ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Method And Algorithm For Calculating The Probabilistic Evaluation Of Stroke Recurrence.

Sudakov O, Gladskikh N, Alexeev N*, Bogacheva E, Kuzmenko N, Fursova E, Kudashova E, and Sudakov D.

Voronezh State Medical University named after N.N. Burdenko

ABSTRACT

This article describes the development of a mathematical model for predicting the risk of developing a repeated cerebral stroke based on the minimum available diagnostic information that is sufficient to achieve the necessary reliability of the diagnosis. The method of Bayes and Kulbak was chosen as a mathematical apparatus. The technique of the choice of informative features for the solution of the task is described. The final algorithm of the model for predicting the risk of recurrent stroke is given.
Keywords: repeated stroke of the brain, mathematical model, Bayesian method

Corresponding author

INTRODUCTION

Any doctor in his work needs "probabilistic thinking" and, in particular, an understanding of the probabilistic approach to diagnosis. Apparently, this approach is largely based on the subconscious process that underlies the diagnosis by an experienced physician taking into account pathological symptoms, frequent symptoms, symptoms not characteristic for this disease or not occurring in it ever.

The application of probabilistic methods does not mean that a doctor must always apply a mathematical apparatus at the patient's bedside. All these calculations need to be done once in determining the treatment plan.

The terms "symptom" are often used as synonyms. However, it is more convenient to understand a certain value (range) of the symptom, for example, the pulse rate is increased to more than 120 beats per minute, the appearance of the "vomiting" sign or the absence of this sign, etc. Therefore, in the subsequent discussion, instead of the words "tag value" or "Range of trait" will often use the term "symptom".

The frequency is the ratio of the number of patients with each disease or having a certain symptom to the total number of patients in the group. This value is sometimes called selective probability, since it is calculated on the basis of an analysis of a certain sample of patients, and not of their entire population. For brevity, the sampling probability is often called probability (although this is inaccurate).

Indeed, probability is judged by the frequency of the event actually recorded in the experiment or clinical observation, that is, by its selective probability. However, the less experiments are done, the more random the detected frequency is and the more strongly it differs from the true probability. The more observations or experiments done, the more frequent the magnitude approaches the probability of a disease or symptom.

The probability is denoted by the letter P and is more often calculated in fractions of a unit, rather than in percentages.

The a priori probability p (A1) is the probability of a second stroke, that is, the probability for any patient, no matter what symptoms it has, have a second stroke.

The a priori probability of the symptom p (Xij) is the probability of the symptom Xij, in the whole group, that is, the probability for any patient in the troupe, regardless of which illness he suffers, is the symptom Xij. This value is the ratio of the number of patients having the symptom Xij to the total number of patients in the group.

The conditional probability of symptom Xij, with the possibility of repeated stroke (hypothesis A1) p (Xij | A1) is the probability of having the symptom Xij, provided that it belongs to the main group. This value is equal to the ratio of the number of patients with a repeated stroke, having the symptom Xij to the total number of patients suffering from this disease.

By analogy, the conditional probability of symptom Xij, with the possibility of repeated stroke (hypothesis A2) p (Xij | A2) is the probability of having a symptom Xij, provided that it belongs to a control group. This value is equal to the ratio of the number of patients without a recurrent stroke having a symptom Xij to the total number of patients suffering from this disease.

In this study, it is advisable to use Bayesian formula to determine probabilistic estimates of stroke recurrence.

The Bayes formula is sometimes called the inverse probability theorem or the hypothesis theorem. This is quite applicable to diagnostic problems: the formula allows one to choose one of several possible diagnostic hypotheses, based on the calculation of disease probabilities according to the probability of the symptoms found in the patient. Using this formula, based on the following data:

1. P (Ak), a priori probability of each disease in the group;
2. P (xilAk), the conditional probability (frequency) of all the symptoms for each of the diseases Akconsistently calculate the probabilities of each diagnostic hypothesis and choose the most probable one.

In the simplest form, the Bayesian formula has the following form:

$$P(A_1 / x_1) = \frac{P(A_1)P(x_1 / A_1)}{P(x_1)} \quad (1)$$

This approach has a number of indications and limitations (like any approach to diagnosis). Here it is necessary to dwell only on one circumstance. In this approach, the a priori probability is automatically taken into account, affecting the number of patients with each disease in the cells of the table. At the same time, unlike the conditional probability of symptoms, the a priori probability of the disease in the group, as already mentioned, is a very variable one. It depends on the choice of group. As a specific group, patients in a given geographical area can be considered; patients, served by some medical institution or even a certain doctor. The composition of the group and the frequency of individual diseases can vary dramatically with time. For example, the frequency of infectious diseases depends on the epidemiological situation, many diseases occur with different frequencies at different times of the year, depend on a number of changing social factors, etc. Therefore, in many cases it turns out beneficial to separately consider in the diagnostic algorithm the probability of the symptoms that, with each disease characterized by a certain resistance and can be once studied, and then used in many studies, and the a priori probabilities of the diseases themselves. The combination of these two values is made only at the time of the diagnosis itself using the Bayes formula. However, in order for these two quantities to be sufficient in the calculation, it is necessary in formula (1) to replace the a priori probability of the symptom P (xlj) in the denominator by the equivalent so-called full probability formula for this symptom:

