A method for automatic medical diagnosis

Igor Grabec [1,*], Eva Švegl [2], Mihael Sok [2]

[1]*Slovenian Academy of Sciences and Arts,* [2]*University Medical Centre, Ljubljana, Slovenia*

Abstract   This research paper presents a new method for the automatic diagnosis of diseases using a personal computer. Forming a basis for the characterization of diseases, a wide set of symptoms is introduced, and a particular disease is characterized by a set of statistical weights assigned to those symptoms. Information about the patient state is provided by a graphic interface in which the user confirms symptom indicators. Agreement between these symptoms and classified symptoms of a particular disease is then estimated by the sum of corresponding weights, where the disease corresponding to the maximal agreement is proposed as the result of the diagnosis. A disease likelihood estimator is calculated and presented to assess the reliability of the diagnosis. With regard to the automatic assessment of the diagnosis the corresponding algorithm and the properties of the computer program are included. Finally, the effectiveness of this method of medical diagnosis is demonstrated through four typical examples involving differently expressed symptoms. The diagnostic system resembles semantically driven sensory-neural network.

1. Introduction

Diagnosing diseases is a primary task of doctors in clinical practice. Over the evolution of medical science a rather complex set of symptoms and corresponding methods for their estimation has been introduced. In the past symptoms were chiefly evaluated by direct observations, while nowadays ever more instrumental techniques, and with them related quantitative variables, are utilized for this purpose [1, 2, 3]. Along with this development various information processing techniques are invented and introduced into the medical practice with the aim of supporting doctors decision making for diagnoses, as well as in cases where people are interested in assessing their health states [2, 3, 4, 5]. Consequently, the research, innovations and new developments in this field are widely supported by various foundations at national and international levels. For instance, a recent call for project proposals devoted to the development of new automatic methods for diagnosis of several outstanding diseases frequently present in western populations was also the main motivation for our research on this field [6]. Our work was mainly oriented to the development of a simple method for assessment of diagnosis that could be carried out automatically on a personal computer based on an exchange of information with a user [4, 5, 7]. The aim of this article is to describe this method and present the performance of the corresponding system.

In accordance with the conditions specified in the call for project proposals we predominantly consider data about a human's state that can be specified descriptively, although data of quantitative character are not excluded [4, 5]. The basic problem on the path to an automatic assessment of diagnosis is then to transform semantically specified data of a patient into a quantitative form that could be further utilized in a computer program [7, 8].

---

*Correspondence to: Igor Grabec (Email: igor.grabec@amanova.si) Kantetova ul. 75, SI-1000 Ljubljana, Slovenia.

An equivalent problem also relates to properly describing disease characteristics scattered across the vast medical literature and contained within the professional knowledge of doctors [1, 7, 8]. To overcome these problems we first specify descriptively an acceptable fixed set of symptoms and then perform its encoding into quantitative variables. Utilization of such variables further enables automatic execution of a medical diagnosis by a personal computer [4, 5].

A diagnosis generally represents a mapping of the examined patient's data to symptoms characterizing diseases and an optimal assessment of the corresponding disease [8, 9, 10]. Therefore, the basic problem in the development of an automatic method for assessment of the diagnosis is to specify a mathematical algorithm for optimal mapping [8, 11]. One might expect that applying statistical methods from medicine supported by corresponding modeling and prediction of natural laws as well as various methods developed in the field of artificial intelligence could be applicable for this purpose [1, 12, 13, 14, 15, 16]. The goals of this article, however, are to introduce a relatively simple algorithm for generating diagnoses and to demonstrate the performance of the corresponding computer program [4, 5]. Although various web pages and even cell phones already offer a variety of such applications [1, 12, 14, 16, 17], the corresponding algorithms are often not explained. For the specification of the algorithm we consider the performance of a doctor when assessing a diagnosis based on a conversation with the patient being treated. A doctor primarily asks the patient about his symptoms and then compares that obtained data with known symptoms describing properties of various diseases. We try to specify an optimal comparison by characterizing the significance of various symptoms in terms of statistical weights. The result of a doctor performing this comparison is here described by the estimator of agreement between the patient's symptoms and diseases [4, 5]. Its determination can be left to the computer, together with the presentation of the result of the diagnosis. As a result of the process of diagnosis, the disease with the maximal agreement is finally proposed. To support this diagnosis decision the likelihood of the disease, given the data, is described by a corresponding statistical estimator [4, 5].

In the following paragraphs we introduce variables, databases and the algorithm for the automatic estimation of diagnosis. Based on these requisites, we then describe the composition of the computer program and demonstrate its performance with four characteristic examples of variously exhibited symptoms.

Table 1. Selected diseases

d	Disease name - $D(d)$
1	Anemia
2	Urinary tract infection
3	Diabetes, type 2
4	Atrial fibrillation
5	Acute hemorrhagic stroke
6	Obstructive sleep apnea
7	Tuberculosis
8	Chronic obstructive pulmonary disease
9	Pneumonia
10	Otitis media
11	Leukocytosis
12	Hepatitis A
13	Hypertension
14	Throat inflammation
15	No disease

2. Bases of automatic diagnosis

2.1. Description of symptoms and diseases

In accordance with the call for project proposals [6] we select the 14 most frequent diseases in the western world. Their names $D(d)$ are given in Table 1. Each sample is denoted by the disease index d. In this set the state of "*no disease*" is also included, so that it contains $N_d = 15$ samples in total.

The properties of diseases are characterized by the set of $N_s = 46$ symptoms $S(s)$ presented in Table 2 and denoted by the symptom index $1 \leq s \leq N_s$. The properties of a particular symptom $S(s)$ are described by several possible indicators, for example: $I(s, i) = \{\text{yes}, \text{no}, \text{fever}, \ldots\}$. In the complete set of symptoms 29 samples are described by only two indicators: $I(s, i) = \{\text{yes}, \text{no}\}$, while the maximal number of indicators can be $N_i = 9$. All options are presented in Table 3. Among them the empty element $\emptyset$ is also treated as an indicator, although it provides no information and does not contribute to the diagnosis.

