Information Technology and Management Science
doi: 10.1515/itms-2014-0001
2014 / 17

Decision Support Models and Software for the Differential Immunophenotype Diagnostics of Leukosis and Lymphomas

Natalia Novoselova [1], Igor Tom [2], Michael Belevtsev [3],
[1, 2] *United Institute of Informatics Problems*, [3] *Belarusian Research Center for Pediatric Oncology, Hematology and Immunology*

***Abstract* – This article describes the software and underlined decision support models for the immunophenotype diagnostics of leukosis (leukemia) and lymphomas adjusted for the marker or human leukocyte antigen (CD-antigen) coexpressions. Using the model knowledge base, the decision inference algorithm allows computing the degree of manifestation of the disease subtypes for the input immunophenotype features. Software provides the two-stage diagnostics of the leukemia subtypes and the lymphoma diagnostics using the set of the developed rules, possibility to observe the diagnostic results and corresponding reference information. The patient data are organized according to the unified registration card, which provides the possibility to work at the different diagnostic levels: diagnostics of the extended groups of leukosis, diagnostics of the leukemia subtypes, diagnostics of the adult and child lymphomas.**

***Keywords* – Diagnostic rules, decision support, immunophenotyping, modeling.**

I. Introduction

Acute leukosis and lymphomas are among the most widespread oncological diseases. In the childhood, these types of the oncological disease take the first place with a frequency of 4.3 cases for the population of 100,000 [1]. In Belarus, the survival rate for these diseases comprises more than 80%, which is comparable with the best world practice, but the situation with the adult patients is slightly worse. The general approach to improve the survival rate is the treatment personalization, based on the enhancement of the diagnostic accuracy.

The basic diagnostic method for these types of disease is the identification of the differentiation stage and maturity of the tumor cells using the immune marker analysis, based on the specified combination of the monoclonal antibodies, which recognize the CD-antigens on the cell surface. At present, more than 30 leukemia and lymphoma subtypes are described with no less than 40 markers, which in different combinations (presence/absence) are used for the diagnostics. The standard marker combination is presented only in less than a half number of cases; the rest is characterized by the presence of the coexpressions, nonrelevant to the particular disease and the absence of some core markers. The diagnostic accuracy depends on the knowledge of all the available marker combinations and the possible coexpressions, which require the higher level of training and deep background in the domain. Therefore, to aid in the diagnostic process we have developed the decision support models and the corresponding software for the differential diagnostics of the leukemia and lymphoma subtypes adjusted for the marker coexpressions. Using the automation of the data processing, the proposed software system helps to differentiate the disease subtypes and thereafter to raise the quality and diagnostic timeliness for the adequate treatment. The first attempts in this direction were made for the childhood leukemia on the basis of the standard combinations of marker expressions with a subsequent identification of coexpressions [1]. In the proposed software system, the decision-making is implemented on the basis of the diagnostic rules, extracted from the database and presented as the weighted marker combinations and the specialized decision inference algorithm for the identification of the degree of manifestation for each disease subtype. The treatment personalization is provided by more accurate diagnostic rules, accounting for marker coexpressions. The application of the software system for the personal training will allow raising the educational level of the specialists dealing with differential diagnostics and reducing the training costs.

II. Decision Support Models for the Differential Immunophenotype Diagnostics of Leukosis and Lymphomas

The decision support models [2] consist of the following components (Fig. 1):

- Input data module;
- Knowledge base module;
- Decision-making module;
- Output data module.

Input data module provides the following functions:

- inputting and storing the values of the immunophenotype features for the diagnostic cases;
- input data processing necessary for the implementaton of the differential diagnostics. The processing includes discretization of the expression values of the markers according to the specified thresholds. The markers with the values greater than the threshold are considered to be expressed and their new value is equal to 1, otherwise the markers are latent and the value is equal to 0.

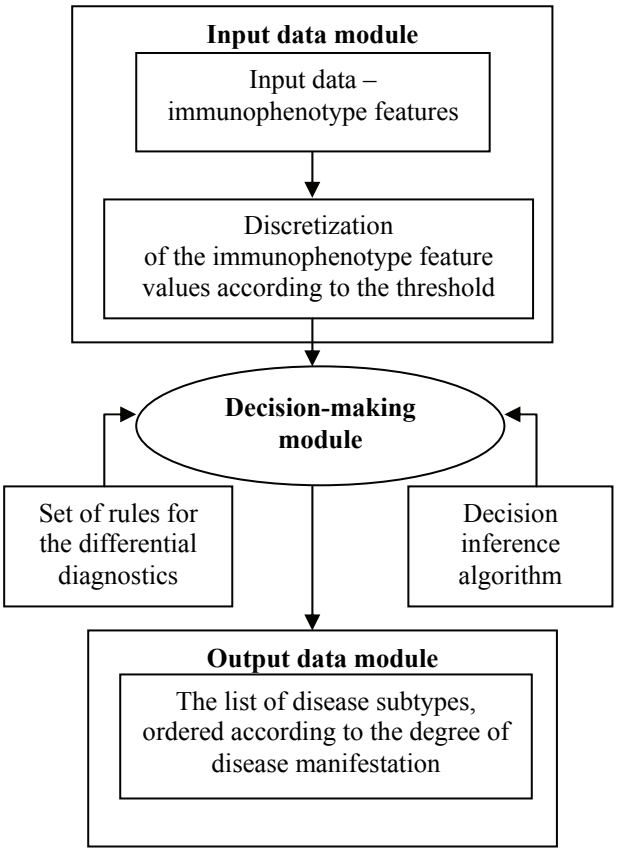

Fig. 1. The general scheme of the decision support model.

Decision-making module is the main part of the model allowing one to conduct the disease diagnostics on the basis of the available knowledge base and decision inference algorithm.

Knowledge base [3] includes the standard rules and rules developed for all the individual disease subtypes. The developed rules are based on the statistical analysis of the immunophenotype features using the cross-tabulation method.

Decision inference algorithm allows computing the degree of manifestation of the disease subtypes for the input diagnostic case using the model knowledge base.

Output data module provides the following functions:

– output data conversion necessary for the formation of the list of disease subtype ordered according to the degree of manifestation.

– output of the list of the disease subtypes with corresponding values of their manifestation.

The models for the diagnostics of leukemia and lymphomas have the identical structure with the differing knowledge base content and different organization of the decision-making module.

The model output list allows estimating the degree of manifestation for each individual disease subtype, where the subtype at the top of the list presents the most probable diagnosis.

