BIO-ALGORITHMS
AND
MED-SYSTEMS

MEDICAL SCIENCES

ENGINEERING

PHYSICS

COMPUTER SCIENCES

COMMERCE

LAW

Contents

STRUCTURE MODELING BASED COMPUTER AIDED T-CELL EPITOPE DESIGN

GOPICHANDRAN SOWMYA[1], AYYAPPAN VAISHNAVI[1], PANDJASSARAME KANGUEANE[1,2]

[1]*Biomedical Informatics, Pondicherry, 607402, India;* [2]*AIMST University, Kedha 08100, Malaysia. Email –
kangueane@bioinformation.net; Phone: +91 413 2633589, Fax: +91 413 2633722*

Abstract: Progress in computational modeling for structural biology has motivated the use of molecular mechanics calculations for synthetic peptide design as potential T-cell epitopes (peptides inducing immunogenicity). Short antigen peptides from virus/bacteria/parasite are recognized by host specific human leukocyte antigen (HLA) molecules for T-cell sensitive cellular immune response. However, HLA molecules are highly polymorphic at the sequence level among ethnic population (American Indian, Australian aboriginal, Black, Caucasoid, Hispanic, Mixed, Oriental, Pacific Islander and Unknown ethnicity). The binding of peptides to host HLA molecules are both specific and sensitive. The use of computer aided molecular modeling principles has been shown for the design of T-cell specific epitopes as potential vaccine candidates. Application of computational techniques such as molecular dynamics simulation (MDS), self consistent ensemble optimization (SCEO), free energy (FE) estimation, computational combinatorial ligand design (CCLD), 3D quantitative structure activity relationship (3D-QSAR) and structure based virtual pockets (SBVP) in HLA-peptide binding prediction is discussed. The ability of modeling and design to predict peptide binding to a wide array of defined HLA alleles finds application in proteome wide scanning of bacteria/virus/parasite proteomes towards cocktail peptide vaccines.

HLA Immunology & T-cell epitopes

Structure of HLA-peptide complexes determined by X-ray crystallography **[1]** and competitive binding assay data **[2]** indicates the binding of HLA molecules to short antigenic peptides during T-cellular immune response. It is known that the specific binding of peptides to HLA molecules is the crucial step in T-cell mediated immune response. The binding of peptide to HLA alleles during T-cell immunity is illustrated in **Figure 1.** High polymorphism among HLA molecules in different ethnicity determines specificity and sensitivity of peptide binding. Sequence polymorphism is frequently concentrated at the peptide binding groove of both class I and II HLA molecules (**Figure 2**). The current International ImMunoGeneTics information system (IMGT)/human leukocyte antigen (HLA) database (January 2010) contains over 4447 HLA alleles **[3].** The binding of these alleles to overlapping short antigen peptides (9-20 residues) is combinatorial in selection and degeneration. The binding of peptide to HLA alleles through position specific accommodation of peptide residues within the binding groove is shown in **Figure 3.** The binding of peptides to allele specific HLA alleles is both specific and sensitive determined through the physics and chemistry of the pockets accommodating the position specific peptide residues **(Figure 4)**. Hence, the prediction of human leukocyte antigen (HLA) peptide binding finds application in epitope (antigen peptide with immunogenicity) design for the development of vaccines and diagnostics of diseases associated with T-cellular immunity. Peptide vaccines are cocktail peptides from virus/bacteria/parasite with specificity to maximum number of known HLA alleles. The use of molecular mechanics calculations and predictions in the design of T-cell epitopes for specificity to maximum HLA allele coverage is discussed.

Molecular mechanics in T-cell Epitope design

Progress in computational structural biology during the last decade has motivated many to assess the usefulness of molecular mechanics and modeling calculations for HLA-peptide binding predictions. The methodology often practiced in the design of HLA specific peptides as potential T-cell epitopes is illustrated in **Figure 5.** This involves the modeling and assessment of HLA-peptide complexes for potential design of T-cell epitopes. The procedure involves the modeling of both HLA receptors and
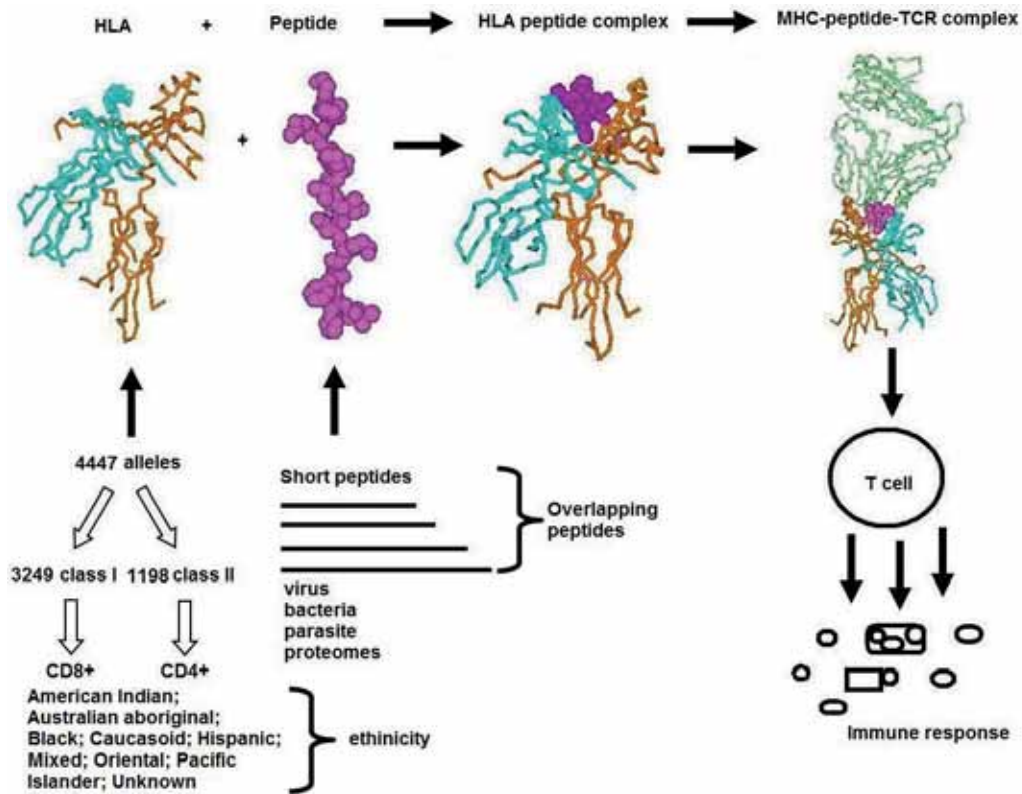
Fig. 1. T-cell immune response as a result of HLA-peptide binding is shown. HLA allele specific binding of antigen peptides in the context of ethnicity is schematically illustrated. T cell response is either T cytotoxic (CD8+) or T-helper (CD4+). HLA = Human Leukocyte Antigen; TCR = T-cell receptor
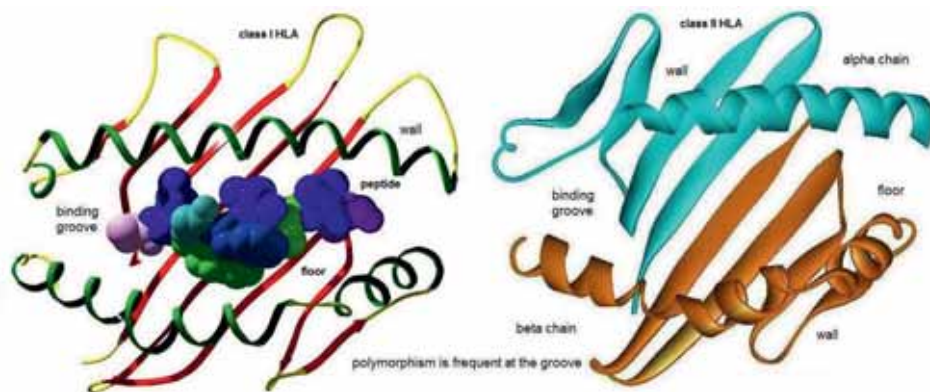
Fig. 2. The peptide binding groove in class I and II HLA molecules is shown. Class I and II molecules have structurally different peptide binding grooves. The class I molecule groove is formed between domains α1 and α2 of α chain. The class II molecule groove is formed between α and β subunits. Peptides bond to class I molecules are under the constraints of the groove. Peptides bond to class II molecules have an extended conformation at the groove

peptide ligands within the binding groove. All known HLA molecules have at least 70% sequence identity to one or more known HLA structures in Protein Data Bank (PDB). It should be noted that homologous proteins share a common fold, even when the overall sequence identity is less than 10% **[4].** Molecular mechanics based calculations are commonly performed to predict the structures of all known HLA alleles. Modeling the HLA receptor is a trivial issue given the advancement in computational structural biology over the last two decades. However, peptide modeling is non-trivial due to non-availability of peptide templates given the diverse nature of peptide backbone conformation in the binding groove. The backbone conformations of bound peptides are generally not conserved in the binding groove. The bound peptides are thus flexible and the middle part of the peptides usually bulges out of the binding groove (**Figure 3**). This bulging part allows the backbones to take different patterns at the groove. The generic peptide structure determination meth-

ods like molecular dynamics simulation (MDS), self consistent ensemble optimization (SCEO), free energy (FE) estimation computational combinatorial ligand design (CCLD), 3D quantitative structure activity relationship (3D-QSAR) and structure based virtual pockets (SBVP) are have been shown for HLA-peptide binding prediction. However, these methods have both merits and demerits. The other challenge in HLA-peptide binding prediction is scoring method for binding. The scoring of peptide binding with the HLA receptor using a suitable interaction function or weighted matrix either for hydrophobicity or hydrophilicity is still a great challenge.

Molecular dynamics simulation (MDS)

The approach based on molecular dynamics simulation (MDS) of HLA-peptide complexes shown by Rognan *et al.,* (1994) **[5]** allows a crude discrimination of binders from non-binders. The approach
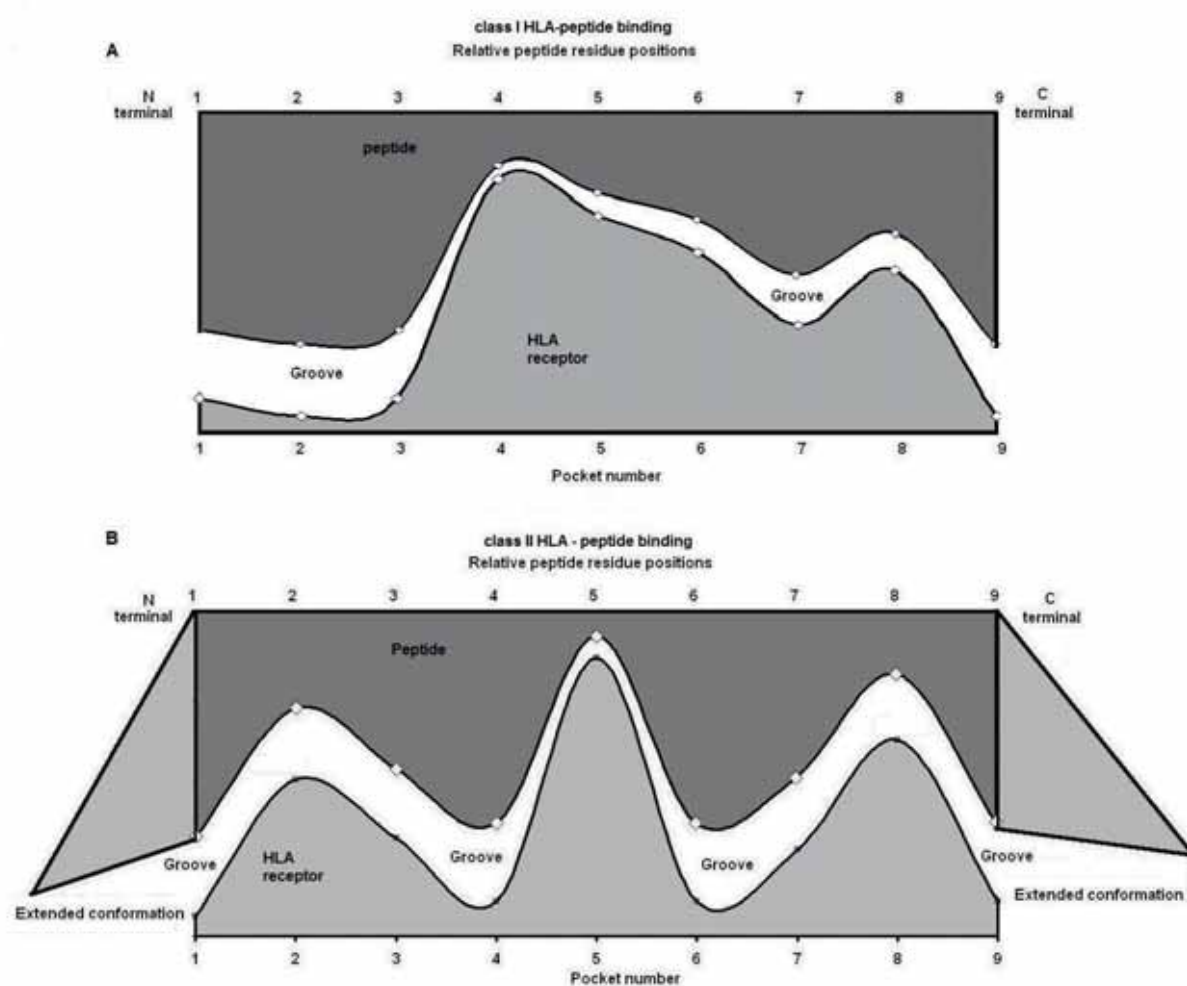
Fig. 3. The binding of peptides to the HLA receptor in both class I (A) and class II (B) molecules is schematically illustrated. The positional anchor specific binding of peptides to the HLA receptor binding groove is shown. The class I peptide interaction anchor is generated based on a dataset of 22 class I – peptide structural complexes from Protein databank (PDB) compiled by Zhao *et al.,* (2003) **[28]**. The class II peptide interaction anchor is generated using a dataset of 15 class II – peptide structural complexes from Protein databank (PDB) compiled by Mohanapriya *et al.,* 2009 **[29]**

calculates change in free energy during simulation of the HLA-B*2705 complex using AMBER force field described by Ponder & Case **[6]**. This method is highly efficient for prediction at an atomic scale despite heavy computational time for capturing the simulation trajectory. Structural and dynamical properties of the solvated protein-peptide complexes (atomic fluctuations, solvent-accessible surface area and hydrogen bonding pattern) were found to be in qualitative agreement with the available binding data in these cases. However, this approach is not suitable for high throughput screening of proteome.

