Development of diagnosis model based on intersecting fuzzy sets method

E N Korovin\textsuperscript{1}, E A Nazarenko\textsuperscript{1}, E I Novikova\textsuperscript{1}, M A Sergeeva\textsuperscript{1} and E V Kondaurova\textsuperscript{1}

\textsuperscript{1}Voronezh State Technical University, Voronezh, Russia

E-mail: ken170576@mail.ru

Abstract. The article discusses one of the approaches to the analysis and decision-making on the patients with thyroid diseases diagnosis, based on a multi-criteria method - the method of fuzzy alternatives sets. The paper provides a theoretical analysis of the situation in the world in the field of endocrine diseases and the rationale for the method choice. The advantage of using the fuzzy sets intersection method in the field of medicine is shown. A visual calculation of the case with the input data of the patient was carried out and the diagnosis was derived. Calculations were carried out in the same way with a test sample of patients. Based on the results, high predictive abilities were obtained for using the model. Mathematical models were derived, which can later be implemented in information and software that increases the efficiency of the myocardial infarction complications predicting process. The implementation of the developed system will allow medical organizations to increase the efficiency and accuracy of patients' preliminary diagnostics.

1. Introduction

Diseases of the thyroid gland rank second in endocrinology. Hypothyroidism is a widespread disease; its incidence is 1: 4000 people in the world. If the first signals of the disease are missed, it can develop into a more complex form and pathology.

The diagnosis of thyroid diseases is difficult, and they are easily confused with other diseases not related to endocrinology. Therefore, at the initial stages, the doctor's decision-making is carried out under a number of uncertainties conditions, and the patient's characteristics are individual.

Automation of the diagnostic data processing by using computer technology is a means to take into account a large number of incoming diagnostic signs, taking into account their individual coefficient of significance. And at the same time, it allows to exclude errors connected with subjective factors, such as the doctor's lassitude, underestimation of the individual symptoms significance, prejudgement of opinion regarding certain diagnoses, etc.

Diseases of the thyroid gland are one of the most common endocrinopathies in the world. According to the WHO, the number of cases of thyroid gland pathologies detection in the world is increasing by 5% per year. The reason for this is the deteriorating environment, the presence of hereditary factors, unbalanced nutrition. The work will consider three main diseases: thyrotoxicosis (ICD code - E05), hypothyroidism (ICD code - E03), subacute thyroiditis (ICD code - E06).

2. Theoretical analysis

Modern medical organizations work with a huge amount of information. Medical input data can be represented by various types of variables, therefore, work with some of them cannot be carried out within the framework of classical mathematical operations.

Diseases of the thyroid gland are very similar to neurological disorders and sometimes the doctor cannot clearly assess the symptoms manifested in the patient, therefore, the inaccuracy of signs, as well as diagnoses, can manifest itself in the fuzziness of expert assessments.

Currently, the theory of fuzzy logic and fuzzy sets is increasingly used in medicine.

As can be seen from the above mentioned, the topic related to the development of models for decision-making based on multicriteria methods is relevant today, therefore this work is devoted to the consideration of such methods in a wide area of endocrine diseases.
In this work, the method of fuzzy intersection of fuzzy sets is used, which is the basis for creating a convenient tool for practical use by endocrinologists.

The inherent advantages of the fuzzy sets intersection method, such as the inclusion of qualitative variables in the analysis, operating with fuzzy input data, linguistic variables, rapid modeling of complex dynamical systems and comparing them with a given degree of accuracy, make it possible to justify its choice as an effective and promising mathematical method for making decisions by a doctor - establishing a diagnosis.

3. Methods

From the compiled matrix of the diseases symptomatology indicated above, based on 80 clinical charts, with the help of 7 experts (who were the doctors of the medical and preventive treatment institutions), the most significant signs of diseases were identified.

The evaluation of signs was carried out by using a 6-point scale, where the score "6" corresponded to the maximum importance of the symptom, and the score "1" is the minimum one. The symptoms with an average expert rating of at least "3" were selected to be used as criteria for choosing the best alternative (diagnosis). A list of these criteria and their indicators/characteristics for a specific disease is shown in Table 1.

As can be seen from Table 1, all the data have different dimensions, therefore, it is advisable to use the foundations of fuzzy systems. Thus, reducing the problem of establishing a diagnosis to the concepts of fuzzy logic, we have the following considerations:

- a) many alternatives \( A \{ a_1, a_2, a_3, a_4 \} \), where \( a_1 \) is hypothyroidism, \( a_2 \) is subacute thyroiditis, \( a_3 \) is thyrotoxicosis, \( a_4 \) is another pathology;

- b) a set of criteria \( F \{ F_1, F_2, \ldots, F_{14} \} \), where the value of the criterion corresponds to the names of the criteria from Table 1.

