Ministry of Public Health of Ukraine
"Ukrainian Medical Stomatological Academy"

"APPROVED"
at the meeting of the Department
of Medical Informatics, Medical Biophysics
«27» august 2020
Minutes №. 1 «27» august 2020
Head of department _____ O.V.Silkova

METHODICAL GUIDANCE

for students' self-directed work when preparing and during the practical session

Academic Subject	Medical Information Science
Module No 2	Medical knowledge and decision making in medicine and dentistry
Topic	Methods of support in taking decisions. Strategies for obtaining medical knowledge. Expert systems
Year of study	2
Speciality	Foreign Student Training (Medcine/Stomatology)
Number of academic hours	2

Poltava – 2020

1. **Relevance of the topic**:

The topic is very important for future doctors in their professional activity, positively influences the students in their attitude to the future profession, forms professional skills and experience as well as taking as a principle the knowledge of the subject learned.

Studying the complex processes which appear in the nature, in a human body or at carrying out of research experiments, we not always can take into account all existing factors: from them more powerful and the some people it is possible to ignore some. Thus models of such processes, the phenomena are developed, which capable to substitute for them completely and at which studying we can receive the new information on them.

2. **The specific aims**:

To have general knowledge of the topic studied

To understand, to remember and to use the knowledge received

To form the professional experience by reviewing, training and authorizing it

To be able to carry out laboratory and experimental work

To know to analyze ES, to classify ES's.

3. **Basic knowledge and skills necessary to study the topic (inter-disciplinary integration).**

Previous (providing disciplines)	**Obtainable skills**
Informatics bases	- To create algorithm of medical tasks; To set sequence of actions to form knowledge for use in decision-support system;

4. **The tasks for students' individual work**

4.1.The list of basic term, parameters, characteristics, which student should master while preparin for the class.

Term	Definition
Expert system	A computer program that uses domain knowledge to perform a specific task usually human experts perform.
Knowledge engineering	The discipline of acquiring, encoding and using human domain knowledge to develop a computer application.
Algorithms of limited search	It calculate frequencies of combinations of simple logic events in subgroups (classes) of the data
Knowledge base	A store of both factual and heuristic knowledge.
Data Mining	Technologies of excavation of the data

4.2 Theoretical questions for the class (to the topic):

1. What is expert system?
2. Types of expert system?
3. List tasks and methods of knowledge engineering.
4. Structures of expert system
5. Data Mining: features and use.
6. What are artificial neural networks?
7. List advantages of expert systems.
8. List disadvantages of expert systems.
9. List and describe basic components of Expert System.

4.3 Practical tasks pertaining to the topic and to be completed during the class:

Test

1. 1. BUILDING AN EXPERT SYSTEM IS KNOWN AS:

a) Knowledge engineering
b) Building
c) KDD
d) IA
e) Computing

2. EXPERT SYSTEMS ALSO CALLED

a) Task domain
b) Knowledge-based
c) Artificial Intelligence
d) Algorithm
e) None of above

3. EXPERT SYSTEM ARE USED IN

a) Mathematics
b) Languages
c) Engineering
d) Various problem arias
e) All variants correct

4. PRINCIPAL PART OF ES IS

a) Knowledge acquisition subsystem
b) Block of explanation
c) Knowledge base
d) User interface
e) Decision making

5. IT IS A SUBSYSTEM THAT EXPLAINS THE SYSTEM'S ACTIONS:

a) Explanation subsystem
b) Reasoning engine
c) Knowledge acquisition subsystem
d) Knowledge base
e) None of above

Content of the topic:

Expert system is a version of computer systems based on the appropriate: representation of knowledge, in particular medical. As against the considered above diagnostic systems, the medical expert systems to some extent simulate processes of thinking of the doctor at statement of the diagnosis.

In order to prevent terminological different interpretations, to expert we shall carry systems, which: process a plenty of knowledge; represent knowledge in the simple unified form; have the independent mechanism of logic conclusions; can explain results received during processing of knowledge.

Thus the expert system does not act in a role of the teacher, and simply there is a heuristic training of the use at, the expense of granting to it of new opportunities. It is natural, that the medicine, as a sphere of activity, is ideal environment for creation and application of expert systems.