Self consistent ensemble optimization (SCEO)

Lee and McConnell (1995) proposed a general model of class II–associated invariant chain peptide (CLIP) binding to HLA DR3 antigens using molecular modeling calculations **[7]**. In this approach, modeling calculations were performed with self-consistent ensemble optimization (SCEO) **[8, 9]** using the program CARA in the molecular visualization/modelling software LOOK (Molecular Application Group, Palo Alto, CA). SCEO described by Lee & Subbiah (1991) and Lee (1994) predicts side chain conformation by packing optimization **[8, 9]**. The HLA structure from Protein databank (PDB) was used as a template for building the HLA receptor structures. The test peptide was modeled within the peptide-binding groove by using an automated procedure that predicts the coordinates of side chains upon a fixed backbone framework as described elsewhere by Lee, Subbiah & Levitt **[7, 9, 10]**. This method was used by Ren *et al.,* (2000) to test the binding of minor histo-compatibility antigen (mHag) peptides HA-1[H] (**VLHDDLLEA**) and HA-1[R] (**VLRDDLLEA**) to HLA A alleles in the context of bone marrow transplantation (BMT) **[11]**. The binding of peptides HA-1[H] and HA-1[R] to HLA A*0201 is shown in **Figure 6**. It shows the smooth binding of peptide **VLHDDLLEA** to HLA-A*0201 as against **VLRDDLLEA**. The compatible binding of **VLHDDLLEA** at pocket D without any unfavorable atomic clash is shown. Kangueane *et al.,* (2000) also used the same approach to test 126 peptides binding to 12 HLA alleles (A*0201; A*0202; A*0203; A*0206; A*0301; A*1101; A*3101; A*3301; A*6801; A*8802; B*0801; B*5301) as described elsewhere **[12]**. To build the HLA-A*0201-peptide model, the peptide residues (10 mer) were modelled on the Hepatitis B nucleocapsid peptide backbone coordinates from 1HHH as template without disturbing the HLA-A*0201 coordinates. Structural similarity between the class I co-crystals solved to date allows modelling procedures to be carried out for other HLA alleles, using known structural templates. In fact, within the peptide binding domain the class I alleles share 80%–90% identity with HLA-A*0201. Crystal

structures of HLA-A*6801 (pdb code 1TMC), HLA-B*3501 (pdb code 1A9E), and HLA-B*5301 (PDB code 1A1M) were used as template for constructing the structures for HLA-A68, HLA-B35 and HLA-B53 alleles, respectively. Similarly, the corresponding allele specific peptides were modeled based on backbone templates chosen from subsets grouped by Schueler-Furman and co-workers (1998) **[13]**. This experiment established the usefulness of the modeling procedure to large scale HLA-peptide binding predictions towards EPITOPE design.

Free energy (FE)

Margalit, Altuvia and Schueler-Furman (1995; 1997; 1998; 2000) showed the use of structure modeling and free energy calculations for HLA-peptide binding prediction and scoring **[13-16]**. An approach based on free energy involving threading of the peptides using known templates followed by the evaluation of their binding by statistical pair-wise potentials was described **[13-16]**. However, direct prediction of binding affinity values is not possible in this approach. This is because as it is difficult to develop a universal free energy function for HLA-peptide binding. However, it should be noted that there are few attempts described elsewhere by Rognan *et al.,* (1999) **[17]** for free energy estimations.

Knowledge based free energy

The knowledge based scoring method using pair-wise contacts use solved structures for HLA-peptide binding calculations. The physical, chemical compatibility between peptide and HLA groove is estimated using pair-wise potential tables. The binding score is obtained by adding all pair-wise values for residues in the pocket with the corresponding peptide residues at every position in 3D perspective. This enables the ranking of peptides for HLA-peptide binding as described by Altuvia *et al.,* (1997) **[15]**; Schueler-Furman *et al.,* (1998) **[13]**. This method is based on: (1) the availability of appropriate peptide structural templates; and (2) the choice of a pair-wise potential tables.

Various knowledge based pair-wise potentials have been derived from known protein structures **[18-19]**. A basic approximation underlying these potentials is that total 'free energy' of a protein can be expressed as a sum of independent pair-wise interactions. The frequencies of residues pairs in the structures are assumed to represent the interaction preference between different types of residues. This interaction preference between two amino acids is expressed by its comparison with their affinity to a 'reference state'. Various matrices have been published using distinct reference states **[19]**. Miyazawa and Jernigan (1985; 1996) used solvent as
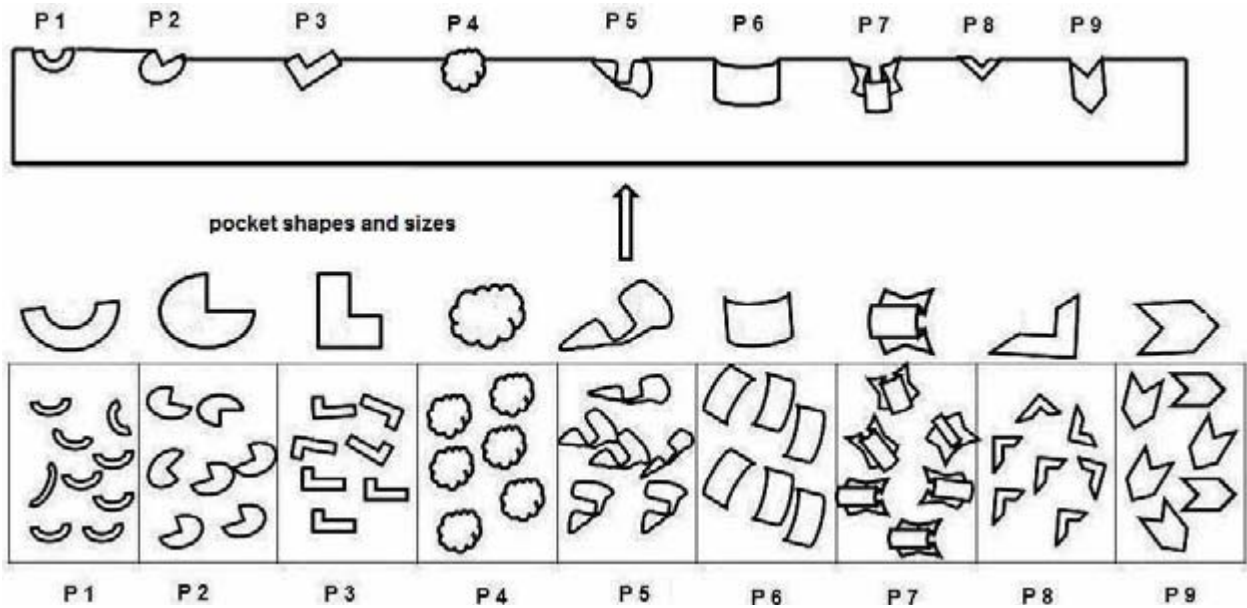
Fig. 4. The binding groove contains binding pockets to accommodate position specific peptide residues. Different alleles have varying pockets with changing pocket shapes and sizes which is deterministic of specificity and sensitivity among ethnicity. The selection of pockets with variable shapes and sizes for each allele as result of HLA sequence polymorphism is schematically shown
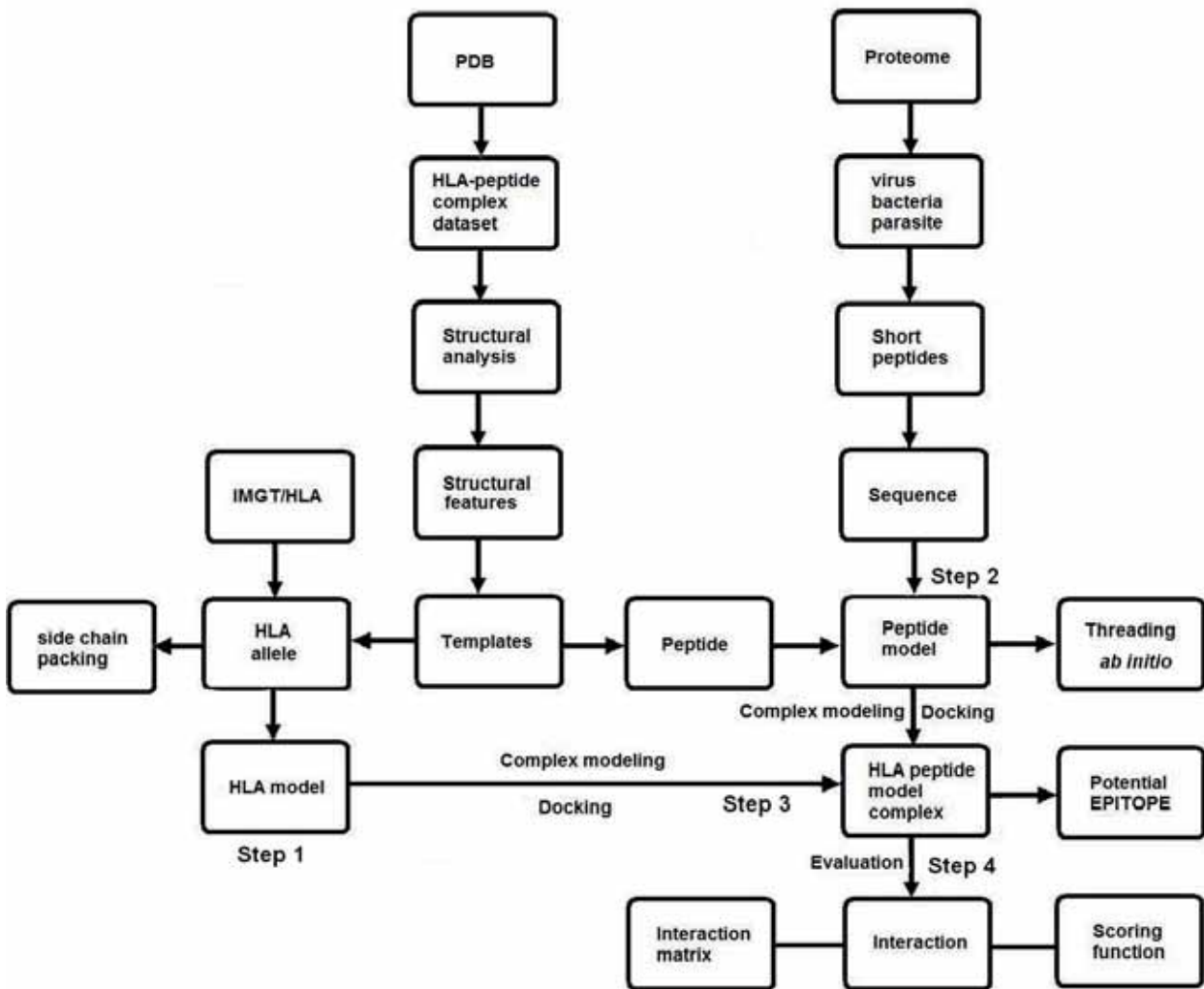
Fig. 5. T-cell epitope design workflow using structural and modeling principles. The workflow is illustrated in four steps: (1) Step 1 – HLA model building; (2) Step 2 – peptide model building; (3) Step 3 – HLA-peptide complex model; (4) Step 4 – Binding/Interaction scoring. PDB = Protein Data Bank; HLA = Human Leukocyte Antigen; IMGT = International ImMunoGeneTics

