The development of the forecasting information system of HIV and tuberculosis coinfection’s outcome in inmates

S B Ponomarev¹, E V Dyuzheva², M E Vostroknutov¹, V E Polishchuk¹ and D S Ponomarev¹

¹ Federal State Institution of the Federal Penitentiary Service, 125130, Moscow, Russia
² Research Scientist of Research Center No.1 of the Research Institute of the Federal Penitentiary Service, 125130, Moscow, Russia

E-mail: ponomarev.dmitry1990@mail.ru

Abstract. The research describes the development of the forecasting information system that can predict the outcome of a disease in inmates with HIV-associated tuberculosis. The aim of the research is to develop an additional high-precision diagnostic criterion giving the possibility of the timely correction of diagnostic, treatment and organizational measures for patients with HIV and tuberculosis co-infection that helps to improve the quality of medical care. The research material is based on the data from clinical cases of patients with HIV and tuberculosis co-infection who were hospitalized in a tuberculosis hospital providing specialized medical care to inmates from 2012 to 2018. (367 people). The study methodology was developed in several stages with the use of the methods of system analysis and mathematical modeling (logical-statistical method of optimally reliable partitions, the method of analysis of hierarchies, artificial neural network, methods of statistical grouping). The result of a complex multi-stage research makes possible the development of the prognostic index forecasting outcome of HIV-associated tuberculosis in prisoners. For the automated calculation of the developed index a software package was created. Interpretation of the received data allows timely correcting of diagnostic and treatment tactics in order to improve the quality of medical care and reduce the hospital mortality rate. The developed information system does not require complex and expensive diagnostic measures, it is easy to use and can be suggested to use as a screening method.

TB control measures taking place in the Russian penitentiary system are of great importance in stabilizing the epidemic situation in the country. Because of the fact that inmates with active tuberculosis (TB) released from prison have a direct impact on the epidemic situation in the civilian public health sector [1, 2]. Despite the stabilization of the tuberculosis’ epidemiological situation in general, in the penitentiary public health sector the epidemiological indicators of HIV-associated tuberculosis (TB/HIV co-infection) are increasing due to an increase in the number of HIV-infected inmates. The said increase in the proportion of cases TB/HIV co-infection in the structure of morbidity makes possible to predict an increase in the tuberculosis epidemiological indicators in the future. It also gives the great significance to the quality of care for patients with HIV-associated TB for the epidemiological welfare of the whole country.
Effective infectious diseases’ forecasting models will provide the necessary information for the timely correction of diagnostic and treatment measures, which further will help to reduce infectious diseases prevalence and hospital mortality in general. It should be noted that the methods of system analysis and mathematical modeling have already been used in predicting the outcome of the disease in patients with TB/HIV co-infection among inmates [3, 4]. For example, developments based on an artificial neural network have shown adequate and reliable results [5]. However, in this research for the first time there are applied the complex of machine learning methods and a hierarchy analysis method.

1. Methods
The forecasting method is based on the complex multi-stage research made with the use of modern methods of mathematical modeling (hierarchy analysis method, artificial neural network method, statistical grouping methods) and machine learning. The named logical-statistical method of optimally reliable partitions developed in the RAS Computing Center named after A.A. Dorodnitsyna and in the RAS Institute of Biochemical Physics named after N.M. Emanuel helps to determine optimal boundaries of the studied trait’s value decomposition in the compared groups with high statistical certainty [6, 7, 9].

The research material was the data of the anamnesis and the results of clinical and laboratory, radiological and bacteriological medical research methods of clinical cases of patients with HIV and tuberculosis co-infection hospitalized in a tuberculosis hospital providing specialized medical care to inmates from 2012 to 2018. (367 people).

A sufficient amount of the studied data provides the reliability of the research results as well as the used methods for statistical data analysis: correlation analysis, calculation of arithmetic mean deviations and standard deviations, calculation of standard errors of the mean (p), Student test with 95% confidence interval, Chi-square test ($\chi^2$), Fisher test as well as performing permutation tests.

2. Model
To weight the established predictors of an adverse outcome for TB/HIV co-infection it was used the analytic hierarchy process (AHP) [8] with laid dawn mathematical foundations and psychological aspects. In the method application’s process the non-formalizable problem is divided into simple components by constructing a scheme in the form of an inverted tree. The problem under study is at the top of it. Below there are the intermediate factors placed from top to bottom in descending order of importance.

**Figure 1.** The example of a scheme in the form of an inverted decision tree for dividing a conditional problem into simple components.
Every feature (element) of a group in a hierarchical line was compared in pair with the use of a specially developed scale of relations (table 1).