Table 2. Symptoms used for characterization of diseases

s	Symptom name - $S(s)$	s	Symptom name - $S(s)$
1	no symptoms	23	thirst, dry mouth
2	sudden onset of illness	24	nervousness, depression
3	hyperthermia	25	throat irritation
4	malaise, tiredness	26	pharyngeal dryness
5	age over 70	27	hoarseness
6	headache	28	obstruction in the throat
7	sudden headache	29	jaundice of skin and sclera
8	cough	30	light feces
9	dizziness	31	frequent infections
10	shortness of breath,	32	ability to lift both arms
	trouble breathing	33	difficulty when speaking
11	pain	34	difficulty moving one side
12	loss of weight in	35	sudden infertility
	the last 6 months	36	difficult verbal expression
13	loss of appetite	37	vomiting
14	irregular pulse	38	snoring during sleep
15	sweating	39	smelly urine
16	difficulty concentrating	40	heart rate
17	sleep disturbance,	41	arterial hypertension
	frequent night waking	42	pulse rhythmicity
18	distorted vision	43	oxygen saturation
19	frequent urination		- pulse oximetry
20	tinnitus	44	breathing frequency
21	bleeding	45	body height
22	dark urine in last 2 weeks	46	body weight

Properties of a particular disease can be characterized by the presence or absence of various symptom indicators. With this aim we can utilize descriptions of the selected diseases $\{D(d); d = 1, \ldots, N_d\}$ from medical literature and indicate by the values 0 or 1 whether an indicator of a particular symptom is characteristic for a selected disease. Although such indications offer simple characterizations of diseases, they are not sufficient to determine an acceptable assessment of a diagnosis [4, 5].

To proceed towards an acceptable diagnosis we consider as an example a patient with disease $d = 14$ - *throat inflammation*. For this disease, the characteristic symptoms are *pain in the pharynx* and *hyperthermia with fever*

Table 3. Sets of possible indicators for the identification of symptoms

s	Possible indicators $I(s, :)$
3	{fever, y, n}
6	{y, n, in the morning, daily persistent}
8	{y-in the morning, y-with sputum, y-daily persistent, y-without sputum,n}
10	{at rest, on exertion, n}
11	{chest,ear,abdomen,pharynx,at swallowing,at speaking,at urinating,shoulder,n}
15	{daily persistent, nocturnal,n}
18	{y, n, one eye only}
19	{y, n, at night}
21	{from the nose, in the sputum, in the urine, n}
31	{urinary tract, tongue, corners of the mouth, lung, n}
40	{<50, 51-70, 71-90, >90}
41	{<100, 100-140, >140}
42	{rhythmic, arrhythmic}
43	{≤ 90, >90}
44	{<10, 11-15, >15}
45	{<150, 151-160, 161-170, 171-180, 181-190, >190} cm
46	{<50, 51-60, 61-70, 71-80, 81-90, 91-100, >100} kg

All other possible indicators are specified by: $I(s, :) = \{\text{yes}, \text{no}\} = \{\text{y}, \text{n}\}$. Components of indicators that are not denoted in Table 3 correspond to empty space $\emptyset$, for example: $I(s, :) = \{\text{fever}, \text{y}, \text{n}\} \equiv \{\text{fever}, \text{y}, \text{n}, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$.

that are described by *symptom* - 11, *indicator* - 4 and *symptom* - 3, *indicator* - 1, respectively. When doctors are looking for a diagnosis, they consider the first indicator to be more important than the second one. A similar consideration is characteristic for a patient with disease $d = 7$ - *tuberculosis*. For this patient a *cough over entire days* is characteristically described by the *symptom* - 8, *indicator* - 3. The fact that tuberculosis is usually not developed suddenly is characterized by *symptom* - 2, *indicator* - 2. As in the first example, in this case the second symptom indicator is usually considered to be less important than the first one. In order to describe the importance of symptom indicators we assign to each disease indicator $I(d, s, i)$ a quantitative statistical weight $W(d, s, i)$. Its value can be determined statistically either by using specialists' suggestions or by more extensive clinical research. In the cases mentioned above, doctors would suggest assigning approximately *three times* greater importance to the first mentioned symptom than to the second one. According to the meaning of the weight, a non important, i.e. empty indicator, is assigned the weight $W = 0$.

Characterization of symptom importance by statistical weights was also performed in our research project [4, 5]. The corresponding data obtained from the clinical environment are presented in Table 4. For example, in the previously mentioned cases the weights of the first and second indicator have the values 3 and 1, respectively. This means that *pain in the pharynx* (or *indicator of cough*) is three times more important for the diagnosis of *throat inflammation* (or *tuberculosis*) than the indicator of *hyperthermia with fever* (or *slow onset*).

With respect to the meaning of weights we conclude that they can even be negative, as is for example: $W(8, 8, 5) = -3$. The negative value of this weight exhibits the absence of the disease with index 8. The disease with index $d = 8$ is *chronic obstructive pulmonary disease*. The symptom with $s = 8$ is *cough*. The indicator with index $i = 5$ is *no*. When this indicator is confirmed, we interpret it as "The absence of cough indicates with weight 3 the absence of chronic obstructive pulmonary disease". However, such examples are rather rare.

