PROGRAM APPLICATION FOR THE DETERMINATION OF THE STATE OF THE PATIENTS WITH JUVENILE RHEUMATOID ARTHRITIS IN THE REMOTE PERIOD

The research considers the process of the elaboration of the program application for the assessment of the remote state of the patients, who in their childhood were made a diagnosis of the juvenile rheumatoid arthritis. The complexity of the given task is that the patients were initially evaluated at an early age, then they got the corresponding treatment, which must be prescribed individually for each patient, as a result of the treatment the symptoms of the disease must disappear or be minimal. For the assessment of the degree of the disease indices JADI-A (joint arthritis damage index-articular) and JADI-E (joint arthritis damage index-extra-articular) are used. Data of 289 patients are used in the research.

Methods of linear programming are proposed to be used for the construction of the system of regression equations which allow to predict the numerical characteristics of the patient’s state in the adult period, taking into account the performed treatment. The structure of the models system consists of the models, which evaluated the state JADI-A and JADI-E in the remote period and the additional control parameters. These are indices that must be limited in a certain way. The presented models can also be used for the selection of the patient-specific treatment strategy, which, taking into account the initial state of the patient will allow to minimize the number of the articular and extra-articular damages. These models were tested for accuracy by means of the determination coefficient, all of them showed accuracy greater than 0.9 on the test data, this proves their high quality. All the models were used as the basis of the program application, which can be used for the storage of the patients’ data at all the stages of the examination and for the assessment of the patients’ state. This program product has been developed on the programming language R, using the framework Shiny. Using the application doctors will be able to observe the patient, store the results of the control and perform the assessment of the patient’s state.

Key words: juvenile rheumatoid arthritis, prediction model, regression analysis, methods of the operations study.

Introduction

The juvenile rheumatoid arthritis (JRA) is widely spread infant rheumatologic disease. Frequency of JRA cases in different regions of the world is from 0.05 to 0.8 %, morbidity rate is from 2 to 16 – 20 cases a year per 100 000 of the child population, death rate – 0.5 – 1 % [1]. In Ukraine the incidence of JRA is 0.2 – 0.4 per 1000 of child population. Incidence rate of JRA differ in various regions of Ukraine. Disease is mainly registered in Chernigiv Region (0.48), city of Kyiv (0.42), Dnipropetrovsk Region (0.34) [3]. The diagnosis in part of the patients is established during the first visit to the doctor, after a year of monitoring possible diagnostic errors also contain fourths cases of the disease, that determines the complexity of the clinical diagnostics of this pathology in children. In half of the patients the disability develops after 3 – 5 years of the disease [1].

Patients with JRA require constant supporting therapy. As the disease starts in childhood, main problem is the selection of such treatment strategy that takes into consideration the state of the patient in his childhood and enables to decrease the manifestation of the disease in the adult age. One of the possible variants of the solution of the problem is the development of the optimization models which will allow to assess the possible state of the patient in the adult age at the stage of the treatment selection. Similar problem is reduced to the search of the minimum of certain mathematical equation, when some variables we can change within certain limits.

Objective: To develop program application that enables to assess the possible state of the patient, taking into account the treatment, carried out.

Materials of the research. Database of JRA patients was used as the material for the research. Total amount of patients in the database was 289 persons, namely, 148 females and 141 males. Age of
the patients was within the limits 16 – 57 years, average age – 26.1. All the patients were subject to
general clinic studies in order to assess the type of the arthritis and detect the symptoms. Database
contained characteristics of the patients in the early period, characteristics of the treatment and state of
the patients in the remote adult period.