Table 1. Relative importance scale when comparing two objects.

| Intensity of relative importance | Definition                                           | Interpretation                                           |
|----------------------------------|------------------------------------------------------|----------------------------------------------------------|
| 0                                | Unable to compare                                     | The expert finds it difficult to compare                  |
| 1                                | Equal value                                           | Equal contribution of two attributes                     |
| 3                                | The slight superiority of one over the other          | Expert gives a slight superiority of one feature over another |
| 5                                | Significant superiority of one over the other         | Expert gives a strong superiority of one attribute over another |
| 7                                | Strong superiority of one over the other              | The expert gives superiority to one attribute that suppresses the significance of another |
| 9                                | Absolute superiority                                 | The obvious superiority of one attribute over another    |
| 2,4,6,8                          | Intermediate decisions between two adjacent estimates of relative importance | Applied in case of expert’s compromise solutions           |

The studied spectrum of possible 69 parameters divided into 9 groups was used to identify predictive mortality factors. 7 leading specialists in prison healthcare were the experts in the research. Final comparative assessments of factors were established in an agreed discussion. The sign opposite to the one with the given priority received the opposite value (1/X), where X is the sign evaluation.

Characteristic feature comparison results are introduced into the matrix of pairwise comparisons. The numerical dimension of the comparison matrix is determined by the number of subsidiary factors of the hierarchical element. Expert assessments’ mathematical calculations are to determine the relative importance of the factors involved in the hierarchy expressed numerically in the form of a priority vector. In the hierarchy structure the normalized matrix priority vector (NMP) or the factor’s weight coefficient is the total grade of weighted score indicators for each of the criteria.

The calculation of weighting factor coefficient in the hierarchy structure (q) is carried out in several stages. At the first stage the eigenvectors of the matrices are calculated by determining the geometric mean of expert assessments (x_i) located in the rows of the matrix (ω_i) (multiplying n elements of the matrix row and extracting a root of n degrees equal to the number of multiplied elements) (1):

\[ w_i = \sqrt[n]{x_{i1}x_{i2} \cdots x_{in}} \]  

(1)

Further, the normalizing factor (R) that presents the sum of the geometric mean of all rows of the matrix of pairwise comparisons is determined for the calculated values (2):

\[ R = w_1 + w_2 + w_3 + \cdots + w_n \]  

(2)

NMP is calculated by dividing each eigenvector of the matrix by a normalizing factor (q), where i can take integer numerical values (3):

\[ q_i = \frac{w_i}{R} \]  

(3)

NMP in a matrix of pairwise comparisons represents the weight coefficient corresponding to the factor.
At the next stage of the research the final coefficient of the each factor significance was calculated as the product of the weighting coefficient of significance of an individual predictive mortality factor and the weighting coefficient of significance of the corresponding factor in the matrix of paired comparisons located on one level higher in the hierarchy. Subsequently, the final index of fatal outcome risk was formed by summing up the identified factors of fatal outcome at the time of hospitalization.

After that statistical methods were used to group data with the construction of intervals according to the Stojers formula for to evaluate the developed index of the mortality risk for a hospitalized patient with TB/HIV co-infection from among inmates (4):

\[ k = 1 + 3.322 \cdot \log(N) \]  

where \( k \) is the number of values’ intervals of the variable, \( N \) is the number of cases of observation in the study group. During the research the fractional value of \( k \) was rounded to the integer value of the intervals. The range of one interval (\( H \)) was determined by the formula (5):

\[ H = \frac{Z_{\text{max}} - Z_{\text{min}}}{k} \]  

where \( Z_{\text{max}} \) and \( Z_{\text{min}} \) are the maximum and minimum values of the mortality risk index in points, \( k \) is the obtained number of intervals.

The ratio of favorable and unfavorable cases of observation in each interval was calculated after dividing the maximum value of the prognostic index into the optimal number of intervals.

The obtained values of the ratios formed curves that graphically reflect the degrees of the predicted frequencies of favorable and lethal outcomes. X-axis in charting reflects the value of prognostic index in points. Y-axis represents the proportion of lethal or favorable outcomes in values from 0 to 1. One curve reflects the proportion of adverse outcomes in each interval (\( \mu \)). Accordingly, the second curve represents the proportion of favorable outcomes in the obtained intervals of the prognostic index values and takes the value 1 - \( \mu \).