By statistical weights the properties of diseases that are descriptively presented in medical science are in fact quantitatively represented. Such presentation mathematically resembles a definition of a function on the set of symptoms. For each disease there generally corresponds a different specific function that is denoted by the index d, and it turns out that the introduced set of statistical weights $\{W(d, s, i); 1 \leq d \leq 15; 1 \leq s \leq 46; 1 \leq i \leq 9\}$ is sufficient for the development of an automatic diagnosis method [4, 5]. However, different diseases could have

different numbers of symptom indicators as well as different weights. In order to compare functions corresponding to different diseases between themselves, it is reasonable and necessary to normalize the weights so that the sum of their positive values equals 1 and negative values equals -1. With this aim we first determine for each disease $D(d)$ the total positive and negative weight by the following sums:

$$W_{t+}(d) = \sum_{s=1}^{N_s} \sum_{i=1}^{N_i} W(d,s,i)H(W(d,s,i)) \quad ; \quad W_{t-}(d) = \sum_{s=1}^{N_s} \sum_{i=1}^{N_i} W(d,s,i)H(-W(d,s,i)) \qquad (1)$$

Here, the Heaviside function $H(x)$, having values: $H(x) = 0$ for $x \leq 0$ and $H(x) = 1$ for $x > 0$, accounts for taking just positive or negative weights in the summation process. A normalized weight $W_n(d,s,i)$ is then determined by dividing previously introduced weights by the corresponding total weights:

$$W_n(d,s,i) = \frac{W(d,s,i)H(W(d,s,i))}{W_{t+}(d)} + \frac{W(d,s,i)H(-W(d,s,i))}{|W_{t-}(d)|} \qquad (2)$$

By this normalization the importance of symptom indicators is described relatively inside the framework pertaining to the disease $D(d)$, and consequently, each disease is further treated in a similar way.

The set of normalized weights is here considered to be a proper quantitative basis to make an automatic assessment of a diagnosis [4, 5]. For this purpose we have to determine which function, or more to the point, which index d_o, optimally corresponds to the state of the examined patient. The corresponding transformation of symptoms into the disease index is defined in the next subsection.

Table 4. Statistical weights describing importance of symptom indicators

d	The values denote: (symptom index s, indicator index i, weight $W(d,s,i)$)
1	(2,2,1) (3,3,1) (4,1,1) (6,1,1) (9,1,1) (10,1,1) (14,1,1) (16,1,1) (17,1,1) (20,1,1) (31,1,1) (31,3,2) (40,4,1) (44,3,2)
2	(2,1,1) (3,1,3) (3,2,1) (11,3,2) (11,8,4) (19,1,3) (22,1,3) (39,1,3)
3	(2,2,2) (3,3,2) (4,1,1) (12,2,2) (18,1,1) (19,1,2) (31,1,2)
4	(2,1,3) (3,2,3) (5,1,3) (10,2,2) (11,1,2) (14,1,3) (41,2,1) (41,3,3) (42,2,3)
5	(2,1,2) (3,3,1) (7,1,2) (9,1,1) (18,3,2) (32,2,2) (33,1,2) (34,2,2) (35,1,2) (36,1,2) (37,1,1)
6	(2,2,1) (4,1,1) (6,3,1) (16,1,1) (17,1,3) (23,1,1) (24,1,2) (41,3,1) (46,6,2) (46,7,3)
7	(2,2,1) (3,2,1) (4,1,1) (8,3,3) (11,1,1) (12,1,1) (13,1,1) (15,2,3) (21,2,3)
8	(2,2,2) (3,3,1) (8,1,3) (8,2,3) (8,5,-3) (31,4,2)
9	(2,1,2) (3,1,3) (3,2,1) (4,1,2) (8,2,1) (8,3,1) (8,4,1) (10,1,1) (10,2,1) (11,1,1)
10	(2,1,2) (3,2,1) (4,1,1) (6,4,2) (11,2,3) (29,1,1)
11	(2,2,1) (3,2,1) (4,1,1) (9,1,1) (10,1,1) (10,2,1) (12,1,1) (13,1,1) (15,1,1) (18,1,1)
12	(2,1,2) (3,2,1) (4,1,3) (11,3,2) (13,1,2) (29,1,3) 22,1,3) (30,1,3) (37,1,2)
13	(2,2,1) (3,3,2) (5,1,1) (6,4,2) (9,1,1) (21,1,1) (41,3,6)
14	(2,1,2) (3,1,1) (11,4,3) (25,1,2) (26,1,1) (27,1,1) (28,1,1)
15	(1,1,1) (2,2,1) (3,3,1) (4,2,1) (6,2,1) (8,5,1) (11,9,1) (13,2,1) (16,2,1) (18,2,1) (21,4,1) (22,2,1) (23,2,1) (24,2,1) (25,2,1) (26,2,1) (27,2,1) (28,2,1) (29,2,1) (30,2,1) (31,5,1) (32,2,1) (33,2,1) (34,2,1) (35,2,1) (36,2,1) (37,2,1) (38,2,1) (39,2,1)

Weights at other indices are 0.

2.2. *Diagnosis algorithm and its corresponding process*

If we want to assess the diagnosis of a certain patient, we must obtain information about his/her symptoms. This is most simply carried out through a questionnaire containing names of symptoms and their indicators

which are confirmed by the patient. Specific confirmed information is then represented by a term of value 1 in the response matrix $\{R(s,i); s = 1 \dots N_s; i = 1 \dots N_i\}$, while all other non-confirmed terms are 0. The value $R(s,i) = 1$ denotes that the corresponding symptom indicator $I(s,i)$ is confirmed in the questionnaire, while $R(s,i) = 0$ denotes the opposite. Although this matrix quantitatively describes the state of the considered patient, it does not take into account the importance of different symptom indicators. To accomplish this part of the diagnosis process it must be properly related to the description of diseases in terms of weights. To proceed in this direction let us consider the performance of a doctor who is trying to make a diagnosis of the patient based on the questionnaire completed by the patient. With this aim the doctor considers only those indicators that are confirmed by the patient and compares them sequentially with symptoms of all diseases. In order to decide whether the patient has a certain disease $D(d)$ the doctor compares the confirmed indicators with the indicators of that selected disease. In the case of agreement the doctor assigns to the confirmed symptom indicator a corresponding normalized statistical weight and then collects all weights to obtain the total weight for a patient's condition associated with the disease $D(d)$. It is given by the expression:

$$A(d) = \sum_{s=1}^{N_s} \sum_{i=1}^{N_i} W_n(d,s,i) R(s,i) \tag{3}$$

and represents the estimator of the agreement between the patient's symptoms and the symptoms of the disease $D(d)$. This procedure, when repeated on the complete set of diseases, yields the set of values $\{A(d); d = 1, \dots, N_d\}$. In the case where all symptoms of a particular disease coincide with the symptoms confirmed by the patient, the value of the agreement estimator is $A = 1$ (or 100%). When assessing the diagnosis, the doctor normally selects the disease with the maximal value of agreement $A(d)$ as the conclusion of diagnosis.