The expert systems distinguish on a way of representation of knowledge. Thus the algorithmic

model of representation of knowledge is widely used. The modem principles of optimization of diagnostic process assume syndromic the analysis of diseases, choice of decisive attributes and creation of diagnostic algorithm. Syndromic the principle of the analysis of diseases allows to narrow diverse symptomatology of disease up to some small amount of information blocks. Syndrome is considered as set of attributes (or large attribute), which can be observed at diseases of bodies and systems, irrespective of them etiology and pathogenesis, and also localization of pathological process. Such phenomenological the approach to treatment considerably facilitates development on its basis diagnostic is scarlet of a rhythm. Creation of algorithm on syndromic a principle having ensured inclusion in it of all diseases and pathological condition, described by the data syndrome. The choice of decisive attributes substantially raises efficiency and efficiency of diagnostics. One of sources of selection of decisive attributes is the saved experience. As a result of the carried out thus expert, work the medical knowledge is represented as algorithmic model. In practice of programming such model describe with the help algorithmic language of programming. The block diagrams of algorithms allowing to present given model in an evident in evident and kind are widely used also to difficult designs from the concrete programming languages. As an example the algorithm of recognition comatose of condition at the patients' sugar diabetes can serve.

Lack of such model is its static character. Entering new or the updating already of available knowledge assumes change of structure of all algorithm and, accordingly, program, realizing it especially, if to take into account that the similar changes at creation of machine diagnostic systems occur repeatedly.

The beginning of works above creation of the first expert systems concerns to 70 years of XX century owing to rapid development of computer science and occurrence of such branch, as an artificial intellect.

Expert Systems are computer programs that are derived from a branch of computer science research called Artificial Intelligence (Al). Al's scientific goal is to understand intelligence by building computer programs that exhibit intelligent behavior. In other words, Al is conditional definition of cybernetic systems modeled some of human intellectual activity sides - logical and analytical thought. It is concerned with the concepts and methods of symbolic inference, or reasoning, by a computer, and how the knowledge used to make those inferences will be represented inside the machine. Expert system is a computer program that uses domain knowledge to perform specific task usually human experts performs.

Of course, the term intelligence covers many cognitive skills, including the ability to solve problems, learn, and understand language; Al addresses all of those. But most progress to date in Al has been made in the area of problem solving - concepts and methods for building programs that reason about problems rather than calculate a solution.

Al programs that achieve expert-level competence in solving problems in task areas by bringing to bear a body of knowledge about specific tasks are called knowledge-based or expert systems. Often, the term expert systems is reserved for programs whose knowledge base contains the knowledge used by human experts, in contrast to knowledge gathered from textbooks or non-experts. More often than not, the two terms, expert systems (ES) and knowledge-based systems (KBS), are used synonymously. Taken together, they represent the most widespread type of Al application.

The area of human intellectual endeavor to be captured in an expert system is called the task domain. Task refers to some goal-oriented, problem-solving activity. Domain refers to the area within which the task is being performed. Typical tasks are diagnosis, planning, scheduling, configuration and design.

Expert systems are the specialized software, each of which is the expert in some narrow subject

domain. As against usual mathematical programs which manipulate the data, expert systems have base of knowledge and in addition manipulate knowledge.

Expert systems help and allow to solve intellectual problems on the basis of the accrued base of knowledge reflecting an operational experience of experts in considered area. Now expert systems are used in various problem areas. More often expert systems are applied in business, in manufacture and in medicine. Expert systems are applied to the decision of difficult problems, but on quality of accepted decisions they do not concede to experts - experts and have "transparency" of decisions. On the other hand, they are capable to fill up the bases of knowledge in interrelation with the expert.

The basic conditions determining expediency and an opportunity of creation of expert systems:

1. There should be the experts having significant experience and capable to verbalize and to explain techniques used by them in the given subject domain;
2. The solved problem should be sufficient is difficult, and its decision to be based on reasoning;
3. The solved problem should concern to the structured area in which the basic concepts should be allocated: definitions, relations and ways of reception of the decision of the given problem.

Building an expert system is known as **knowledge engineering** and its practitioners are called knowledge engineers. Knowledge engineering is the discipline of acquiring, encoding and using human domain knowledge to develop a computer application. The knowledge engineer must make sure that the computer has all the knowledge needed to solve a problem. The knowledge engineer must choose one or more forms in which to represent the required knowledge as symbol patterns in the memory of the computer - that is, he (or she) must choose a knowledge representation. He must also ensure that the computer can use the knowledge efficiently by selecting from a handful of reasoning methods.

