International Medicine 2020; 2(2): 107-118

International Medicine

https://www.theinternationalmedicine.com

Original Article

Use of the complex of models of regression for analysis of the factors that determine the severity of bronchial asthma

Oleh M. Pihnastyi[1], Olga S. Kozhyna[2]

[1]*Department of Distributed Information Systems and Cloud Technologies, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine*

[2]*Department of Fundamentals of Pediatrics No.2, Kharkiv National Medical University, Kharkiv, Ukraine*

Received: 24 November 2019 / Accepted: 11 January 2020

Abstract

Background: According to an International Study of Asthma and Allergies in Childhood (ISAAC), the prevalence of asthma in children of 6-7 years old has increased by 10%, and at the age of 13-14 years by 16% over the last decade. Determining the factors that are keys to the occurrence of the disease and its severity is important in explaining the pathogenesis of bronchial asthma.

Methods: Analyzed 142 indicators of clinical and paraclinical examination of 70 children with asthma. To select factors that could be significant in the formation of severe asthma, applied the method of logistic regression with step-by-step inclusion of predictors. Both quantitative and qualitative characteristics were selected. Each qualitative attribute was coded “1” if the child had this characteristic, or “0” if this characteristic had not been established. The formation of a severe asthma course was accepted according to (1) and the absence of a severe asthma flow formation as (0).

Results: Analyzed the model of paired regression, the boundary value of thymic stromal lymphopoietin was established, exceeding which indicates the high probability of the presence of severe bronchial asthma. Increasing the value of thymic stromal lymphopoietin by 10 pg/mL suggests an increase in the likelihood of severe asthma by 10%.

Conclusions: A complex of steam regression models has been developed to determine the factors characterizing the severity of bronchial asthma. The risk of developing severe bronchial asthma in children has been determined and 15 factors have been identified that affect severe asthma.

Keywords: asthma, child, matched-pair analysis, thymic stromal lymphopoietin

Introduction

Nowadays, there are 339 million people suffering from bronchial asthma in the world [1]. Despite many years of research, asthma is still the most common chronic disease within the children population in different countries of the world, and the incidence increases with every passing year [2]. Asthma remains one of the most common causes of disability in pediatric ages and takes the 4th place in the structure

Address for Correspondence: Oleh M. Pihnastyi, Department of Distributed Information Systems and Cloud Technologies, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine. E-mail: pihnastyi@gmail.com

DOI: 10.5455/im.74961

of common disabilities among 10-14 years old children [3-5]. Patients' quality of life significantly decreases and is followed by notable economic expenditures both for the family and for the society [6]. The first symptoms of the disease often appear in childhood and it is still difficult to prognosticate further course of the disease. Prognostication of asthma course and disease in children is still a difficult problem, because of the multifactority of the disease [7]. Numerous clinical studies of uncontrolled asthma course approve the necessity to analyze the factors influencing severe forms of the disease [8-10]. One such factor can be thymic stromal lymphopoietin (TSLP). There is a relationship between TSLP dysregulation and its role in the pathogenesis of atopic diseases, such as atopic dermatitis, asthma, allergic rhinitis and eosinophilic esophagitis [11,12]. Detection of cases at early stages with the risk of severe asthma development in children is one of the main problems. This gives us the opportunity to use individual methods of therapy and observation in these cases. Selection and substantiation of regression models which define the characterizing severity of bronchial asthma are the aims of the study.

Methods

Let's take a look at the method of construction of the complex of regressive models containing K number of inputs (regressors) x_k, ($k = \overline{1,K}$) and M number of outputs y_m, ($m = \overline{1,M}$).

Let's present the dependence of the inputs and outputs in this form

$$y_m = F_m(x_1, x_2, ..., x_{K-1}, x_K) + \varepsilon_m$$

where ε_m - is a random error. The random error is determined by the fact that measured experimental values of inputs and outputs deviate from its initial values due to random errors. We will construct the regressive model for the case when the random error in the equation of regression ε_m is subject to the normal distribution law.

Let's assume that after examining of i-patient parameter x_k will equal x_{ki}, ($i = \overline{1,N}$). Value of y_m equals to y_{mi}.

$$y_{mi} = F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki}) + \varepsilon_{mi}$$

which is distributed by the normal law with the mathematical expectation

$$F_{mi} = F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})$$

and standard deviation

$$\sigma_{mi} = \sigma_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki}),$$

that characterizes the error of measurement. Let's assume that values of the factors obtained after examining every particular i-patient (clinical examining data) were received with the same accuracy that allows us to write

$$\sigma_{m1} = \sigma_{m2} = ... = \sigma_{mN} = \sigma_m = \text{const}.$$

Then probability distribution of y_{mi} can be presented as

$$\varphi_{mi}(y_{mi}) = \frac{1}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2}\left(\frac{y_{mi} - F_{mi}}{\sigma_m}\right)^2\right) = \frac{1}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{mi}}{\sigma_m}\right)^2\right). \quad (2)$$

The probability that the value of output y_m for i-patient is in this range $(y_{mi}, y_{mi} + dy_m)$ can be calculated with the following equation

$$\varphi_{mi}(y_{mi})dy_m = \frac{dy_m}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{mi}}{\sigma_m}\right)^2\right).$$

Since the events under which y_m factor equals y_{mi} are independent then the possibility that values y_{mi} of the explained factor y_m for the first patient lays within $(y_{m1}, y_{m1} + dy_m)$, for the second patient lays within $(y_{m2}, y_{m2} + dy_m)$,..., for N-patient lays within $(y_{mN}, y_{mN} + dy_m)$ is the multiplication of possibilities.

$$\prod_{i=1}^{N}\varphi_{mi}(y_{mi})dy_m=\prod_{i=1}^{N}\frac{dy_m}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2}\left(\frac{y_{mi}-F_m(x_{1i},...,x_{Ki})}{\sigma_m}\right)^2\right)\to\max. \quad (4)$$