Although the above-described decision appears quite acceptable, it is convenient to introduce a quantitative variable that supports it. To this end we first determine the mean value $<A> = \sum_{d=1}^{N_d} A(d)/N_d$ and the standard deviation $\sigma_A = \sqrt{\text{var}(A)}$ of the distribution of the agreement estimator over the set of diseases $\{A(d); d = 1, \dots, N_d\}$. As the variable for our test, we next consider the relative deviation of the agreement estimator from the mean value: $\Delta_A(d) = (A(d) - <A>)/\sigma_A$. If we observe $\Delta_A(d_o) \gg 1$ in the occurrence of a certain d_o we can conclude that the corresponding $D(d_o)$ is significantly outstanding and should therefore be considered to be a reliable quantitative result of the diagnosis.

The above reasoning is possible when the symptoms of a particular disease are well exhibited. However, this is not always the case, since several diseases share similar symptoms. This leads to similar values of the agreement estimator and small deviations from the mean value. Consequently, a question arises as to how to justify the decision on a diagnosis when $0 < \Delta_A(d_o) < 1$. To address this possibility we divide the agreement estimator $A(d)$ by the sum over the set of all diseases: $\sum_{d=1}^{N_d} A(d)$ and define the disease likelihood estimator $L(d)$ as:

$$L(d) = \frac{A(d)}{\sum_{d'=1}^{N_d} A(d')} \tag{4}$$

This estimator describes the reliability of our decision. A high value serves as a firm argument for selecting the corresponding disease as a reliable result of the diagnosis. It is important to note that $A(d)$ as well as $L(d)$ are quantitative in character and thus could be considered as supplementary to other quantitative data in clinical tests.

2.3. Interpretation of diagnostic system as a sensory-neural network

The diagnostic system introduced by Tables 1-4 can be interpreted as a description of a semantically driven *sensory-neural network* in which a particular sensor corresponds to a symptom from the Table 2, while a neuron corresponds to a disease from the Table 1 [5]. The possible states of a particular sensor $S(s)$ are described by the indicators $I(s,:)$ in Table 3, while the properties of synapses on neurons are characterized by the weights $W(d,s,i)$ given in Table 4. Accordingly, the response matrix $R(s,i)$ represents the signals generated by the sensors during detection of a patient state. When transferred over synaptic weights $W(d,s,i)$ to the neurons, these signals excite them as determined by the agreement estimator $A(d)$. In this context the normalization of weights $W(d,s,i)$ corresponds to a specification of the neural response function having a sigmoid character [13] with the span $(-1, +1)$. The

transition from the agreement $A(d)$ to the likelihood estimator $L(d)$ then corresponds to mutual interaction of neurons which yields the common output of neurons equal to 1. To simplify our description we here utilize the set of synaptic weights specified by doctors based on their knowledge; however, the interpretation of the diagnostic system as a sensory- neural network provides for this purpose various training methods developed in the framework of artificial intelligence and neural networks [1, 13].

3. Computer program for diagnosis assessment

3.1. Basic tasks of the program

The values of estimators A and L can be automatically calculated by a personal computer. To accomplish this the databases of symptoms, their indicators and statistical weights have to be stored in a corresponding computer program, while the data for forming the response matrix R of the patient have to be input through a corresponding channel. The last task can be performed simply over a graphic interface that offers the user the description of symptoms $S(s)$ and their indicators $I(s,i)$ and can be confirmed using a keyboard [4, 5]. The program has to transform the input data into the response matrix $R(s,i)$ and then use it in the calculation of the values $\{A(d); 1 \leq d \leq N_d\}$, $\{L(d); 1 \leq d \leq N_d\}$, $< A >$ and σ_A. These values form the output representing the result of the automatic diagnosis. For this purpose it is convenient to display the diagrams of $A(d)$ and $L(d)$ versus $1 \leq d \leq N_d$ and to offer the user the corresponding files containing numerical data. In addition to this the user should have access to the list of the patient's answers and files containing statistical weights W_n used in the calculation of $A(d)$ and $L(d)$. The program should also allow a simple repetition of the complete procedure as well as testing the diagnosis procedure based upon sets of formal symptom indicators and modification to the weights by a medical specialist.

3.2. Description of program performance

In our project we have developed a program in the environment that interacts with a user over a graphic interface [4, 5]. At the start of its operation the graphic interface presents the user with three options: 1) Identifying a diagnosis based on completing the questionnaire, 2) Testing the program performance based on internally stored data on sets of formal symptoms for all diseases, 3) Changing the weights of symptom indicators. After finishing any option, the complete procedure can be repeated.

In the first option of operation a new window with instructions for the user appears together with a window to enter the patient's name. After accepting the name, the program shows sequentially 46 windows with the symptoms and indicators of the diseases. In these windows the symptom data that are confirmed are used for forming the response matrix R. From this matrix, the values of A and L estimators are calculated. The results are transferred to the user over various channels. The most informative is the displayed diagram of $A(d)$ and $L(d)$ distribution versus the disease index d. The lines at the levels of $< A >$, $< A > +\sigma_A$ and $< A > +2\sigma_A$ are used as references for a visual assessment of the diagnosis. An example is shown in Fig. 1. In addition to the diagram displayed in Fig. 1, the files with the patient's responses and corresponding numerical data are available for printing.

In the second option the program displays the set of diseases. After one of them is selected, the program applies the corresponding set of formal symptoms stored among its databases and uses it instead of the patient's symptoms in the same procedure as in the first option. This step shows the user the optimal example of the selected disease diagnosis.