$$P(x_1) = \sum_k P(A_k)P(x_1 / A_k) \quad (2)$$

After such a replacement, the Bayesian formula takes the following form:

$$P(A_1 / x_1) = \frac{P(A_1)P(x_1 / A_1)}{\sum_л P(A_k)P(x_1 / A_k)} \quad (3)$$

In this form it is called the full Bayes formula or the hypothesis theorem.

However, the patient can be found simultaneously symptoms. How to calculate the probability of a recurrence of a stroke in this case?

Whether we have data on the number of patients who have a complex of symptoms in diseases A1 and A2, the calculation of the probabilities of recurrent stroke in the presence of these symptoms can be calculated on the basis of the modified Bayes formula.

The use of the modified Bayes formula (4) makes it possible to calculate the probability from a set of informative features, using the frequency of manifestation of a particular trait and the a priori probabilities of occurrence of a second stroke.

$$P(A_1 / x_1 x_2 ... x_n) = \frac{P(A_1)P(x_1 / A_1)P(x_2 / A_1)...P(x_n / A_n)}{\sum_k P(A_k)P(x_1 / A_k)P(x_2 / A_k)...P(x_n / A_k)} \quad (4)$$

In this study, p (A1) is the a priori probability of recurrent stroke occurrence, p (A2) is the a priori probability of no recurrent stroke, p (Xij | Ak) is the conditional probability (frequency) of the appearance of the characteristic when belonging to the main or control group.

Essentially, the diagnostic task is to establish a diagnosis using the minimum available diagnostic information that is sufficient to achieve the required reliability of the diagnosis. This usually requires the use of not one symptom, but a set of symptoms (symptom complex). Such an approach can be called a "multidimensional" approach to diagnosing diagnosis, since many features are used simultaneously with it.

It should also be noted that the question of indications for the use of various variants of the Bayes formula requires further research. At the present time, the formulas presented in this study and various algorithms based on the assumption of independence of features are used in a much more widespread way, since a more sophisticated approach to diagnosis involving the use of formulas requires a special study of probabilities different combinations of symptoms.

The procedure for calculating the probability of repeated stroke consists of the following stages:

1. Formation of feature space, which allows to fully identify the state of the modeling object.
2. Formation of a dictionary of informative features on the basis of the Kulbuk criterion.
3. Calculation of the probability of repeated stroke based on the modified Bayes formula (4).
4. Formulation of recommendations for further treatment and prevention.

The algorithm of the proposed procedure is shown in Fig. 1.

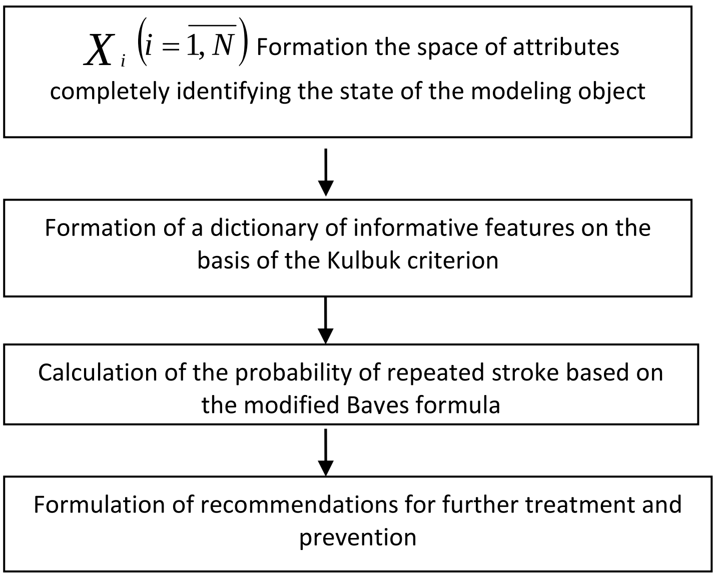

Fig 1: Algorithm of the methodology for determining probabilistic estimates of stroke recurrence.