Let's determine the average $F_m(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki})$ so that expression (4) would comply with the maximum value. In view of the fact that

$$\prod_{i=1}^{N}\frac{dy_m}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2}\left(\frac{\varepsilon_{mi}}{\sigma_m}\right)^2\right)=\frac{dy_m}{\sqrt{2\pi}\ \sigma_m}\exp\left(-\frac{1}{2\sigma_m^2}\sum_{i=1}^{N}\varepsilon_{mi}{}^2\right). \quad (5)$$

It follows that the possibility (4) will be maximal if the sum of the squares of errors ε_{mi} is minimal

$$\sum_{i=1}^{N}(\varepsilon_{mi})^2=\sum_{i=1}^{N}\left(y_{mi}-F_m(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki})\right)^2\to\min. \quad (6)$$

Expression (6) complies with the fact that there are two dependencies $F_m(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki})$, for which the set of experimental values y_{mi} of the output parameter y_m is the most possible. Emphasize that the values of factors determined after examination of every particular i-patient were received with the same accuracy. In another case expression (6) should be written in this form

$$\sum_{i=1}^{N}\left(\frac{\varepsilon_{mi}}{\sigma_{mi}}\right)^2=\sum_{i=1}^{N}\left(\frac{y_{mi}-F_m(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki})}{\sigma_m(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki})}\right)^2\to\min.$$

The function $F_m(x_1,x_2,...,x_{K-1},x_K)$ can be depicted as a superposition Z of known functions

$$F_m(x_1,x_2,...,x_{K-1},x_K)=\sum_{z=1}^{Z}a_z g_z(x_1,x_2,...,x_{K-1},x_K)$$

with a_z coefficients, that are derived from the case (6)

$$\sum_{i=1}^{N}\varepsilon_{mi}\frac{\partial\varepsilon_{mi}}{\partial a_z}=-\sum_{i=1}^{N}\left(y_{mi}-\sum_{v=1}^{Z}a_v g_{vi}\right)\cdot g_{zi}=0$$

where

$$\frac{\partial\varepsilon_{mi}}{\partial a_z}=-g_{zi}\qquad g_{zi}=g_z(x_{1i},x_{2i},...,x_{(K-1)i},x_{Ki}).$$

Let's show the received system of equations in the following form

$$\sum_{v=1}^{Z}a_v\sum_{i=1}^{N}g_{vi}\cdot g_{zi}=\sum_{i=1}^{N}y_{mi}\cdot g_{zi}. \qquad (z=1..Z) \quad (8)$$

Unknown coefficients a_z are determined by the solution of the system of linear equations (8) with the defined system of functions $g_z(x_1,x_2,...,x_{K-1},x_K)$ which is in many cases an orthogonal system of functions on the interval of changes of values of the parameters x_k. In that regard let's present the dependence $F_m(x_1,x_2,...,x_{K-1},x_K)$ as an orthogonal range of Wiener [13]:

$$F_m(x_1,x_2,...,x_{K-1},x_K)=a+\sum_{k=1}^{K}b_k x_k+\sum_{k=1}^{K}\sum_{k2=1}^{K}c_{kk2}x_k x_{k2}+\sum_{k=1}^{K}\sum_{k2=1}^{K}\sum_{k3=1}^{K}d_{kk2k3}x_k x_{k2}x_{k3}+... \quad (9)$$

This submission allows expressing coefficients a_z at functions $g_z(x_1, x_2, ..., x_{K-1}, x_K)$ through initial moments of random errors X_k, Y_m. Define the values $a, b_k, c_{kk2}, d_{kk2k3}...$, at which the left side of the expression (6) accepts the minimal value. For this we find extreme value for expression (6) regarding the rates $a, b_k, c_{kk2}, d_{kk2k3}...$ with the dependence of input and output factors (9):

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right)\frac{\partial F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})}{\partial a} = 0, \quad (10)$$

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right)\frac{\partial F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})}{\partial b_{k1}} = 0,$$

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right)\frac{\partial F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})}{\partial c_{k1k2}} = 0,$$

Differentiate the function $F_{mi} = F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})$ respectively on $a, b_k, c_{kk2}, d_{kk2k3}...$, obtain

$$\frac{\partial F_{mi}}{\partial a} = 1, \qquad \frac{\partial F_{mi}}{\partial b_k} = x_{ki}, \qquad \frac{\partial F_{mi}}{\partial c_{kk2}} = x_{ki}x_{k2i}. \quad (11)$$

Substitute the values of partial derivatives (11) of the function $F_m(x_1, x_2, ..., x_{K-1}, x_K)$ into the system of equations (10), we obtain the system of equations for identification of unknown rates $a, b_k, c_{kk2}, d_{kk2k3}...$:

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right) = 0, \quad (12)$$

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right)x_{ki} = 0,$$

$$\sum_{i=1}^{N}\left(y_{mi} - F_m(x_{1i}, x_{2i}, ..., x_{(K-1)i}, x_{Ki})\right)x_{ki}x_{k2i} = 0,$$