OPENING PAPER
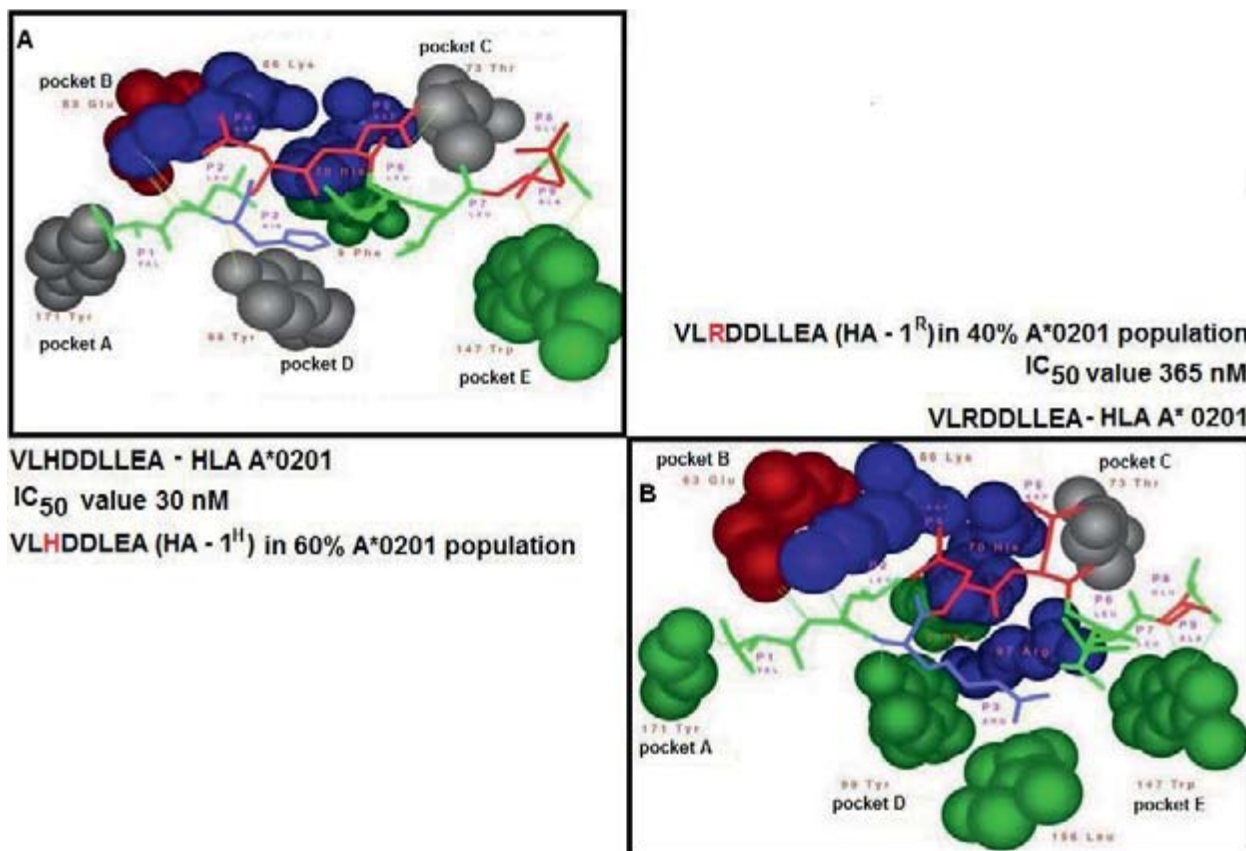
Fig. 6. Modeling interaction of peptides **(A)** HA-1[H] (VLHDDLLEA) and **(B)** HA-1[R] (VLRDDLLEA) binding to HLA A*0201 is shown. The binding of HA-1[H] to HLA-A*0201 is 12 fold higher than that of HA-1[R] as described elsewhere by Dan et al., (1998) **[28]**. The molecular modeling calculations of HA-1[H] (VLHDDLLEA) and HA-1[R] (VLRDDLLEA) to different HLA-A alleles has been shown elsewhere by Ren *et al.,* (2000) **[10]**

reference state and developed a matrix with emphasis on hydrophobic interactions **[20-21]**. Betancourt and Thirumalai (1999) modified the table by changing the reference state to a defined, single solvent-like molecule **[22]**. The resulting matrix represents hydrophilic interactions. Altuvia *et al.* (1995; 1997) used the potential tables of Miyazawa and Jernigan (1985; 1996) **[20-21]** to rank modeled HLA-peptide structures **[14-15]**. This procedure failed to predict hydrophilic interactions **[15]**. The pair-wise potential table of Betancourt and Thirumalai (1999) **[22]** have been reported to have successfully selected hydrophilic interactions as described by Schueler-Furman *et al.,* (2000) **[16]**.

Free energy scoring function

A number of free energy scoring functions has been developed for different purposes and these functions have been used for HLA-peptide binding predictions. One of the approaches for HLA-peptide binding prediction is a tailor made free energy function (FRESNO) combined with homology modeling as described by Logean *et al.,* (2001) **[23]** and Rog-

nan *et al.,* (1999) **[17].** Starting from the primary sequence of the protein antigen, individual 3D structures of all possible HLA-peptide complexes are constructed by homology modeling. The FRESNO scoring function has been used to calculate binding free energy of HLA-peptide interactions. The approach allows for the prediction of absolute binding affinities in a high throughput manner. An extension of this work is EpiDock described by Logean & Rognan (2002) **[24]** which is shown to predict potential T-cell epitopes from viral proteomes and can roughly predict unknown peptide binding motifs. However, the critical issue in this approach is the selection of appropriate peptide templates for HLA-peptide structure prediction.

Computational combinatorial ligand design (CCLD)

Computational combinatorial ligand design (CCLD) described by Zeng *et al.,* **(2001) [25]** has been used for placing amino-acids in specific pockets. It uses 3D information from the crystal structure of HLA molecules. Using chemical fragments as models

for amino-acids, a set of peptides predicted to bind the HLA groove were produced. The results showed that CCLD is sensitive to capture important features of peptide binding in sequence and structure.

3D QSAR

A method based on three-dimensional quantitative structure affinity relationship (3-D QSAR) of HLA-peptide complexes by Doytchinova & Flower, (2001) **[26]** has also been described. Three-dimensional quantitative structure activity relationship **(3D QSAR)** studies have been applied to explore the molecular interactions between HLA and peptides. They provide coefficient contour maps identifying areas of the peptides that require a particular physicochemical property to increase binding. Flower and Doytchinova **[26]**, **[27]** applied comparative molecular similarity indices analysis (CoMSIA) for HLA–peptide-binding prediction. CoMSIA uses the interaction potential around aligned sets of 3D peptide structures to describe the contributions to binding. Five types of similarity index (steric bulk, electrostatic potential, local hydrophobicity, and hydrogen-bond donor and hydrogen bond acceptor abilities) were calculated, using a common probe atom with 1 Å radius, charge +1, hydrophobicity +1, and hydrogen-bond donor and acceptor properties +1 in this method. CoMSIA can predict the binding affinity of a peptide with a residue not presented in the initial training set. However, it cannot assess the contribution of residues at each position and the interactions between them. CoMSIA also returns 3D representations for visual inspection. This approach is partly data-driven and is dependent on the quality of binding data.

Structure based virtual Pockets (SBVP)

Kangueane and colleagues have been studying the structural features of HLA-peptide binding since 2000 **[11, 12, 28, 29, 31].** This helped in the understanding the nature of structural pockets accommodating the peptide binding to the groove. Zhao *et al.,* (2003) [28] and Mohanapriya *et al.,* (2009) **[29]** used a dataset of structures to define structure based virtual pockets for peptide binding towards the development of a T-EPITOPE designer **[32].** In this method, nine structural pockets are defined, and the binding affinity between HLA (sequence defined) and peptide (sequence defined) is given by the sum of residue – residue compatibility between peptide residues and corresponding pockets. The residue positions in HLA forming the pockets are conserved among HLA alleles and polymorphism among them disturbs only the side chains leaving the back bone intact. The quantification of the interaction between the HLA pocket and peptide residue pair is calculated by the application of the Q matrix

(a 20 * 20 property matrix developed using 237 physico-chemical amino-acids properties) described by Mathura & Braun (2001) using Eigen vectors **[32]**. This method produced a good positive predictive value (PPV) and the procedure is highly suitable for high throughput screening of proteomes in the identification and design of potential T-cell Epitopes with specificity for all known HLA alleles.

Issues and caveats

The key issue in the design of T cell epitope is the physical binding of peptide to HLA molecules through chemical features. The prediction of HLA–specific peptide binding involves the building of peptide structure in the groove followed by the estimation of peptide binding to allele-specific HLA structure. Technical difficulties include the lack of sufficient templates for building peptide structures in the groove. This is in addition to the unavailability of an appropriate interaction function or matrix for HLA peptide–binding score calculations. It should be noted that the building of molecular models for HLA peptide–binding prediction is also computationally expensive because of atom level calculations using molecular mechanics equations.

Conclusion

The current challenge in synthetic vaccine design is the development of a methodology to identify and test short antigen peptides as potential T-cell epitopes. Many allele specific data driven statistical models and 3D structure based procedures are available for HLA-peptide binding prediction and calculations. Data driven statistical methods such as Artificial Neural Network (ANN), Hidden Markov Model (HMM), Support Vector Machine (SVM) and matrix methods are generally available for A*0201, A*1101, A*0301, B*0801, DR1, and DR4. The merits and limitations of data-dependent methods have been extensively documented elsewhere. The non-availability of HLA class specific peptide binding data for most alleles is the difficulty in data-driven methods. The 3D structure based procedures are generally adopted for any given allele as the prediction is based on molecular energy functions developed using first principles. The use of MDS, SCEO, FE, CCLD, 3D QSAR and SBVP is discussed for HLA–peptide-binding prediction. The complete mapping of HLA-specific short antigenic peptides in a given viral/bacterial genome will ultimately result in the rational identification of immunogenic epitopes. Although, peptide–HLA specificity plays an important role in the generation of the cell-mediated immune response, other coupled and undoubtedly important mechanisms such as antigen processing, peptide transport, loading of peptides to HLA molecules, and the phenomenon of T-cell receptor (TCR) rep-

OPENING PAPER

ertoires need to be considered. A better understanding of HLA specificity and the multitude of variables associated with the cell-mediated immune response will lead to the development of a methodology for the generation of a library of epitopes *in silico* that could be tested for HLA–peptide binding and immuno-genicity. The successful sampling of short antigenic peptides from a pool of viral/bacterial genome sequence using computational tools will aid in faster and cost-effective means of identifying immunogenic peptides that could be further tested using in vivo models, for consideration as vaccines and therapeutics. The application of mathematical models using computational tools is rapidly advancing to uncover the hidden mystery behind cell-mediated immune response. However, caveats regarding model refinement and cross-validation have to be addressed in future.

