STATISTICAL METHODS AND SOFTWARE IN MEDICINE

M.S. Abramovich, e-mail: Abramovichms@bsu.by

Research Institute for Applied Problems of Mathematics and Informatics
Belarusian State University, Minsk, Belarus

Computer data analysis in medical diagnostics

Coronary ischemic disease is characterized by significant changes of the blood flow turbulence and, as a result, by the appearance of high-frequency sounds that are caused by the stenosis of arteries. The acoustic blood flow signal consists of 3 fragments: systole, flap of the mitral valve damper and diastole. The most informative part of this signal, which is used for diagnostics of the coronary ischemic disease, is the diastole. We consider approaches to the diagnostics of the coronary ischemic disease based on the parametric discriminant analysis in the different informative feature spaces: statistics calculated from the wavelet coefficients, covariance functions and parameters of the Markov chains.

Methods of robust discriminant analysis and robust logistic regression were used at the Belarusian Research Institute of Oncology and Medical Radiology for diagnostics of malignant neoplasms in the major sites (lung cancer, gastric cancer, colon cancer) on the basis of biochemical blood assay tests. To construct quadratic discriminant functions, complexes (sets) of informative tests were formed by using several statistical approaches. Since the biochemical blood assay tests have outliers at the terminal stage of malignant neoplasm the robust decision rules, stable for outliers, were used for patients classification. Use of robust discriminant analysis for diagnostics of malignant neoplasms allowed to increase accuracy of diagnostics by 4-5% as against to use of the classical decision rules.

The problem of metastatic appearance risk for melanoma disease classification has been successfully solved with the developed statistical method. The method for each patient gives one of the three decisions: 1) the patient belongs to the class of successfully cured; 2) the patient needs further treatment to prevent high risk of metastatic appearance; 3) the information on the patient is not sufficient, and some more analysis is required. This method is realized in statistical software and showed increase of effectiveness by 20 %.

Spatio-temporal cluster analysis of disease

The methods of spatial and spatio-temporal cluster analysis are widely used in medicine for studying of geographic dissemination of different types of diseases.

Two common types of spatial and spatio-temporal clustering methods are constructed: global clustering methods allow to evaluate the presence of clustering throughout the territory; local clustering methods are used to find out the location and the size of possible clusters. The method based on a statistic which allows to detect clusters with different flexible shapes is considered. In the case of existing outliers among observations so-called robust statistical methods of the spatial cluster analysis were considered. A cluster detection research was held for thyroid carcinoma diagnostical data among children up to 14 years old from 1989 to 2009 in the Republic of Belarus. Spatio-temporal cluster analysis has confirmed that there was a significant increase in the number of thyroid carcinoma cases among children throughout the territory of Gomel region in the 1990s.

The models for the dynamics of clusters in the spatio-temporal cluster analysis of incidence levels were constructed. The methods for the cluster dynamics monitoring based on sequential statistical tests were developed and realized in the algorithmic form. These results are used for the spatio-temporal cluster analysis of malignancies in children and adolescents in Belarus during the post-Chernobyl period. The methodology can be extended to different applications to solve the problems of the quickest detection of spatial changes in clusters or the changes of the characteristics within clusters.