Development of the point scoring system of medical parameters for HIV severity monitoring and modeling in numerical terms

D S Ponomarev¹, S B Ponomarev¹, E L Averyanova², V E Polishchuk¹ and A A Burt¹

¹ Research Institute of the Federal Penitentiary Service of Russia, Narvskaya ul. 15a building 1, Moscow, 125130, Russia
² Pskov State University, ul. Lenina 2, Pskov, 180000, Russia

E-mail ponomarev.dmitry1990@mail.ru

Abstract. The article considers the possibility of using a point scoring system of medical parameters for HIV severity monitoring and modeling in numerical terms. There were investigated the features of the use of medical parameters in the form of a point scoring system as well as the possibility of applying a numerical model for determining the HIV severity when tuberculosis is joined. To solve the problem, it was developed a step-by-step algorithm for applying the point scoring system. Conclusions were drawn. The possibility to predict in numerical terms the outcome of the disease (in particular the AIDS progression) using the number of systematically collected medical parameters in HIV carriers is of great interest. The information system development allows to determine the most important properties of the process and to separate its non-essential characteristics. Often the information system and modeling allow us to formulate new hypotheses and gain new knowledge about the object, those which were not available during the research.

1. Introduction
There are many ways to model and diagnose the stages of HIV. However, the problem of HIV spread, and determination of the disease severity remains of a great importance. It should be said that AIDS severity evaluation in numerical terms gives us the advantage due to several reasons, these are: the ability to reduce the expenses of medical experimental research; study of HIV course over time; the possibility to build prognosis in the form of time series; the possibility to identify relationships and dependencies between the parameters; the choice of optimal treatment strategies for patients; the research based on theoretical studies makes can be conducted without risk of negative consequences for patient health.

The possibility to predict in numerical terms the outcome of the disease (in particular the development of AIDS) using the number of systematically collected medical parameters in HIV carriers is of great interest. The information system development allows to determine the most important properties of the process and to separate its non-essential characteristics. Often the information system and modeling allow us to formulate new hypotheses and gain new knowledge about the object, those which were not available during its research. Development of models based on the formalization of relations between its elements is of great interest. The numerical (point)
expression of the main parameters of the HIV course can be used not only to assess the patient’s condition, but also to develop a model (for example, using machine learning methods).

All of the above allows us to eliminate knowledge gaps about the object, as well as to reveal new qualitative issues that initially could not be determined in a practical way. It should also be noted that many researches in this area has already been carried out already and this fact confirms the research relevance [1-4]. The use of artificial neural networks is of the greatest importance and interest. [5, 6]. However, for the further speps in this direction it is necessary to identify the correlation between the parameters and substantiate the choice of them.

It should be said that experimental studies of predicting the HIV course and the AIDS progression is a rather complicated process that needs to attract significant resources. Therefore, it seems more rational to conduct research at a theoretical level.

Summarizing all of the above, it should be noted that a method based on system analysis, machine learning, and mathematical statistics helps to create a simple non-invasive and accurate prognosis of AIDS development in HIV-infected patients which is based on the most informative medical parameters. And to achieve this, first of all it should be formed a precisely established scoring system for measuring those medical parameters.

2. Research methods
The studies [1-3] reveal a connection between the progression from HIV to AIDS and the stages of concomitant tuberculosis. But the proposed prognosis methods do not always take into account the whole range of parameters characterizing the patient’s state. Today, there are many methods identifying statistical links. However, the most used statistical method that helps to solve this problem is the correlation analysis.

In this study, to establish statistical links [7] between the main medical parameters and AIDS progression parameters, we used data generated on the basis of more than two years monitoring of HIV-infected inmates in the Federal Penitentiary System of Russia.

There were selected the main parameters gathered by the regular and systematic monitoring in HIV-infected in the Federal Penitentiary System of Russia. Each of these parameters was assigned a X value. A Y value was assigned for the parameter that reflects the risk of progression from HIV to AIDS (AIDS progression parameter). It was decided that the parameters X are considered as input signals, and Y are considered as output signals.

3. Conducting research
The most important parameters from the analyzed 30 parameters (X) were identified via methods of mathematical statistics in clinical research, optimal partitioning of data [8, 9] and correlation analysis.

Due to the fact that we are trying to determine a new model of HIV progression, as the zero value (0 points) we take the first stage of the disease and not the absence of the disease. As a result we have the following point scoring distribution of the HIV progression stages: Stage 1 - 0 points, Stage 2 - 1 point, Stage 3 - 2 points, Stage 4 A-B - 3 points, Stage 4 B-B - 4 points (Figure 1). This parameter has been assigned the value $X_2$.

![Figure 1. Point scoring system of HIV progression.](image-url)
Let's look at the definition of tuberculosis point scoring stages. Unlike the above model of HIV progression the absence of tuberculosis should be taken as a null value (0 points). It is proposed the following point scoring distribution of the tuberculosis stages: absence of tuberculosis - 0 points, tuberculosis in the anamnesis, metatuberculous changes in the lungs - 1 point, tuberculosis of thoracic lymph nodes - 2 points, focal tuberculosis - 3 points, infiltrative tuberculosis - 4 points, disseminated tuberculosis - 5 points, tuberculosis with pleurisy - 6 points, generalized tuberculosis - 7 points, tuberculosis recurrence - 8 points (Figure 2). This parameter has been assigned the value $X_3$.