References

1. Govindarajan K.R., Kangueane P., Tan T.W., Ranganathan S. (2003), MPID: MHC-Peptide Interaction Database for sequence-structure-function information on peptides binding to MHC molecules. *Bioinformatics* **19:** 309.
2. http://www.syfpeithi.de/.
3. http://www.ebi.ac.uk/imgt/hla/.
4. Abagyan R.A., Batalou S. (1997), Do aligned sequences share the same fold? *J. Mol. Biol.* **273:** 355.
5. Rognan D., Scapozza L., Folkers G., Daser A. (1994), Molecular dynamics simulation of MHC-peptide complexes as a tool for predicting potential T cell epitopes. *Biochemistry* **33:** 11476.
6. Ponder J.W., Case D.A. (2003), Force fields for protein simulations. *Adv. Prot. Chem.* **66:** 27.
7. Lee C., McConnell H.M. (1995), A general model of invariant chain association with class II major histocompatibility complex proteins. *Proc. Natl. Acad. Sci.* **92:** 8269.
8. Lee C., Levitt M. (1991), Accurate prediction the stability and activity effects of site-directed mutagenesis on a protein core. *Nature* **352:** 448.
9. Lee C. (1994), Predicting protein mutant energetics by self-consistent ensemble optimization. *J. Mol. Biol.* **236:** 918.
10. Lee C., Subbiah S. (1991), Prediction of protein side-chain conformation by packing optimization. *J. Mol. Biol.* **217:** 373.
11. Ren E.C., Kangueane P., Kolatkar P., Lin M.T., Tseng L.H., Hansen J.A. (2000), Molecular modeling of the minor histocompatibility antigen HA-1 peptides binding to HLA-A alleles. *Tissue Antigens* **55:** 24.
12. Kangueane P., Sakharkar M.K., Lim K.S., Hao H., Lin K., Chee R.E., Kolatkar P.R. (2000), Knowledge-based grouping of modeled HLA peptide complexes. *Hum. Immunol.* **61:** 460.
13. Schueler-Furman O., Elber R., Margalit H. (1998), Knowledge based structure prediction of MHC class I bound-peptides: a study of 23 complexes. *Fold Des.* **3:** 549.
14. Altuvia Y., Schueler O., Margalit H. (1995), Ranking potential binding peptides to MHC molecules by a computational threading approach. *J. Mol. Biol.* **249:** 244.
15. Altuvia Y., Sette A., Sidney J., Southwood S., Margalit H. (1997), A structure-based algorithm to predict potential binding peptides to MHC molecules with hydrophobic binding pockets. *Hum. Immunol.* **58:** 1.
16. Schueler-Furman O., Altuvia Y., Sette A., Margalit H. (2000), Structure-based prediction of binding peptides to MHC class I molecules: application to a broad range of MHC alleles. *Protein Sci.* **9:** 1838.
17. Rognan D., Lauemoller S.L., Holm A., Buus S., Tschinke V. (1999), Predicting binding affinities of protein ligands from three-dimensional models: application to peptide binding to class I major histocompatibility proteins. *J. Med. Chem.* **42:** 4650.
18. Jones D.T., Thornton J.M. (1996), Potential energy functions for threading. *Curr. Opin. Struct. Biol.* **6:** 210.
19. Skolnick J., Jaroszewski L., Kolinski A., Godzik A. (1997), Derivation and testing of pair potentials for protein folding. When is the quasichemical approximation correct? *Protein Sci.* **6:** 676.
20. Miyazawa S., Jernigan R.L. (1985), Estimation of effective interresidue contact energies from protein crystal structure, quasi-chemical approximation. *Macromolecules* **18:** 534.
21. Miyazawa S., Jernigan R.L. (1996), Residue-residue potentials with a favorable contact pair term and an unfavourable high packing density term, for simulation and threading. *J. Mol. Biol.* **256:** 623.
22. Betancourt M.R., Thirumalai D. (1999), Pair potentials for protein folding; choice of reference states and sensitivity of predicted native states to variations in the interaction schemes. *Protein Sci.* **8:** 361.
23. Logean A., Sette A., Rognan D. (2001), Customized versus universal scoring functions: application to class I MHC-peptide binding free energy predictions. *Bioorg. Med. Chem. Lett.* **11:** 675.
24. Logean A., Rognan D. (2002), Recovery of known T-cell epitopes by computational scanning of a viral genome. *J. Comput. Aided Mol. Des.* **16:** 229.
25. Zeng J., Treutlein H.R., Rudy G.B. (2001), Predicting sequences and structures of MHC-binding peptides: a computational combinatorial approach. *J. Comput. Aided Mol. Des.* **15:** 573.
26. Doytchinova I.A., Flower D.R. (2001), Toward the quantitative prediction of T-cell epitopes: coMFA and coMSIA studies of peptides with af-

finity for the class I MHC molecule HLA-A*0201. *J. Med. Chem.* **44:** 3572.

27. Doytchinova I.A., Flower D.R. (2002), A comparative molecular similarity index analysis (CoMSIA) study identifies an HLA-A2 binding supermotif. *J. Comput. Aided Mol. Des.* **16:** 535.

28. Zhao B., Mathura V.S., Rajaseger G., Moochhala S., Sakharkar M.K., Kangueane P. (2003), A novel MHCp binding prediction model. *Hum. Immunol.* **64:** 1123.

29. Mohanapriya A., Lulu S., Kayathri R., Kangueane P. (2009), Class II HLA-peptide binding prediction using structural principles. *Hum. Immunol.* **70:** 159.

30. Den Haan J.M., Meadows L.M., Wang W., Pool J., Blokland E., Bishop T.L., Reinhardus C., Shabanowitz J., Offringa R., Hunt D.F., Engelhard V.H., Goulmy E. (1998), The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. *Science* **279:** 1054.

31. Kangueane P., Sakharkar M.K., Kolatkar P.R., Ren E.C. (2001), Towards the MHC-peptide combinatorics. *Hum. Immunol.* **62:** 539.

32. Venkatarajan M.S., Braun W. (2001), New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical-chemical properties. *J. Mol. Model.* **7:** 445.

Bioinformatics

LIGANDS DATABASE

DAMIAN MARCHEWKA[1,2], IRENA ROTERMAN[1]

[1] *Department of Bioinformatics and telemedicine, Collegium Medicum – Jagiellonian University, Lazarza 16, 31-530 Krakow, Poland*
[2] *Faculty of Physics, Astronomy and Applied Computer Science, Reymonta 4, 30-059 Krakow, Poland*

Abstract: The database collecting proteins complexed to any kind of ligand is presented. The search for selected, specific protein including enzymes, protein-protein complexes is possible. The server is publicly available. The development of the database is planned according to users' experiences. The development of any form of ligand databases is of high importance nowadays particularly when medicine is facing the important challenge of individual therapy projected for each patients individually taking into account his/her specificity.

Introduction

The interaction of proteins with any kind of non-protein molecules is critical for their biological activity. The high specificity of this type of interaction ensures the low level of errors and eliminate the random complexation (this is the basic knowledge available in each biochemistry manual [1]). The computer aided drug design is based on the analysis of the ligand-protein complexation creating new molecules able to be complexed with the target protein molecule and modifying their biological activity. The aim-oriented modification of the protein activity is the main goal for molecules called drugs in medicine. The main problem with drug molecules is their low specificity. The expected target molecule is not the only one which is able to complex the drug molecule.

This is why the global search for natural protein ligands is of high importance for new drug design. The analysis and characteristics of the cavities able to complex the ligands may suggest the possible unexpected new ligand (drug) binding [1].

The aim for the constructed database is to monitor the structures of protein complexes with ligands to allow the large scale comparative analysis in case of new ligand synthesis. The prediction of the possible unexpected (not related to the planned therapy) complexation reaction may help to elaborate the techniques of the protective procedure before the drug in injected into the organism. The examples of the unproper metabolits of some drugs (aspirin) may be the source of discomfort and even the allergic reaction [2].

The database in the presented form is in its initial status which is the tool t search for the proteins complexed to particular ligand. The development of the database is planned according to the experiences claimed by the users expecting possible new tools for the database managing as well as the calculation of hydrophobicity distribution based identification of ligand binding localization [3].

Database description

The original data base is the Protein Data Bank (PDB) [4]. Before the local database may be constructed the record of information shall be defined. Particularly the key word for the search must be defined. In this case the information record is as follows:

Local ID | PDB ID | protein name | biological activity | EC code | ligand name | short name | type of ligand |

Record's ID – the specific parameter expected to identify (the number valid in the local database).

PDB – ID

The standard PDB database ID – introduction of his identification allows differentiation of the protein complexed to different ligands [4].

Name of protein

The search for the particular type of protein is possible using its traditional name for example: "ubiquitine". After introduction of the protein name the following information appears available:

Biological function

The definition of the biological functions produces the selection of the proteins belonging to the particular biological activity. For example: reductase produces the list of all reductases present in PDB.

EC ID

The enzyme code according to the enzyme coding system. If the protein is not any enzyme this position in the record stays empty.

Ligand name

The name of the ligand molecule given in this dialog window produces the list of proteins complexed to particular ligand

Short name of ligand

The short name (chemical formula – for example CI is the abbreviation of the CHLORIDE ION) put in dialog window produces the list of proteins complexed to the selected ligand.

Ligand type

The chemical characteristics of the ligand (hydrocarbons, nucleic acids, vitamins etc.)
The scheme of the dialog window allowing the search for particular protein in PDB according to the name of ligand is show in Fig.1.

General description of the database

The results – the list of proteins selected to satisfy the defined conditions – are shown in Fig.2
Tab.1. shows the summary for ligand – and protein-oriented search In form of the relation between search criterion and the final results of the search.
The database is produced on the basis of MySQL although the application was performer using PHP language [5]. The search for the current version of the PDB deposits aimed on the New structures search is programmed using the C++ language [6] The graphic interface is done on the basis of standard scheme available on [7].

Other ligand databases

PDBsum

Many databases are available in internet. One of the most interesting is the PDBsum, which collects information about all proteins in the form as they are deposited in PDB. The very useful option available in this database is the information about
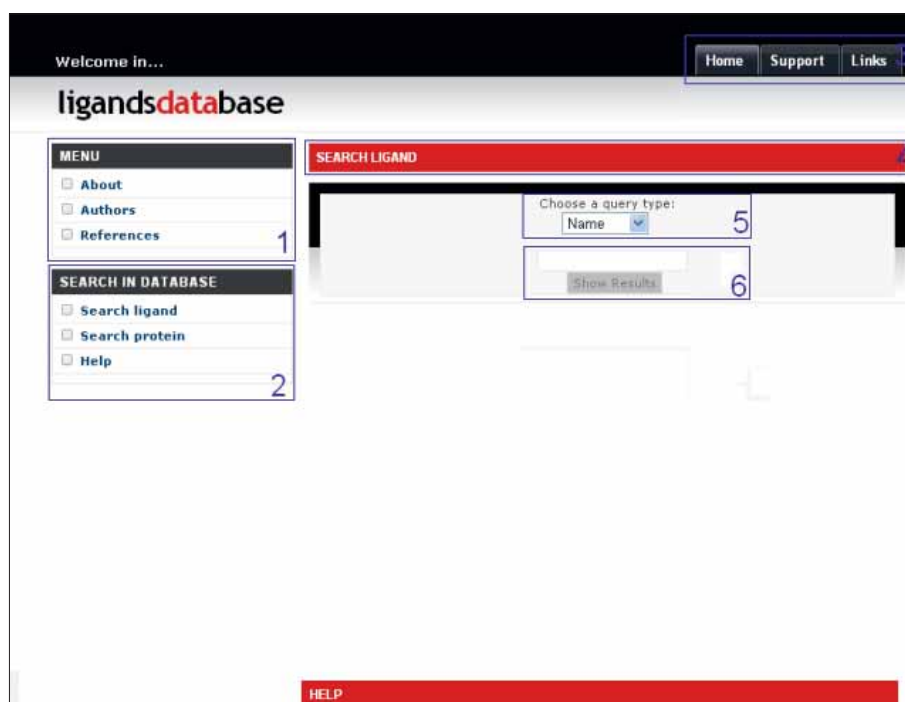
Fig. 1. The schematic presentation of the dialog window, which is available to the user when the search procedure (accordant to Category) is prepared: (1) – main menu, (2) – selection of the search method, (3) – additional menu, (4) – information describing the search, (5) – selection of the search category (PDB ID, name, EC code), (6)– criterium for the search

Tab. 1. The list of proteins with the information presenting the search criterion and appropriate results depending on the key word applied for the search

Ligand search		
Parameters for selection		
Name	ID (PDB)	Category
Results of selection		
– Ligand ID (PDB) – Category List of proteins with the description: – Protein name – Protein ID (PDB) – Biological function – EC code (if applicable)	– Ligand name – Category List of proteins with the information: – Protein name – Protein ID (PDB) – Biological function – EC code (if applicable)	List of ligands belonging to particular type of ligand: – Ligand name – Ligand ID (PDB)
Protein serach		
Parameters for selection		
Name		EC Code
Search results		
– Biological function – EC code (if applicable) Format of data: – Ligand name – Ligand ID (PDB) – Category		– Peptide name
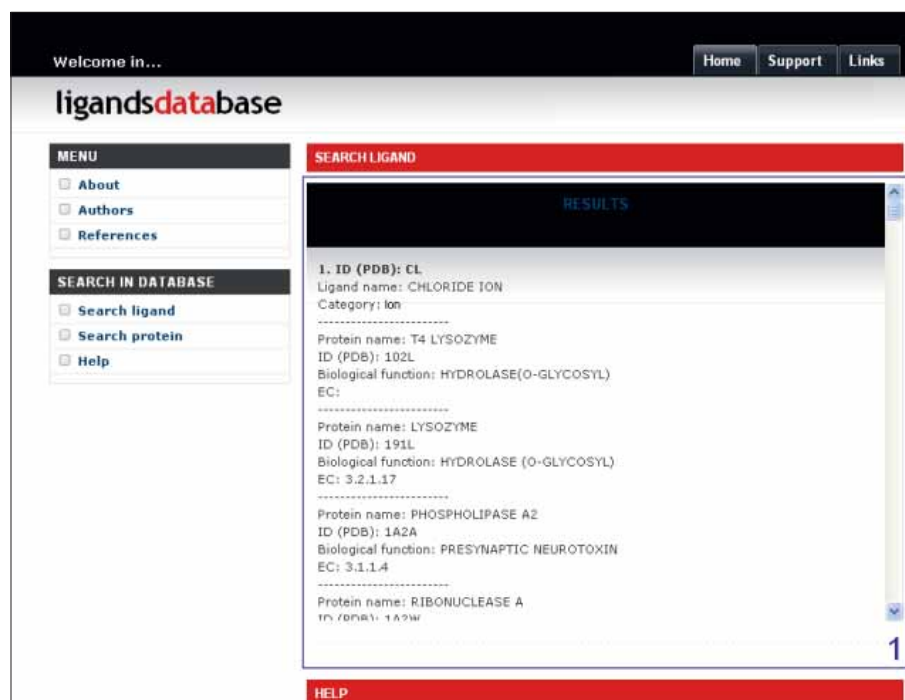
Fig. 2. The schematic presentation of the dialog window which appears after correctly performed search for protein complexed to selected ligand with the list of proteins selected for Ligand name "Chloride ion" and category "Ion"

Bioinformatics

inter-molecular interaction. The list of residues engaged in the complexation is given in quite precise way [8].

