td>Ecological Studies: Basic Principles</td>
<td>Lijana Zaletel-Kragelj, Ivan Eržen</td>
<td>289</td>
</tr>
<tr>
<td>1.4.4</td>
<td>Cross-sectional Studies</td>
<td>Lijana Zaletel-Kragelj, Ivan Eržen</td>
<td>309</td>
</tr>
<tr>
<td>1.4.5</td>
<td>Case-control Studies</td>
<td>Slavenka Janković</td>
<td>333</td>
</tr>
<tr>
<td>1.4.6</td>
<td>Cohort Studies</td>
<td>Slavenka Janković</td>
<td>345</td>
</tr>
<tr>
<td>1.4.7</td>
<td>Introduction to Intervention (Experimental) Studies</td>
<td>Tatjana Pekmezović, Lijana Zaletel-Kragelj</td>
<td>359</td>
</tr>
<tr>
<td>1.5</td>
<td>Qualitative Study Designs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5.1</td>
<td>Acquiring Qualitative Skills for Public Health Research: Using Interviews to Generate Data</td>
<td>Danica Rotar Pavlič</td>
<td>381</td>
</tr>
<tr>
<td>1.5.2</td>
<td>Qualitative Methods: Focus Groups</td>
<td>Rok Fink, Andreja Kukec, Mojca Jevšnik</td>
<td>403</td>
</tr>
<tr>
<td>1.5.3</td>
<td>Delphi Analysis</td>
<td>Neda Milevska-Kostova, William N. Dunn</td>
<td>423</td>
</tr>
</tbody>
</table>

Chapter 2: SPECIAL EPIDEMIOLOGICAL AND OTHER METHODS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Environmental and Occupational Health Epidemiology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1.1</td>
<td>Principles and Methods of Environmental Epidemiology: An Overview</td>
<td>Ivan Eržen, Lijana Zaletel-Kragelj</td>
<td>439</td>
</tr>
<tr>
<td>2.1.2</td>
<td>The Integrated Exposure Uptake Biokinetic (IEUBK) Model for Lead in Children</td>
<td>Rok Fink, Andreja Kukec, Matej Ivartnik, Ivan Eržen</td>
<td>465</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Epidemiological Indicators of Environmental Health</td>
<td>Alexandra Cucu, Maria Nitescu</td>
<td>497</td>
</tr>
<tr>
<td>2.1.4</td>
<td>Environmental Health Risk Assessment Studies</td>
<td>Dragan Gjorgjev, Vladimir Kendrovski, Fimka Tozija</td>
<td>525</td>
</tr>
<tr>
<td>2.1.5</td>
<td>The Geographic Information System (GIS) Use in Analysis of Traffic Air Pollution</td>
<td>Andreja Kukec, Rok Fink, Saša Erlih, Ivan Eržen</td>
<td>547</td>
</tr>
<tr>
<td>2.1.6</td>
<td>Basic Occupational Health Indicators on Sick Leave</td>
<td>Marjan Bilban, Lijana Zaletel-Kragelj</td>
<td>573</td>
</tr>
</tbody>
</table>
2.1.7 Workplace Risk Assessment
Elisaveta Stikova, Neda Milevska-Kostova, Petar Bulat
Doncho Donev, Neda Jocić

2.2 Infectious Diseases Epidemiology
2.2.1 Surveillance
Irena Klavs

2.2.2 Outbreak Investigation
Maja Sočan

2.3 Oral Health Epidemiology
2.3.1 Oral Health Indicators in Europe
Barbara Artnik

2.4 Quality of Life Measurements
2.4.1 Health Related Quality of Life and General Quality of Life – Concepts and Measurement
Gorka Vušetić Mavrinac

Chapter 3: METHODS OF PUBLIC HEALTH INTERVENTIONS

3.1 Health Communication
3.1.1 The Potential of Public Service Advertising in the Field of Public Health
Dobriana Sidjimova, Mariana Dyakova, Zaharina Savova