At the initial stage of development of expert (intellectual) systems the majority of theorists and developers considered, that the problem of noegenesis can be successfully solved by dialogue of the engineer on knowledge with the expert, the expert in applied area. But here the well-known paradox was showed now - the expert is more qualified, the less it is capable to explain the reasoning. On the other hand, developers of intellectual systems have faced a number of problems where the expert basically can describe in formal language "mechanism" of decision-making, but does not wish to open the professional secrets serving as a source of his well-being and prosperity.

The marked reasons have caused now a priority of other approach to creation of knowledge bases of the expert systems, received the name “detection of knowledge in databases” (knowledge discovery in databases - KDD). For today this approach is considered the most actual. If the person, the expert in applied area, cannot state a course of the ideas at decision-making let for him the computer program will try to make it. Thus from the expert demonstration of samples as data sets, with the resulting judgments of the expert "adhered" to them (training sample) is required only.

KDD process includes some stages. Accumulation of the crude data, selection, preparation, transformation of the data, and search of laws in the data, an estimation, generalization and structuration of the found laws here concerns.

Modern requirements to data processing with the purpose of detection of knowledge:

- The data have unlimited volume;
- The data are heterogeneous (quantitative, qualitative, categorical);
- Results should be concrete and clear;
- Tools for processing the "crude" data should be simple in use.
- Basic Components of Expert System
- Every expert system consists of two principal parts: the knowledge base; and the reasoning, or inference, engine.

Expert systems must have 3 constituents (functional assemblies) (fig. 1, A; arrows mark information flows):

1) knowledgebase;

2) “inference engine” (inference machine, reasoning engine) - some means of using of knowledge;

3) “human-computer interface” - means of communication with the user.

By other point of view, expert systems must have 4 functional blocks (fig.1, B):

1) knowledge base;

2) block of conclusion (deduction);

3) block of explanation of deduction;

4) block of accumulation and edition of knowledge base.

Second and third blocks are components of reasoning engine.

In medical expert systems there are necessary block of patient data input and block of concluding information output for user.

Knowledge base: A store of both factual and heuristic knowledge. That is, knowledge base is a set of rules and facts describing the domain of an application. An ES tool provides one or more knowledge representation schemes for expressing knowledge about the application domain. Some tools use both frames (objects) and IF-THEN rules. The knowledge is represented as logical statements frequently.

Examples:

The reliable sign of rheumatism is formed heart trouble - FACT

Erythrocytes live in human bloodstream near 30 days - FACT

If heart noise is discovered, it must be analyzed with phonocardiography - RULE

Factual knowledge is that knowledge of the task domain that is widely shared, typically found in textbooks or journals, and commonly agreed upon by those knowledgeable in the particular field.

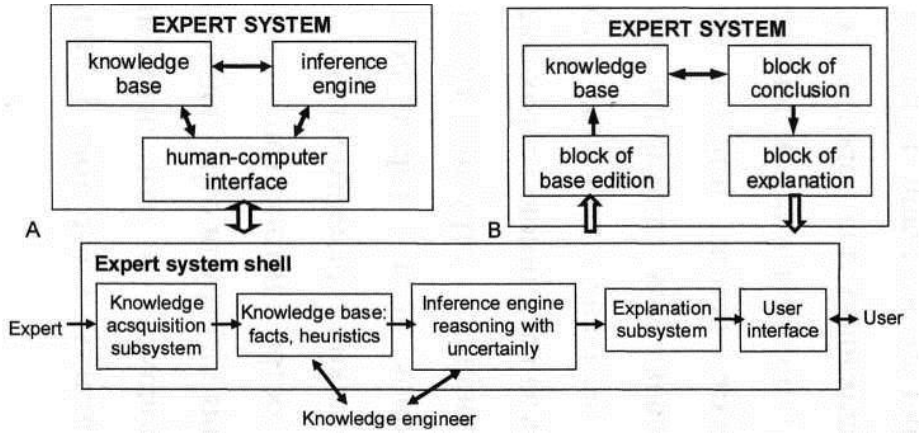

Figure 1. Basic Components of Expert System Tools.

Various levels of structure complexity are shown.

Heuristic knowledge is the less rigorous, more experiential, more judgmental knowledge of performance. In contrast to factual knowledge, heuristic knowledge is rarely discussed, and is largely individualistic. It is the knowledge of good practice, good judgment, and plausible reasoning in the field. It is the knowledge that underlies the "art of good guessing."

