# Creation of the software for the acute leukemia automatic differentiation

**Belevtsev MV[1], Kulaga VV[2], Spiridonov SV[2], Savva NN[1], Aleinikova OV[1]**

[1] Belarussian Research Center for Pediatric Oncology and Hematology, Minsk, Belarus
[2] United Institute of Informatics Problems, Minsk, Belarus

## Abstract

**Purpose:** The aim of our works was created the software for automatic classification of the acute leukemia by immunological subtypes on the basis of the immune marker analysis of blasts, which can used as the "Decision Making Support System".

**Material and methods:** At The Belarussian Center for Pediatric Oncology and Hematology (Belarus, Minsk) one of the most wide diagnostic immune marker panels is used, that is recommended by the research group (EGIL-95). The panel contains the following monoclonal antibodies: Control, CD45, CD14, CD1a, CD2, CD5, CD7, CD3, CD4, CD8, CD19, CD20, CD22, CD10, HLA-DR, CD34, CD13, CD33, CD15, CD117, CD11c, mIgM, cy IgM, CD79a, MPO, TdT. Positiveness and negativeness of these monoclonal antibodies on the tumor cells is defined by the flow cytometry method using the corresponding devise. The search algorithm is realized by the modern optymalization algorithms application.

**Results:** To avoid the logical algorithm for finding the solution has been worked out, which allows to find correct solutions by partial estimation of information about the markers. On the basis of the modern knowledge in the area of the differential diagnostics of acute leukosis in children, where the definite combination of positive (1) and negative (0) markers corresponds to each type of the acute leukosis. The developed algorithm forms the basis of the calculation program. The essence of the algorithm consists in the fact, that influence of only the definite markers set is typical for the definite groups of diseases with possible co-expressions. Consider this in more detail and describe the algorithm of diagnostics for every type of acute leukosis, multistage process of classification with the purpose to achieve criteria of defeat in the lymphomas.

**Conclusions:** Hence the task of acute leukemia immune marker diagnostics automation can be surely referred to the decision making support tasks, which imply conformation or refute of this or that hypothesis. The final decision is made by a person with consideration of recommendations obtained as a result of the software operation. Besides this, the created software can be used as a teaching program by specialists for acute leukemia diagnostics.

**Key words:** childhood leukemia, immunophenotyping, automatic, subtype.

## Introduction

Contemporary classification of acute leukemia is based on the tumor cells immune marker definition [1,5,7]. Utilization of the wide panel of monoclonal antibodies (over 25) for this purpose raises the diagnostics self-descriptiveness, but also increases number of positive and negative expression combinations, that substantially complicates comparison of the combination obtained with the required subtype of acute leukemia. Therefore, the immune marker diagnosis requires deep and long-time period of training specialists, who, except knowledge in immune marker techniques, have also got knowledge in the area of positive and negative monoclonal antibodies expression combination comparison with the leukemia subtype [2-4,6]. The aim of our works was created the software for automatic classification of the acute leukemia by immunological subtypes on the basis of the immune marker analysis of blasts, which can used as the "Decision Making Support System".

## Material and methods

At The Belarussian Center for Pediatric Oncology and Hematology (Belarus, Minsk) one of the most wide diagnostic

ADDRESS FOR CORRESPONDENCE:
Michael Belevtsev
Belarusian Center for Pediatric Oncology and Hematology,
P.O. Lesnoye-2, Minsk district 223052, Belarus,
Tel: +375 17 2024089, Fax: +375 17 2024222
e-mail: HYPERLINK mailto: Belevtsev_m@mail.ru

immune marker panels is used, that is recommended by the research group (EGIL-95). The panel contains the following monoclonal antibodies: Control, CD45, CD14, CD1a, CD2, CD5, CD7, CD3, CD4, CD8, CD19, CD20, CD22, CD10, HLA-DR, CD34, CD13, CD33, CD15, CD117, CD11c, mIgM, cy IgM, CD79a, MPO, TdT. Positiveness and negativeness of these monoclonal antibodies on the tumor cells is defined by the flow cytometry method using the corresponding devise. The search algorithm is realized by the modern optimization algorithms application.