Engage the initial moment, expressed through the frequencies of appearances of covariates X_k, Y_m

$$\frac{1}{N}\sum_{i=1}^{N}x_{ki} = \alpha_1[X_k] = \alpha_{01}[X_k, X_{k2}] = m_{xk}, \quad \frac{1}{N}\sum_{i=1}^{N}x_{ki}^2 = \alpha_2[X_k] = \alpha_{02}[X_k, X_{k2}], \quad (13)$$

$$\frac{1}{N}\sum_{i=1}^{N}x_{ki}x_{k2i} = \alpha_{11}[X_k, X_{k2}], \qquad \frac{1}{N}\sum_{i=1}^{N}y_{mi} = \alpha_1[Y_m] = \alpha_{01}[X_k, Y_m] = m_{ym},$$

$$\frac{1}{N}\sum_{i=1}^{N}x_{ki}y_{mi} = \alpha_{11}[X_k, Y_m].$$

Values for initial moments calculated with the formula (13) will be the same possible as the relevant values for the general population as we increase the number of experiments N. Considering (13) we obtain a linear $a, b_k, c_{kk2}, d_{kk2k3}...$ system of equations

$$\begin{cases} a & +\sum_{k=1}^{K} b_k \alpha_{01}[X_k, X_{k2}] & +\sum_{k1=1}^{K}\sum_{k2=1}^{K} c_{k,k2}\alpha_{11}[X_k, X_{k2}] & +... = \alpha_{01}[X_k, Y_m], \\ a\sum_{k=1}^{K}\alpha_{10}[X_k, X_{k2}] & +\sum_{k=1}^{K}\sum_{k2=1}^{K} b_k \alpha_{11}[X_k, X_{k2}] & +... & +... = \alpha_{11}[X_k, Y_m], \\ a\sum_{k=1}^{K}\sum_{k2=1}^{K}\alpha_{11}[X_k, X_{k2}] & +... & +... & +... = \alpha_{21}[X_k, Y_m], \\ ... & +... & +... & +... = \alpha_{31}[X_k, Y_m]. \end{cases} \quad (14)$$

In case of selection of linear relationship for parameters x_1, y_1 $(K = 1,\ M = 1)$ expression (9) become

$$y_1 = a + b_1 x_1$$

with the value of coefficients a and b_1, that are determined by the system of equations (14):

$$\begin{cases} a & +b_1 \cdot \alpha_{01}[X_1, X_1] & = \alpha_{01}[X_1, Y_1], \\ a \cdot \alpha_{10}[X_1, X_1] & +b_1 \cdot \alpha_{11}[X_1, X_1] & = \alpha_{11}[X_1, Y_1]. \end{cases}$$

Using the definition of dispersion and covariance of a random variable

$$D_{x1} = \frac{1}{N}\sum_{i=1}^{N}(x_{1i} - m_{x1})^2 = \alpha_{11}[X_1, X_1] - (m_{x1})^2, \quad (15)$$

$$K_{x1y1} = \frac{1}{N}\sum_{i=1}^{N}(x_{1i} - m_{x1})(y_{1i} - m_{y1}) = \alpha_{11}[X_1, Y_1] - m_{y1} \cdot m_{x1},$$

the expression for coefficients a and b_1 can be presented as

$$a = m_{y1} - m_{x1}\frac{K_{x1y1}}{D_{x1}}, \qquad b_1 = \frac{K_{x1y1}}{D_{x1}},$$

that determines the equation of the model of matched regression

$$y_1 = \left(m_{y1} - m_{x1}\frac{K_{x1y1}}{D_{x1}}\right) + \frac{K_{x1y1}}{D_{x1}}x_1. \quad (16)$$

Increasing of covariance rate via the absolute value leads to enhancement of the dependence between x_1 and y_1 parameters. If we turn to dimensionless variables

$$\xi_1 = \frac{x_1 - m_{x1}}{\sqrt{D_{x1}}}, \qquad \eta_1 = \frac{y_1 - m_{y1}}{\sqrt{D_{y1}}}, \quad (17)$$

we obtain a dimensionless equation of model of couple regression

$$\eta_1 = r_{x1y1} \cdot \xi_1, \qquad r_{x1y1} = \frac{K_{x1y1}}{\sqrt{D_{x1}D_{y1}}}, \qquad -1 \le r_{x1y1} \le 1, \quad (18)$$

that looks more compact comparing to (16). The correlation coefficient r_{x1y1} between factors x_1 and y_1, characterizes the degree of the linear relationship between them. The module of value r_{x1y1} defines the degree of density of linear relationship. For the factors which acquire non-random values, $|r_{x1y1}| = 1$. Equality $r_{x1y1} = 0$ means that there is no linear relationship between factors x_1 and y_1.

Unprocessed data of experimental research are presented in [14]. The study respects for human rights in accordance with the current legislation of Ukraine, follows international ethical requirements and does not violate any scientific ethical norms as well as standards of biomedical research. Individual data of the patients are encoded. Before the experiment conduction, factors that could affect the severity of bronchial asthma are identified. The following 142 factors are categorized [Table 1], allowing to unify the data analysis and structure the relationships between them [15] [Figure 1–6].

Table 1. Classification of factors [2]

№	Name	Common information
1	**Course type**	Severe persistent, moderate persistent, mild persistent, intermittent
2	**Principal diagnosis (case taking)**	Allergic rhinitis, atopic dermatitis, deflection of the nasal septum, secondary cardiomyopathy
3	**Anamnesis of disease**	Number of years from the first symptoms
4	**Anamnesis of life**	Bronchial asthma in mother, allergic rhinitis in mother, bronchial asthma in father, allergic rhinitis in father, bronchial asthma in relatives of the second generation
5	**Complete blood count**	RBC, HBC, WBC, eosinophils%, basophils%, band neutrophils%, segmented neutrophils%, lymphocytes%, monocytes%, ESR
6	**Urinalysis test**	Color, appearance, specific gravity, pH, leukocytes, transitional epithelial cells, mucus
7	**Skin allergy test (pollen allergens)**	Chenopodium, birch, sagebrush, sunflower, timothy, cock's-foot, rye, elytrigia, ryegrass, alder, fescue grass, walnut, linden, plantain, nettle, dandelion, sumpweed, foxtail, ragweed, chestnut, cypress, maple
8	**Skin allergy test (household allergens)**	Domestic dust, pillow feather, daphnia, rabbit hair, cat hair, dog hair, sheep wool
9	**Food allergy test**	Egg white, carp fish, pollack fish, milk, lemon, apple, raspberries, beetroot, cabbage, beef, pork, chicken, banana, watermelon, orange, tangerine, grape, cocoa, soybeans, black tea, tomato, beans, rice groats, buckwheat grits, wheat groats, oat grits, corn grits, rye grits
10	**Ig E**	Serum immunoglobulin E
11	**Spirogram**	VC, FVC, PEF, FEF(MEF) 25%, FEF(MEF) 50%, FEF(MEF) 75%, FEV 25-75%, FEV1, FEV/FVC %, MVV.
12	**Immunological status**	Leukocytes, neutrophils (%;10^3 cells), lymphocytes (%;10^3 cells), CD3 (%;10^3 cells), CD4 (%;10^3 cells), CD8 (%;10^3 cells), CD16 (%;10^3 cells), CD22 (%;10^3 cells), CD25 (%;10^3 cells), IgA, IgM, IgG, phagocytosis of latex, %, phagocytic number, total complement (CH 50), CIC with 3.5% PEG, units., spontaneous NBT tests, %, spontaneous IAN tests, units., stimulated NBT test, %, stimulated IAN test, units, lysosomal cationic proteins, units
13	**TSLP**	The level of serum thymic stromal lymphopoietin
14	**Age**	Age of the examined children from 6 to 17 years
15	**Gender**	Surveyed 56 boys and 14 girls