III. Knowledge Base and Decision-Making Module for the Leukosis Diagnostics

The general scheme of the decision-making module is presented in Fig. 2. According to Fig. 2, the two-stage diagnostics is proposed, where at the first stage the biphenotype (BAL) cases are revealed and the most probable extended leukemia group is determined (B-cell Acute Lymphoblastic Leukemia (B-ALL), T-cell Acute Lymphoblastic Leukemia (T-ALL), Acute Myeloid Leukemia (AML)).

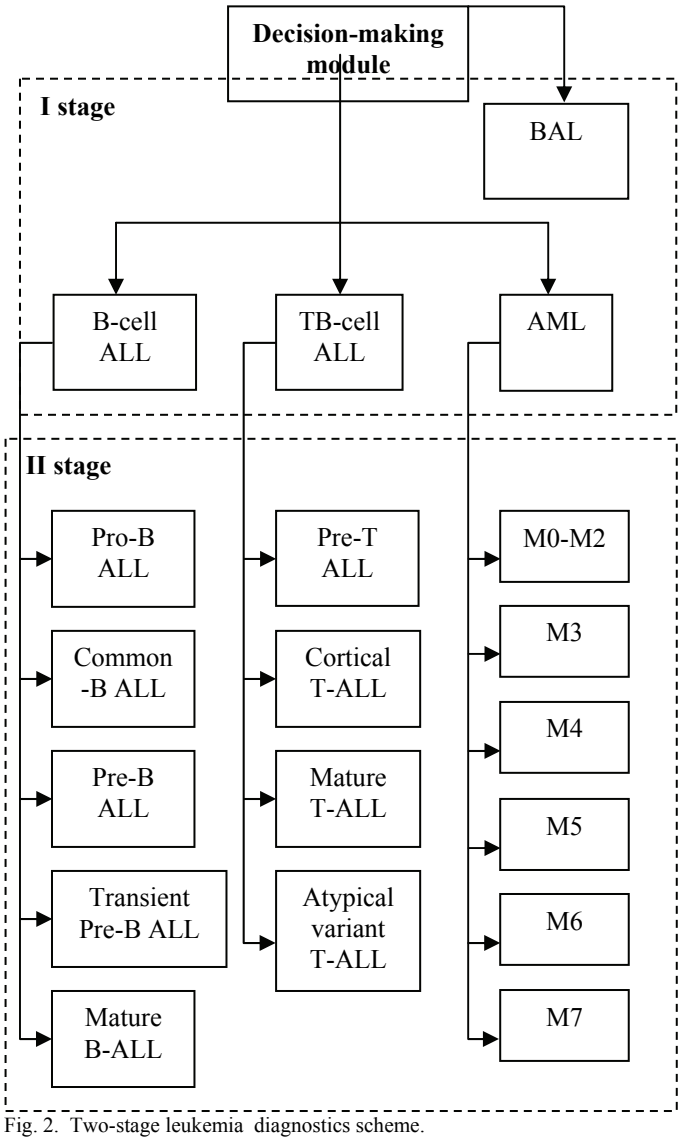

Fig. 2. Two-stage leukemia diagnostics scheme.

At the second stage, the specific disease subtype is identified for the determined general group, i. e., the in-depth analysis is pursued. The content of the model knowledge base is presented in Table I.

The set of rules, which includes the number of markers (MPO, CD19, CD3 etc.) according to WHO classification [4–5] for the biphenotype diagnostics, is the following:

Rule 1: If ((**MPO=1)** and **(**SUM3>=2)) and ((**CD19=1)** and **(**SUM1>=2)) then «Biphenotype B-ALL + AML».

Rule 2: If ((**MPO=1)** and **(**SUM3>=2)) and ((**s/cyCD3=1)** and **(**SUM2>=2)) then «Biphenotype T-ALL + AML».

Rule 3: If ((**CD19=1)** and **(**SUM1>=2)) and ((**s/cyCD3=1)** and **(**SUM2>=2)) then «Biphenotype T-ALL + B-ALL»,

where

SUM1= **SUM(**CD10, s/cyCD22, cyCD79a)

SUM2= **SUM(**CD4, CD8, CD5, CD1a, CD2, TCRα/β, TCRγ/δ)

SUM3= **SUM(**CD13, CD33, CD117, CD15)

The diagnostic rules for the extended groups of leukosis and the corresponding subtypes are presented as the set of weights and importance factors for each marker, which define the profile of the corresponding group or subtype. Weights are received as a result of the statistical analysis of the contingency tables [6]–[7], constructed for the two nonmetric features. In our case, the two features are immunophenotype feature and disease group or subtype. The example of the weights for the extended leukemia groups is presented in Table II. The weight values indicate the influence of the marker for the diagnostics of the particular leukemia group.

TABLE I

THE LIST OF DIAGNOSTIC RULES FOR THE LEUKEMIA

Rules	Specification
Identification of the presence of biphenotype	Identification according to EGIL
	Identification according to modified WHO classification
Identification of the group of leukemia : B-ALL , T-ALL , AML [8–10]	
Identification of the leukemia subtype for B-ALL	
Identification of the leukemia subtype for T-ALL	
Identification of the leukemia subtype for AML	

TABLE II

MARKER WEIGHTS FOR DIFFERENT LEUKEMIA GROUPS (GROUP PROFILES)

CD-antigen	**B-ALL**	**T-ALL**	**AML**
CD2	0.04298	0.851064	0.262821
CD3	0.008671	0.404255	0.050633
CD5	0.028329	0.979167	0.080537
CD7	0.022663	1	0.457317
CD4p_CD8m	0.002833	0.086957	0.268293
CD4m_CD8p	0.008547	0.5	0.037037
CD4p_CD8p	0	0.434783	0.006289
CD1a	0.023881	0.543478	0.103093
tdt	0.933824	0.717949	0.105263
CD19	0.997199	0.021739	0.078313
CD20	0.384181	0	0.104167
CD22	0.97191	0	0.022556
CD10	0.91573	0.234043	0.124224
cy_CD79a	0.968858	0.357143	020339
s_lgm	0.017442	0	0
cy_lgm	0.151408	0.02381	0.08547
hla_dr	0.988372	0.068182	0.64557
CD13p	0.310734	0.104167	0.825301
CD33p	0.194842	0.106383	0.939759
CD117p	0.023121	0.136364	0.882353
CD14p	0.058824	0	0.188889
CD15p	0.109422	0.02439	0.590062
MPOp	0.143885	0175	0.772414
CD11bp	0.760486	0.139373	0.76392
CD11cp	0.189201	1.255625	1.011119
CD34	0.735955	0.404255	0.533333
CD45p_CD14m	0.715543	0.955556	0.968553

IV. KNOWLEDGE BASE AND DECISION-MAKING MODULE FOR THE LYMPHOMA DIAGNOSTICS

The general scheme of the decision-making module is presented in Fig. 3. According to Fig. 3, the diagnostics is performed separately for children and adults, where for each group the different set of lymphoma subtypes is considered.