3.1.2 Communication and Behaviour in Dental Practices
Zaharina Savova, Dobriana Sidjimova

Chapter 4: METHODS OF PLANNING AND EVALUATION

4.1 Evaluation of Health Outcome Change
4.1.1 Measuring the Burden of Disease: Disability Adjusted Life Years (DALY)
Doncho Donev, Lijana Zaletel-Kragelj, Vesna Bjegović, Genc Burazeri

4.2 Screening and Diagnostic Tests Evaluation
4.2.1 Test Validity Measures and Receiver Operating Characteristic (ROC) Analysis
Jadranka Božikov, Lijana Zaletel-Kragelj
4.3 Evaluation of Economic Efficiency

4.3.1 Basic Concepts in Health Economics and Methods for Economic Evaluation
Doncho Donev

4.3.2 Economic Evaluation in Healthcare: Practical Approach
Pia Vračko, Lijana Zaletel-Kragelj

4.4 Public Health Planning

4.4.1 SWOT Analysis
Andrej Plesničar, Lijana Zaletel-Kragelj

4.4.2 Rapid Assessment and Response – RAR
Enida Imamović, Dragana Nikšić

4.4.3 Measurement, Monitoring and Evaluation of Public Health Systems: Assessment of Essential Public Health Functions
Fimka Tozija, Draga Gjorgjev, Dance Gudeva Nikovska

4.4.4 Strategic Planning in Health Care – General Approach
Doncho Donev, Neda Milevska-Kostova, Adriana Galan

4.4.5 Priority Setting in Health Care
Vladimir Lazarevik, Doncho Donev, Dance Gudeva Nikovska, Blaško Kasapinov

Chapter 5: SUPPORTIVE METHODS/TOOLS/TECHNOLOGIES

5.1 Information and Communication Technology

5.1.1 Databases and Their Organization
Josipa Kern, Slavica Sović

5.2 Capacity Building

5.2.1 Public Health Capacity Building: Adult Education Principles and Methods
Gordana Pavleković, Lijana Zaletel-Kragelj, Anja Kragelj

5.2.2 Tele-education as a New Method of Medical and Public Health Education
Izet Mašić

5.2.3 Designing and Planning Educational Programmes in Public Health
Gordana Pavleković, Lijana Zaletel-Kragelj, Anja Kragelj, Nataša Škerget

Index
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 Akademische 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:


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
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, MCPPP, 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. Pavlekovíc 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 Pavlić 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
<table>
<thead>
<tr>
<th>Title</th>
<th>TEST VALIDITY MEASURES AND RECEIVER OPERATING CHARACTERISTIC (ROC) ANALYSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module: 4.2.1</td>
<td>ECTS (suggested): 0.20</td>
</tr>
</tbody>
</table>
| 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 (positive 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.

<table>
<thead>
<tr>
<th>Test</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>Positive</td>
<td>a</td>
</tr>
<tr>
<td>Negative</td>
<td>c</td>
</tr>
</tbody>
</table>

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.

![Decision Matrix](image)

**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).

![Complete Decision Matrix](image)

**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).

  \[
  \text{sensitivity} = \frac{TP}{TP + FN}
  \]  
  \text{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).

  \[
  \text{specificity} = \frac{TN}{TN + FP}
  \]  
  \text{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).

  \[
  \text{FPR} = \frac{FN}{TP + FN} = 1 - \text{sensitivity}
  \]  
  \text{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).

  \[
  \text{FNR} = \frac{FN}{TP + FN} = 1 - \text{specificity}
  \]  
  \text{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):

$$sensitivity = P(O \mid B) = \frac{P(O \cap B)}{P(B)}$$  \hspace{1cm} \text{Equation 5.}

$$specificity = P(O' \mid B') = \frac{P(O' \cap B')}{P(B')}$$  \hspace{1cm} \text{Equation 6.}

$$FNR = P(O' \mid B) = P(O \mid B)$$  \hspace{1cm} \text{Equation 7.}

$$FPR = P(O \mid B') = P(O' \mid B')$$  \hspace{1cm} \text{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.