The expediency of using fuzzy logic applies to this problem, which is characterized by a large number of imprecise or undefined values. The use of fuzzy logic methods in modeling the process of establishing a diagnosis of thyroid disease will make it possible to assess the cumulative effect of each parameter on the final result of calculations.

### Table 1. Comparative characteristics of indicators/characteristics of signs for the thyroid gland diseases

| №  | Diseases sign (criterion) | Indicators/characteristics for the disease: |
|----|--------------------------|-------------------------------------------|
|    |                          | hypothyroidism | subacute thyroiditis | thyrotoxicosis |
| 1  | TSH level, \( \muIU/ml \) | \( >4.0 \)     | 1.6 – 4.0          | \( <1.6 \)     |
| 2  | T3 level, nmol/l         | \( <1.17 \)    | >2.18, in some cases – norm [1.17 – 2.18] | >2.18          |
| 3  | T4 level, nmol/l         | \( <62 \)     | >141, in some cases – norm [62 – 141]   | >141           |
| 4  | ESR level, mm/h          | Reduced or increased | Increased           | Increased       |
| 5  | Thyroid volume, ml       | Reduced or increased | Increased           | Increased       |
| 6  | Appetite                 | Reduced or norm | Reduced            | Increased       |
| 7  | Body temperature, °C     | Reduced        | Increased           | Norm            |
| 8  | Body weight, kg          | Increased      | Reduced            | Reduced         |
| 9  | Thyroid pain             | Mild           | Strong              | Moderate        |
| 10 | Neck pressure            | Absent         | High                | Slight          |
| 11 | Urinary frequency        | Seldom         | Norm                | Frequent         |
| 12 | Hand tremor              | No             | No                  | Yes             |
| 13 | Swelling                 | Yes            | No                  | No              |
| 14 | Skin health              | Dry            | Watery              | Watery          |

The degree of membership in the set of alternatives, which are generalized characteristic functions, is termed the membership function.

For the considered problem of establishing a diagnosis, graphs of the dependences \( \mu_A(x) \) on a certain symptom of the disease were constructed. The values of the membership function can take a value from the range \([0; 1]\), while 0 means no belonging to the considered set of diseases (hypothyroidism, subacute thyroiditis, thyrotoxicosis, other pathology), and 1 means complete belonging to this disease.
Figures 1, 2 show examples of such a construction results. Figure 1 shows the graphs of membership functions for the $F_1$ criterion "TSH level", where the values of the TSH level measured in μIU/ml are plotted along the abscissa. Figure 2 shows the $μA(x)$ plots for the $F_2$ criterion "T3 level", where the abscissa shows the values of the T3 level measured in nmol/l. In addition, graphs were built for the rest of the signs.

![Figure 1](image1.png)

**Figure 1.** Values of membership functions for the diseases under consideration, depending on the TSH level

![Figure 2](image2.png)

**Figure 2.** The value of membership functions for the diseases under consideration, depending on the T3 level

4. Experimental part

To test the reliability and suitability of the constructed model, 15 people with suspected diagnosed diseases were tested. The model test on one of 15 examples is presented below.

Patient K., 34 years old, complains on: a) palpitations; b) notes weight loss up to 5 kg; c) constant trembling of hands; d) severe swelling of the legs; e) decreased appetite; f) frequent urination; g) sometimes there is pain and pressure in the neck.

Considers herself to be ill for 5 years.

The examination revealed: pulse 120 beats/min, blood pressure 142/90 mm Hg, which corresponds to hypertension, exophthalmos, the thyroid gland is enlarged evenly with a smooth surface, soft-elastic consistency, slightly painful, displaced when swallowing, body temperature 36.8 °C, the skin is watery at times (according to the patient’s statement).

After that, the endocrinologist sent her for an ultrasound of the thyroid gland and prescribed the analysis for the thyroid hormones study, the general blood test.

The result of ultrasound was: a diffuse enlargement of the gland of the 3rd degree (by 11 ml) with an increased blood supply.

The study of thyroid hormones showed: the increase in the level of T3 hormones (3.58 nmol/l), T4 (185 nmol/l), the decrease in the content of thyroid stimulating hormone (1.1 μIU/ml). From the general blood test, the ESR level was determined - 5.5 mm/h.

In accordance with all the examinations results, the patient was diagnosed with thyrotoxicosis.