The set of diagnostic rules, which comprise the model knowledge base, is presented in Table III.

The same as for leukosis, the diagnostic rules are the set of

weights and importance factors for each particular marker, which define the profile of the corresponding lymphoma subtype.

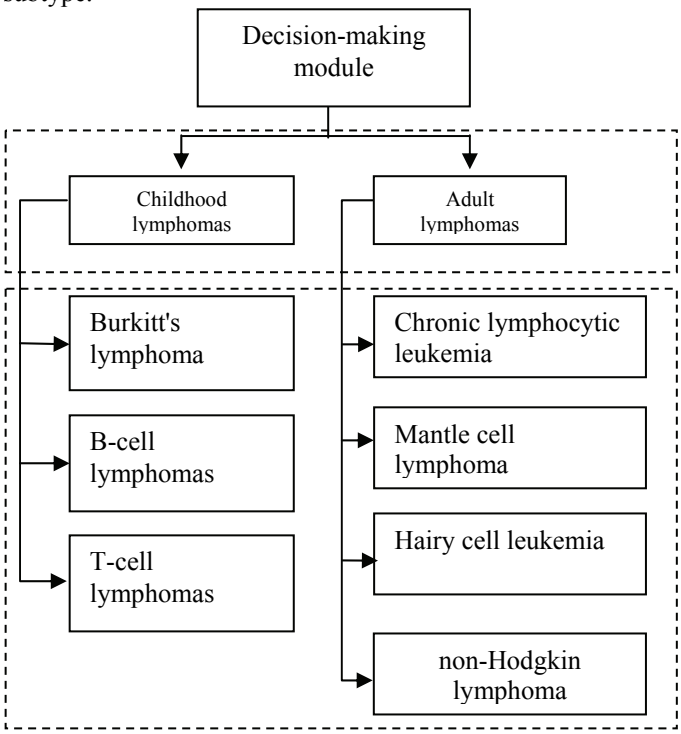

Fig. 3. Lymphoma diagnostics scheme.

TABLE III

THE LIST OF DIAGNOSTIC RULES FOR THE LYMPHOMAS

Rules	Specification
Identification of the lymphoma subtype for children	Subtypes: Burkitt's lymphoma, B-cell lymphomas (Pro-B, common-B, Pre-B), T-cell lymphomas (Pro-T, Pre-T, cortical-T).
Identification of the lymphoma subtype for adults	Subtypes: chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL), hairy cell leukemia (HCL), Non-Hodgkin lymphoma (NHL).

The example of the weights for the adult lymphomas is presented in Table IV.

TABLE IV

MARKER WEIGHTS FOR ADULT LYMPHOMA GROUP

CD-antigen	CLL	MCL	HCL	NHL
CD19+	1.00	1.00	0.47	1.00
FMC7	0.05	0.95	1.00	0.89
CD23	1.00	0.05	0.00	0.22
CD5	1.00	1.00	0.00	0.07
CD20	0.65	1.00	1.00	0.96
CD43	0.99	0.11	0.00	0.15
CD24	1.00	0.95	0.06	0.89
CD79b	0.30	1.00	1.00	0.81
CD11c	0.00	1.00	1.00	0.67
CD25	0.00	0.00	1.00	0.00
CD103	0.00	0.00	1.00	0.00
kappa	0.37	0.39	0.42	0.67
lambda	0.26	0.63	0.58	0.31

V. DECISION INFERENCE ALGORITHM FOR THE DIAGNOSTICS OF LEUKOSIS AND LYMPHOMAS

The diagnostic decision is made using the specially developed algorithm. The model knowledge base composed of the set of rules and the binary (discretized marker values) profile of the diagnostic cases is used as the inputs to the algorithm. According to the algorithm, each disease subtype is assigned the value of the degree of manifestation *PF*. The degree value reflects the correspondence between the disease subtype profile and the marker profile of the patient $x = (x_1, \quad , x_N)$. To calculate the *PF*, the following expression is used:

$$PF = \frac{\sum C_n}{N}. \tag{1}$$

where *PF* is the degree value for the particular disease subtype, C_n stands for coefficients of the individual markers for the particular disease subtype ($0 \le C_n \le 1$), $n = \overline{1, N}$, N is the number of the markers with the nonempty values. The coefficients C_n are calculated using the marker weights ω_{ij} and importance factors k_{ij} as follows:

$$C_n^i = \begin{cases} k_{ij} \cdot \omega_{ij}, & \text{если} \quad x_j = 1 \\ k_{ij} \cdot (1 - \omega_{ij}), & \text{если} \quad x_j = 0 \end{cases}. \tag{2}$$

where C_n^i stands for marker coefficients for the *i*th disease subtype, ω_{ij} is weight of the *j*th marker for the *i*th disease subtype, k_{ij} is importance values for the *j*th marker for the *i*th disease subtype, x_j is the discretized value of the *j*th marker.

The importance of the individual specific markers for the diagnostic of the particular disease subtype is considered with the importance factor k, e.g., if $k = 2$ the value of the marker coefficient C_n is doubled.

For the differential diagnostics, all the values of the immunphenotype features, which characterize the patient, are considered. The decision is also made when not all the marker values are present. However, in the absence of the information for the most specific markers the diagnostics will be inaccurate. The diagnostic result is the list of the disease subtype with the corresponding degrees of manifestation $PF_1, PF_2, \ldots, PF_M$. Finally, the degree values are divided by the maximal value to receive the result in the interval [0.1]:

$$PF_m = \frac{PF_m}{\max(PF_1, PF_2, \ldots, PF_M)}, \quad m = \overline{1, M}. \tag{3}$$

The example of the decision-making process for the B-ALL diagnostic case with the input values in Table V is considered. The first stage of the leukemia diagnostics is performed, i. e., the recognition of the extended groups: B-ALL, T-ALL, AML. According to Table V, 17 markers are negative (expressed on less than 20% cells), 8 – positive and for 2 markers the values are absent.

To calculate the value of the degree of manifestation *PF*

consider the weights for the B-ALL profile in Table II. The importance factors of the markers are presented in Table VI.