3. Results

As a result of the research as it was shown in details above there were of TB/HIV co-infection was selected as the input signal of the model. determined 23 predictive factors of an adverse outcome in hospitalized patients with TB/HIV co-infection from among inmates [9]. The next step of the research it was used mathematical modeling methods for to set the weighting factor for each predictor and it was developed a prognostic index. The outcome of TB/HIV co-infection was chosen as the input signal of the model.

Using methods of statistical data grouping it was developed a scale for to evaluate the obtained values of the index (Y) which is the outgoing signal of the forecasting information system of the TB/HIV co-infection outcome (figure 2).

The rating scale for the developed prognostic index is divided into the following intervals. Group I - “0-11” points (prognosis bona) - probability of adverse outcome 0% - no risk of adverse outcome. Group II - "11,1-35” points (prognosis dubia) - probability of an adverse outcome 0-50% - low risk of a lethal (adverse) outcome. Group III - «35,1-55» points (prognosis mala) - the probability of an adverse outcome 50,1-99,9% - higher risk of adverse outcome. Group IV - over 55.1 points (prognosis letalis) - probability of an adverse outcome 100% - absolutely unfavorable prognosis of the disease.
Figure 2. The values evaluating scale of prognostic index of TB/HIV co-infection adverse outcome in the forecasting information system of TB/HIV co-infection outcome.

For each group of the predicted outcome of the disease there were proposed diagnostic and organizational measures for to improve the quality of health care. To optimize the practical application of the developed methodology for predicting the adverse outcome of TB/HIV co-infection it was created the computer program “The program for calculating the lethal risk index of a hospitalized patient with HIV-associated tuberculosis from among inmates in penitentiary institutions” (registration No. 20169613944 dated 03/26/2019), which automatically calculates the prognostic index. The use of the developed forecasting information system for the TB/HIV co-infection outcome allows a short-term prognosis (within one year of dispensary observation) of the outcome of HIV-associated TB in patients from among inmates. This information system is intended to assess the disease outcome in patients hospitalized in specialized tuberculosis hospitals, as well as at the prehospital stage.

4. Conclusion

Thus, the main goal of the forecasting information system for the TB/HIV co-infection outcome is to improve the quality of health care for patients with HIV-associated tuberculosis from among inmates as well as reducing mortality from this disease by obtaining an additional highly accurate diagnostic criterion that helps in timely correction of treatment, prophylactic and organizational measures.

The advantage of the proposed forecasting information system for the TB/HIV co-infection outcome is simplicity, convenience and evaluation speed, possibility to evade the invasive methods use, minimum cost due to the absence of necessity for additional expensive laboratory and instrumental diagnostic methods. In addition, the said qualities of the proposed method allow using it as a screening test.

References

[1] Sterlikov S, Ponomarev S and Averyanova E 2018 Epidemic situation of HIV tuberculosis co-infection in the penitentiary system of the Russian Federation. Ural Medical Journal 8 95-97
[2] Vasilieva I, Sterlikov S, Belilovsky E, Borisov S and Ponomarev S 2018 Tuberculosis in penitentiary institutions in Europe and the Russian Federation Tuberculosis and lung diseases 5 6-13
[3] Ponomarev D, Sterlikov S, Ponomarev S and Averyanova E 2019 Using an artificial neural network in model prognosis in patients with a combination of HIV infection and tuberculosis in prisons. Tuberculosis and lung diseases 5 78-9
[4] Blagodatsky G, Gorokhov M, Ponomarev D, Ponomarev S and Vologdin S 2017 Informatization of the quality of medical care in the management of the medical service of the Federal penitentiary service of Russia. Collected: Information Technologies in Science,
Management, Social Sphere and Medicine 481-4

[5] Ponomarev D, Gorokhov M and Averyanova E 2019 Development of a model based on an artificial neural network for to predict the development of AIDS in inmates In the collection: penitentiary medicine in Russia and abroad Moscow 98-9

[6] Guliev R, Senko O, Zateyshchikov D, Nosikov V, Uporov I, Kuznetsova A and other mathematical biology and bioinformatics 2016 Application of optimal partitions for multi-parameter data analysis in clinical trials 1 46-63

[7] Dyuzheva E, Kuznetsova A and Senko O 2017 Determination of risk factors for cardiovascular mortality in penitentiary institutions using machine learning methods Doctor and information technology 2 25-49

[8] Saati T. 1993 Decision Making. The method of analysis of hierarchies Radio and communications 56-138

[9] Vostroknutov M, Dyuzheva E and Ponomarev S 2018 Lethal risk indicators in patients with comorbid HIV infection and tuberculosis in prisons. Ural Medical Journal. 8 29-32