A large number of factors that determine the presence of a certain sign are coded with 0 or 1 (e.g. [4. Bronchial asthma in mother]=1, if this sign characterizes the patient and [4. Bronchial asthma in mother]=0, if the sign is absent). There are factors in the analysis above that can accept values related to several situations. For instance, factor [8. Pillow feather], Figure 6, characterizing the degree of intensity of a diagnostic skin test is coded so every degree of intensity relates to a certain number. ([8. Pillow feather]=0 if the sign is absent, [8. Pillow feather]=1 or 2 if the reaction is weakly positive, 3 and 4 if it is positive. Factor [8. Domestic dust] also relates to such kind of factors (0- if there is no reaction to dust, 1 and 2 – mild reaction, 3 and 4 – positive, Figure 2). Factors that can turn continuous variables after measurement are characterized by the unit of measurement ([3. Number of years from the first symptoms], [12. CD8 10*3 cells]), [13.

TSLP], [Figure 1-4] [14]. The number before the name of the factor determines the category to which this factor relates to, (Table 1: TLSP factor, category №13) [15].

Factors that were identified after clinical examination in not more than two patients were excluded from the input unprocessed data [Table 2]. The list of excluded factors is in Table 2. It's expected that excluded factors don't affect the result of the research too much so they can be omitted. It allowed shortening the number of factors for analysis by ~10%. So, 130 factors were used for the regressive model.

Table 2. List of excluded factors [15]

№	Name	The number of patients with the present factor	Description of the category of factor [Table 1]
1	**Vasomotor rhinitis**	2	2. Principal diagnosis (case taking)
2	**Allergic rhinosinusitis**	2	
3	**Corn**	2	7. Skin allergy test (pollen allergens)
4	**Acacia**	2	
5	**Willow**	1	
6	**Sorrel**	2	
7	**Hazel**	2	
8	**Egg yolk**	2	9. Food allergy test
9	**Hake**	2	
10	**Cucumber**	2	
11	**Carrot**	2	
12	**Potatoes**	2	

The next step of the research was dividing the data into two different parts. The first part was used for the construction of the regressive model [16], the second part – the testing one, for verification of the regressive model adequacy.

Results and Discussion

Conducted literature review characterizes bronchial asthma as a heterogeneous disease with an increasing number of severe cases [17-19]. Numerous clinical research of low-controlled severe bronchial asthma confirms the necessity of studying new biological markers for appropriate therapy onset [20]. For understanding the pathophysiology of severe bronchial asthma, the analysis of factors affecting the beginning of severe forms is needed. Severe bronchial asthma will be characterized by explained factors [1. Severe persistent]. We have already analyzed the factor "thymic stromal lymphopoietin" [13. TSLP] [21-24], which is used for assessment of the severity of bronchial asthma. Its high numbers are usually for severe bronchial asthma [25-27]. Researchers Soumelis [28] and Lee [29] proved the importance of TSLP in the pathogenesis of atopic dermatitis.

The relationship between the researched factors is determined by the correlation coefficient [16], which complies with the level of the linear relationship between factors. As a consequence of that, it's advisable to choose a combination of factors for the construction of the complex of the models of matched regression those to have both the highest and the lowest values of the correlation coefficient via absolute value. The group of factors with the highest value of the correlation coefficient is supposed to have significant relations between factors in the model of matched regression. By contrast, a group of factors with the lowest value of the correlation coefficient is supposed to have no relations between factors. Both groups are needed for the research. The first group allows defining the set of factors that can be used to identify the severity of bronchial asthma. In our research, we assume that there is a relationship between explained factor and regressor (the weak relationship is also taken into account), if the correlation coefficient is

$$\left|r_{x1y1}\right| \geq 0.21. \tag{19}$$

The full list of factors with the value of the correlation coefficient (19), used for the construction of the model of matched regression is given in [16]. The second group of factors is characterized by the absolute value of the correlation coefficient

$$\left|r_{x1y1}\right| < 0.21. \tag{20}$$

It's assumed that for this group there is no relationship between the explained factor and the regressor. Such an approach allows us to reduce the dimensionality of the model. In our case, it's only 15 factors out of 130 (including excluded ones) for which the condition is met

(19). Equations of the regressive models for the factors defined by restriction (19) are given in Table 3. The list of factors for which the lowest value of the correlation coefficient $|\ r_{x1y1}\ |$ < 0.02 is given in Table 4. The extended list of factors satisfying the restriction (20) is in [16].