CSA

The Catalytic Site Atlas – the database oriented on the classification of enzymes with the precise description of the residues engaged in the catalytic reaction. This information is given using experimental data described in the publications [9].

Summary

Although there are many databases oriented on ligand classification the every new one is important due to the large differentiation of the problems related to the basic research as well as all forms of applied biological sciences including pharmacology and medicine. The recognition of the possible ligand complexation mechanisms is of high importance for the pharmacology particularly interested in minimization of unfavorable site effects. Medicine is facing the new challenge which is the individual therapy. Each person differs versus any other. The drug therapy for one person is not working for another one sometimes. The influence on the small (SNP – single nucleotide polymorphism) difference in genetic material influences the protein activity. The reaction of the proteins to the drug seems to be difficult to be generalized. This is why the global analysis of proteome (understood as the collection of all proteins present in particular organism) is of critical meaning for the future medicine including pharmacology in particular.

The Ligand data base was presented formerly [10] although the aim of that database is different. The ligand before being applied to calculations should be described in the sense of their parameterization. The data base [10] was planed to be the tool for the specialists calculating the partial charges distribution, vdW parameters, torsional parameters and others. The data base presented in this paper is aimed to be applied to the analysis of the ligand interactions with proteins. The PDBSum is the most popular and powerful tool for this subject. Although the analysis presented there is based on the standard models. The home made model describing the ligand-protein interaction is planned to be verified on large scale calculation. This is why the search for proteins complexed to the common ligand is necessary. So far the search for proteins selected form PDB according to the criteria which are related to the ligand molecule is presented.

Availability: http://crick.cm-uj.krakow.pl/ligandsdb/.

References

1. Berg J.M, Tymoczko J.L, Stryer L. (2007), Biochemistry. W.H. Freeman and Company.
2. Jurkowski W., Obtułowicz K., Porebski G., Roterman I. (2009), Molecular Dynamics Study of Acetylsalicylic Acid Metabolites complexed with human serum Albumin. *Current Drug Metabolism* 10, 448-458.
3. Bryliński M., Konieczny L., Roterman I. (2006), Ligation site in proteins recognized in silico. *Bioinformation* 1, 127--129.
4. PDB: http://www.rcsb.org/pdb/home/home.do.
5. Welling L., Thomson L. (2005), PHP e MySQL Desenvolvimento Web. Polish Edition Translation: D. Kaczmarek, P. Gonera, Helion, Gliwice.
6. Grębosz J. (2006), Symfonia C++ Standard, Edition 2000, Kraków.
7. http://styleshout.com.
8. PDBSum: http://www.ebi.ac.uk/pdbsum/.
9. CSA: http://www.ebi.ac.uk/thornton-srv/databases/CSA/.
10. Broniatowska E., Stępień O., Roterman I. (2006), The ligand database Server. *Bio-Algorithms and Med.-Systems* 2, 21-24.

Bioinformatics

ASSOCIATION BETWEEN THE N-ACETYLATION GENETIC POLYMORPHISM AND ASPIRIN-INDUCED URTICARTIA

GRZEGORZ PORĘBSKI[1], KRYSTYNA OBTULOWICZ[1], MONIKA PIWOWAR[2],

IRENA ROTERMAN-KONIECZNA[2]

[1] *Jagiellonian University, Medical College, Department of Clinical and Environmental Allergology, Jagiellonian University, Sniadeckich 10, 31-531 Cracow, Poland*
[2] *Jagiellonian University, Medical College, Department of Bioinformatics and Telemedicine, Jagiellonian University, Łazarza 16, 31-530 Cracow, Poland*

Abstract: N-acetylation represents one of the key pharmacogenetic traits in metabolism of xenobiotics, including drugs. Dominance of 'slow-acetylators' has been demonstrated in atopic patients, contact allergy and has been investigated in patients with sulfonamide hypersensitivity. The aim of this survey was to assess genotype of N-acetyltransferase2 (NAT2) polymorphism in patients with aspirin-induced urticaria (AIU).

To the study we included 24 consecutive patients with medical history of urticaria induced by aspirin and 123 healthy controls. In the patients' group oral provocation tests (OPT) up to 565mg cumulative dose of acetylsalicylic acid were performed. The NAT2 alleles (*4-wilde type, *5, *6 and *7) were determined by polymerase chain reaction-restriction fragment length polymorphism method with DNA extracted from peripherial blood.

It was observed 53.7% of slow acetylators in the control group, and respectively 66.7%, 83.3% and 100% in the patients with AIU, in patients with positive OPT and in patients with positive reaction in OPT assessed as severe. Statistical significance association between NAT2*5/NAT2*6 genotype and aspirin-induced urticaria was revealed.

The linkage of AIU and slow acetylation (NAT2*5/NAT2*6) seems to be highly probable in the patients with medical history of skin hypersensitivity to aspirin confirmed by positive oral provocation test.

Key words: urticaria, acetylsalicylic acid, NAT2, polymorphism

Background

Drugs are one of the most common xenobiotics affecting human health nowadays. Beside positive effects, they can cause wide variety of hypersensitive reactions ranging from maculopapular exanthema to anaphylactic shock which may be fatal (1). On the other hand, more 'over the counter' drugs are available in uncontrolled condition, including aspirin (ASA) and other nonsteroidal anti-inflammatory drugs. The two frequent clinical presentations of aspirin hypersensitivity are: aspirin-induced bronchial asthma/rhinosinusitis (AIA) and aspirin-induced urticaria/angioedema (AIU) (2). In contrast to AIA, progress in understanding urticaria sensitive to aspirin has been slow (3). The therapeutic and the certain adverse actions of acetylsalicylic acid are related to its ability to inhibit cyclooxygenase (COX) in various tissues and to participate in transacetylation reactions. Acetylation of COX results in irreversible inhibition of this enzyme and anti-inflammatory effects, but acetylation of different circulating proteins may also result in a hypersensitivity response (4). Acetylation polymorphism was shown to be associated with atopic diseases (5), especially bronchial asthma. N-acetylation status was explored as

a susceptibility factor for the development of allergy to para-phenylenediamine (6) and sulphamethoxazole hypersensitivity (7). N-acetylation polymorphism is one of the oldest and most intensively studied pharmacogenetic traits. Hepatic N-acetyltransferase 2 (NAT2) takes part in the detoxication of some drugs and xenobiotics (8, 9). Variation in the elimination of xenobiotics, as well as in NAT2 activity in the liver and other tissues is genetically determined and responsible for fast and slow acetylators phenotype (10, 11). The human NAT2 gene contains an 870 base pairs protein-coding region (12). To date, one allele with a code for fast acetylation (wild-type) and about 20 mutated alleles coding impaired n-acetyltransferase have been discovered (13, 14, 15). In white subjects three of all the NAT2 allelic variants (NAT2*5, NAT2*6, NAT2*7) have been demonstrated to account for majority of the slow NAT2 acetylator genotypes (15). Because of the high frequency of mutations, xenobiotic acetylation defect is observed in about 50% of white individuals (11, 16).

Genetic and environmental factors jointly influence the risk of developing many human diseases e.g. drug hypersensitivity that focuses attention of researchers nowadays (17). Because drug hypersensitivities belong to complex diseases consistent

with genetic variation and gene-environment interactions (18) we decided to study the N-acetylation genetic polymorphism in the group of patients with aspirin–induced urticartia derive from Polish population.

Associations analysis (Case-control studies of unrelated subjects, unhealthy and healthy) of particular allelic variants of NAT2 gene with observed clinical symptoms were one of the main aim of our research. Statistical analysis that is now widely used to study the role of genetic susceptibility and gene-environment interactions in the etiology of complex diseases was applied to assess it (19).

Material and methods

We studied 24 consecutive, adult patients (18 women and 6 men; age range: 28 to 64 years, mean ± SD: 42,8±13,7) with anamnesis of urticaria induced by aspirine. The patients were diagnosed at the hospital (Allergology Division, University Hospital, Cracow, Poland). The thorough evaluation of subjects included oral provocation test (OPT) with aspirin, as well as routine biochemical blood analysis, physical signs and symptoms. Oral, single-blind, placebo-controlled diagnostic challenge tests with aspirin were performed following *European Academy of Allergology and Immunology* guidelines referring to contraindications and drug withdrawal times (21). Four exponentially increasing doses of aspirin (40 mg, 75 mg, 150 mg, 300 mg) were administered every 1,5 h until a cumulative dose of 565 mg was reached. The day before challenges with placebo according to the same schedule were performed. Any skin eruptions were carefully monitoring throughout the challenge procedure and for 12 h after the last dose of aspirin. The challenge procedure was interrupted if cutaneous reactions appeared. OPT was considered as positive in case of occurrence pruritus and/or wheals within monitored period. Reactions were assessed as severe if pruritus and/or wheals were intense (>50 wheals/the period of challenge or large confluent areas of wheals). Cutaneous adverse symptoms are relived by the administration of antihistaminics. The control group was composed of 123 healthy subjects (80 women and 43 men; age range: 18 to 77 years, mean ± SD: 27,7±14,6). All subjects were white and came from Polish population. The study was approved by the local bioethics committee.

There was analyzed polymorphism of N-acetyltransferase 2 (NAT2) gene (AC: NM_000015) located on 8p22 humane chromosome among patients coming from Polish population with anamnesis of urticaria caused by aspirine. Genetic variability – single nucleotide polymorphisms NAT2 gene were analyzed. Data were taken from central repository for both single base nucleotide substitutions and short deletion and insertion polymorphisms – SNP database (20). From among all identified alleles that occur in NAT2 gene in human genetic material there only selected single nucleotides were chosen to measure association between particular alleles/genotype and disease related phenotype. These were G590A (NAT2*6), A803G (NAT2*5), and G857A (NAT2*7) point mutations (Table 1).

NAT2-genotyping was performed according to method described before by Gawronska–Szklarz et al (5). Genotypes were determined by the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assays (22). The RFLP method was used to identify the point mutations in the NAT2 gene (NAT2*5, NAT2*6, NAT2*7). The following restriction endonucleases – *Kpn*I (MBI, Fermentas), *Taq*I (MBI, Fermentas), *Dde*I (Gibco BRL), and *Bam*HI (MBI, Fermentas) were used to detect the presence or absence of mutations at the positions C481T, G590A, A803G, and G857A, respectively. The wild-type allele (NAT2*4) was recognized by the occurence of the restriction sites for the *Kpn*I, *Taq*I, *Dde*I, and *Bam*HI endonucleases.

Statistical analysis. The NAT2*4 allele was considered to be the fast allele, and NAT2*5, NAT2*6, NAT2*7 were considered to be the slow alleles. The distribution of genotype in patients with aspirin-induced urticaria was compared with healthy individuals and statistically evaluated by use of the chi-squared test (χ^2), with Yate's correction and Fisher's exact test for the small groups (StatSoft, Inc. 2003. STATISTICA data analysis software system, version 6). Secondly, accession NAT2 genetic data was performed by using CASECONTROL procedure run in SAS software (SAS 9.1 version; module SAS/GENETICS) that is designed for interpretation of marker data to help point out the genes that affect susceptibility to a disease. There were applied three statistical methods for testing association of particular markers with a disease locus. It was Armitage's trend test, the allele case-control test and the genotype case-control test to measure association particular alleles/genotypes with disease phenotype.

Results

We genotyped the control population of 123 individuals and 24 patients with aspirin-induced urticaria for NAT2 polymorphism. In 12 patients OPT with aspirin was considered as positive. In 8 of them skin reaction was assessed as sever. Slow acetylators were identified in 53.7% of healthy subjects, in 66.7% of AIU-patients (Table 2) and in 83.3% of patients with positive OPT with aspirin (Table 3). All patients with severe reaction in OPT were slow acetylators. Comparing with the control population, we observed the certain preponderance of slow acetylators in the group of patients with AIU (Table 2). The more evident, but not statistically significant preponderance of slow acetylators was observed in patients with positive OPT (Table 3). In subgroup with severe reaction in OPT, slow acetylators significantly predominated comparing with the control group (p = 0,009).