The third option allows specialists to examine how a variation of indicator weights influences the diagnosis process. By adjusting the weights and further testing the diagnosis using the second option, the performance of the complete program can be gradually improved. With this aim the program displays the set of diseases in the first window and allows modification of their weights in the second window. The file with the changed weights is also accessible to the user.
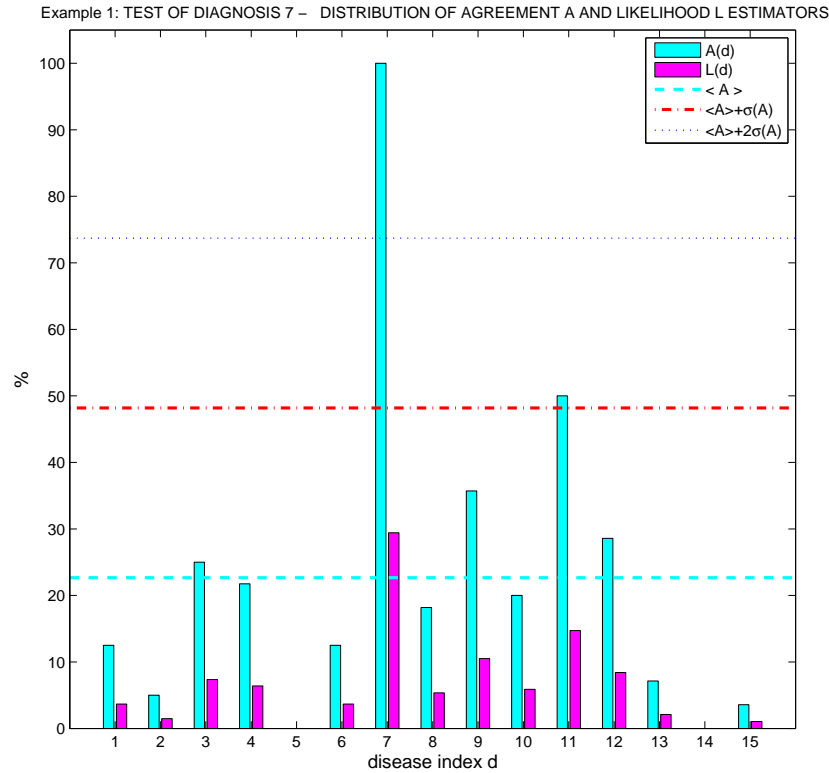
Figure 1. Distribution of estimators A and L in testing the program performance based on the set of optimal symptoms of *tuberculosis* - disease with index 7.

3.3. Testing of the program using formal symptom indicators

To demonstrate the program's performance we first utilize results of its testing performed with the formal indicators of the selected diseases. The first example shows the program's performance in the case where the disease $D(7)$ - *tuberculosis* is selected. In this test we get the diagram shown in Fig. 1. The value of the corresponding optimal agreement estimator is in this case $A_o(7) = 100\%$, and it surpasses the mean value $< A >$ for $3.3\,\sigma_A$. This outstanding deviation from the mean value and other values of A indicates a correct assessment of the diagnosis. But the value of A corresponding to disease *Leukocytosis* with the index $d = 11$ also surpasses the mean value $< A >$ by more than one σ_A. This outcome indicates that the symptoms of both diseases are in a sense similar. In spite of this property, the distribution of $A(d)$ suggests selecting *tuberculosis* as the resulting disease of the diagnosis. Although the agreement with its symptom indicators is $A_o = 100\%$, the likelihood value of *tuberculosis* disease is only $L_o = 29\%$; the other values, however, are still appreciably smaller: $L(d) \ll L_o; d \neq d_o$. It is also interesting that the value $A(15) = 4\%$ indicates the possibility of the state *no disease* but only with low likelihood $L(15) = 1\%$. This is due to the fact that some indicators which are characteristic of the first disease are also characteristic of the absence of a disease, among the considered set of diseases.

Similar properties of the program's performance found in testing the diagnosis of *tuberculosis* disease are also observed when testing is performed using formal sets of other diseases. The corresponding optimal likelihood values $\{L_o(d); d = 1, \dots 15\}$ are:

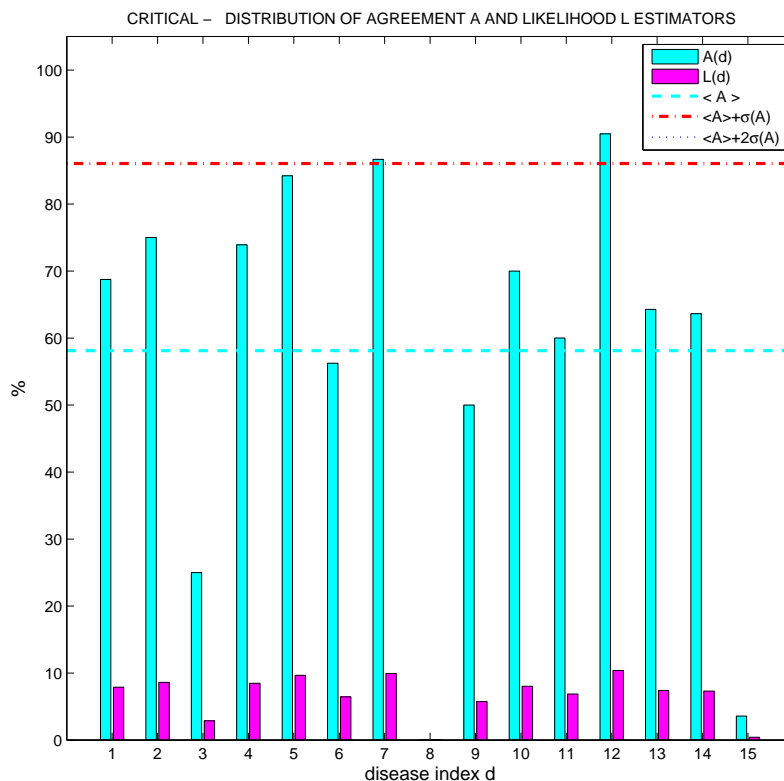$$L_o = (29, 32, 34, 32, 37, 33, 29, 48, 24, 32, 28, 28, 35, 48, 41)\% \tag{5}$$

Figure 2. Distribution of estimators A and L in the case of a patient in a critical health state.