TABLE V

THE MARKER PROFILE FOR THE CONSIDERED EXAMPLE

No.	Marker name	Value 1 – positive, 0 –negative, −1 – absent value	No.	Marker name	Value 1 – positive, 0 – negative, −1– absent value
1	CD2	0	15	s_lgm	0
2	CD3	0	16	cy_lgm	0
3	CD5	0	17	hla_dr	1
4	CD7	0	18	CD34	0
5	CD4p_CD8m	0	19	CD45p_CD14m	1
6	CD4m_CD8p	0	20	CD13p	0
7	CD4p_CD8p	0	21	CD33p	0
8	CD1a	0	22	CD117p	0
9	tdt	1	23	MPOp	0
10	CD19	1	24	CD15p	0
11	CD20	1	25	CD11bp	−1
12	CD22	1	26	CD11cp	0
13	CD10	1	27	CD14p	−1
14	cy_CD79a	1			

The value of *PF* consists of two summands: C_1 – the sum of marker coefficients with a positive input value, C_2 – the sum of marker coefficients with a negative input value. C_1 and C_2 are calculated subject to the marker importance factors. For the considered example, $C_1 = 6.87$ and $C_2 = 14.98$.

TABLE VI

IMPORTANCE COEFFICIENTS FOR THE B-ALL PROFILE

No.	Marker name	Importance	No.	Marker name	Importance
1	CD2	1.00	15	s_lgm	1.00
2	CD3	1.00	16	cy_lgm	1.00
3	CD5	1.00	17	hla_dr	2.00
4	CD7	1.00	18	CD34	1.00
5	CD4p_CD8m	1.00	19	CD45p_CD14m	1.00
6	CD4m_CD8p	1.00	20	CD13p	1.00
7	CD4p_CD8p	1.00	21	CD33p	1.00
8	CD1a	1.00	22	CD117p	1.00
9	tdt	2.00	23	MPOp	1.00
10	CD19	2.00	24	CD15p	1.00
11	CD20	1.00	25	CD11bp	1.00
12	CD22	2.00	26	CD11cp	1.00
13	CD10	2.00	27	CD14p	1.00
14	cy_CD79a	2.00			

Hence $PF_{B\text{-}ALL} = \dfrac{C_1 + C_2}{25} = 0{,}874$. The values of $PF_{T\text{-}ALL} = 0{,}493$ and $PF_{AML} = 0{,}494$ are calculated just the same. The diagnosis is made in accordance with the maximal value $PF^* = \max\{PFB_{B\text{-}ALL}, PF_{T\text{-}ALL}, PF_{AML}\}$ and corresponds to the B-ALL leukemia group. The diagnosis predicted with the model is equal to the true diagnosis.

VI. SOFTWARE FOR THE DIAGNOSTICS OF LEUKEMIA S AND LYMPHOMAS

The proposed models and decision inference algorithm are implemented as software in Visual Studio environment [11]. The software interface for the data input and editing is created dynamically on the basis of the internal electronic presentation of the patient registration card. The process of the decision-making is performed using the electronic diagnostic rules, which are stored in the XML files [12]. The main functions, which are implemented, are the following:

− registration and storage of the results of the immune marker analysis on the basis of the XML template of the patient registration card;

− storage of the diagnostic rules for the leukemia and lymphoma subtypes on the basis of the XML templates;

− editing and checking the filling completeness of the patient registration card;

− performing the disease diagnostics according to the developed decision-making models for leukemia and lymphomas;

− adjustment of the disease diagnostics level: leukemia or lymphoma subtypes;

− delivery of the diagnostic results as an ordered list of the disease subtypes with the corresponding degrees of manifestation;

− delivery of the information according to the selected diagnosis and the list of the corresponding coexpressions;

− observing and storing the patient registration card;

− getting the reference information for the leukemia and lymphoma subtypes.

Figure 4 presents the screenshot with the initial patient data and opened control tab – B-ALL marker type.

Software allows making the adjustment of the diagnostic level to select the disease type for further decision-making process (Fig. 5). After the selection of the certain disease type, the corresponding diagnostic rules are loaded into the computer program. The results of the two-stage leukemia diagnostics are depicted in Fig. 6, where the initial patient data with known diagnosis Common B-ALL are used.

Fig. 7 presents the diagnostic results for the adult patient with hairy cell leukemia.

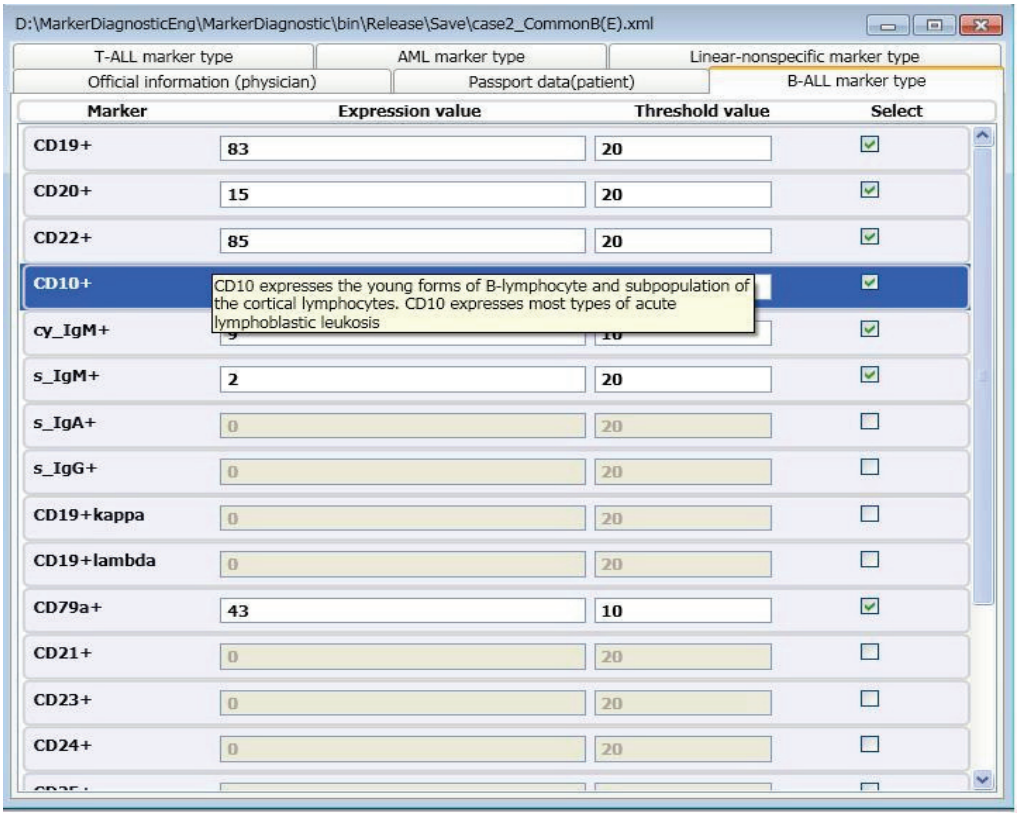

Fig. 4. Patient input data with the marker expression values.