The distribution of genotypes coding slow and fast acetylation was similar in the AIU-patients and the control group, except frequency of genotype NAT2*5/NAT2*6, that was significantly increased among patients with AIU (χ^2 = 13.93; p < < 0,001). It is interesting, that there were no homozygous fast acetylators (2 wild-type alleles) detected among patients with AIU. Data used to statistical analysis (by using SAS/Genetics procedures) had satisfactory factors. Polymorphic information content (PIC) was 0.5557, heterozygosity (Het) was 0.75 and p-value for Hardy-Weinberg equilibrium test was 0.6169. Data incorporated to genotype test were revealed statistical significance association between genotype and aspirin-induced urticaria (p=0.018, α=0.05) among examined patients. Genotypes frequencies were: NAT2*4/NAT2*5 = = 0.2083; NAT2*4/NAT2*6 = 0.1250; NAT2*5/NAT2*5 = = 0.2083; NAT2*5/NAT2*6 = 0.3750; NAT2*6/NAT2*6 = 0.0417; NAT2*6/NAT2*7 = 0.0417. It suggests association between

NAT2*5/NAT2*6 genotype and aspirin-induced urticaria (what confirmed results obtained by using χ^2 test). Particular alleles association with clinical symptoms seems to be negligible (p=0.161, α=0.05). Trend test also did not deliver important results from statistical point of view (p=0.174, α=0.05).

Discussion

The frequency of homozygous fast, heterozygous fast, and homozygous slow acetylators in the control group was comparable to those obtained in previous studies in healthy individuals from Polish population (23, 24). Similar results were also found in other European populations (25). Cascorbi et al investigated 844 unrelated German subjects and showed genotypes that accounted for slow and fast acetylation in, respectively, 58.9% and 41.1% of the population (15). Bouchardy et al studied 172 control individuals in France and found the following distribution of NAT2 genotypes: 5.8% homozygous fast genotypes, 41.3% heterozygous fast genotypes, and 53.0% homozygous slow genotype (26).

The distribution of the investigated genotypes has been found to be not much different from the values observed in other white subjects (15, 24, 26). The only difference was observed in NAT2*5/NAT2*6 genotype frequency, that was found significantly more frequent in AIU-patients comparing with healthy individuals.

In majority of studies, aimed at prediction of the phenotype in clinical settings, the analysis of the most prevalent mutations that produce a defective NAT2 function is generally accepted as sufficient (26). In this study the NAT2 genotyping analysis was based on the determination of 3 main mutated allelic variants, although approximately 20 have been described to date (13).

Association between acetylation polymorphism and different kinds of hypersensitivity was quite extensively studied. Kawakubo et al showed that N-acetylation status may be a susceptibility factor for the development of allergy to para-phenylenediamine (6). Nacak and coworkers found increased frequency of slow acetylators among extrinsic asthma patients compared with control subjects (27). The other study was undertaken to investigate slow acetylator phenotype and genotype in HIV-positive patients with sulphametoxazol hypersensitivity. Results suggested that acetylation defect is unlikely to predispose to sulphametoxazol hypersensitivity, although its minor role was not excluded (7). Gawronska-Szklarz et al observed the significant predominance of slow acetylators among patients with atopic allergy compared with healthy subjects (5).

Numerous studies indicate that hypersensitivity reactions triggered by aspirin are related to inhibition of prostaglandin synthesis with overproduction of leukotrienes and are associated with inhibition of cyclooxygenase (COX), specifically COX-1 (2). However several reports suggest that the salicyl or acetyl moieties of aspirin may be responsible for development of immuno-antigenicity in aspirin hypersensitivity (28, 29). The acetylation of proteins, lipids and other biomolecules by aspirin has also been suggested as a basis of aspirin hypersensitivity (30, 31). These authors have detected antibodies against aspirin-altered serum albumin in the sera from aspirin-intolerant patients and those with rheumatoid arthritis, who have consumed aspirin. But reaction of aspirin-acetylated albumin

in the pathogenesis of aspirin sensitivity has not been clearly elucidated up to now. In our study we investigated association between the NAT2 genetic polymorphism determining the acetylation phenotype and aspirin–induced urticartia, one of most common clinical manifestation of aspirin hypersensitivity. We did not observe statistically significant differences between patients with episodes of AIU in medical anamnesis and healthy controls (respectively 66.7% and 53.7% of slow acetylator). However, if aspirin hypersensitivity is confirmed by diagnostic 'gold standard': positive oral provocation test, percentage of slow acetylators increases to 83.3%. Then again, taking into consideration the patients with severe positive reactions in OPT, this percentage increases to 100.0%. Therefore there is noticeable tendency: the stronger clinical reaction to aspirin, the higher number of patients with slow acetylation phenotype.

Conclusion

In conclusion, the linkage of AIU and slow acetylation seems to be highly probable in the patients with medical history of skin hypersensitivity to aspirin confirmed by positive oral provocation test. Slow acetylation phenotype is likely to predispose to AIU in these patients. This statement has to be verified in subsequent studies with large groups of patients.

Acknowledgments

References

1. Demoly P., Bousquet J. (2002), Drug allergy diagnosis work up. Allergy 57, Suppl. 72: 37-40.
2. Szczeklik A., Nizankowska-Mogilnicka E., Sanak M. (2007), Hypersensitivity to Aspiryn and Other NSAIDs: Mechanisms, Clinical Presentation and Management, in: Drug Hypersensitivity, W.J. Pichler (ed.). Basel, Karger, pp. 340-349.
3. Mastalerz L., Setkowicz M., Sanak M., Szczeklik A. (2004), Hypersensitivity to aspirin: common eicosanoid alterations in urticaria and asthma. J. Allergy Clin. Immunol. 113: 771-5.
4. Rainsford K.D. (2004), Side Effects and Toxicology of the Salicylates, in: Aspirin and Related Drugs, K.D. Rainsford (ed.). London and New York, Taylor & Francis, pp. 367--554.
5. Gawronska-Szklarz B., Luszawska-Kutrzeba T., Czaja--Bulsa G., Kurzawski G. (1999), Relationship between acetylation polymorphism and risk of atopic diseases. Clin. Pharmacol. Ther. 65: 562-9.
6. Kawakubo Y., Merk H.F., Masaoudi T.A., Sieben S., Blome-ke B. (2000), N-acetylation of paraphenylenediamine in human skin and keratinocytes. J. Pharmacol. Exp. Ther. 292: 150-5.

7. Alfirevic A., Stalford A.C., Vilar F.J., Wilkins E.G., Park B.K., Pirmohamed M. (2003), Slow acetylator phenotype and genotype in HIV-positive patients with sulphamethoxazole hypersensitivity. Br. J. Clin. Pharmacol. 55: 158-65.

8. Blum M., Demierre A., Grant D.M., Heim M., Meyer U.A. (1991), Molecular mechanism of slow acetylation of drugs and carcinogens in humans. Proc. Natl. Acad. Sci. USA. 88: 5237-41.

9. Weber W.W., Hein D.W. (1985), N-acetylation pharmaco-genetics. Pharmacol. Rev. 37: 25-79.

10. Hickman D., Sim E. (1991), N-acetyltransferase polymor-phism. Comparison of phenotype and genotype in humans. Biochem. Pharmacol. 42: 1007-14.

11. Clark D.W. (1985), Genetically determined variability in acetylation and oxidation. Therapeutic Implications. Drugs. Apr 29(4): 342-75.

12. Grant D.M. (1993), Molecular genetics of the N-acetyl-transferases. Pharmacogenetics. 3: 45-50.

13. Vatsis K.P., Weber W.W., Bell D.A. et al. (1995), Nomencla-ture for N-acetyltransferases. Pharmacogenetics. 5: 1-17.

14. Vatsis K.P., Martell K.J., Weber W.W. (1991), Diverse point mutations in the human gene for polymorphic N-acetyl-transferase. Proc. Natl. Acad. Sci. USA. 88: 6333-7.

15. Cascorbi I., Drakoulis N., Brockmoller J., Maurer A., Sper-ling K., Roots I. (1995), Arylamine N-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Cau-casian individuals: correlation with phenotypic activity. Am. J. Hum. Genet. 57: 581-92.

16. Pontes Z.B., Vincent-Viry M., Gueguen R., Galteau M.M., Siest G. (1993), Acetylation phenotypes and biological variation in a French Caucasian population. Eur. J. Clin. Chem. Clin. Biochem. 31: 59-68.

17. Hunter D.J. (2005), Gene-environment interactions in hu-man diseases. Nat. Rev. Genet. Apr 6(4): 287-98. Review.

18. Martinez F.D. (2005), Gene-environment interactions in asthma and allergies: a new paradigm to understand dis-ease causation. Immunol. Allergy Clin. North Am. 25: 709--721.

19. Spinka C., Carroll R.J., Chatterjee N. (2005), Analysis of case-control studies of genetic and environmental factors with missing genetic information and haplotype-phase am-biguity. Genet. Epidemiol. Sep 29(2): 108-27.

20. http://www.ncbi.nlm.nih.gov/SNP.

21. Nizankowska-Mogilnicka E., Bochenek G., Mastalerz L. et al. (2007), EAACI/GA2LEN guideline: aspirin provocation tests for diagnosis of aspirin hypersensitivity. Allergy. 22.

22. Spurr N.K., Gough A.C., Chinegwundoh F.I., Smith C.A. (1995), Polymorphisms in drug-metabolizing enzymes as modifiers of cancer risk. Clin. Chem. 41: 1864-9.

23. Gawronska-Szklarz B., Gornik W., Pawlik A. et al. (1997), N-acetylation and hydroxylation polymorphisms in type II diabetics with microvascular disturbances. Eur. J. Clin. Pharmacol. 51: 431-5.

24. Zielinska E., Niewiarowski W., Bodalski J., Stanczyk A., Bolanowski W., Rebowski G. (1997), Arylamine N-acetyl-transferase (NAT2) gene mutations in children with allergic diseases. Clin. Pharmacol. Ther. 62: 635-42.

25. Vatsis K.P., Weber W.W. (1993), Structural heterogeneity of Caucasian N-acetyltransferase at the NAT1 gene locus. Arch. Biochem. Biophys. 301: 71-6.

26. Bouchardy C., Mitrunen K., Wikman H. et al. (1998), N-acetyltransferase NAT1 and NAT2 genotypes and lung cancer risk. Pharmacogenetics. 8: 291-8.

27. Nacak M., Aynacioglu A.S., Filiz A. et al. (2002), Associa-tion between the N-acetylation genetic polymorphism and bronchial asthma. Br. J. Clin. Pharmacol. 54: 671-4.

28. Schwartz M.A., Amidon G. (1966), Reaction of aspiryn with amines. Potential mechanism. J. Pharm. Sci. 55: 1464-5.

29. Kallos P., Schlumberger H.D. (1978), Immunogenic impuri-ties' in acetylsalicylic acid. J. Pharm. Pharmacol. 30: 67-8.

30. Minden P., Farr R.S. (1967), Human antibodies against acetylsalicylic acid altered human serum albumin. Arthritis Rheum. 10: 399.

31. Flemstorm G., Marsden N.V., Richter W. (1976), Passive cutaneous anaphylaxis in guinea pigs elicited by gastric absorption of dextran induced by acetylsalicylic acid. Int. Arch. Allergy Appl. Immunol. 51: 627-36.

Tab. 1. Polimorphisms in exon2 of NAT2 gene. Data were taken from SNP database (20)

dbSNP rs# cluster id	Hetero-zygosity	Function	dbSNP allele	Protein residue	Codon position	Amino acid position	Nukleotyde number
rs1799929	0.418	synonymous	T	Leu [L]	1	161	481 C→T (NAT*5A)
	0.418	contig reference	C	Leu [L]	1	161	
rs1799930	0.384	nonsynonymous	A	Gln [Q]	2	197	590 G→A (NAT2*6)
	0.384	contig reference	G	Arg [R]	2	197	
rs1208	0.430	nonsynonymous	A	Lys [K]	2	268	803 A→G (NAT2*5B)
	0.430	contig reference	G	Arg [R]	2	268	
rs1799931	0.121	nonsynonymous	A	Glu [E]	2	286	857 G→A (NAT2*7)
	0.121	contig reference	G	Gly [G]	2	286	

Tab. 2. Distribution of patients with aspirin-induced urticaria and control subjects according to NAT2 genotypes

Classes of NAT2 genotypes	Genotypes	Urticaria (n = 24)		Control subjects (n = 123)		p value*
		n	%	n	%	
Homozygous fast acetylator	NAT2*4/NAT2*4	0	-	8	6,5	NS
Heterozygous	NAT2*4/NAT2*5	5	20,8	37	30,1	NS
	NAT2*4/NAT2*6	3	12,5	9	7,3	NS
	NAT2*4/NAT2*7	0	-	3	2,4	NS
total		8	33,3	57	46,3	
Homozygous slow acetylator	NAT2*5/NAT2*5	5	20,8	30	24,4	NS
	NAT2*5/NAT2*6	9	37,5	11	9,0	p< 0,001
	NAT2*5/NAT2*7	0	-	9	7,3	NS
	NAT2*6/NAT2*6	1	4,2	13	10,6	NS
	NAT2*6/NAT2*7	1	4,2	3	2,4	NS
total		16	66,7	66	53,7	NS

NS, No significant differences
* chi-squared test

Tab. 3. Distribution of patients with positive oral provocative test with aspirin and control subjects according to NAT2 genotypes