The value $L_o(15) = 41\%$ appears a bit surprising. It is obtained when the first indicator "*no symptoms*" in the questionnaire is associated with the answer no, but also all other symptoms are associated with answers no. Since indicators of several diseases are also associated with answers no, their estimators A and L need not equal zero, and, consequently, the value $L_o(15) \neq 100\%$ is obtained. The value $L_o(15) = 41\%$ thus indicates that the answers given on the questionnaire indicate with likelihood 41% that the patient's state cannot be described by the set of considered diseases. However, if the first question "*no symptoms*" is confirmed by the answer yes, then the corresponding $A_o(15)$ and $L_o(15)$ estimators have the value 100%, while all other values are equal to 0. This case corresponds to a "*healthy state*".

The mean value of the set of optimal values $\{L_o(d); d = 1, \ldots, 15\}$ is even higher than in the case of *tuberculosis*, amounting to $< L_o >= 34, 6\%$. The corresponding standard deviation is $\sigma_L = 7\%$ while the maximal and minimal values are $L_{o,max} = L_o(8) = L_o(14) = 48\%$ and $L_{o,min} = L_o(9) = 24\%$, respectively. These data indicate that the likelihood value of ≈ 0.35 yields a rather firm quantitative argument for accepting the result of the automatic diagnosis.

The results obtained by testing the program's performance with the formal sets of symptom indicators show that the diagnosis performed based on our description of symptoms and properties of diseases by statistical weights is quite reliable in the case of well exhibited symptoms. In accordance with this, the quantitative estimator of the agreement between given symptoms and formal symptoms of various diseases could be considered to be complementary to quantitative data obtained by other clinical tests.

The final conclusion proves even more significant when many symptoms are given which could even be contradictory. This case is demonstrated in the second example, where the distribution of the agreement estimator is

Table 5. Symptom indicators confirmed by Patient X in the questionnaire

s	1	2	3	4	5	11	25	26	27	28	31	38	40
i	2	1	2	1	1	4	1	1	1	1	2	1	2

Table 6. Statistical weights characterizing disease 14 and the responses of patient X.

	(symptom - s, indicator - i, weight - $W(d,s,i)$)
$D(14)$	(2,1,2) (3,1,1) (11,4,3) (25,1,2) (26,1,1) (27,1,1) (28,1,1)
patient X	(2,1,2) (11,4,3) (25,1,2) (26,1,1) (27,1,1) (28,1,1)

Weights at other indices are 0.

determined from data provided by a patient in a rather bad or critical health state who confirms nearly all symptom indicators given in the questionnaire. The graph corresponding to such a patient is shown in Fig. 2. In this example several values of $A(d)$ are similar and may surpass the mean value $< A >$. This property of the distribution $\{A(d); d = 1, \ldots, N_d\}$ indicates that it is rather unreliable to select any of them as the result of diagnosis. For example, although the maximal value $A(12)$ surpasses the mean value $< A >$ by more than σ_A in this case, the low value of its likelihood $L(12) \sim 12\%$, which is considerably less than the mean value $< L_o > = 34,6\%$ and even the minimal value $L_{o,min} = 24\%$ obtained from testing the program's performance, confirms this conclusion.

3.4. Clinical tests

To demonstrate the program's performance on examples from clinical practice we perform diagnoses of diseases $D(14)$ - *throat inflammation* and $D(10)$ - *otitis media* based on data obtained from patients. The diagram of the first case is shown in Fig. 3. The set of symptom indicators confirmed by patient X in the questionnaire is shown in Table 5, while all other indicators were left blank. Patient X suffered from pain in the throat. Statistical weights characterizing $D(10)$ - *throat inflammation* are given in the first line of Table 6, while the second line shows these weights when associated to the indicators confirmed by the patient X. Six indicators confirmed by patient X agree with the indicators of $D(14)$; the most important agreement is between symptom indicators *pain* ($s = 11, i = 4$) and *throat irritation* ($s = 25, i = 1$), associated with the weights $W = 3$ and $W = 2$, respectively. The sum of the weights associated with the patient's answers equals 10, while the total sum of weights characterizing $D(14)$ equals 11. This yields the value of relative agreement $A(14) = 10/11 \sim 91\%$. As shown in Fig. 3, this value surpasses the mean value $< A >$ of the agreement distribution by 2.96 standard deviations σ_A. All other values $\{A(d); d \neq 14\}$ are essentially below the value $A(14)$; therefore, we accept as the result of diagnosis the disease 14: *throat inflammation*. This conclusion is supported by the rather high value of the likelihood estimator $L(14) = 29$.

The next example demonstrates the diagnosis of patient Y suffering from pain in the inner part of the ear. The set of symptom indicators confirmed by this patient is given in Table 7, while the corresponding diagram of agreement A and likelihood L is shown in Fig. 4. The mean value and standard deviation of the agreement estimator are $< A > = 30.0\%$ and $\sigma_A = 23.1$, respectively. As in the previous example, here the maximal agreement properly takes place at $d_o = 10$. This property is the consequence of agreement between several indicators confirmed by the patient and those corresponding to the disease $D(10)$. The first line of Table 7 shows the indicators and weights that characterize $D(10)$, while the second line shows coinciding confirmed indicators. All other indicators are associated with the weight $W = 0$ when considering this disease. Confirmed indicators agree rather well with the formal indicators of the disease, but here the most important agreement is between symptom indicators of *pain* ($s = 11$) having weight $W(10, 11, 2) = 3$ with *sudden onset of illness* ($s = 2$) and *headache* ($s = 6$), associated with the weights $W(10, 2, 1) = 2$ and $W(10, 6, 4) = 2$, respectively. The symptom indicator with index 29 does not agree with the set of formal disease indicators associated with the weight $W(10, 29, 1) = 1$. Consequently, the total
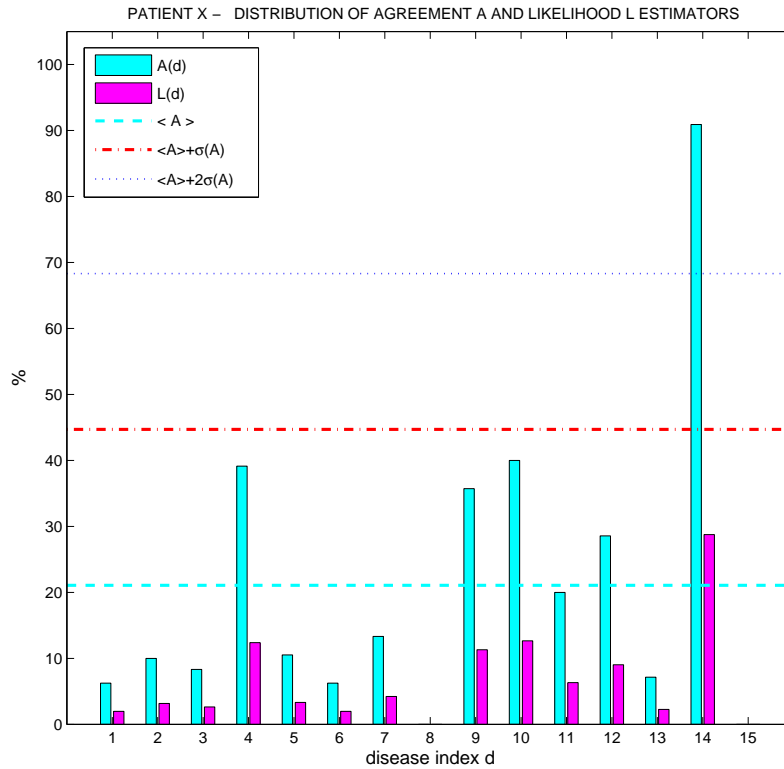
Figure 3. Distribution of estimators A and L based on the questionnaire of a patient suffering from *throat inflammation* disease with index 14.

