Development of the information analytical system for childhood oncohematology: project ISTC #B-522

Tom IE[1], Krasko OV[1], Potapnev MP[2], Uglova TA[2]

[1] United Institute of Informatics Problems of The National Academy of Sciences of Belarus, Minsk, Belarus
[2] Belarussian Center for Pediatric Oncology and Hematology, Minsk, Belarus

Abstract

Propose: The main objectives were development of an information-analytical system (IAS) for processing personal, clinical, laboratory and other medical data, applying IAS in acute childhood leukemia particularly to select prognostic risk factors (PRFs) and to predict outcome of induction therapy based on the patient's profile of PRFs. Selection of adequate therapy was based on the results of analysis of individual profiles of PRFs for children with acute leukemias.

Materials and methods: The medical data were obtained in Belarussian Research Center for Pediatric Oncology and Hematology (BRCPOH). Altogether, data from 189 patients with acute lymphoblastic leukemia (ALL) and 74 patients with acute myeloid leukemia (AML) were analyzed. To verify the results of application the analytical system testing sample was obtained in Medical High School Hannover (collaborator of the project Prof. K.Welte).

Results: The IAS for selecting PRFs and for predicting outcome of induction therapy based on the patient's profile of PRFs. It can predict the outcome of induction therapy for protocols ALL BFM 90 M (Minsk, BRCPOH), ALL MB 2002 (Minsk, BRCPOH), ALL BFM 2000 (Hannover, MHSH), AML MM-2000 (Minsk, BRCPOH), AML MM-2003 (Minsk, BRCPOH).

Conclusions: The new PRFs and their combinations were explored by using IAS, the combinations were selected to predict outcome of induction therapy.

Key words: information-analytical system, prognostic risk factors, outcome of induction therapy.

ADDRESS FOR CORRESPONDENCE:
Igor E. Tom
Laboratory of bioinformatics
United Institute of Informatics Problems
The National Academy of Sciences of Belarus
6 Surganov Street, 220012, Minsk, Belarus
e-mail: tom@newman.bas-net.by

Introduction

The main objectives of the project were development of an information-analytical system (IAS) for processing personal, clinical, laboratory and other medical data, applying IAS in acute childhood leukemia particularly to select prognostic risk factors (PRFs) and to predict outcome of induction therapy based on the patient's profile of PRFs. Selection of adequate therapy was based on the results of analysis of individual profiles of PRFs for children with acute leukemias.

The main objectives have been reached by using the most progressive informational technologies for database design, access application development and the intelligent methods of medical data analysis.

The basis of an elaborated IAS is a three-tier (client – Web-server – database server) architecture that allows to part data storage and data handling and provides a security access to the data. Such an architecture means that only user dialogues is proceed on client's computer while business logic is checked by Web-server applications.

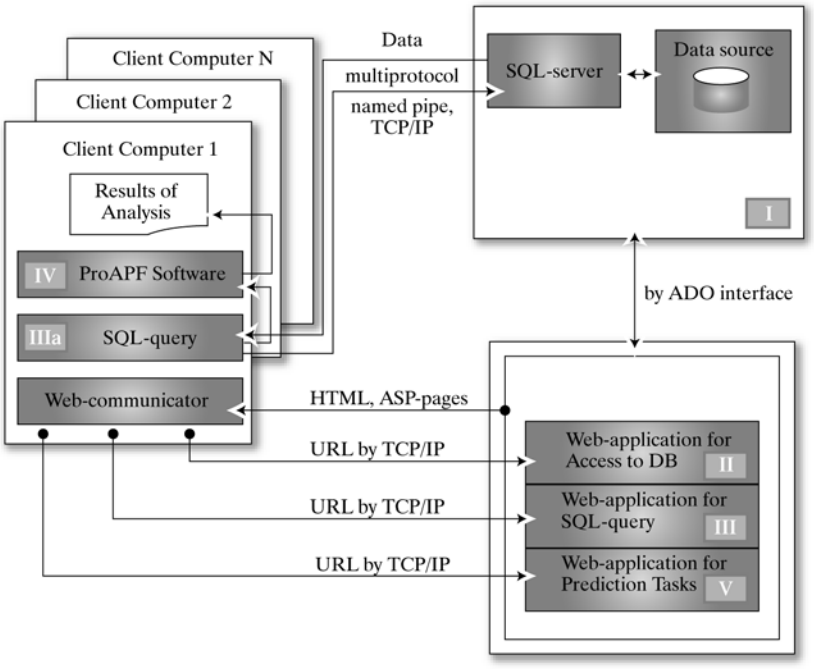
Material and methods

The medical data were obtained in Belarussian Research Center for Pediatric Oncology and Hematology (BRCPOH). Altogether, data from 189 patients with ALL and 74 patients with AML were analyzed. To verify the results of application the analytical system testing sample was obtained in Medical High School Hannover (collaborator of the project Prof. K.Welte).

Results

SQL-server was appropriated to store the data. We have used ASP-technology and Internet Information Server (IIS 5.0) to create user Web-application with ADO-components to

Figure 1. **The structure of IAS**



access the database server. The structure of IAS are as follows (see *Fig. 1*).