Classes of NAT2 genotypes	Genotypes	positive OPT (n = 12)		Control subjects (n = 123)		p value*
		n	%	n	%	
Homozygous fast acetylator	NAT2*4/NAT2*4	0	-	8	6,5	NS
Heterozygous	NAT2*4/NAT2*5	2	16,7	37	30,1	NS
	NAT2*4/NAT2*6	0	-	9	7,3	NS
	NAT2*4/NAT2*7	0	-	3	2,4	NS
total		2	16,7	57	46,3	NS
Homozygous slow acetylator	NAT2*5/NAT2*5	4	33,3	30	24,4	
	NAT2*5/NAT2*6	5	41,7	11	9,0	p< 0,001
	NAT2*5/NAT2*7	0	-	9	7,3	NS
	NAT2*6/NAT2*6	1	8,3	13	10,6	NS
	NAT2*6/NAT2*7	0	-	3	2,4	NS
total		10	83,3	66	53,7	NS**

NS, No significant differences
* chi-squared test
** chi-squared test with Yate's correction

BIOINFORMATIC ENVIRONMENT FOR ANALYZING LARGE SCALE GENOME SEQUENCE

Krzysztof Sarapata[1], Marek Sanak[2]

1 – Department of Bioinformatics and Telemedicine, Jagiellonian University School of Medicine
2 – Department of Medicine, Jagiellonian University School of Medicine

Abstract: Presented application for large scale genome analysis is useful to test whole chromosomes with Genscan program. Implemented solutions manage with following problems:
- Genscan program, as context dependent gene seeking program, makes impossible simple division of long sequences into subsequences and simple sum of results
- optimization of input subsequences to obtain high level of accuracy with Genscan program
- calculation of the longest input subsequences which Genscan can analyze (in comply with hardware limitations).
The main result of the work is the ability of scanning large scale sequences to find genes without redundancy error or losses, while analyzing result from dividing ranges. This paper focuses on conditions of using Genscan program. Analyzed sequences came from human genome sequences.
Keywords: genome analysis, large scale genome analysis, nucleotide sequences distribution, Genscan wrapper

Streszczenie: Niniejsza praca prezentuje rozwiązanie informatyczne umożliwiające analizę dużych sekwencji nukleotydowych na poziomie całych chromosomów z wykorzystaniem programu Genscan. Powstałe rozwiązania dotyczą następujących problemów:
- zastosowanie kontekstowego algorytmu w programie Genscan uniemożliwiającego prosty podział dużych sekwencji na mniejsze i prostego sumowania wyników skanowania,
- optymalizacja wejściowych subsekwencji dla uzyskania najlepszych rezultatów analizy sekwencji za pomocą programu Genscan (*high levels of accuracy program*),
- wyznaczenie najdłuższej subsekwencji do bezpośredniej i jednokrotnej analizy programem Genscan, wynikającej z ograniczeń sprzętowych.
Uzyskanym efektem pracy jest możliwość skanowania dużych sekwencji z zapewnieniem braku redundancji wyników analizy oraz braku utraty danych w miejscach podziału dużej sekwencji. W pracy zwrócono uwagę na kryteria i warunki zastosowania programu Genscan. Analizowane dane dotyczyły genomu ludzkiego.
Słowa kluczowe: wielkoskalowa analiza genomu, dystrybucja nukleotydów w sekwencji, Genscan wrapper

Introduction

Haploid human genome consists of 3 billions nucleotides arranged in 23 (24 in males) chromosomes – strings of 50 to 263 millions bp. Whole genome analysis is a demanding task for a user limited with the current computer's processor speed and random access memory. By a typical approach, any genomic task is tested on fragmented chromosomal sequences and ascertainment of the results approximations are inferred to the whole genome scale. Large nucleotide sequences can be split and fractional results summed up if the results have additive properties or a context depended rules can be applied.

Nature of human genome composition is heterogeneous. Distinct properties might be calculated for distinct C+G compositional regions (isochores) of the human genome. Distinct properties are connected with intron length and gene density depending from isochores. In consequence, several various sets of model parameters of Genscan (e.g. initial probability) were derived from simple classification input sequences for different groups of isochores. Initial probabilities express the length of intergenic region and probability of: introns, 5′UTR and 3′UTR regions. Sequences of learning sets in Genscan were centered on genes and typical ones contain little or no flanking intergenic DNA. Consistently high levels of accuracy of results are observed when sequences are similar to the learning sets. As a consequence of the effects mentioned above users should take care of proper input sequences and subsequences before using Genscan for large scale sequences.

We present a software environment for the program Genscan[1,2], a computational tool for analysis of nucleotide sequence. The program predicts location of genes and points to these typical constituents e.g. promoter, exons including the first and the last one, and polyadenylation signal. Genescan process the sequence using the following scheme[3]:
1. The sequence and its parameters are read and stored in dynamically allocated arrays.
2. The sequence is scored using the probabilistic models of coding/non-coding regions, donor and acceptor splicing

sites, etc., which characteristics are provided in the parameter file. The resulting scores are also stored in dynamically allocated arrays.

3. Modified Viterbi, forward and backward recursions are performed which allow determination of the most likely gene structure in the sequence, probability of each predicted exon, and (optionally) all suboptimal exons with a probability above the given cutoff.

4. The predicted gene structure(s), suboptimal exons, associated score information and the corresponding predicted peptides are then printed to stdout (the text output). Optionally, a PostScript output file which displays the locations of all predicted exons is also created.

5. The graphical output can be created.

6. The program displays the time it took to run and exits.

Chromosomal sequences are analyzed by various methods that belong to pattern recognition with identical and similarity vectors models. These methods are: window-based methods[4], bayesian analysis of isochores[5], wavelet multiresolution[6], entropic segmentation method[7].

Our model allows realizing window based method scan. Moved along sequences the window has dynamically modified width overlapped of each subsequences.

Genescan stratifies the analyzed sequence into four isochoric groups based on its guanine-cytosine pairs content:

Isochore	1	2	3	4
GC%	0-43	43-51	51-57	57-100

The best prediction as compared to experimental results is achieved if the input sequence contains only one gene, however, a context based use of Genscan software enables to predict the presence of many genes in the input sequence.

Empiric estimate of the maximal input sequence length for Genescan is as follows:

$$Size_{max}[Mbp] = 2 \cdot RAM[GB]$$

For typical personal computers according to this formula possible sequence length is:

1GB – 2Mbp, 2GB – 4Mbp. On the basis of our own experience these acceptable lengths are even overestimated and it is practically possible to input at most 50% of the estimated sequence lengths. For comparison, total human chromosome sequence is within the interval of 250Mbp (chromosome 1) – 46.5Mbp (chromosome 21).

Thus, the first possible approach to overcome hardware limits is to split the long sequence into subsequences within acceptable hardware limits. These partial sequences are not optimal because one can expect loss of information at the borders of partition. Alternatively, subsequences can be created with an overlap, for example using a moving window with a constant size overlapping ends. This method has some drawbacks in its application because there is a tendency to shorten the size of overlapping sequence as the sequence analysis by Genescan performs optimally if only one gene is located within such a fragment. As a consequence, number of subsequences grows rapidly giving redundant results. In the worst but possible scenario, analysis of the corresponding overlaps detects different genes. In extension to this solution, a re-analysis of patches which link subsequences can arbitrarily exclude falsely detected genes. Analysis of the sequence splitted into fragments should follow the addition rule i.e. the total results have to correspond results obtained by analysis of the whole sequence.

Only certain points of the sequence division allow to obey the addition rule (Figure1). While difficult to find, points of division are expected in the genomic regions with a low density of genes. A series of poitns of division can exist within a genomic region creating a segment of division.
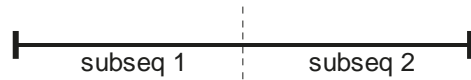
Fig. 1. Point of the sequence division

If points of division are not found, empirically overlapping regions have to be adjusted to contain a common gene with an identical organization of its coding region (promoter, introns/exons, polyadenylation signal) (see Figure 2).

Total sequence has to be split into fragments much larger than the size of a gene. Overlapping segments have to be defined with one nucleotide precision.

Material and Methods

Presented program is developed as a set of scripts in Ruby language, which scans a large genomic sequence (total chromosome sequence), splits the input sequence into subsequences, runs external Genscan program, saves results of Genescan

Fig. 2. Overlapping regions with common gene

```
                      ┌──────────────┐
                      │    START     │
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │ list of AGP  │
                      │    files     │
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐        ╭────────────────────────╮
                      │  AGP file    │        │ dmax = 100000 [bp] max  │
                      │  parsing     │        │ distance between genes  │
                      └──────┬───────┘        │ gnum_max = 5 amount of  │
                             ▼                │ genes in overlap region │
                      ┌──────────────┐        │ gnum_min = 2 min. amount│
                      │ Chromosome   │        │ of genes in overlap genes│
                      │ data save    │        ╰────────────────────────╯
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │  AGP file    │        ╭────────────────────────╮
                      │  iteration   │        │ BISECTION METHOD        │
                      └──────┬───────┘        │ c = 10000 {bp] max      │
                             ▼                │ convergence             │
                      ┌──────────────┐        │ max_w = 3000000 [bp]    │
                      │  Left site   │        │ max window's width      │
                      │  calculate   │        ╰────────────────────────╯
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │  Right site  │
                      │  calculate   │
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │   create     │
                      │subfile(left_site,│
                      │  right_site) │
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │  GenScan     │        ╭────────────────────────╮
                      │  processing  │        │ Hierarchical system     │
                      └──────┬───────┘        │ searching common gene or│
                             ▼                │ noncoding range in overlap│
                         genscan              │ region                  │
                         result?              ╰────────────────────────╯
                             ▼
                      ┌──────────────┐
                      │ join genscan │
                      │   results    │
                      └──────┬───────┘
                             ▼
                      ┌──────────────┐
                      │ save genscan │
                      │ result in db │
                      └──────┬───────┘
                             ▼
                         AGP file end?
                             ▼
                      ┌──────────────┐
                      │     END      │
                      └──────────────┘
```
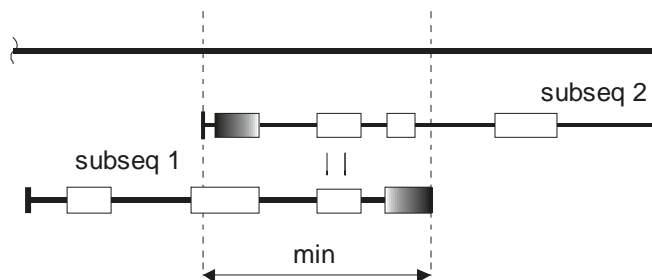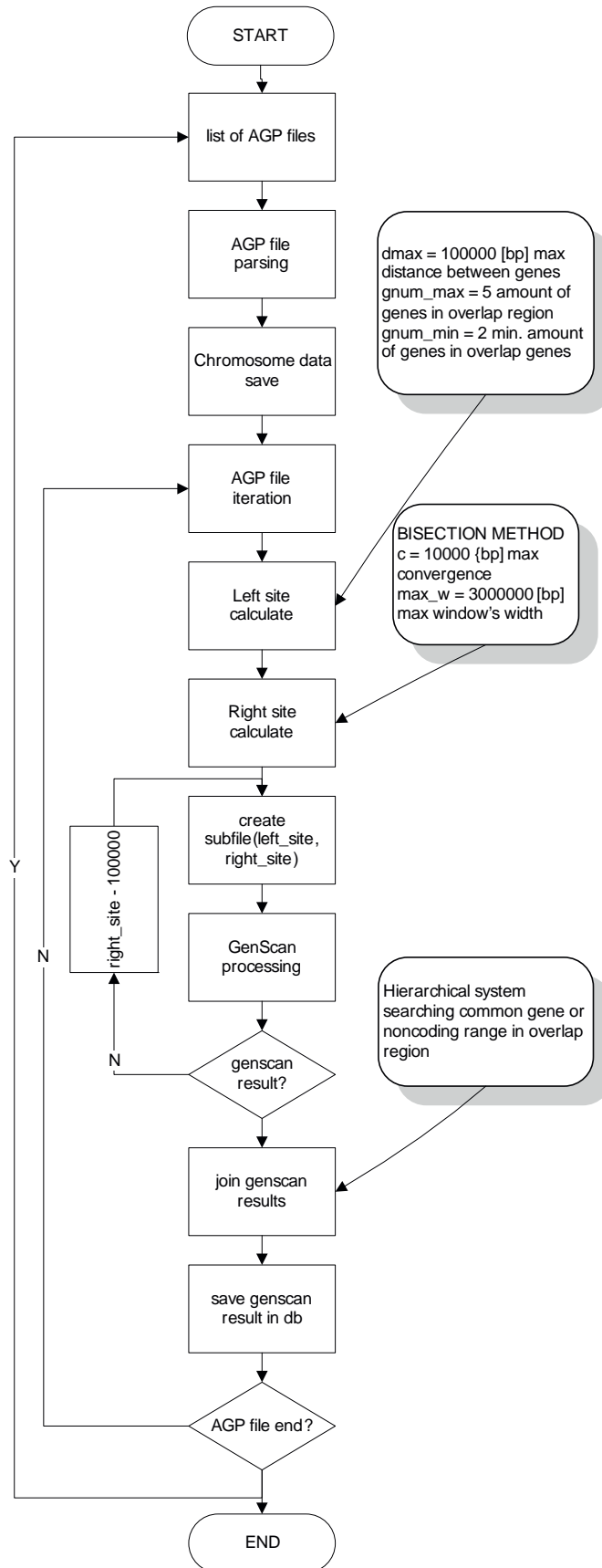
Fig. 3. Algorithm of the program

and attempts to reassemble partial results into optimized total one. Input data tested were reference sequences of total chromosomes of human genome (hs_ref_chr.fa[8]) and detailed data from *.agp files, which permits initially exclude chromosomal regions not containing genes (centromers, telomers, acrocentric regions or unknown sequence regions). The algorithm of the program is presented at the diagram (Figure 3):

Some parameters were assumed to test the model program:

Left boundary of the subsequence starts from the fifth gene at the end of the previous subsequence if the distance between the genes is less than maximal allowed (d_{max})

d_{max} = 100000
gnum_max = 5
gnum_min = 2

The right boundary of the subsequence is found by bisection empirically adjusted to parameters of external program call:

genscan HumanIso.smat fileout –cds
timeout = 60 [s]

Analysis of large sequences requires detection of: memory run out (memory leak) caused by uncontrolled use of libraries e.g. Bio::FastaFormat which allocates whole chromosome sequence in different formats. This was resolved by the use of a module garbage collection (GC) when appropriate, which enforces defragmentation of freed memory but in cost of decreased processing speed. Program Genscan is called as an external procedure (IO[9]), along with its process identifier to control timeouts and other states of its execution.