![Figure 2. Point scoring system of the tuberculosis stages.](image)

Body temperature is one of the important parameters. In accordance with generally accepted standards for human body temperature it was developed the following point scoring system: 0 points – temperature from 36.6 °C to 37.0 °C, 1 point - temperature from 37.0 °C to 38.0 °C; 2 points - temperature above 38.0 °C (high temperature) (Figure 3). For convenience of further calculations this parameter has been assigned the value $X_5$.

![Figure 3. Point scoring system of body temperature measuring for the purpose of monitoring and modeling the AIDS severity in numerical terms.](image)

Since HIV-Tuberculosis co-infection is the object of the study, the presence or absence of chronic obstructive pulmonary disease was taken into account. In the presence of obstructive pulmonary disease it was assigned 1 point, and in its absence it was assigned 0 points. This parameter has been assigned the value $X_7$. 
In addition, in the research it has been taken into account the size of the liver. The measurement of liver size was carried out according to the well-known method developed by M.G. Kurlov. Liver size was given in centimeters ($X_9$). Due to the rather large gradations of this parameter, it was decided not to reduce it to a point form and to leave this parameter in actual values.

For the same reason the number of CD4 (another parameter that was taken into account) was measured in thousands of cells per milliliter (thousand cells / ml).

For AIDS progression parameter (Y) there were obtained the following correlation coefficients for the above parameters with moderate and strong correlation of values: $R_{X2}=0.5$; $R_{X3}=0.5$; $R_{X5}=0.5$; $R_{X7}=0.6$; $R_{X9}=0.6$; $R_{X12}=0.5$.

It was conducted regression analysis of the obtained data that was based on the identified correlations between the parameters [10]. As a result, regression equation coefficients were established. The results are presented in Table 1.

| Parameters | Coefficients |
|------------|--------------|
| Y-intersection | 1.3 |
| $X_2$ – HIV stage, points; | 0.3 |
| $X_3$ – tuberculosis stage, points; | 0.2 |
| $X_5$ – body temperature, points; | 0.1 |
| $X_7$ – chronic obstructive pulmonary disease (points); | 0.3 |
| $X_9$ – the size of the liver according to the method developed by M.G. Kurlov, centimeters; | 0.1 |
| $X_{12}$ – CD4, thousand cells / ml | $0.3 \times 10^3$ |

It should be said that the value of the outgoing parameter Y was divided into point equivalents. There were set restrictions determining that the value of Y less than 2 corresponds to a low risk of AIDS development, the value of Y from 2 to 3 corresponds to a medium risk of AIDS development, the value of Y more than 3 corresponds to a high risk of AIDS development.

Based on the obtained model it was developed a computer program that makes possible to predict the course of a disease. The program assumes that the user enter data in a point equivalent for the parameters considered above ($X_2$ – HIV stage, points; $X_3$ – tuberculosis stage, points; $X_5$ – body temperature, points; $X_7$ – chronic obstructive pulmonary disease (points); $X_{12}$ – CD4, thousand cells / ml). Further, the program calculates the point equivalent (in this case parameter Y) and the disease progression according to the formula that was obtained as a result of regression analysis.

4. Results
The study presents the development of the point scoring system of medical parameters for HIV severity monitoring and modeling in numerical terms. It was established the correlation between such parameters as the tuberculosis stage, deviation of body temperature, chronic obstructive pulmonary disease (presence or absence thereof), liver size and the AIDS progression in patients; there were
developed scoring point equivalents of these parameters allowing the use of statistical methods and machine learning methods.

The results of the study can be used for further development of the information system that allows deeper understanding of the processes of AIDS progression. The coefficients obtained by means of the regression analysis are the basis for the model that enables to prognose the risk of AIDS progression. The developed method has been tested in practice. It was also developed a computer program based on the results of the correlation analysis and the regression model. The implementation of this program will help to improve the quality of monitoring of the disease progression in inmates.

References
[1] Sizyakina L P and Andreeva I I 2005 A method for diagnosing the stages of HIV infection patent №2251701 RU (Moscow: Rospatent)
[2] Denisenko V B and Simovanyan E M 2010 A method for predicting the course of HIV infection in children patent №2444299 RU (Moscow: Rospatent)
[3] Ponomarev S B and Averyanova E L 2016 A method for predicting the development of AIDS in HIV-carriers patent №2597805 RU (Moscow: Rospatent)
[4] Vostroknutov M E Dyuzheva E V and Ponomarev S B 2018 Risk indicators for the death of patients with comorbid HIV infection and tuberculosis that occur in prisons Ural Medical Journal 8 (163) 29-32
[5] Ponomarev D S Gorokhov M M and Averyanova E L 2019 Development of models based on an artificial neural network for predicting the AIDS progression in inmates Federal Penitentiary Institution PKU Research Institute "Penitentiary Medicine in Russia and Abroad" pp 98-9
[6] Ponomarev D S Sterlikov S A Ponomarev S B and Averyanova E L 2019 The use of the artificial neural network in modeling predictions in patients with a combination of HIV infection and tuberculosis in prisons Tuberculosis and lung diseases 97(5) 78-9
[7] Förster E and Rönz B 1983 Correlation methods and regression analysis Finance and Statistics 151-218
[8] Guliev R R Senko O V Zateyshchikov D A Nosikov V V Uporov I V and Kuznetsova A V 2016 Application of optimal norms for multi-parameter data analysis in clinical studies Mathematical Biology and Bioinformatics 1 46-63
[9] Sergienko V I and Bondareva I B 2006 Mathematical statistics in clinical studies Geotar-Media 1 304
[10] Draper N and Smith G 2007 Regression analysis. Multiple regression Dialectics 3 912