Table 3. Complex of the pair regression models

№	r_{x1y1}	**Explainable factor,** y_1	**Regressor,** x_1	**Regression model (16)**
1	- 0.26	13. TSLP	6. Color	$y_1 = 33.36 - 15.16x_1$
2	-0.26		12. CD8 10*3 cells	$y_1 = 60.27 - 44.6x_1$
3	-0.24		5. ESR	$y_1 = 39.4 - 3.61x_1$
4	0.22		4. Bronchial asthma in relatives of the second generation	$y_1 = 23.75 + 24.67x_1$
5	0.25		14. Age	$y_1 = 3.83 + 2.0x_1$
6	0.26		2. Allergic rhinitis	$y_1 = 17.9 + 15.16x_1$
7	0.28		8. Domestic dust	$y_1 = 9.37 + 7.23x_1$
8	0.30		8. Pillow feather	$y_1 = 16.73 + 10.8x_1$
9	0.37		2. Atopic dermatitis	$y_1 = 23.03 + 49.04x_1$
10	0.40		8. Sheep wool	$y_1 = 15.76 + 17.67x_1$
11	0.45		8. Rabbit hair	$y_1 = 16.94 + 15.26x_1$
12	0.77		1. Severe persistent	$y_1 = 16.63 + 68.04x_1$
1	-0.26	1. Severe persistent	12. CD8 10*3 cells	$y_1 = 0.52 - 0.5x_1$
2	-0.25		12. CD 25%	$y_1 = 0.40 - 0.01x_1$
3	-0.23		12. CD 25 10*3 cells	$y_1 = 0.29 - 0.24x_1$
4	-0.21		10. Ig E	$y_1 = -0.03 + 0.19x_1$
5	0.24		8. Rabbit hair	$y_1 = 0.08 + 0.09x_1$
6	0.25		5. Eosinophils	$y_1 = 0.03 + 0.02x_1$
7	0.32		8. Sheep wool	$y_1 = 0.04 + 0.16x_1$
8	0.32		8. Domestic dust	$y_1 = -0.08 + 0.09x_1$
9	0.27		7. Foxtail	$y_1 = 0.08 + 0.19x_1$
10	0.30		7. Cypress	$y_1 = 0.07 + 0.15x_1$
11	0.30		2. Allergic rhinitis	$y_1 = 0.03 + 0.2x_1$
12	0.32		3. Number of years from the first symptoms	$y_1 = -0.01 + 0.02x_1$
13	0.34		8. Pillow feather	$y_1 = 0.02 + 0.14x_1$
14	0.43		2. Atopic dermatitis	$y_1 = 0.1 + 0.65x_1$
15	0.77		13. TSLP	$y_1 = -0.09 + 0.01x_1$

Table 4. Regressors for which lack of connection with the explained factor is assumed

№	**Regressor,** x_1	№	**Regressor,** x_1
1.	9. Black tea	**9.**	6. Specific gravity
2.	6. Appearance	**10.**	9. Apple
3.	7. Walnut	**11.**	7. Sunflower
4.	12. Total complement (CH 50)	**12.**	6. Transitional epithelial cells
5.	7. Sagebrush	**13.**	12. IgA
6.	12. Lymphocytes, 10*3 cells	**14.**	8. Daphnia
7.	2. Deflection of the nasal septum	**15.**	5. Segmented neutrophils, %
8.	4. Bronchial asthma in mother		

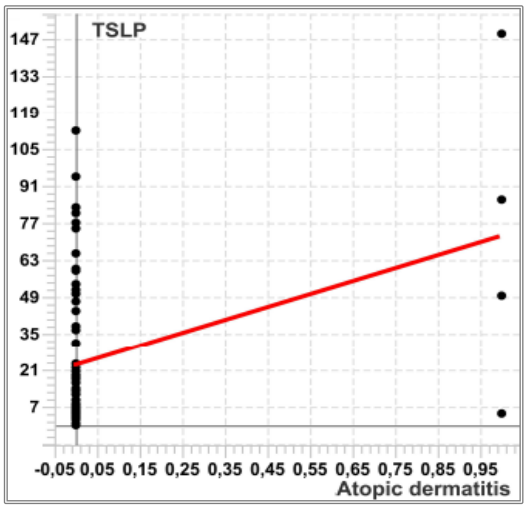

Figure 1. The pair regression model [13. TSLP] and [2. Atopic dermatitis]

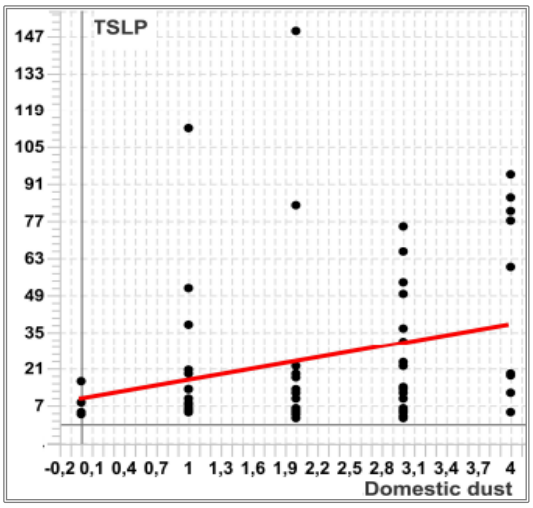

Figure 2. The pair regression model [13. TSLP] and [8. Domestic dust]

In Figure 1, there is a model of matched regression that defines the relationship between the explained factor [13. TSLP] and the regressor [2. Atopic dermatitis]. The straight line defines the relationship between the explained factor and the regressor. The points designate the experimental data, received after clinical research. As it was mentioned above, enhancement of the severity of the disease leads to an increase in TLSP level. In the works of Ziegler and Artis, the relationship between [13. TSLP] factor and [2. Atopic dermatitis] regressor is defined, however, quantitative characteristics are not considered [30]. The real work complements the research conducted in [30], by quantitative relation. Presence of the sign [2. Atopic dermatitis] increases the possibility of bronchial asthma occurrence. This relation is portrayed in Figure 1. The correlation coefficient of the factor [2. Atopic dermatitis] and the explained factor is 0.37 [Table 3] which implies a weak relation. The average value of [13. TSLP] in case of the absence of the sign [2. Atopic dermatitis] is 23 pg/mL and 72 pg/mL ([2. Atopic dermatitis]=1) in case of its presence. Despite the fact that the presence of the sign [2. Atopic dermatitis] leads to an increase of the factor [13. TSLP] by 40, a low number of patients who have this sign allows to create, in the real work, only approximate model of matched regression that confirms the statement about stochastic increase of the factor [13. TSLP] in case of the presence of the sign [2. Atopic dermatitis], seen in [30].