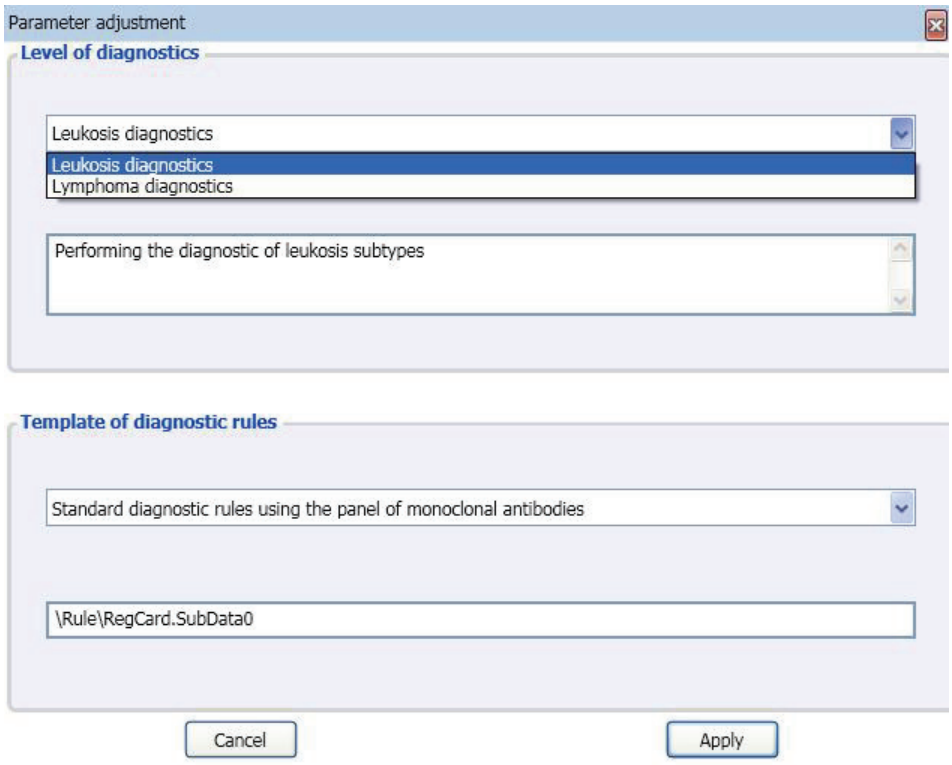

Fig. 5. Selection of the disease type for diagnostics.

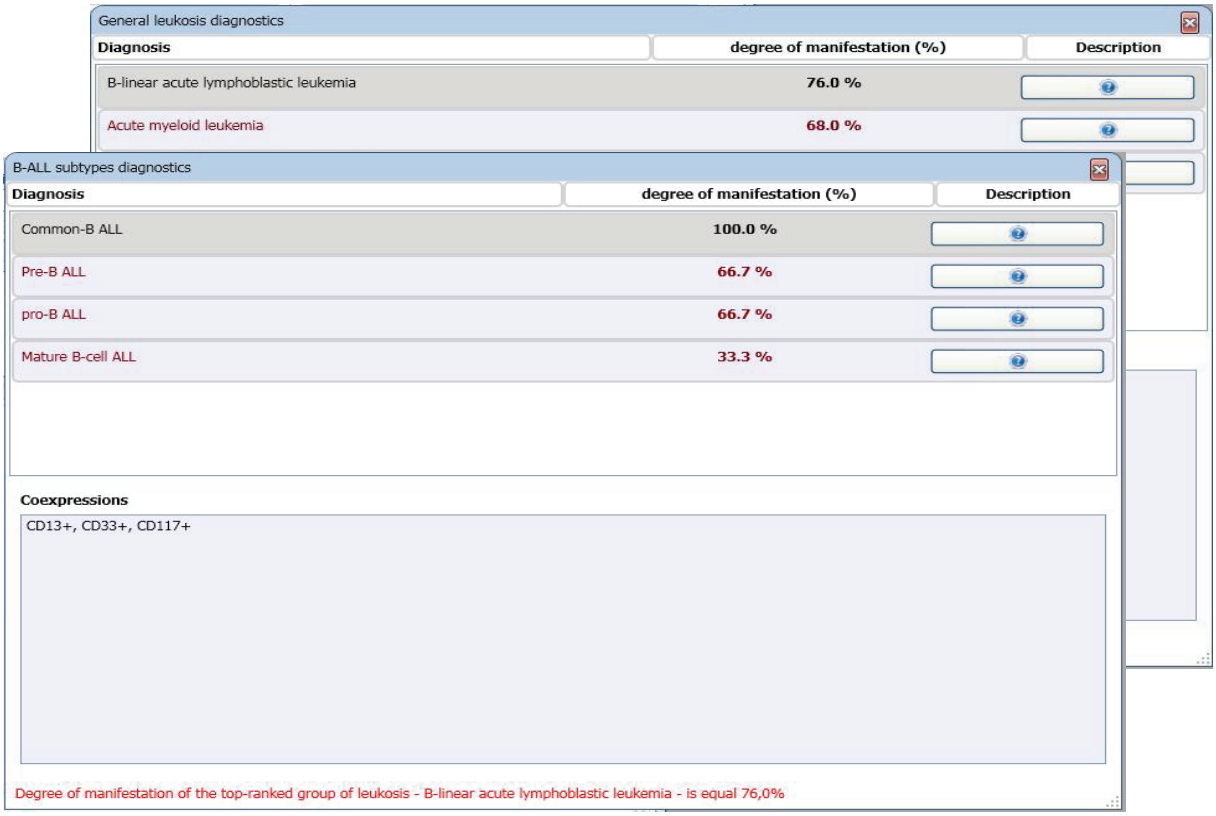

Fig. 6. Results of the two-stage leukemia diagnostics.