Knowledge representation formalizes and organizes the knowledge. One widely used representation is the production rule, or simply rule A rule consists of an IF part and a THEN part (also called a condition and an action). The IF part lists a set of conditions in some logical combination. The piece of knowledge represented by the production rule is relevant to the line of reasoning being developed if the IF part of the rule is satisfied; consequently, the THEN part can be

concluded, or its problem-solving action taken. Expert systems whose knowledge is represented in rule form are called rule-based systems.

Another widely used representation, called the unit (also known as frame, schema, or list structure) is based upon a more passive view of knowledge. The unit is an assemblage of associated symbolic knowledge about an entity to be represented. Typically, a unit consists of a list of properties of the entity and associated values for those properties.

Knowledge acquisition subsystem: A subsystem to help experts build knowledge bases. Collecting knowledge needed to solve problems and build the knowledge base continues to be the biggest bottleneck in building expert systems.

Reasoning engine is a program that imposes a general control strategy on how the system is working. It includes inference mechanisms for manipulating the symbolic information and knowledge in the knowledge base to form a line of reasoning in solving a problem. It includes block of conclusion (deduction) and block of explanation of deduction. Shortly, reasoning engine is a set of rules for making deduction.

Explanation subsystem: A subsystem that explains the system's actions. The explanation can range from how the final or intermediate solutions were arrived at to justifying the need for additional data.

User interface: The means of communication with the user. The user interface is generally not a part of the ES technology, and was not given much attention in the past. However, it is now widely accepted that the user interface can make a critical difference in the perceived utility of a system regardless of the system's performance.

Output of expert system can be medical opinion, recommends for treatment, surgical operations, drug and physiotherapeutic prescriptions. They have consultative character.

Security duties: all information in the database (as in knowledge base, as in set data about analyzed case) must be absolutely correct to protect patient from any mistakes.

Integrating reasoning: an expert system is computer-based system (mainly software) that uses knowledge and facts, and apply an appropriate reasoning technique (inferencing) to solve problems in a given field (domain) that normally require the services of human experts.

The main features of expert system are:

- it is limited to a specific domain (area of expertise); it is a rule based usually;
- it can reason with uncertain information (yes / no / don’t know answer from the user);
- it counsels (deliver advices); it explain its reasoning to the user.

The primary goal of expert systems research is **to make expertise available to decision makers and technicians who need answers quickly**.

These knowledge-based applications of artificial intelligence have enhanced productivity in business, science, and engineering.

Each new deployment of an expert system yields valuable data for what works in what context, thus fueling the Al research that provides even better applications.

Wide introduction of expert systems of diagnostics of diseases can raise economic efficiency of diagnostics:

Creation of expert systems can lower loading on the personnel and enable to make satisfactory decisions to less qualified personnel.

Application of expert systems is possible as the training manual for increase of a skill level by students and the medical personnel that will reduce time for training of young doctors, will allow young doctors-therapists at diagnostics and treatment of patients to achieve level of leading experts.

Application of expert systems will allow to raise overall performance as a whole owing to improvement of quality of the accepted decisions (more exact statement of the diagnosis), reduction

of time of diagnostics of patients problems, an opportunity of forecasting of consequences of the accepted decision, the analysis and reduction of time of a choice of the best decision.

Expert systems can be used for an estimation and probabilistic forecast of dynamics of functional reserves of an organism, rendering of the advisory help in a choice of optimum structure of the profound inspection, ways and methods of correction of a health state, and also special preparation and training of medical staff not only in the clinical purposes, but also in preventive medicine.

Advantages of expert systems

- Provides a second opinion on a rare or complex medical condition.
- Compares data of many of previous cases, giving a clear picture of a situation, offers the best action based on historical cases.
- Collect expert data of many places, which would otherwise be unavailable.
- Helps doctors to evaluate probability of success (surgical of therapeutical).
- Reduce the possibility of errors.
- Helps doctors to make decisions which could save time or money.
- A large database of knowledge can be added to anf kept up-to-date - it can store more knowledge that a person.
- The system cannot “forget" or get facts wrong.
- It survives forever, will not retires.

Disadvantages of expert systems

- The database must be updated regularly (as new methods of a treatment, drugs, and operations appear).
- Quality of the output depends on quality of expert data.
- Quality of the output depends on quality of input data.
- The majority of patients prefer to talk to a doctor rather than search through s computer system.