In Figure 2, there is a model of matched regression, defining the relation between the explained factor [13. TSLP] and the regressor [8. Domestic dust]. The correlation coefficient between the explained factor [13. TSLP] and the regressor [8. Domestic dust] is 0,28 [Table 3] that assumes a weak relation. Learning the dependence of the values of factor [13. TSLP] on the value of the regressor [8. Domestic dust], the values of the explained factor [13. TSLP] were determined (7,0-21,0 pg/mL) at which the skin test in the conducted clinical research is considered weakly positive ([8. Domestic dust]=1 or [8. Domestic dust]=2), in case of significant allergic reaction to this allergen ([8.Domestic dust]=3 or [8.Domestic dust]=4) with maximal TSLP index 38,3 pg/mL. Conducted experimental research say about the tendency to increasing of [13. TSLP] depending on the severity of allergic reaction of factor [8. Domestic dust]. Shifting from one degree of severity of the factor [8. Domestic dust] to another leads to increasing of the prognostic rate of [13. TSLP] proportionally to the value of the coefficient of matched regression $b_1 = 7,23$. The absence of the sign is defined by the value of the coefficient $a = 9,37$.

In Figure 3, there is a model of matched regression that defines the dependence of explained factor [13. TSLP] on the regressor [4. Bronchial asthma in relatives of the second generation]. The correlation coefficient between the explained factor [13. TSLP] and the regressor [4. Bronchial asthma in relatives of the second generation] is 0,22 [Table 3] that assumes the presence of weak relation. On the basis of the analysis of clinical research, an assumption is made that the presence of the sign of the regressor for the considered sample leads to the value of the factor [13. TSLP] that is higher than 50 pg/mL. In case of absence of the sign that characterizes this regressor a prognostic value for [13. TSLP] is 21 pg/mL.

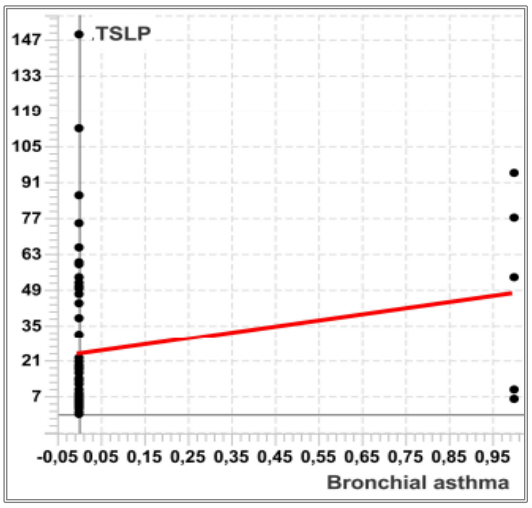

Figure 3. The pair regression model [13. TSLP] and [4. Bronchial asthma in relatives of the second generation]

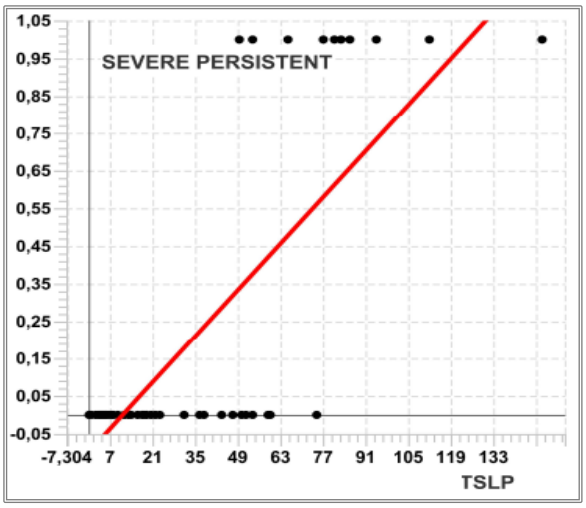

Figure 4. The pair regression model [1. Severe persistent] and [13. TSLP]

In Figure 4, there is a model of matched regression defining the dependence of the explained factor [1. Severe persistent] on the regressor [13. TSLP]. The correlation coefficient between the explained factor and the regressor is 0,73 [Table 3] which implies a strong relation and confirms conclusions made in [16]. The correlation coefficient, in this case, is a few times greater than for regressors mentioned above [2. Atopic dermatitis], [8. Domestic dust], [4. Bronchial asthma in relatives of the second generation], that is an important result of the real work. The presence of a strong relationship allows us to say about the general qualitative tendency of the explained factors depending on the similar regressors. The analysis of the model of a matched regression allows defining a break-off value of the [13. TSLP] factor, at which we can claim about a high possibility of the presence of severe bronchial asthma. Brake-off value of the factor [13. TSLP] is 63 pg/mL. This model of a matched regression can be, at some point, interpreted as a probabilistic model for prognostication of the severity of bronchial asthma, thus the distribution function of random variable determining the value of the [13. TSLP] factor can be roughly written as

$$\begin{cases} F(x) = 0, & [13.TSLP] < 9.0; \\ F(x) = -0.09 + 0.01 \cdot x, & 9.0 \le [13.TSLP] < 109.0; \\ F(x) = 1, & [13.TSLP] \le 109.0. \end{cases}$$

On the basis of the distribution function, $F(x) = 0,$ there was an important assumption made within the clinical experiment, that increase of [13. TSLP] by 10 pg/mL leads to an increase in the possibility of severe bronchial asthma by 10%.