$$PPV = \frac{TP}{TP + FP}$$  \hspace{1cm} \text{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.

\[
= \frac{P(O \mid B) \times P(B)}{P(O \mid B) \times P(B) + (1 - P(O' \mid B')) \times (1 - P(B))} =
\]

\[
= \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}
\]

\[
P(B' \mid O') = \frac{P(O' \mid B') \times P(B')}{P(O' \mid B') \times P(B') + P(O' \mid B) \times P(B)} =
\]

Equation 14.

\[
= \frac{P(O' \mid B') \times P(B')}{P(O' \mid B') \times (1 - P(B)) + (1 - P(O \mid B)) \times P(B)} =
\]

\[
= \frac{\text{specificity} \times (1 - \text{prevalence})}{\text{specificity} \times (1 - \text{prevalence}) + (1 - \text{sensitivity}) \times \text{prevalence}}
\]
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.

\[
ATA = \frac{TP + TN}{TP + FP + FN + TN}
\]

**Equation 15.**

\[
RTA = \frac{\left( \frac{TP}{TP + FP} + \frac{TN}{FN + TP} \right)}{2}
\]

**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](image_url)

**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.](image)

The points in the ROC diagram could be fitted with a curve which can be smoothed (Figure 6).
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 6.** Fitted and smoothed Receiver Operating Characteristic (ROC) curve. LEGEND: TPR = true positive rate, FPR = false positive rate.

**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).

![Graphical representation](image)

**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 $\theta$, is in practice calculated as (Equation 17):

$$
\theta = \theta_{1-0} + \theta_{2-1} + ... + \theta_{(i+1)-i} = \\
= \left[ (\text{FPR}_1 - \text{FPR}_0) \times \frac{\text{TPR}_0 + \text{TPR}_1}{2} \right] + \\
+ \left[ (\text{FPR}_2 - \text{FPR}_1) \times \frac{\text{TPR}_1 + \text{TPR}_2}{2} \right] + ... \\
...+ \left[ (\text{FPR}_{i+1} - \text{FPR}_i) \times \frac{\text{TPR}_i + \text{TPR}_{i+1}}{2} \right]
$$

**Equation 17.**
For the case presented in Figure 8 the Equation 17 reads as (Equation 18):

\[
\theta = \theta_{B-A} + \theta_{C-B} + \ldots + \theta_{F-E} = \\
\left( \text{FPR}_B - \text{FPR}_A \right) \times \left( \frac{\text{TPR}_A + \text{TPR}_B}{2} \right) + \\
\left( \text{FPR}_C - \text{FPR}_B \right) \times \left( \frac{\text{TPR}_B + \text{TPR}_C}{2} \right) + \ldots \\
\left( \text{FPR}_F - \text{FPR}_E \right) \times \left( \frac{\text{TPR}_E + \text{TPR}_F}{2} \right)
\]  
 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):  

\[
J = \max \left\{ \left( \text{sensitivity} + \text{specificity} \right) - 1 \right\}
\]  
 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):

$$J = \max \{(TPR + TNR) - 1\}$$

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):

$$J = \max \{TPR - FPR\}$$

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 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%).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>90</td>
<td>495</td>
<td>585</td>
</tr>
<tr>
<td>Negative</td>
<td>10</td>
<td>9405</td>
<td>9415</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>9900</td>
<td>10000</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Calculation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>100/10000</td>
<td>0.010 (1.00%)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>90/100</td>
<td>0.900 (90.0%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>9405/9900</td>
<td>0.950 (95.0%)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>90/585</td>
<td>0.154 (15.4%)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>9405/9415</td>
<td>0.999 (99.9%)</td>
</tr>
</tbody>
</table>