**Problem set up.** Clinical data, used in the study, were divided into the groups by their content. They
can be represented as the block matrix of the object-property [2]:

\[
X = \begin{bmatrix}
P & S^{\text{in}} & Q^{\text{in}} & S^{\text{out}} & Q^{\text{out}} & U
\end{bmatrix}
\]

where \( p_i, i=1,...,g \): are variables describing the initial parameters of the patient (age, gender, weight,
blood analysis, clinical symptoms of the disease). These variables are important when the treatment
strategy is prescribed but they will not be modeled as state variables after treatment; \( q^{\text{in}}, q^{\text{out}} \): main
characteristic of the disease that can be considered as principle criterion of treatment (minimization of
clinic manifestations). As such criterion in the given study the JADI-A[3] (assessment of articular
damage) and extraarticular damage (JADI-E) index after treatment was chosen; \( S^{\text{in}}_i, S^{\text{out}}_i, i=1,...,d \):
are variables of the patients state, values of which must be taken into account when the state of the
patient is assessed after the treatment. Database contains the values of the patients state before \( S^{\text{in}}_i \),
\( i=1,...,d \) and after the treatment \( S^{\text{out}}_i, i=1,...,d \); \( u_i, i=1,...,h \) are variables, characterizing the applied
treatment (doses and duration of the drug administration) [4].

For the convenience, the data, described above, can be presented by the block matrix of the object-property X:

\[
X = \begin{bmatrix}
P & S^{\text{in}} & Q^{\text{in}} & S^{\text{out}} & Q^{\text{out}} & U
\end{bmatrix}
\]

where \( \mathbf{P} = \begin{bmatrix}
p_{11} & \cdots & p_{1g} \\
\vdots & \cdots & \vdots \\
p_{ng} & \cdots & p_{ng}
\end{bmatrix} \) – is the matrix of the patients parameters, where each row contains the
value \( g \) of the parameters, according to the number of the patients row, \( n \) – is the amount of the patients
in the sample;

\[
\mathbf{S}^{\text{in}} = \begin{bmatrix}
s^{\text{in}}_{11} & \cdots & s^{\text{in}}_{1d} \\
\vdots & \cdots & \vdots \\
s^{\text{in}}_{ng} & \cdots & s^{\text{in}}_{nd}
\end{bmatrix}, \quad \mathbf{S}^{\text{out}} = \begin{bmatrix}
s^{\text{out}}_{11} & \cdots & s^{\text{out}}_{1d} \\
\vdots & \cdots & \vdots \\
s^{\text{out}}_{ng} & \cdots & s^{\text{out}}_{nd}
\end{bmatrix}, \quad d \) – is the amount of the patients state variables;

\[
\mathbf{Q}^{\text{in}} = (q^{\text{in}}_1, q^{\text{in}}_2, ..., q^{\text{in}}_n)^T, \quad \mathbf{Q}^{\text{out}} = (q^{\text{out}}_1, q^{\text{out}}_2, ..., q^{\text{out}}_n)^T \)
\) – are vectors of the criterion values before and
after the treatment;

\[
\mathbf{U} = \begin{bmatrix}
u_{11} & \cdots & u_{1h} \\
\vdots & \cdots & \vdots \\
u_{nh} & \cdots & u_{nh}
\end{bmatrix}, \quad h \) – is the number of variables, having therapeutic action.
Matrix X presents the data of the transition of the object i, being in the state \( q_i^{in} \), \( q_i^{out} \) \( j = 1, \ldots, d \) in the state \( s_{ij}^{in} \), \( s_{ij}^{out} \) \( j = 1, \ldots, d \) by means of the applied values of the treatment impacts \( u_{ij} \), \( j = 1, \ldots, h \). The same for each patient \( i = 1, \ldots, n \) [5].