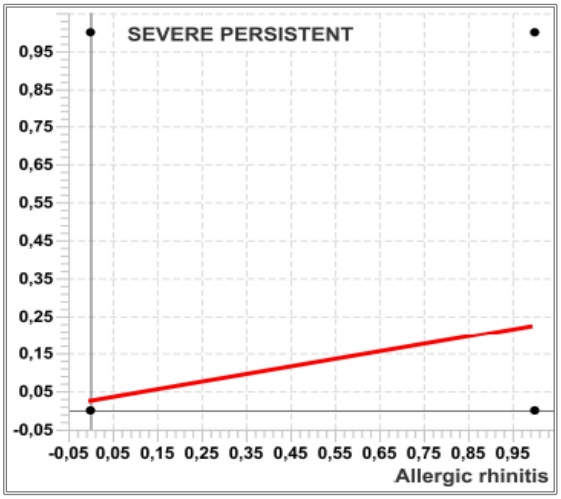

Figure 5. The pair regression model [1. Severe persistent] and [2. Allergic rhinitis]

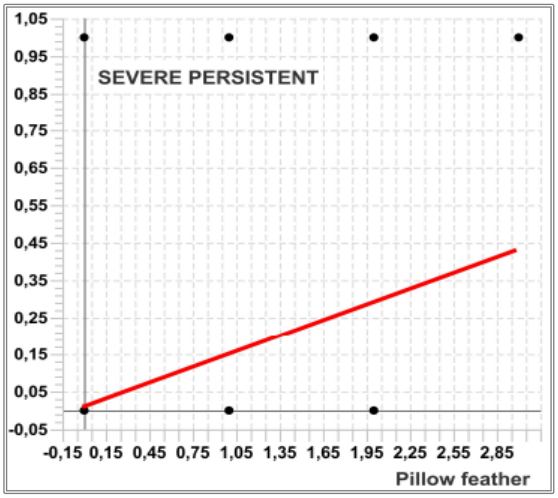

Figure 6. The pair regression model [1. Severe persistent] and [8. Pillow feather]

In Figure 5, there is a model regression that determines the dependence of the explained factor [1. Severe persistent] on the regressor [2. Allergic rhinitis]. The correlation coefficient between the explained factor and the regressor is 0.3 (Table 3), which tells us about a low dependence and confirms conclusions made in [16]. In case of the absence of this factor in the history of the disease (the disease ([2. Allergic rhinitis]=1)) an assumption based on the real work can be made that only 15% of the children may develop severe bronchial asthma. The absence of the factor ([2. Allergic rhinitis]=0) doesn't exclude the possibility of severe bronchial asthma development.

In conclusion, we will analyze the regressive model that determines the dependence of the explained factor [1. Severe persistent] on the regressor [8. Pillow feather]. The regressor is coded with numbers: 0– absence of the sign, 1 and 2 – weakly positive, 3 and 4 – positive. In the conducted research, the regressor accepted values in the interval [0;4]. If the value of the regressor [8. Pillow feather]=3, then the prognostic value for the explained factor is 0.45. It's assumed that if [8. Pillow feather]=4 then the value of the explained factor will probably surpass 0.5.

Conclusion

In this work, the complex of models pair regression for determining the severity of bronchial asthma was presented. The main provisions used for creating the complex of models regression were determined and substantiated. The premises were used, that a random error in the equation of regression is submitted to the normal distribution law, and the values of the factors determined after examination of every patient were received with the same accuracy. It made it possible to elaborate the way of assessment of the risk for the development of severe bronchial asthma in children and identify 15 factors that affect severe bronchial asthma. It was shown that the regressors chosen for the analysis are weakly linked to the explained factors, which makes it possible to determine only qualitative dependence between them. The analysis of factor thymic stromal lymphopoietin [13. TSLP] confirms the conclusions [21-23, 29], that it's the most prominent indicator for the assessment of the severity of bronchial asthma. The correlation coefficient between [1. Severe persistent] and the regressor [13. TSLP] is 0,77 that implies a strong dependence between them. The analysis of the model of matched regression made it possible to identify a border value of the factor [13. TSLP], surpassing of which is considered as a high possibility for the presence of severe bronchial asthma. The increase of the factor [13. TSLP] by 10 pg/mL is considered as 10% higher possibility for severe bronchial asthma.

These conclusions made it possible to formulate the perspectives for further studies. These studies are:

a) construction and analysis of the model with two regressors that allows clarifying the results of the research;

b) verification of the assumption that the values of the factors determined after examination of every patient have the same accuracy.

Conflict of interest

The authors declare no conflicts of interest.

Funding

There was no funding received for this article.