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%).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>54</td>
<td>2</td>
<td>56</td>
</tr>
<tr>
<td>Negative</td>
<td>6</td>
<td>38</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>40</td>
<td>100</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Calculation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence</td>
<td>60/100</td>
<td>0.600 (60.0%)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>54/60</td>
<td>0.900 (90.0%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>38/40</td>
<td>0.950 (95.0%)</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>54/56</td>
<td>0.964 (96.4%)</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>38/44</td>
<td>0.864 (86.4%)</td>
</tr>
</tbody>
</table>

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](image_url)

**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.

![ROC Diagram](image)

**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.

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>TPR (sensitivity)</th>
<th>FPR (1-specificity)</th>
<th>TPR-FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.988</td>
<td>0.012</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.927</td>
<td>0.073</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.927</td>
<td>0.073</td>
</tr>
<tr>
<td>4</td>
<td>0.992</td>
<td>0.747</td>
<td>0.245</td>
</tr>
<tr>
<td>5</td>
<td>0.967</td>
<td>0.517</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>0.967</td>
<td>0.517</td>
<td>0.45</td>
</tr>
<tr>
<td>7</td>
<td>0.932</td>
<td>0.317</td>
<td>0.615</td>
</tr>
<tr>
<td>8</td>
<td>0.89</td>
<td>0.175</td>
<td>0.715</td>
</tr>
</tbody>
</table>
Table 5. Cont.

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>TPR (sensitivity)</th>
<th>FPR (1-specificity)</th>
<th>TPR-FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0.825</td>
<td>0.087</td>
<td>0.738</td>
</tr>
<tr>
<td>10</td>
<td>0.825</td>
<td>0.087</td>
<td>0.738</td>
</tr>
<tr>
<td>11</td>
<td>0.725</td>
<td>0.038</td>
<td>0.687</td>
</tr>
<tr>
<td>12</td>
<td>0.575</td>
<td>0.013</td>
<td>0.562</td>
</tr>
<tr>
<td>13</td>
<td>0.445</td>
<td>0.003</td>
<td>0.442</td>
</tr>
<tr>
<td>14</td>
<td>0.335</td>
<td>0</td>
<td>0.335</td>
</tr>
<tr>
<td>15</td>
<td>0.245</td>
<td>0</td>
<td>0.245</td>
</tr>
<tr>
<td>16</td>
<td>0.245</td>
<td>0</td>
<td>0.245</td>
</tr>
<tr>
<td>17</td>
<td>0.135</td>
<td>0</td>
<td>0.135</td>
</tr>
<tr>
<td>18</td>
<td>0.135</td>
<td>0</td>
<td>0.135</td>
</tr>
<tr>
<td>19</td>
<td>0.098</td>
<td>0</td>
<td>0.098</td>
</tr>
<tr>
<td>20</td>
<td>0.069</td>
<td>0</td>
<td>0.069</td>
</tr>
<tr>
<td>21</td>
<td>0.044</td>
<td>0</td>
<td>0.044</td>
</tr>
<tr>
<td>22</td>
<td>0.027</td>
<td>0</td>
<td>0.027</td>
</tr>
</tbody>
</table>

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.

![ROC Curve](image)

**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>
<thead>
<tr>
<th>Test T</th>
<th>Disease D</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>18</td>
<td>3</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>5</td>
<td>214</td>
<td>219</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>217</td>
<td>230</td>
<td></td>
</tr>
</tbody>
</table>

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.

**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:
   \[
   \frac{18}{23} = 0.78 \ (78\%)
   \]
2. Nosological specificity:
   \[
   \frac{214}{217} = 0.99 \ (99\%)
   \]
3. Diagnostic specificity:
   \[
   \frac{18}{21} = 0.86 \ (86\%)
   \]
4. Diagnostic sensitivity:
   \[
   \frac{214}{219} = 0.98 \ (98\%)
   \]

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.