Data mining technologies

The basic analytical tools satisfying listed requirements today carry to area of technologies Data Mining (excavation of the data). The concept of patterns (models, samples) is put in a basis of these technologies and the dependences reflecting multidimensional mutual relations in the data.

Search of patterns is made by automatic methods which have been not limited to frameworks of aphoristic assumptions of structure of sample and a kind of distributions of values of analyzed parameters.

Important position of Data Mining - non-triviality of searched patterns. It means, that they should reflect unevident, unexpected regularities in the data, the making so-called latent Knowledge (hidden Knowledge). Many experts have realized that the especial analytical toolkit corresponding to their system complexity is necessary for revealing laws in the real vital phenomena. In turn, to a society there has come understanding, that the "crude” data contain a deep layer of knowledge at which competent excavation the present nuggets can be found out.

Systems Data Mining are applied on two basic directions: 1) as a mass product for business - appendices; 2) as tools for carrying out of unique researches (genetics, chemistry, medicine and so forth). The quantity of installations of mass products, by available data, achieves tens thousand.

Term Data Mining which has appeared in 1978 has got high popularity approximately since first half 90th years.

For example, the methods of traditional mathematical statistics making a basis of statistical packages, are useful mainly to check of beforehand formulated hypotheses (verification - driven data mining) and for the “rough” prospecting analysis making a basis of operative analytical data processing (online analytical processing, OLAP). The main reason of the limited efficiency of the

majority of procedures for revealing interrelations in the data included in statistical packages, - the concept of averaging on the sample, resulting in operations above nonexistent sizes (for example, average temperature of patients on hospital, average height of a house in the street, consisting of palaces and hovels, etc.). So-called “multivariate methods” such as discriminant, factorial and other similar kinds of the analysis come to an end result through operations above fictitious vectors of average values, and also covariance and correlation matrixes. Therefore, their results are quite often inexact, sin adjustment and absence of sense.

The software realizing neural-network approach, also quite often carry to category Data Mining. The basic lack of classical neural-network paradigms consists that the neural network represents a "grey" box. First, the topology of neural-network here is set proceeding from heuristic reasons. And, second, in trained neural networks with complex topology of weight of hundreds and thousands interneuronal connections do not give in to the analysis and interpretation by human.

The approach connected to development so-called self-organizing (growing or evolable) Boolean neural networks which structure gives in to decoding as logic statements, corresponds to the purposes and problems Data Mining, but suffers lacks, of the whole inherent in evolutionary algorithms.

Evolable - evolutionary, evolved.

The idea of systems of reasoning on the basis of similar cases (case based reasoning - CBR) is at first sight extremely simple. To make the forecast for the future or to choose the correct decision, these systems find close analogues of a cash situation in the past and choose the same answer which was for them correct. Therefore this method still names a method “the nearest neighbor” (nearest neighbor). Recently distribution was received also with the term “memory based reasoning”" which accents attention, that the decision is accepted on the basis of the information which has been saved up in memory.

CBR systems show quite good formal results in the diversified problems. Their main minus count that is impossible to tell on the basis of what concrete factors are basis on which CBR system build the answers. Other, more serious minus consists in an arbitrariness which is supposed by CBR systems at choice of “affinity” measure. Besides it can be groundless distribution of the common measure of affinity on data selection as a whole.

Methods of search of logic laws in the data satisfy to requirements of Data Mining in the greatest measure. Their results, are expressed as IF - THEN and WHEN - ALSO rules more often.

Formalization of subject knowledge within the framework of expert system is carried out by use of deciding rules as statements of type: IF (condition) - THAT (decision).

With the help of such rules problems of forecasting, classification, recognition of images, segmentations of databases (DB), extraction from the given "latent" knowledge, interpretations of the data, establishments of associations in a DB are solved, etc. Logic methods work in conditions of the diverse information. Their results are effective and transparent for perception. We shall consider the basic approaches to search of logic laws in databases.

THE BASIC APPROACHES TO SEARCH OF LOGIC LAWS

Trees of decisions (decision trees) are the approach most widespread now to revealing and the image of logic laws in the data. Algorithms of limited search calculate frequencies of combinations of simple logic events in subgroups (classes) of the data. Examples of simple logic events: X = C1; X <C2; X> C3; C4 <X <C5, etc., where X - some parameter (field), C, - constants. As restriction the length of a combination of simple logic events (at M.Bongard it was equal 3) serves.