References

1. The global asthma report 2018. Auckland: The Global Asthma Network; 2018 [cited 2019 Sep 17]. Available from: www.globalasthmanetwork.org.
2. Winer RA, Qin X, Harrington T, Moorman J, Zahran H. Asthma incidence among children and adults: Findings from the Behavioral Risk Factor Surveillance System Asthma Call-back Survey—United States, 2006–2008. J Asthma 2012;49:16-22.
3. Beasley R, Semprini A, Mitchell EA. Risk factors for asthma: is prevention possible? Lancet 2015;386(9998):1075-85.
4. Fleming L, Murray C, Bansal A, Hashimoto S, Bisgaard H, Bush A, et al. The burden of severe asthma in childhood and adolescence: results from the paediatric U-BIOPRED cohorts. Eur Respir J 2015;46:1322-33.
5. Chung KF, Wenzel SE, Brozek JL, Bush A, Castro M, Sterk PJ, et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. Eur Respir J 2013;43:343-73.
6. Foster JM, McDonald VM, Guo M, Reddel HK. "I have lost in every facet of my life": the hidden burden of severe asthma. Eur Respir J 2017;50:1700765.
7. Fajt ML, Wenzel SE. Asthma phenotypes and the use of biologic medications in asthma and allergic disease: The next steps toward personalized care. J Allergy Clin Immunol 2015;135:299-310.
8. Fitzpatrick AM, Teague WG, Meyers DA, Peters SP, Li X, Li H, et al. Heterogeneity of severe asthma in childhood: Confirmation by cluster analysis of children in the National Institutes of Health/National Heart, Lung, and Blood Institute Severe Asthma Research Program. J Allergy Clin Immunol 2011;127:382-9.e13.
9. Just J, Gouvis-Echraghi R, Rouve S, Wanin S, Moreau D, Annesi-Maesano I. Two novel, severe asthma phenotypes identified during childhood using a clustering approach. Eur Respir J 2012;40:55-60.
10. Konradsen JR, Nordlund B, Lidegran M, Pedroletti C, Grönlund H, van Hage M, et al. Problematic severe asthma: A proposed approach to identifying children who are severely resistant to therapy. Pediatr Allergy Immunol 2010;22(1-Part-I):9-18.

11. Allakhverdi Z, Comeau MR, Jessup HK, Yoon BR, Brewer A, Chartier S, et al. Thymic stromal lymphopoietin is released by human epithelial cells in response to microbes, trauma, or inflammation and potently activates mast cells. J Exp Med 2007;204:253-8.
12. Reche PA, Soumelis V, Gorman DM, Clifford T, Liu Mr, Travis M, et al. Human thymic stromal lymphopoietin preferentially stimulates myeloid cells. J Immunol 2001;167:336-43.
13. Wiener N. The Extrapolation, Interpolation, and Smoothing of Stationary Time Series. Willey; 1949.
14. Kozhyna OS, Pihnastyi OM. Statistical data of a clinical study of the severity of bronchial asthma in children of the Kharkov region, 2017. Mendeley Data 2019;2.
15. Kozhyna OS, Pihnastyi OM. Data structure of clinical research. Human Health & Disease 2019;3:71- 9.
16. Kozhyna OS, Pihnastyi OM. Covariance coefficients factors from a clinical study of the severity of bronchial asthma in children of the Kharkov region, 2017. Mendeley Data 2019;1.
17. Braken MB, Belanger K, Kukson VO, Trichet E, Kristiani DS, Lederer BP. Genetic and perinatal risk factors for the occurrence and severity of asthma: a review and theoretical analysis. Epidemiol Rev 2002;24:176–89.
18. Bakopulu F, Veltsista A, Vassi I, Gika A, Lekea V, Priftis K, et al. Can we look at asthma in childhood with optimism? Greek cohort study. J Asthma 2009;46:171–4.
19. Bousquet J, Mantzouranis E, Cruz AA. Uniform definition of asthma severity, control, and exacerbations: document presented for the World Health Organization Consultation on severe asthma. J Allergy Clin Immunol 2010;126:926–38.
20. Campo P, Rodríguez F, Sánchez-García S, Barranco P, Quirce S, Pérez-Francés C, et al. Phenotypes and endotypes of uncontrolled severe asthma: new treatments. J Investig Allergol Clin Immunol 2013;23:76-88.
21. Chauhan A, Singh M, Agarwal A, Paul N. Correlation of TSLP, IL-33, and CD4 + CD25 + FOXP3 + T regulatory (Treg) in pediatric asthma. J Asthma 2015;52:868-72.
22. Ying S, O'Connor B, Ratoff J, Meng Q, Fang C, Cousins D, et al. Expression and cellular provenance of thymic stromal lymphopoietin and chemokines in patients with severe asthma and chronic obstructive pulmonary disease. J Immunol 2008;181:2790-98.
23. Ying S, O'Connor B, Ratoff J, Meng Q, Mallett K, Cousins D, et al. Thymic stromal lymphopoietin expression is increased in asthmatic airways and correlates with expression of Th2-attracting chemokines and disease severity. J Immunol 2005;174:8183–90.
24. Fornasa G, Tsilingiri K, Caprioli F, Botti F, Mapelli M, Meller S, et al. Dichotomy of short and long thymic stromal lymphopoietin isoforms in inflammatory disorders of the bowel and skin. J Allergy Clin Immunol 2015;136:413–22.
25. Shikotra A, Choy DF, Ohri CM, Doran E, Butler C, Hargadon B, et al. Increased expression of immunoreactive thymic stromal lymphopoietin in patients with severe asthma. J Allergy Clin Immunol 2012;129:104–11.e9.
26. Bunyavanich S, Melen E, Wilk JB, Granada M, Soto-Quiros ME, Avila L, et al. Thymic stromal lymphopoietin (TSLP) is associated with allergic rhinitis in children with asthma. Clin Mol Allergy 2011;9(1).
27. Harada M, Hirota T, Jodo AI, Doi S, Kameda M, Fujita K, et al. Functional analysis of the thymic stromal lymphopoietin variants in human bronchial epithelial cells. Am J Respir Cell Mol Biol 2009;40:368–74.
28. Soumelis V, Reche PA, Kanzler H, Yuan W, Edward G, Homey B, et al. Human epithelial cells trigger dendritic cell mediated allergic inflammation by producing TSLP. Nat Immunol 2002;3:673–80.
29. Lee EB, Kim KW, Hong JY, Jee HM, Sohn MH, Kim KE. Increased serum thymic stromal lymphopoietin in children with atopic dermatitis. Pediatr Allergy Immunol 2010;2:457-60.
30. Ziegler SF, Artis D. Sensing the outside world: TSLP regulates barrier immunity. Nat Immunol 2010;11:289-93.