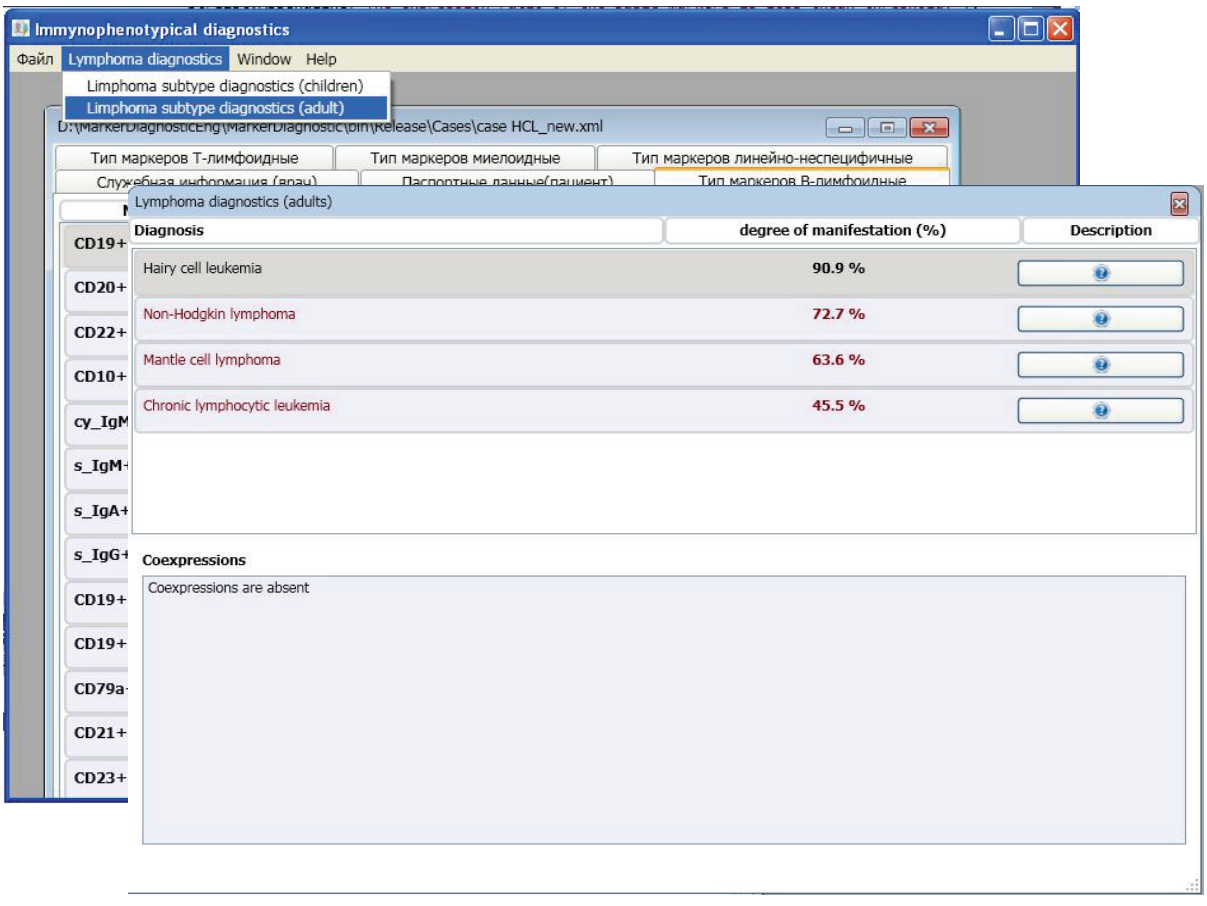

Fig. 7. Diagnostic results for the adult patient.

The constituent part of the software system is the electronic help on the basic types of leukemia and lymphoma. The help system allows dynamically browsing the information for each individual leukemia or lymphoma subtype

VII. Verification of the Decision Support Models

Verification of the decision support model for the diagnostics of the lymphoma subtypes in children was performed using the input dataset with 79 patients, characterized by the expression values of the 31 markers. The same dataset was used for the identification of the individual marker weights for the disease subtypes, i.e., disease subtype profiles, which are used as the diagnostic rules for the disease prediction.

Diagnostic (classification) accuracy for the dataset was equal to 97.5%. Only one data case with Burkitt's lymphoma and one data case with B-cell lymphomas were incorrectly predicted. The classification results are presented in Table VII.

The classification accuracy with the standard combination of markers was equal to 96%.

TABLE VII

Classification Results for the Child Dataset with Lymphomas

Lymphoma subtype		True values		
		Burkitt's lymphoma	T-cell lymphomas	B-cell lymphomas
Prediction	Burkitt's lymphoma	55	0	1
	T-cell lymphomas	0	19	0
	B-cell lymphomas	1	0	3

Verification of the decision support model for the diagnostics of the lymphoma subtypes in children was performed using the input dataset with 180 patients, characterized by the expression values of the 13 markers. Diagnostic accuracy for the dataset was equal to 97.2%. Only two data cases with NHL and 3 cases with CLL were incorrectly classified. The classification results are presented in Table VIII.

The classification accuracy with the standard combination of markers was equal to 94%.

TABLE VIII

Classification Results for the Adult Dataset with Lymphomas

Lymphoma subtype		True values			
		CLL	**MCL**	**HCL**	**NHL**
Prediction	**CLL**	114	0	0	3
	MCL	0	19	0	0
	HCL	0	0	17	0
	NHL	1	1	0	25

Verification of the decision support model for the diagnostics of the leukemia subtypes was performed using the input dataset with 577 patients, characterized by the expression values of the 27 markers. At the first stage of diagnostics, 7 data cases were revealed as biphenotypes, among them 3 cases were attributed to B-ALL+T-ALL biphenotype and 4 cases to B-ALL+AML biphenotype. The description of the binary marker profiles of the detected cases with biphenotypes are depicted in Table IX. For these cases the markers, which are specific to several types of leukemia, are simultaneously expressed (e.g., for the first case the T-ALL markers – CD2, CD3, CD5, CD7 and B-ALL markers – CD19, CD20, CD22, CD79).