1. Subsystem #1: The Database for collecting (according the registration form) epidemiological, clinical and laboratory information in a form of relational database. We have chosen MS SQL Server as a Database Management System (DBMS) which is tight compatible with Windows NT/2000/XP platforms (the last is a preferable operational system in computer network of BRCPOH). The logical structure of the database corresponds the elaborated registration form. The form is clearly structured and contains data grouped by 7 mayor epidemiological, clinical and laboratory categories: personal information, history of disease, pre-diagnosis period, information about treatment, clinical manifestation, laboratory data, complications during induction therapy and toxicity. Registration form is created to fill computerized database for patient both with ALL and AML.

2. Subsystem #2: Web-application for a remote access to the database on purpose to add, view or edit information. According to the logical structure of the database we developed software for the basic function of data access (add, edit, delete) as stored procedures of Database server for different kind of data. We proposed to develop the client application based only on stored procedures, that restrict access to database for users, that means security and reliability of data of electronic register enhanced. The client application is designed in the format of sections of the registration form. We use three-tier "Client – Web-server – Database server" architecture, that grants remote access to data via Internet or Intranet. The user can access to data using only Microsoft Internet Explorer. The middle level (Web-server) ensures security exchange between client and Database server on ASP-platform. We constructed two kind of pages: DHTML-forms and action-pages. DHTML-forms only can read data from database and perform data for user. They also provide data input. Action-pages are handled by Web-server without exchange with client. This Web-application include more 175 ASP-pages.

3. Subsystem #3: Web-application for simplifying complicated SQL-queries generation to get necessary information from Database for further analysis. The subsystem for simplifying complicated SQL-queries generation is made in two versions for "Client – Web-server – Database server" and "Client – Server" architecture. Each of them can be adapted to database logical structure invisible for users. The query tools are based on the interactive query constructing using GUI. The user selects necessary criteria for query, after that subsystem translates them into SQL-query (Transact SQL-language is used in DBMS MS SQL Server 2000). The dialog with user helps to define input and output data. The query tools to electronic register data extract are created for further analysis in the subsystem #4 (ProAPF Soft). Query results are presented in form of a table where each row is a case, each column is a parameter. This table can be saved in various formats (.txt, .doc, .xls, .dbf) for further analysis.

4. Subsystem #4: The special software for multivariate statistical and intelligent of data analysis. The subsystem for multivariate statistical and intelligent data analysis is developed as a special software "Professional Analysis of Prognostic factors" (ProAPF). ProAPF software finds out prognostic factors and their combinations for medical classification and prediction tasks. The main reason in developing ProAPF is to create modern information technology for discovering combination of various prognostic risk factors (PRFs) for induction therapy

response in childhood acute leukemias. The main goal of ProAPF application is to enhance the efficacy of therapy using the estimation of patient's state based on the prognostic risk factors for selection an adequate intensity of therapy. ProAPF software actualizes the information technology of joint application multivariate statistical methods and intelligent neural network analyses of data with coherent input/output on the each stage of analysis.

5. Subsystem #5: Web-application for prediction task of induction therapy outcome for children with acute leukemias. The applied results are presented as subsystem #5 – Web-application for predict of induction therapy outcome for children with acute leukemias. This subsystem allows to review classification trees, scoring models, results of classification for as well as to predict outcome based on patient's PRFs profile (input by user) for the following group of patients:

1. male patients with B-lineage ALL, day 15 – early response
2. male patients with B-lineage ALL, day 33 – remission induction
3. female patients with B-lineage ALL, day 15 – early response
4. patients with T-lineage ALL, day 15 – early response
5. patients with T-lineage ALL, day 33 – remission induction
6. patients with AML, day 14 – early response
7. patients with AML, day 28 – response.

Using this subsystem it can predict the outcome of induction therapy for protocols ALL BFM 90 M (Minsk, BRCPOH), ALL MB 2002 (Minsk, BRCPOH), ALL BFM 2000 (Hannover, MHSH), AML MM-2000 (Minsk, BRCPOH), AML MM-2003 (Minsk, BRCPOH).

Conclusions

1. The information analytical system (IAS) was developed to predict the induction therapy outcome for patients suffering from acute leukemias. IAS was developed with the modern informational technology of database design, Web-access tools and the intelligent methods of medical data analysis. IAS was applied in Belarussian Research Center for Pediatric Oncology & Hematology (BRCPOH, Minsk, Belarus).

2. The database for prognostic risk factors (PRFs), collected at diagnosis and during early treatment course in children with acute lymphoblastic leukemia and acute myeloid leukemia, was created. Altogether, data from 189 patients with ALL and 74 patients with AML from BRCPOH were enrolled for analysis.

3. The information technology of analysis of prognostic risk factors actualized as ProAPF software was developed to identify patient-specific risk group to apply risk group – oriented chemotherapy in patients with ALL or AML.

4. The technique of accounting the individual profile of patient-specific PRFs and their association with response to induction therapy at critical timepoints for patients with ALL or AML was elaborated using statistical and intelligent neural network analysis methods.

Acknowledgement

The financial support of European Community under the contract to The International Science and Technology Center (Project #B-522) is greatly acknowledged.