Table 7. Symptom indicators confirmed by Patient Y in the questionnaire.

s	1	2	3	4	5	6	7	8	9	10
i	2	1	2	1	1	4	2	5	1	3
s	11	12	13	14	15	16	17	18	19	20
i	2	2	2	2	3	2	1	2	2	1
s	21	22	23	24	25	26	27	28	29	30
i	4	2	2	2	2	2	2	2	2	2
s	31	32	33	34	35	36	37	38	39	40
i	3	1	2	2	2	2	2	1	2	2
s	41	42	43	44	45	46				
i	1	1	1	2	4	4				

weight of the patient's answers is equal 9, while the total weight of the formal indicators equals 10. This yields an optimal value of relative agreement $A_o = A(10) = 9/10 = 90\%$. As shown in Fig. 4, this value surpasses the mean value $< A >$ of the agreement distribution by 2.96 standard deviations σ_A. Since all other values $\{A(d); d \neq 10\}$ are greater than one σ_A below the optimal value A_o, we also here conclude that the result of the automatic diagnosis points to the correct disease being 10: *otitis media*.

Table 8. Statistical weights characterizing disease 10 and the responses of patient Y.

	(symptom - s, indicator - i, weight - $W(10, s, i)$)
$D(10)$	(2,1,2) (3,2,1) (4,1,1) (6,4,2) (11,2,3) (29,1,1)
patient Y	(2,1,2) (3,2,1) (4,1,1) (6,4,2) (11,2,3)

Weights at other indices are 0.

Our conclusion is related to the value $L(10) = 20$ of the likelihood estimator, which appears low in comparison to values L_o obtained from program testing. This is rather surprising, since the optimal value A_o significantly surpasses the mean value $< A >$, and most indicators of the patient's symptoms agree with the formal description of the disease. So we have to answer the question: "What is the reason for the rather low value $L(10) = 20$ of the disease likelihood estimator?" With this aim let us look to the diagram in Fig. 4. It reveals that in addition to the maximal value $A(10) = 90\%$ there is another high value $A(15) = 71\%$. This value indicates that the patient's state cannot be characterized by a disease among the considered set of diseases. The corresponding likelihood is $L(15) = 16\%$. Since it is rather high, the value $L(10) = 20$ is small in comparison with the value $L_o(10) = 32\%$ that is obtained when the patient's symptom indicators completely agree with the formal symptoms of disease $D(10)$. These properties can be understood if the formal indicators of the disease *otitis media* and the properties of the patient's answers are considered. Except for 5 of the patient's indicators that coincide with the formal indicators of the disease *otitis media*, the remaining mostly correspond to a *healthy state*. According to these properties, we can conclude that the maximal value $A(10) = 90\%$ could be considered a rather acceptable indicator of the presence of *otitis media*, while the second peak at $d = 15$ indicates that patient Y is, except for the attack of the disease *otitis media*, in a rather good health condition.

4. Discussion and conclusions

The main contribution of this article is the development of a new quantitative method for diagnosing of 14 most frequent diseases of the western society that can be rather simply performed by a personal computer and utilized by professionals as well as a broad population. The presented examples indicate that the selected sets of symptoms and statistical weights of diseases can be regarded as a proper basis for an automatic diagnosis of selected diseases by a personal computer. The advantage of the proposed method is the quantitative estimation of the agreement between patient symptoms and formal symptoms characterizing the properties of the diseases. By using this estimator various subjective errors could be avoided at the assessment of a diagnosis. Moreover, a reliability evaluation of the diagnosis can also be described quantitatively by the estimator of disease likelihood. In addition to these properties, which are important from the professional point of view, the proposed method and the corresponding computer program [4, 5] could be widely applied by users interested in diagnosing their diseases outside the professional environment.

Throughout the development of our method we have followed standard procedure for a doctor when assessing a diagnosis [4, 5]. We are aware that our method corresponds to a rather crude simplification of the professional performance. To improve our method one should take more symptoms as well as diseases into account. However, making such an improvement requires more effective description of the corresponding sets. We expect that for this purpose applying a hierarchic structure could be advantageous. A hierarchic structure could provide increased reliability of the diagnosis assessment.