In accordance with this clinical chart to the constructed membership functions, we defined the following fuzzy sets, represented by equations 1 - 14, which will be used in further calculations:

$$
μ_{F_1} = \left\{ \begin{array}{c} 0 \\ \frac{0.24}{a_1} \\ \frac{1}{a_2} \\ \frac{0.24}{a_3} \\ \frac{1}{a_4} \end{array} \right\}, \quad (1) \\
μ_{F_4} = \left\{ \begin{array}{c} 0 \\ \frac{0.9}{a_1} \\ \frac{0.9}{a_2} \\ \frac{0.6}{a_3} \\ \frac{1}{a_4} \end{array} \right\}, \quad (8)
$$
Let us make a reservation that all quality criteria are of different importance for decision makers. And, proceeding from this, we will make calculations.
The processing of the initial information is carried out in three stages:
b) the membership functions specific values determination according to criteria;
c) a convolution of the available information in order to identify the best alternative.

The first two stages were completed by us in the first engraving of this course work. Formulas 1-14 are the results.

Further, to determine the best alternative, a matrix of criteria evaluations was built. It is based on the information analysis on the importance of symptoms for these diseases, obtained from seven experts. It uses the standard comparison scale shown in Table 2. The pairwise comparison matrix is filled out based on these estimates.

After that, Table 3 presents the formulas and facts of calculations of subsequent actions leading to the result for the eigenvector of the matrix ($\beta$) search.

At the next stage, we determined the set of optimal alternatives, taking into account the calculated coefficient $\beta$, by using the following equation:

$$B = F_1^{\beta_1} \cap F_2^{\beta_2} \cap \ldots \cap F_{14}^{\beta_{14}}$$

As the calculations result, the following values of $B_i$ were obtained for each alternative. From the obtained values, the minimum values were selected, which is displayed in Table 4.

![Table 4](image)

After that, from the obtained minimum values, the maximum value equal to 0.4209 was selected, which is the best alternative and corresponds to the diagnosis of thyrotoxicosis.

5. Results and their discussion

Thus, let us draw conclusions: at the initial stage, the main symptoms of thyroid diseases were selected based on expert estimates. Then the graphs of membership functions for the selected criteria were built, their values were determined. The patient's case was selected and the available information was convoluted in order to identify the best alternative.

In the case of the output data choice, the most probable disease is thyrotoxicosis, which corresponds to the results of the method calculation.

The obtained results of the alternatives multicriteria choice were program implemented in the application by using the Borland Delphi visual editor. The developed application offers the best alternative when deciding the problem of establishing a diagnosis by means of the established algorithms and procedures, that is, a probable diagnosis of three possible ones: hypothyroidism, subacute thyroiditis, thyrotoxicosis. This computer program allows to make a diagnosis to the applied patient with a reliability of 94.12%.

Summarizing the results, we can say that such a methodological approach to reducing the dimension of the feature space facilitates and simplifies the solution of poorly structured problems of multicriteria classification and choice when establishing a diagnosis for a patient with thyroid disease.
References

[1] Zadeh L 1976 The concept of a linguistic variable and its application to making approximate decisions Moscow: Mir

[2] Analysis, synthesis, decision planning in economics Moscow: Finance and Statistics 2000 368 p

[3] Topolyansky A G and Borodulin V N 2007 Big medical encyclopedic dictionary M.: Ripoll classic 960 p

[4] Novikova E I and Rodionov O V 2012 Algorithmization and management of the process of diagnosing gynecological diseases based on multivariate modeling monograph. Voronezh: VSTU 132 p

[5] Novikova E I, Rodionov O V and Korovin E N 2008 Modeling of biomedical systems textbook, Voronezh: VSTU 196 p

[6] Novikova E I 2006 Analysis, algorithmization and management of the process of diagnosis of gynecological diseases based on multivariate modeling Novikova E. I. abstract of the dissertation for the degree of Candidate of Technical Sciences. Voronezh State Technical University. Voronezh

[7] Novikova E I 2006 Analysis, algorithmization and management of the process of diagnosis of gynecological diseases based on multivariate modeling dissertation for the degree of Candidate of Technical Sciences. Voronezh

[8] Novikova E I and Kondaurova E V 2020 Development of a model for diagnosis based on the method of ranking alternatives on a set of linguistic vector estimates System analysis and management in biomedical systems 19(4) 127-131

[9] Novikova E I and Korshunova E A 2016 Development of models to support decision-making in the diagnosis of lung diseases System analysis and management in biomedical systems 15(3) 466-469

[10] Novikova E I, Rodionov O V and Faustova A Yu 2007 Algorithmization and management of the process of diagnosing gynecological diseases based on multivariate modeling System analysis and management in biomedical systems 6(2) 434-438

[11] Novikova E I and Korshunova E A 2016 Development of models to support decision-making in the diagnosis of lung diseases System analysis and management in biomedical systems 15(3) 466-469