In conclusion, it should be noted that at the present time a large number of algorithms for computational diagnostics based on the probabilistic approach to some extent or the like of Bayes formula. Most of these algorithms are designed for the use of computer technology, which allows for much more complex calculations.

REFERENCES

[1] Esaulenko I.E.Genetic and Clinic-Pathogenetic Peculiarities of Prediction of Development and the Effects of Obesity at Young Persons /I.E. Esaulenko, T.N.Petrova, E.N.Kolesnikova, O.V.Sudakov //Research Journal of Pharmaceutical, Biological and Chemical Sciences. 2017; № 8(3): p. 1368-1374.

[2] Shiryaev O.Y. Psychosomatic Aspects of Chronic Heart Failure / O.Y. Shiryaev , V.L Yankovskaya., A.V. Budnevsky// International journal of biomedicine. 2017, № 7(3): p. 248-250.

[3] Tokmachev R.E.The Possibility of Nob-Pharmacological Methods in Increasing Clinical Efficiency of

Treating Patients with Chronic Heart Failure and Metabolic Syndrome / R.E. Tokmachev, A.V.Budnevsky, Kravchenco A.Y. // Research journal of pharmaceutical biological and chemical sciences. 2017, № 8(6): p. 932-839.
[4] Shtankov S.I., Shiryaev O.Y., Sudakov O.V. and others Software Implementation the Methodology for Calculating Integral Indicators Rehabilitation Potential of Patients with Schizofrenia / Research journal of pharmaceutical biological and chemical sciences. 2018, №9(1): p. 950-953.
[5] Mathematical, algorithmic and software support of the information support complex for making medical decisions in patients with diabetes mellitus and arterial hypertension / O.V. Sudakov, N.A. Gladskikh, E.V. Bogacheva, N.Yu. Alekseev, O.A. Androsova // System analysis and management in biomedical systems. 2015. P. 14. № 4. P. 815-819.
[6] Application of neural network modeling to support decision-making in the diagnosis of chronic heart failure / E.A. Fursova, E.I. Novikova, O.V. Sudakov // System analysis and management in biomedical systems.2009.T.8.№ 2.C.410-413.
[7] Modeling of severe forms of acute toxic hepatitis / N.Yu. Alekseev, N.Yu. Kuzmenko, O.V. Sudakov // System analysis and management in biomedical systems.2012.T.11.№ 2.C.481-484.
[8] Construction of the model of the adaptation potential of patients with type 2 diabetes and arterial hypertension / O.V. Sudakov, T.P. Kuchkovskaya, A.V. Sviridov, G.M. Panyushkina, E.A. Studenikina // System analysis and management in biomedical systems. 2014. T..13.№ 2. 2. P. 447-452.
[9] Construction of an adaptive algorithm for the treatment of patients with diabetes mellitus and arterial hypertension / O.V. Rodionov, E.N. Korovin, O.V. Sudakov, E.A. Fursova, D.V. Sudakov // System analysis and management in biomedical systems. No. 3.S. 688-690.
[10] Assessment of the risk of recurrent stroke in patients with type 2 diabetes and hypertension / N.A. Gladsky, O.V. Sudakov, E.V. Bogacheva, N.Yu. Alekseev, E.A. Fursova // System analysis and management in biomedical systems. T.15.№ 1.C.C. 123-127.
[11] Construction of a mathematical model for choosing the type of coronary angioplasty in patients with coronary heart disease and diabetes mellitus / A.I. Borodulin, A.V. Sviridov, O.V. Sudakov, E.A. Fursova, N.Yu. Alekseev // Applied information aspects of medicine. 177. № 2.C.56-58.
[12] Esaulenko I.E. Information system for the selection of diagnostic signs of autonomic syndrome. I.E. Esaulenko, N.A. Gladskikh, S.I. Shtankov, M.A. Zheleznyakov // System analysis and management in biomedical systems. 2011. T. 10. № 4. Pp. 864 -869.
[13] Mathematical and theoretical models of diagnostics of vegetative syndromes. ON. Gladskikh, S.N. Shipilov, E.V. Bogacheva, S.I. Shtankov. // Applied information aspects of medicine. 2011. T. 11. № 2. P. 44-45.
[14] Calculation of diagnostic assessments of the effectiveness of the use of Mexidol in the provision of medical care to victims with TBMT in prehospital and early hospital stages / I.E. Esaulenko, V.L. Radushkevich, N.A. Gladskikh, E.V. Bogacheva // System analysis and management in biomedical systems. 2013. T. 12. № 4. P. 896-901.
[15] Gladskikh N.A. Development of methods of classification-prognostic modeling in the system of staffing of territorial health care / dissertation for the degree of candidate of technical sciences / VSTU. Voronezh. 2008, 211 pp.