At the first stage of the study, from the total set of the parameters, characterizing the patients, the set of the characteristics of the patients who had statistically valid correlation with JADI-A and JADI-E indices, was selected. Besides, the indices, having mutual correlation were rejected. Complete list of the parameters in presented in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 (X1) )</td>
<td>Age at the moment of the examination</td>
</tr>
<tr>
<td>( P_2 (X2) )</td>
<td>Gender f-1, m-2</td>
</tr>
<tr>
<td>( P_3 (X3) )</td>
<td>Age of the patient at the moment of the disease start, yr</td>
</tr>
<tr>
<td>( P_4 (X4) )</td>
<td>Morning stiffness in childhood, min</td>
</tr>
<tr>
<td>( P_5 (X5) )</td>
<td>Erythrocytes sedimentation rate onset, min</td>
</tr>
<tr>
<td>( P_6 (X6) )</td>
<td>Blood glucose in childhood before treatment</td>
</tr>
<tr>
<td>( S_1^{in} (X7) )</td>
<td>Erythrocytes sedimentation rate in childhood before treatment</td>
</tr>
<tr>
<td>( S_2^{in} (X8) )</td>
<td>C-reactive protein (CRP) in childhood before treatment</td>
</tr>
<tr>
<td>( S_3^{in} (X9) )</td>
<td>Pain severity according to visually analogue scale in childhood before treatment</td>
</tr>
<tr>
<td>( U_1 )</td>
<td>Course duration of gluco-corticoids administration</td>
</tr>
<tr>
<td>( U_2 )</td>
<td>Gluco-corticoids, cumulated dose</td>
</tr>
<tr>
<td>( U_3 )</td>
<td>Methotrexat, dose</td>
</tr>
<tr>
<td>( S_1^{out} (X10) )</td>
<td>Erythrocyte sedimentation rate (ESR) in adult age after treatment</td>
</tr>
<tr>
<td>( S_2^{out} (X11) )</td>
<td>C-reactive protein (CRP) in adult age after treatment</td>
</tr>
<tr>
<td>( S_3^{out} (X12) )</td>
<td>Pain severity according to visual analogue scale in adult age after treatment</td>
</tr>
<tr>
<td>( P_8 (X14) )</td>
<td>Occurrence of the symmetric arthritic in childhood</td>
</tr>
<tr>
<td>( P_9 (X15) )</td>
<td>Pain in cervical section in childhood</td>
</tr>
<tr>
<td>( P_{10} (X16) )</td>
<td>Occurrence of dactylitis in childhood</td>
</tr>
<tr>
<td>( P_{11} (X17) )</td>
<td>Occurrence of lymphadenopathy, splenomegaly in childhood</td>
</tr>
<tr>
<td>( P_{12} (X18) )</td>
<td>Pain in the ridge in childhood</td>
</tr>
<tr>
<td>( P_{13} (X19) )</td>
<td>Occurrence of the uveitis in childhood</td>
</tr>
<tr>
<td>( q_1^{out} )</td>
<td>Arthritis Damage Index (ADI)</td>
</tr>
<tr>
<td>( q_2^{out} )</td>
<td>Extraarticular Damage Index (ADI-E)</td>
</tr>
</tbody>
</table>

**Solution of the set problem.** The set problem was solved by means of programming language R.
and development environment R Studio. Regression models for the criterion $q_1^{\text{out}}$ and $q_2^{\text{out}}$, and limitations $s_1^{\text{out}}$, $s_2^{\text{out}}$, $s_3^{\text{out}}$ were composed in the execution of the optimization problem solution. For this purpose the set matrix was expanded by means of adding to variables of $x$ type variables of the types: $x^2$, $\frac{1}{x}$, $x_i * x_j$, $\frac{1}{x^2}$, $\frac{1}{x_i * x_j}$, and also the products of these variables on the control criteria $U_1$, $U_2$ and $U_3$.

Result of JADI A modeling is given below:

$$q_1^{\text{out}} = 0,012 * X7 * X8 - 0,000001241 * U2 * X7 * X8 - 0,00003347 * U3 * X4 * X9 + 0,006 * U3 * X9X5 + 0,715 * X17 + 0,001 * U3 * X7X1 - 0,185 * X8X3 + 0,006 * U2X62 - 3,924 * U3X4 * X9 + 0,683 * X19 - 0,528 * X2 + 504,709 * 1X7 * X8 - 0,028 * U1 * X4X7 + 0,674 * X4X7 - 7,210$$

Determination coefficient for the model $q_1^{\text{out}}$ is 0.963, this proves high accuracy of the results.

Diagram of residuals – it is a graph that shows the residuals on the vertical axis and an independent variable on the horizontal axis. If the dots on the residual graph are randomly distributed around the horizontal axis then the model of the linear regression sufficiently describes the data from the table, in other case non-linear model is more acceptable [6]. In our case this graph can be used for the assessment of the fact if the input data matrix was sufficiently expanded.