On the basis of comparison of the calculated frequencies in various subgroups of the data it is concluded utility of this or that combination for an establishment of association in the data, for classification, forecasting and so forth.

Evolutionary algorithms. Many researchers see a way of development of analytical methods to development of evolutionary algorithms. Among them the most popular are the genetic algorithms, trying to model mechanisms of a heredity, variability and selection in wildlife.

Genetic algorithms are attractive that them is convenient to parallelizate (to multisequence). For example, it is possible to break generation on some groups and to work from each of them independently, providing from time to time an intergroup exchange of several chromosomes. Exist as well other methods of multisequencing of genetic algorithms. At the same time, these algorithms for today have serious disadvantages.

Primarily the term neural network had been used to refer to a network or circuit of biological neurons now artificial neural networks, which are composed of artificial neurons or nodes, are programming constructs that mimic the properties and behavior of biological neurons.

Artificial neural networks have been applied successfully to speech recognition, image analysis and adaptive control, in order to construct autonomous robots and so on. Most of the currently employed artificial neural networks for artificial intelligence are based on statistical estimation, optimization and control theory.

Neural networks (NN) are:

- mathematical models that resemble nonlinear regression models, but are also useful to model nonlinearly separable spaces;
- "knowledge acquisition tools" that learn from examples.

Tasks for self-check:

Task 1:

6. 1. FIRST EXPERT SYSTEMS BEGUN TO DEVELOP IN:
   a) 1950
   b) 1920
   c) 1970
   d) 1990
   e) 2016
7. THIS ALGORITHM CALCULATE FREQUENCIES OF COMBINATIONS OF SIMPLE LOGIC EVENTS IN SUBGROUPS
   a) Evolutionary algorithm
   b) Trees of decisions
   c) Algorithms of limited search
   d) Genetic algorithms
   e) Key algorithm
8. THE SYSTEM "THE FINDING OF LOGIC LAWS ON THE BASIS OF REPRESENTATIONS OF LOCAL GEOMETRY" IS CALLED
   a) Wiz Why
   b) Deep Data Diver
   c) See5
   d) NN
   e) WWW
9. ONE OF THE MAIN CHARACTERISTICS OF THE ES:
   a) The practicality
   b) Performance
   c) Completeness
   d) Effectiveness

e) Wisdom

10. MEDICAL KNOWLEDGE IN THE ES REPRESENTED AS:

a) algorithmic model
b) a set of unrelated features
c) a set of features, some of which are interconnected
d) the text model
e) table

References:

Basic.

1. Olenets S.Yu. Medical informatics [Text]:Tutorial guide / Olenets S.Yu.: HSEE of Ukraine "UMSA". – Poltava: TOV "ASMI", 2017. – 160 p.:im.
2. Handbook of Medical Informatics. Editors: J.H. van Bemmel, M.A. Musen. – http://www.mieur.nl/mihandbook; http://www.mihandbook.stanford.edu
3. Mark A. Musen B. Handbook of Medical Informatics // Електронний ресурс ftp://46.101.84.92/pdf12/handbook-of-medical-informatics.pdf
4. Edward H., Shortliffe J., Cimino J. Biomedical Informatics, 2014 // Електронний ресурс: http://www.rhc.ac.ir/Files/Download/pdf/nursingbooks/Biomedical%20Informatics%20Computer%20Applications%20in%20Health%20Care%20and%20Biomedicine-2014%20-%20CD.pdf
5. Коровіна Л.Д. Медична інформатика : навчальний посібник для студентів вищих медичних навчальних закладів / Л. Д. Коровіна - Полтава : РВВ УМСА, 2008. – 144 с. – англ. мовою.
6. Marzeniuk, V.P. Biophysics and medical informatics : Manual for Students of the Higher Medical Schools of the III-IV Degree of Accreditation / V.P. Marzeniuk, V.D. Didukh, D.V. Vakulenko at al. – Ternopil : Ukrmedknyha, 2004. Vol. 1: – 479 c. :

Additional.

1. www.imia.org (Міжнародна Асоціація Медичної Інформатики)
2. www.mihandbook.stanford.edu (Медична інформатика, Стенфордський університет)
3. www.ncbi.nlm.nih.gov (Національна бібліотека медицини США)
4. www.cochrane.ru (Розділ Кохранівського співтовариства)

The methodical guidance has been completed by **S.Y. Olenets**