<table>
<thead>
<tr>
<th>Number of symptoms</th>
<th>Diseased</th>
<th>Healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>2</td>
</tr>
</tbody>
</table>

Sketch the ROC curve (point) if the positive test is considered as the presence of 3 or more symptoms of the disease.\(^\text{24}\)

**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.

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TPR</td>
<td>FPR</td>
<td>TPR</td>
<td>FPR</td>
<td>TPR</td>
<td>FPR</td>
</tr>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.023</td>
<td>0.000</td>
<td>0.019</td>
<td>0.000</td>
<td>0.027</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>0.051</td>
<td>0.001</td>
<td>0.036</td>
<td>0.001</td>
<td>0.046</td>
<td>0.001</td>
</tr>
<tr>
<td>4</td>
<td>0.067</td>
<td>0.002</td>
<td>0.068</td>
<td>0.002</td>
<td>0.069</td>
<td>0.002</td>
</tr>
<tr>
<td>5</td>
<td>0.072</td>
<td>0.003</td>
<td>0.084</td>
<td>0.004</td>
<td>0.082</td>
<td>0.003</td>
</tr>
<tr>
<td>6</td>
<td>0.089</td>
<td>0.004</td>
<td>0.093</td>
<td>0.005</td>
<td>0.098</td>
<td>0.004</td>
</tr>
</tbody>
</table>