![Fig. 1. Graphs of «JADI A» model](image)

It is seen in the graphs that the residual values are distributed from -0.10 to 0.20 for other fixed variables. It is seen from the histogram graph for the residual that greater part of the residual is concentrated near the average value with almost normal curve, the curve has thin intervals at the ends.

Results of the regression for JADI E are given below.
\[ q_{2}^{\text{out}} = 0,010 \times X5 \times X9 - 1,060 \times X19 - 0,006 \times X5 \times X7 - 0,007 \times U1 \times X5 - 0,121 \times X9 + 0,00001297 \times U3 \times X7 - 0,175 \times U3 + 0,203 \times X14 + 0,017 \times U3 \times X4X8 + 0,231 \times U3X3 \times X8 + 0,0000003582 \times U2 \times X5 \times X8 + 0,013 \times U1 \times X9X8 - 0,05 \times U2X92 - 0,001 \times U3 \times X9X4 + 4,916 \]  

(2)

Determination coefficient for the model \( q_{2}^{\text{out}} \) is 0.945, this also proves high accuracy of the results.

It is seen in the graphs that the residual values are distributed from -0.10 to 0.5 for all different variables of the fixed variable. The graph of the histogram for the residual shows that greater part of the residual is concentrated near the average value with almost normal curve, and the curve has thin intervals at the ends.

In the same way the models for the control parameters were obtained, namely: \( s_{1}^{\text{out}} \) – ESR, \( s_{2}^{\text{out}} \) – CRP, \( s_{3}^{\text{out}} \) – VAS. These models are given in the formulas (3 – 5). For \( s_{1}^{\text{out}} \) the value of the determination coefficient – 0.937 was obtained, for \( s_{2}^{\text{out}} \) – 0.916 and for \( s_{3}^{\text{out}} \) – 0.9145.

\[ s_{1}^{\text{out}} = U10,129 \times X3X9 - 21,830 \times 1X5 \times X7 + U2 - 0,001 \times X5X9 + 27,111 \times 1X7 \times X8 + 0,000009618 \times X7 \times X8 + U3 - 0,000006995 \times 1X32 + 0,569 + 0,058 \times X5 \times X6 - 1,562 \times X2 - 0,980 \times X17 + 10,275 \]  

(3)

\[ s_{2}^{\text{out}} = U1 - 1,042 \times 1X6 + U20,001 \times 1X4 \times X5 + U3(0,013 \times X4X6 + 0,003 \times X3X7 + 7,314 \times 1X5 \times X6) + 0,008 \times X8 \times X9 - 2,682 \times X8X5 + 1,646 \times X15 - 1,335 \times X19 + 9,904 \]  

(4)
\[ s_{3}^{out} = U1 - 0,729 \ast 1X6 + 3,206 \ast 1X3 \ast X5 \ast U20,029 \ast 1X1 \ast X3 + U3 - 5,8433 \ast 1X4 \ast X5 + 0,009 \ast X8 \ast X9 - 0,012 \ast X1 \ast X3 + 2,860 \ast X16 + 2,357 \ast X18 \ast 0,001 \ast X4 \ast X7 + 8,783 \]

Quality adequacy of this set of models was evaluated on real data of the patients. For this purpose from the initial set of data 10 % of the observations were randomly selected. This array did not participate in modeling and was used only for the evaluation of the model adequacy by comparing the table and model data. Accuracy was determined by means of the correlation analysis, which enabled to obtain the value of 0.86 for JADI А and 0.92 for JADI Е.

These models were used in the program application, which was developed on the programming language R, using framework Shiny. Program application consists of the module of the patient input, module of treatment strategy assessment and data base.

Fig. 3 illustrates model dialog “Forecasting result”, it contains the results of forecasting the state of JRA patient determination in the remote period for the individual patient.
Conclusions

The obtained models enable to assess the efficiency of the treatment of JRA patients according to JADI A and JADI E indices. The obtained models have been evaluated according to the indices of the determination coefficient, that took the values of 0.963 and 0.945, correspondingly. According to the results of modeling the program application was developed, it can be used in real medical practice for the patient-specific analysis of the patients state in the remote period, taking into account his initial state.

REFERENCES