TABLE IX

Binary Marker Profiles for Cases with Biphenotypes (Bif_BT Corresponds to B-ALL+T-ALL, Bif_BM Corresponds to B-ALL+AML)

	Bif_BT	Bif_BT	Bif_BT	Bif_BM	Bif_BM	Bif_BM	Bif_BM
CD2	1	1	1	0	0	0	0
CD3	1	1	1	0	0	0	0
CD5	1	1	1	0	0	0	0
CD7	1	1	1	0	0	0	0
CD4p_CD8m	0	0	0	0	0	0	0
CD4m_CD8p	0	0	0	0	0	0	0
CD4p_CD8p	0	0	0	0	0	0	0
CD1a		0	0	0	0	0	0
tdt	1			1	1	1	1
CD19	1	1	1	1	1	1	1
CD20	1	1	1	1	0	0	0
CD22	1	1	1	1	1	1	1
CD10		1	1	1	1	1	1
cy_CD79a	1			1	1		1
s_lgm	0	0	0	0	0	0	0
cy_lgm	1			0	0	0	0
hla_dr	1	1	1	1	1	1	1
CD13p	0	0	1	1	1	1	1
CD33p	0	0	0	1	1	1	0
CD117p	0	0	0	0	0	0	1
CD14p							
CD15p		0	0	0	0	0	0
MPOp	0			1	1	1	1
CD11bp				0		0	
CD11cp		0	0				
CD34	1	1	0	0	1	0	1
CD45p_CD14m		0	1	1	1	1	1

At the first stage of diagnostics, the diagnostic accuracy was 98.8%. Four AML cases and 3 T-ALL cases were incorrectly classified. The classification results are presented in Table X. The classification accuracy with the standard combination of markers was equal to 98.2%.

TABLE X

Classification Results for the Leukemia Cases

Acute leukemia group		True values		
		B-ALL	T-ALL	AML
Prediction	B-ALL	359	2	1
	T-ALL	0	45	3
	AML	0	1	163

At the second stage of diagnostics, the diagnostic accuracy was 88.3 % for B-ALL subtypes, 82.6 % for T-ALL subtypes. The most number of errors was due to the classification of the AML subtypes (classification accuracy ~70%). Such a result can be explained by the great number of common antigens, expressed on the cell surface of the myeloid and monocytic lines; therefore, the final diagnosis must take into account the results of morphological and cytochemical analysis.

VIII. Conclusion

We have developed and implemented as software the decision support models for the immunophenotypical diagnostics of leukosis and lymphomas. These models allow performing the diagnostics using both the standard marker combinations and the diagnostic rules constructed as a result of the analysis of the real dataset stored at the State Institution "Belarusian Research Center for Pediatric Oncology, Hematology and Immunology".

According to the verification on the real datasets of patients with leukemia and lymphomas, the use of the generated diagnostic rules gives slightly better classification accuracy in comparison with the standard combinations.

The software is assigned to support the decision-making for the differential diagnostics of leukemia and lymphomas on the basis of immune marker analysis. The system provides the possibility to register, edit and store the patient data in the XML-format.

Applying the software will allow for more accurate disease diagnostics on the basis of the personalization of the diagnostic process, refining the diagnostics using the specialized decision support models adjusted for marker coexpressions. The software can also be considered an educational tool for training the specialists of the clinical-diagnostic medical units in order to raise their professional level.

References

[1] Visualization technologies in medicine. Diagnostics of the acute leukosis and the tumors of the central nervous system in children / O. Aleynikova etc. Minsk: UIIP NAS Belarus, 2007, 244 p.

[2] N. F. Matsatsinis, Y. Siskos. *Intelligent support systems for marketing decisions*, Kluwer Academic Publishers, 2002. http://dx.doi.org/10.1007/978-1-4615-1147-2

[3] S. Krishna. *Introduction to Database and Knowledge-base Systems*. World Scientific Publishing, 1992, p. 328. ISBN 981-02-0619-4. http://dx.doi.org/10.1142/1374

[4] X. F. Zhao, I. Gojo, T. York, Y. Ning and M. R. Baer, "Diagnosis of biphenotypic acute leukemia: a paradigmatic approach," Int. J. Clin. Exp. Pathol, vol. 3(1), pp. 75–86, 2010.

[5] M.C.Béné, "Biphenotypic, bilineal, ambiguous or mixed lineage: strange leukemias," Haematol., vol. 94(7), pp. 891–893, 2009. http://dx.doi.org/10.3324/haematol.2009.007799

[6] S. Glanz, Medical-biological statistics. M: Praktika. 1999, 459 p.

[7] A. Agresti, *An introduction to categorical data analysis,* 2nd ed., JohnWiley & Sons, Inc., 2007, 371 p.

[8] J. W. Vardiman, "The World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues: an overview with emphasis on the myeloid neoplasms," *Chem. Biol. Interact*., vol. 184(1–2), pp. 16–20, 2010. http://dx.doi.org/10.1016/j.cbi.2009.10.009

[9] G. Basso, B. Buldini, L. De Zen and A. Orfao, "New methodologic approaches for immunophenotyping acute leukemias," *Haematologica*, vol. 86(7), pp. 675–692, 2001.

[10] A. C. Seegmiller, S. H. Kroft, N. J. Karandikar and R. W. McKenna, "Characterization of Immunophenotypic Aberrancies in 200 Cases of B Acute Lymphoblastic Leukemia," *Am. J. Clin. Pathol*., vol. 132(6), pp. 940–949, 2009. http://dx.doi.org/10.1309/AJCP8G5RMTWUEMUU

[11] A. Hejlsberg, M. Torgersen, S. Wiltamuth, P. Golde, *The C# Programming Language (Covering C# 4.0),* 4th ed.. Microsoft Corporation, 2011, 815 p.

[12] E. T. Ray. *Learning XML*, 2nd ed., O'Reilly Media, 2003, 418 p.

Natalia Novoselova received the Doctoral degree in Computer Sciences from the United Institute of Informatics Problems, National Academy of Sciences of Belarus (NASB) in 2008. Since 2000 she has been a Senior Scientific Researcher at the Laboratory of Bioinformatics, United Institute of Informatics Problems NASB in Minsk, Belarus. Her research interests include data mining and statistical methods, algorithms for multivariable data analysis, especially in the field of biology and medicine. She is the author of more than 20 research publications in these areas.
Address: Laboratory of Bioinformatics, United Institute of Informatics Problems, Surganova Str. 6, Minsk, 220012, Belarus. Phone: +375-17-2842092
E-mail: novosel@newman.bas-net.by.

Igor Tom received the Doctoral degree in Computer Science from the Institute of Engineering Cybernetics, National Academy of Sciences of Belarus (NASB) in 1986. Since 1999 he has been a Chief of the Laboratory of Bioinformatics at the United Institute of Informatics Problems of NASB. His current research interests are in the fields of the development of intelligent methods of data analysis and information technologies for medical and industrial applications. He is the author and co-author of more than 180 scientific publications, including 50 full papers in journal.
Address: Laboratory of Bioinformatics, United Institute of Informatics Problems, Surganova Str. 6, Minsk, 220012, Belarus. Phone: +375-17-2842153
E-mail: tom@newman.bas-net.by.