## Results and discussion

To solve the task directly is extremely complicated, as it requires selecting a single one from 771 possible solutions by the numerical value of 27 markers. Utilization of some special algorithms can only slightly decrease the computational complexity. Therefore, for finding this task solution in the real time it is required to apply some restrictions and admissions in the solution. The analysis that has been carried out allows defining the value for every marker, which has been named "the threshold value". In the case of exceeding this value it is possible to affirm with the great credibility value, that this marker has the great influence on the general process of diagnostics. During this hypothesis verification it was stated, that introduction of these restrictions does not effect anyhow the task solution accuracy. As a result we can substantially reduce the task computational complexity by application of the information, represented in the binary system, during the task solving process. It means, that in case of exceeding the positiveness criterion by the marker (20% and for some of them 10%), it is assigned a unit value, otherwise – zero value. Although such problem formulation substantially simplifies the solution searching process, however, it is also very complicated and difficult in formalization. It is caused by the fact, that when searching the solution "directly" one has to solve the complicated combinatorial task of estimation of 134217728 possible combinations of the markers and selection of a single solution. The process of writing the software for estimation of all possible variants can take much time. To avoid this logical algorithm for finding the solution has been worked out, which allows finding correct solutions by partial estimation of information about the markers. On the basis of the modern knowledge in the area of the differential diagnostics of acute leukosis in children, we have constructed the table (*Tab. 1*), where the definite combination of positive (1) and negative (0) markers corresponds to each type of the acute leukosis. The developed algorithm forms the basis of the calculation program. The essence of the algorithm consists in the fact, that influence of only the definite markers set is typical for the definite groups of diseases with possible co-expressions. Consider this in more detail and describe the algorithm of diagnostics for every type of acute leukosis.

For diagnostics of the B-linear ALL with possible co-expressions the section corresponding to the markers of the B-linear ALL (CD19, CD20, CD22, CD10, cy IgM, s IgM, CD79a, CD45+CD14, CD34, HLA-DR, TdT) and markers, which can be met in the form of co-expressions (CD3, CD4+CD8-, CD4-CD8+, CD4+CD8+, CD5, CD7, CD1a, CD11c, CD14, CD13, CD33,CD117, MPO) is chosen from the general *Tab. 1*. Further every possible combination (sequence) of 0 and 1 is recorded as a binary number, which then is converted into the decimal one. Positions during the conversion are represented by the sequential numbers of markers. For example, the sequence, corresponding to the Pro-B ALL, recorded in the binary system, is equal to 00000000111000100000000001111, that corresponds to the number 125847296 in the decimal system. Such manipulations are made for all possible combinations of the selected markers. As a result 31 numbers are obtained corresponding to the B-linear ALL. After it has been defined, that the obtained sequence unambiguously belongs to the selected group, a search of possible co-expressions in *Tab. 1* is carried on.

During diagnostics of the T-linear ALL with possible co-expressions the section corresponding to the markers of the T-linear ALL (CD2, CD3, CD4+CD8-, CD4-CD8+, CD4+CD8+, CD5, CD7, CD1a, TdT, CD34, CD45+CD14) with all possible co-expressions (CD19, CD20, CD22, CD10, cy IgM, s IgM, CD79a, CD14, CD13, CD33, CD117) is formed from the general *Tab. 1*. Every possible combination (sequence) of 0 and 1 is recorded as a binary number, which is then converted into the decimal one in the same manner as it was described earlier. As a result of the transformations 84 numbers are obtained corresponding to the T-linear ALL.

For diagnostics of the AML with possible co-expressions of the lymphoid markers the markers typical to the AML (CD13, CD33, CD15, CD117, CD11b, CD11c, CD14, MPO, CD45+CD14, CD34, HLA-DR) with possible co-expressions of B-lymphoid (CD19, CD20, CD22, CD10, cy IgM, s IgM, CD79a, TdT) and T-lymphoid markers (CD2, CD3, CD4+CD8-, CD4-CD8+, CD4+CD8+, CD5, CD7, CD1a, TdT) is chosen from the general *Tab. 1*. Every possible combination (sequence) of 0 and 1 is recorded as a binary number, which is then converted into the decimal one in the same manner as it was described earlier. As a result of the transformations 72 numbers are obtained corresponding to the AML. After definition of the fact, that the obtained sequence unambiguously belongs to the selected group, the search of possible co-expressions is carried on. Diagnosis of the acute bi-phene leukosis is formed from the previous groups in the case, if the number of co-expressions exceeds 4 by any separate variant of coincidence.

## Conclusions

Hence the task of acute leukemia immune marker diagnostics automation can be surely referred to the decision making support tasks, which imply conformation or refute of this or that hypothesis. The final decision is made by a person with consideration of recommendations obtained as a result of the software operation. Besides this, the created software can be used as a teaching program by specialists for acute leukemia diagnostics. Creation of the (universal) software second version is planned, which could be applied at any modification of the monoclonal antibodies diagnostic panel.