Operation of our program resembles the operation of a sensory neural network [4, 5, 13]. In such a network a patient's responses excite signals from sensors that further excite neurons such that their outputs represent agreement between input signals and formal indicators of diseases stored in their memories. An acceptable operation is obtained when the responses of neurons to stimulation from sensors are determined by the statistical weights of the diseases.
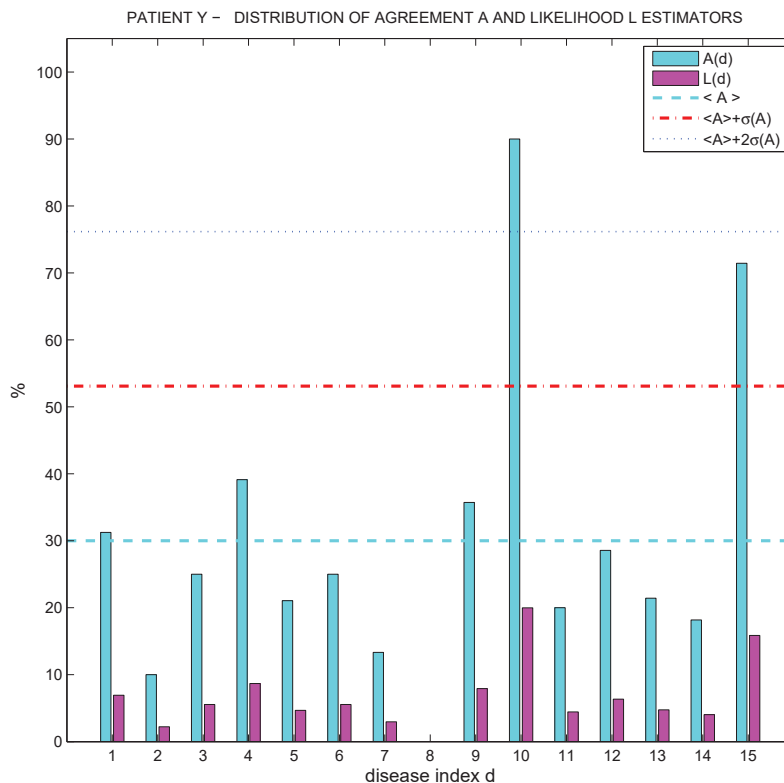
Figure 4. Distribution of estimators A and L based on the questionnaire of a patient suffering from *otitis media* - disease with index 10.

In the development of our method we have utilized symptom weights determined by doctors [4, 5]. However, in the interest of refining the diagnosis the corresponding data could also be automatically created and even improved during the application of the corresponding computer program. Various methods developed for training artificial neural networks could be applied for this purpose [8, 13]. Such an adaptation would in fact allow for the acquisition of new medical knowledge and also for its storage.

REFERENCES

1. I. Kononenko, *Machine learning for medical diagnosis: history, state of the art and perspective*, Artificial Intelligence in Medicine, vol. 23, no. 1, 89-109, 2001.
2. H. W. Chiu, Y. C. J. Li, *Toward precise and preventive health-care with computational tools*, Computer Methods and Programs in Biomedicine 153 (2018) A1.
3. C. Li, L. Hou, B. Y. Sharma, H. Li, C. S. Chen, Y. Li, X. Zhao, H. Huang, Z. Cai, H. Chen, *Developing a new intelligent system for the diagnosis of tuberculous pleural effusion*, Computer Methods and Programs in Biomedicine vol. 153, 211?25, 2018.
4. M. Sok, E. Švegl, and I. Grabec *Statistical diagnosis of diseases*, Proc. Int. Conf. Applied Statistics, Ribno (Bled) Slovenia, Sept. 18-21, 2016, *Abstracts and Program*, p30.
5. M. Sok, E. Švegl, and I. Grabec *Sensory-neural Network for Medical Diagnosis*, Int. Conf. Evolving and Adaptive Intelligent Systems - EAIS 2017, Ljubljana, Slovenia, May 31 ?June 2, 2017, http://msc.fe.uni-lj.si/eais2017, Publisher: https://www.ieee.org
6. *Call for project proposals*, https://en.wikipedia.org/wiki/Tricorder_X_Prize, 2013.
7. S. El-Sappagh, M. Elmogy, A.M. Riad, *A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis*, Artificial Intelligence in Medicine, vol. 65, no. 3, 179-208, 2015.

8. K. B. Wagholikar, V. Sundararajan, and A. W. Deshpande, *Modeling paradigms for medical diagnostic decision support: a survey and future directions,* Journal of Medical Systems, vol. 36, no. 5, 3029-3049, 2012.
9. *Merriam-Webster's Dictionary*, http://learnersdictionary.com/definition/diagnosis.
10. *Wikipedia*, https://en.wikipedia.org/wiki/Medical diagnosis.
11. E. P. Balogh, B. T. Miller, and J. R. Ball, Eds; Committee on Diagnostic Error in Health Care, Board on Health Care Services; Institute of Medicine; The National Academy of Sciences, Engineering and Medicine, *Improving Diagnosis in Health Care*, The National Academic Press, Washington DC, 2015.
12. X. H. Zhou, N. A. Obuchowski, and D. K. McClish, *Statistical Methods in Diagnostic Medicine*, Wiley series in probability and statistics, 2nd Edition, 2011.
13. I. Grabec, and W. Sachse, *Synergetics of Measurement, Prediction and Control*, Springer, Berlin, 1997.
14. A. Sadikov, V. Groznik, J. Žabkar, M. Možina, D. Georgiev, Z. Pirtošek, I. Bratko, *ParkinsonCheck smart phone application*, In: T. Schaub, G. Friedrich, and B. O'Sullivan, Eds., Proc. of ECAI Frontiers in Artificial Intelligence and Applications, p. 263, IOS Press, Amsterdam, 2014.
15. T. Shaikhina, N. A. Khovanova, *Handling limited datasets with neural networks in medical applications: A small-data approach*, Artificial Intelligence in Medicine, vol. 75, 51-63, January, 2017.
16. C.-L. Chi, W. N. Street, D. A. Katz, *A decision support system for cost-effective diagnosis*, Artificial Intelligence in Medicine, vol. 50, no. 3, 149-161, 2010.
17. https://www.google.si/?gws_rd=ssl#q=diagnosis+of+diseases+by+symptoms, http://www.healthline.com/symptom-checker.