Michael Belevtsev received the Doctoral degree in Biological Sciences in 2003. Since 2003 he has been the Head of the Laboratory of the Immunological Research at the Belarusian Research Center, for Pediatric Oncology, Hematology and Immunology. Since 2012 he has also been a Deputy Director. His research interests include innate defects of the immune system, cancer immunology. He has more than 100 publications in the fields of immunology, oncology.
Address: Pos. Lesnoe-2, Minsk district, 223040, Belarus.
E-mail: belevtsev_m@mail.ru.

Nataļja Novoselova, Igors Toms, Mihaels Belevcevs. Lēmumu pieņemšanas modeļi un programmnodrošinājums diferenciālajai imūnajai fenotipiskajai leikožu un limfomu diagnostikai

Rakstā apskatīti lēmumu pieņemšanas modeļi un atbilstošais programmnodrošinājums, kas paredzēti, lai palīdzētu speciālistam veikt leikožu un limfomu diagnostiku atbilstoši imunoloģiskajai fenotipizācijai. Modeļu zināšanu bāzes iekļauj likumu kopas, kas paredzētas leikožu un limfomu diagnostikai un kas izveidotas, balstoties uz statistiskās analīzes, kas veikta diagnostisko gadījumu datubāzē, kas uzkrāta valsts uzņēmumā „Republikas bērnu onkoloģijas, hematoloģijas un imunoloģijas zinātniski praktiskais centrs", kā arī izmantojot standarta likumus. Pamata leikožu, limfomu un to atbilstošo apakštipu slimību diagnostikas likumi ir svara koeficientu un svarīguma koeficientu kopas katram diagnostiskajam marķierim, kas nosaka atbilstošās grupas vai apakštipa profilu. Svara koeficienti iegūti, balstoties uz saistīto tabulu statistisko analīzi. Diagnostiskie lēmumi tiek pieņemti, izmantojot īpaši izstrādātu algoritmu, kas aprēķina slimības apakštipu izteiktības pakāpi, balstoties uz imunoloģiski fenotipisko rādītāju ievades vērtībām un izmantojot modeļa zināšanu bāzi. Leikožu un limfomu diagnostiskajiem modeļiem ir kopīga moduļu struktūra, bet to diagnostisko likumu saturs un lēmumu pieņemšanas moduļa organizācija atšķiras. Modeļa izejas dati tiek attēloti, kā ranžēts slimības apakštipu saraksts, kas ļauj novērtēt katra slimības apakštipa izteiktības pakāpi katram konkrētajam diagnostikas gadījumam. Saraksta augšgalā atrodas visiespējamākie apakštipi. Specializētais programmnodrošinājums realizē izstrādātos lēmumu pieņemšanas modeļus, nodrošinot

divpakāpju leikožu diagnostikas un limfomu diagnostikas veikšanu, izmantojot izstrādāto diagnostisko likumu kopu, kā arī diagnostikas rezultātu un atbilstošās uzziņu informācijas atainošanu. Ieejas dati ir organizēti atbilstoši vienotajai pacienta reģistrācijas kartei, kas nodrošina darbu dažādos diagnostikas līmeņos: izvērstajā leikožu diagnostikas līmenī, padziļinātajā leikožu diagnostikas līmenī un bērnu un pieaugušo limfomu diagnostikas līmenī. Izstrādātās sistēmas izmantošana veicinās precizētas slimības kartes iegūšanu, ņemot vērā pacienta individuālos ģenētiskos datus, kā arī uzlabos diagnostiku, izmantojot specializētus lēmumu pieņemšanas modeļus leikožu un limfomu imunoloģiski fenotipiskajai diagnostikai, ņemot vērā paralēlo ekspresiju. Tāpat sistēmu var izmantot par mācību līdzekli studenti, zinātnisko grādu pretendenti un praktizējošie speciālisti, apgūstot leikožu un limfomu diagnostikas pamatmetodes, kas izmanto imunoloģiski fenotipiskos rādītājus.

Наталья Новоселова, Игорь Том, Михаель Белевцев. Модели принятия решений и программное обеспечение для дифференциальной иммунофенотипической диагностики лейкозов и лимфом

В статье рассматриваются модели принятия решений и соответствующее программное обеспечение, позволяющее помочь специалисту при диагностике лейкозов и лимфом по результатам иммунофенотипирования. Базы знаний моделей включают в себя наборы правил по диагностике лейкозов и лимфом, которые построены на основе статистического анализа базы диагностических случаев, накопленной в государственном учреждении «Республиканский научно-практический центр детской онкологии, гематологии и иммунологии», а также стандартные правила. Правила для диагностики основных групп лейкозов, лимфом и соответствующих им подтипов заболевания представляют собой наборы весовых коэффициентов и коэффициентов важности для каждого диагностического маркера, которые определяют профиль соответствующей группы или подтипа. Весовые коэффициенты были получены на основе статистического анализа таблиц сопряженности. Диагностическое решение выносится с использованием специально разработанного алгоритма, позволяющего рассчитать степени выраженности подтипов заболевания на основе входных значений иммунофенотипических показателей с использованием базы знаний модели. Диагностические модели для лейкозов и лимфом имеют общую модульную структуру с отличающимся составом диагностических правил и организацией модуля принятия решений. Выходные данные модели представляются в виде упорядоченного списка подтипов заболевания, что позволяет оценить степень выраженности каждого из подтипов заболевания для конкретного диагностического случая. Вверху списка располагается наиболее вероятный подтип. Специализированное программное обеспечение реализует разработанные модели принятия решений, обеспечивая проведение двухуровневой диагностики лейкозов и диагностику лимфом с использованием набора разработанных диагностических правил, просмотр результатов диагностики и соответствующей справочной информации. Входные данные организованы согласно унифицированной регистрационной карте пациента, обеспечивающей возможность работы на разных уровнях диагностики: уровень расширенной диагностики лейкозов, уровень углубленной диагностики лейкозов, уровень диагностики лимфом для детей и взрослых. Использование разработанной системы будет способствовать получению уточненной картины заболевания с учетом индивидуальных генетических данных пациента, совершенствованию диагностики за счет использования специализированных моделей принятия решений для иммунофенотипической диагностики лейкозов и лимфом с учетом коэкспрессий. Система также может быть использована в качестве обучающего средства для студентов, аспирантов, практикующих специалистов по основным методам диагностики лейкозов и лимфом по набору иммунофенотипических показателей.