**Acknowledgments**

This work is supported by the ISTC B-736 grant.

*Table 1. Marker overflows*

| Marker | Pro-B ALL | Common B ALL | Pre-B ALL | Transitory Pre-B ALL | Mature-B ALL | Probably Co-expression | Pre-T ALL | Cortical-T ALL | Mature-T ALL | Atypical-T ALL | Probably Co-expression | Early and Intermediately AML | Mature AML | AML | Probably Co-expression | B-lymphoid +Myeloid | T-lymphoid +Myeloid |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CD2 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1 | 1 | 1/0 | - | 0 | 0 | 0 | 1/0 | 0 | 1 |
| CD3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1 | 1/0 | - | 0 | 0 | 0 | 0 | 0 | 1/0 |
| CD4+ CD8- | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | - | 0 | 0 | 0 | 1/0 | 0 | 1/0 |
| CD4- CD8+ | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 1/0 | 1/0 | - | 0 | 0 | 0 | 0 | 0 | 1/0 |
| CD4+ CD8+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | - | 0 | 0 | 0 | 0 | 0 | 1/0 |
| CD5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 1/0 | 0 | 1 |
| CD7 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 1/0 | 1/0 | 1 |
| CD1a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | - | 0 | 0 | 0 | 0 | 0 | 1/0 |
| CD19 | 1 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CD20 | 1/0 | 1/0 | 1/0 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 |
| CD22 | 1 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CD10 | 0 | 1 | 1 | 1 | 0 | - | 0 | 1/0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 |
| cy IgM | 0 | 0 | 1 | 1 | 0 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 |
| s IgM | 0 | 0 | 0 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CD79a | 1 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| CD13 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1 | 1 | - | 1 | 1 |
| CD33 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 0 | 0 | 0 | 1/0 | 1 | 1 | 1 | - | 1 | 1 |
| CD15 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1/0 | 1/0 | - | 1/0 | 1/0 |
| CD117 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 0 | 0 | 0 | 1/0 | 1 | 0 | 1/0 | - | 1/0 | 1/0 |
| CD11b | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1/0 | 1/0 | - | 1/0 | 1/0 |
| CD11c | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 1/0 | 1/0 | - | 1/0 | 1/0 |
| CD14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | - | 0 | 0 |
| MPO | 0 | 0 | 0 | 0 | 0 | 1/0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1 | 1 | - | 1 | 1 |
| CD45+ CD14- | 1/0 | 1/0 | 1/0 | 1 | 1 | - | 1/0 | 1 | 1 | 1 | - | 1/0 | 1 | 1 | - | 1/0 | 1/0 |
| CD34 | 1 | 1/0 | 1/0 | 0 | 0 | - | 1 | 1/0 | 0 | 1/0 | - | 1 | 0 | 1/0 | - | 1 | 1 |
| HLA-DR | 1 | 1 | 1 | 1 | 1 | - | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | - | 1 | 1/0 |
| TbT | 1 | 1 | 1/0 | 1/0 | 0 | - | 1 | 1 | 0 | 1/0 | - | 0 | 0 | 0 | 1/0 | 1 | 1 |

## References


1. Borowitz MJ. Immunophenotyping of acute leukemia by flow cytometry. Clin Immunol Newsletter, 1993; 13: 53-73.
2. Jennings CD, Foon KA. Recent advances in flow cytometry: application to the diagnosis of hematologic malignancy. Blood, 1997; 90: 2863-92.
3. Launder TM, Bray RA, Stempora L, et al. Lymphoid-associated antigen expression by acute myeloid leukemia. Am J Clin Pathol, 1996; 106: 185-91.
4. Piatoska-Jakubas B et al. The diagnostic significance of blast immunophenotype assay in patients with acute leukemia. Acta Haematol Pol, 1996; 27: 33-41.
5. Barnard DR, et al. Morphologic, immunologic and cytogenetic classification of acute myeloid leukemia and myelodysplastic syndrom in childhood: a report from the Children's Cancer Group. Leukemia, 1996; 10(1): 5-12.
6. Verwer BJ, Terstappen LW. Automatic lineage assignment of acute leukemias by flow cytometry. Cytometry, 1993;14(8): 862-75.
7. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, Laster TA, Bloomfield CD. The World Health Organization classification of neoplasms of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting – Airlie House, Virginia, November, 1997. Hemat Journ, 2000; 1(1): 53-66.