\(^{24}\) **Answer to Task 2:**

![ROC curve diagram]
<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TPR</td>
<td>FPR</td>
<td>TPR</td>
<td>FPR</td>
<td>TPR</td>
</tr>
<tr>
<td>7</td>
<td>0.097</td>
<td>0.005</td>
<td>0.098</td>
<td>0.006</td>
<td>0.115</td>
</tr>
<tr>
<td>8</td>
<td>0.104</td>
<td>0.006</td>
<td>0.105</td>
<td>0.007</td>
<td>0.121</td>
</tr>
<tr>
<td>9</td>
<td>0.107</td>
<td>0.007</td>
<td>0.112</td>
<td>0.008</td>
<td>0.135</td>
</tr>
<tr>
<td>10</td>
<td>0.116</td>
<td>0.008</td>
<td>0.117</td>
<td>0.009</td>
<td>0.142</td>
</tr>
<tr>
<td>11</td>
<td>0.120</td>
<td>0.009</td>
<td>0.129</td>
<td>0.010</td>
<td>0.159</td>
</tr>
<tr>
<td>12</td>
<td>0.125</td>
<td>0.010</td>
<td>0.131</td>
<td>0.013</td>
<td>0.169</td>
</tr>
<tr>
<td>13</td>
<td>0.133</td>
<td>0.011</td>
<td>0.138</td>
<td>0.015</td>
<td>0.174</td>
</tr>
<tr>
<td>14</td>
<td>0.142</td>
<td>0.012</td>
<td>0.138</td>
<td>0.016</td>
<td>0.181</td>
</tr>
<tr>
<td>15</td>
<td>0.145</td>
<td>0.013</td>
<td>0.142</td>
<td>0.017</td>
<td>0.186</td>
</tr>
<tr>
<td>16</td>
<td>0.148</td>
<td>0.014</td>
<td>0.170</td>
<td>0.019</td>
<td>0.193</td>
</tr>
<tr>
<td>17</td>
<td>0.168</td>
<td>0.018</td>
<td>0.182</td>
<td>0.020</td>
<td>0.201</td>
</tr>
<tr>
<td>18</td>
<td>0.173</td>
<td>0.019</td>
<td>0.186</td>
<td>0.022</td>
<td>0.210</td>
</tr>
<tr>
<td>19</td>
<td>0.201</td>
<td>0.022</td>
<td>0.207</td>
<td>0.025</td>
<td>0.222</td>
</tr>
<tr>
<td>20</td>
<td>0.207</td>
<td>0.024</td>
<td>0.210</td>
<td>0.028</td>
<td>0.235</td>
</tr>
<tr>
<td>21</td>
<td>0.208</td>
<td>0.025</td>
<td>0.224</td>
<td>0.029</td>
<td>0.247</td>
</tr>
<tr>
<td>22</td>
<td>0.228</td>
<td>0.029</td>
<td>0.235</td>
<td>0.033</td>
<td>0.251</td>
</tr>
<tr>
<td>23</td>
<td>0.233</td>
<td>0.031</td>
<td>0.256</td>
<td>0.037</td>
<td>0.259</td>
</tr>
<tr>
<td>24</td>
<td>0.235</td>
<td>0.032</td>
<td>0.261</td>
<td>0.040</td>
<td>0.270</td>
</tr>
<tr>
<td>25</td>
<td>0.247</td>
<td>0.035</td>
<td>0.263</td>
<td>0.042</td>
<td>0.292</td>
</tr>
<tr>
<td>26</td>
<td>0.255</td>
<td>0.038</td>
<td>0.277</td>
<td>0.050</td>
<td>0.304</td>
</tr>
<tr>
<td>27</td>
<td>0.265</td>
<td>0.043</td>
<td>0.294</td>
<td>0.057</td>
<td>0.318</td>
</tr>
<tr>
<td>28</td>
<td>0.277</td>
<td>0.047</td>
<td>0.305</td>
<td>0.061</td>
<td>0.340</td>
</tr>
<tr>
<td>29</td>
<td>0.292</td>
<td>0.054</td>
<td>0.340</td>
<td>0.070</td>
<td>0.378</td>
</tr>
<tr>
<td>30</td>
<td>0.319</td>
<td>0.067</td>
<td>0.352</td>
<td>0.077</td>
<td>0.399</td>
</tr>
<tr>
<td>31</td>
<td>0.370</td>
<td>0.092</td>
<td>0.378</td>
<td>0.088</td>
<td>0.422</td>
</tr>
<tr>
<td>32</td>
<td>0.378</td>
<td>0.101</td>
<td>0.403</td>
<td>0.101</td>
<td>0.449</td>
</tr>
<tr>
<td>33</td>
<td>0.436</td>
<td>0.132</td>
<td>0.427</td>
<td>0.115</td>
<td>0.476</td>
</tr>
<tr>
<td>34</td>
<td>0.488</td>
<td>0.167</td>
<td>0.462</td>
<td>0.133</td>
<td>0.502</td>
</tr>
<tr>
<td>35</td>
<td>0.593</td>
<td>0.266</td>
<td>0.506</td>
<td>0.162</td>
<td>0.547</td>
</tr>
<tr>
<td>36</td>
<td>0.712</td>
<td>0.437</td>
<td>0.557</td>
<td>0.207</td>
<td>0.594</td>
</tr>
<tr>
<td>37</td>
<td>0.864</td>
<td>0.652</td>
<td>0.597</td>
<td>0.250</td>
<td>0.681</td>
</tr>
<tr>
<td>38</td>
<td>0.924</td>
<td>0.774</td>
<td>0.739</td>
<td>0.417</td>
<td>0.789</td>
</tr>
<tr>
<td>39</td>
<td>1.000</td>
<td>0.996</td>
<td>0.830</td>
<td>0.574</td>
<td>0.923</td>
</tr>
<tr>
<td>40</td>
<td>1.000</td>
<td>1.000</td>
<td>0.911</td>
<td>0.731</td>
<td>0.989</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td></td>
<td>0.967</td>
<td>0.912</td>
<td>1.000</td>
</tr>
<tr>
<td>42</td>
<td></td>
<td></td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

From the data presented in Table 8\textsuperscript{25}, make the following:

\textsuperscript{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**

---

\(^{26}\) 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: \(J = 0.405\)

RECOMMENDED READINGS