During the scan of too long sequence, Genescan can behave unpredictably which is terminated by 'kill TERM' and 'kill KILL'. Detection of these states includes incomplete results due to out of memory, 'nan' sequences in the output file or repeated values '0.000'.

Split of input sequence into subsequences

The beginning of a subsequence

The left boundary of the subsequence is obtained using the following algorithm:

1. The last gene in the previous subsequence has to be found. The next subsequence start is arbitrarily d_{max} = 100000bp before this gene.
2. If any other gene or its fragment is found in the d_{max} region, the first step of the algorithm is repeated. This iteration is repeated maximally gnum_max = 5 times or when d_{max} region does not contain any gene.
3. Subsequence can contain maximally gnum_max(5) genes which allows to easily find at least one common with the adjacent subsequence at the step of analysis of overlapping sequences. Alternatively, subsequence may represent a non-coding region which length equals d_{max}, which is sufficient to maintain an unique context for subsequent analysis.

The end of subsequence

The right boundary of the subsequence is obtained by the method of bisection[10], which finds the longest sequence possible to analyze with Genescan. Two parameters are available: the maximal precision of the right boundary (assumed: 10000bp) and length of the longest subsequence found empirically, which cannot be correctly analyzed by the program (assumed 3000000bp). This value also limits the maximal length of a subsequence, which is a very useful option if adequate segments are to be splitted.

Analysis of overlapping regions

The aim is to find at least one identical gene in the overlapping sequences of adjacent segments. This is searched in respective results sets as their crossed values. If this condition fails, within the overlapping subsequences the longest noncoding region is searched which is shared by both of them. This seeks of common genes within overlapping subsequences is done using a gradation:

level	Promoter		exons		polyA	
	Locus	score	Locus	score	locus	score
5	x	x	X	x	x	X
4			X	x		
3			X			
nc						

'nc' – the longest common non-coding region

Examples:
The last gene in the subsequence 1 is at the distance *d* (d > d_{max}) from the previous one. Genes filled with white are correctly ascertained (context genes). Black color highlights incorrectly identified genes; the color lightens with the distance of a gene from the subsequence boundary. In this case no result set crosses but the common region is satisfactory enough to make a section at the *w* site (see Figure 4).

A similar example, in which after second iteration an interval having a sufficient length was found and contains a shared gene (see Figure 5).

Figure 6 presents the most common example of a full iteration which enables bisection before the last five genes of the subsequence 1.

Assembly of the results of subsequences analysis

The way Genscan results are assembled is dependent on algorithm finding overlapped regions of subsequences. If a shared gene is present, results contain all genes found in subsequence 1, including the shared gene, and genes from subsequence 2 starting from then next one after the shared gene.
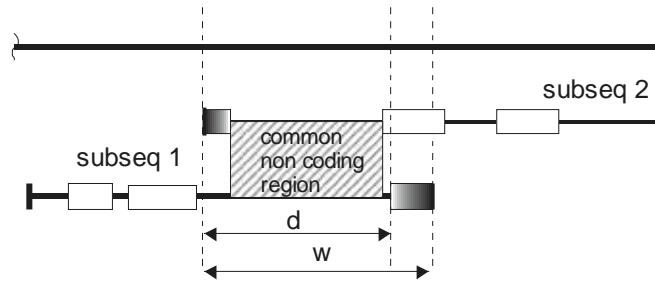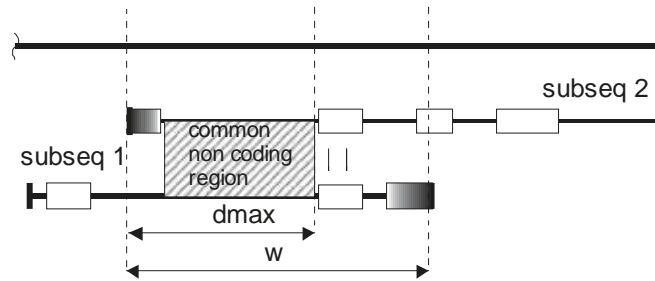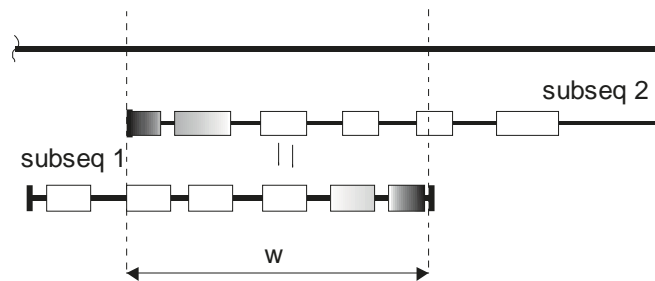
Fig. 4.

Fig. 5.

Fig. 6.

Validation test of the method

Model parameters: d_{max} – shared length of overlapped non-coding region, w – length of overlap, gnum_max – maximal number of genes present in the overlapped region, were found empirically and are declared values.

d_{max} = 100000
gnum_max = 5
gnum_min = 2

The test consisted of the analysis of a long input genomic sequence using the same declared model values but changing the constraint on the longest subsequence length. The same sequence of human chromosome 22 was used for testing.

The test was completed for a maximal subsequence length limited by:
1. Computer hardware
2. 1500000bp
3. 1000000bp

Results

Using the model, satisfactory results could be obtained during assembly of continuous sequence from the subsequence output for different length of the fragments. Some difficulties were encountered if the class of isochores had to be changed between 1 and 2 for neighboring subsequences, because no shared genes were found. A manual analysis in some cases showed though that genes had some identical elements like length of some exons, but the program chose to proceed by finding a common longest non-coding region.

Each run of chromosome 22 analysis brought different number of genes found in a consequence of their different classification by Genescan program using subsequences. These overlapped regions, which showed differences in genes identification are show in Figure 8.

Summary:

Chromosome id: 1

id	from	to	length	genes_range	save_genes	left_site	r_o_lap	comp	level
1	14430001	15077850	647850	1..40	1..40				
2	15227851	16727851	1500001	1..57	1..55	5	130542	gene	5
3	16597310	18097310	1500001	1..60	4..56	5	76422	gene	3
4	18020889	18889431	868543	1..50	2..50				
5	18939432	20439432	1500001	1..62	1..59	5	199892	gene	5
6	20239541	21739541	1500001	1..65	3..63	5	130718	gene	5
7	21608824	23108824	1500001	1..53	4..50	5	169891	gene	5
8	22938934	24438934	1500001	1..40	3..38	5	369870	gene	5
9	24069065	25569065	1500001	1..33	4..31	5	198586	gene	5
10	25370480	26870480	1500001	1..42	4..39	5	70646	gene	5
11	26799835	28299835	1500001	1..46	3..45	5	245970	gene	5
12	28053866	29553866	1500001	1..41	5..38	5	176686	gene	5
13	29377181	30877181	1500001	1..37	3..35	5	189286	gene	5
14	30687896	32187896	1500001	1..50	4..45	5	188809	nc	
15	31999088	33499088	1500001	1..63	1..58	5	93984	nc	

Fig. 7. Example of output summary scan of large sequence. Start ("from") and end ("to") position subsequence on whole sequence, sub-sequence length ("length"), amount finding genes ("genes_range") and those genes, which were saved ("save_genes"), amount of genes in left site in overlapped range ("left_site"), width of overlapped range ("r_o_lap"), kind of matches ("comp"): gene "gene" or common non coding region "nc", level of match overlapping area ("level")
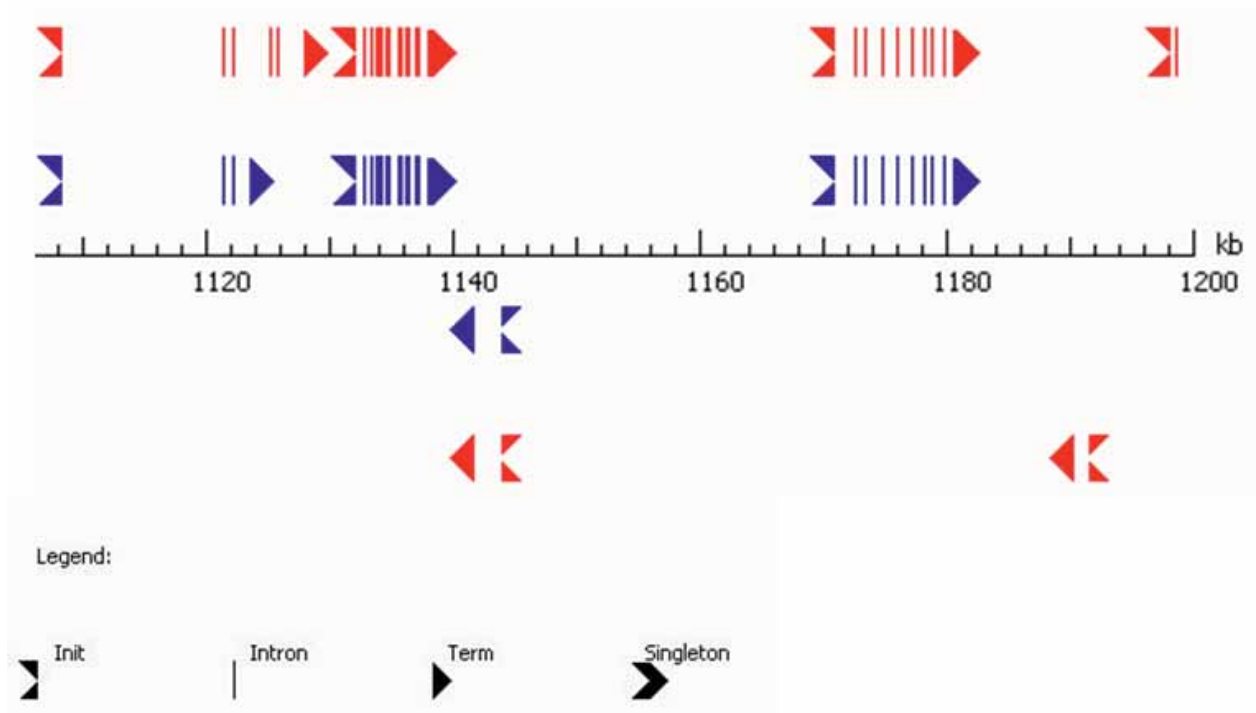

Fig. 8. Genscan predicted genes in two subsequences: 18020889-18889431 (red color on picture, isochore 2, 50.87% C+G) and 16953473-18715496 (blue color on picture, isochore 3, 54.96% C+G)

Conclusion and Perspectives

The model developed by us enables to modify parameters in search of the optimal outcome. This can be achieved by increased number of shared genes in the overlaps between subsequences.

Use of the program (Genscan) optimized to analyze preferably one gene during a long sequence analysis requires appropriate dissection of segments. Two possible solutions are suggested in the publications: a search of the gene promoter following CpG island, which helps to define a break point for subsequence containing this particular gene[11]; or using breakpoints cues based on expressed site tags (ESTs)[12] pointing to known coding sequences of genes.

Our results suggest that following optimization length of subsequences correlates negatively with the density of genes in the analyzed genomic region, if they were found using bisection method. If a tool were available to test the context of the genes, preferably by search in biological databases for their homology, heterogeneity or other parameters using systemic biology tools, this would greatly improve validity of the results.

Further perspective to expand our model requires using different methods of analyzing composition sequences with respect to isochore groups and division of sequences according to obtained parameters.

References

1. Burge C. (1997), Identification of complete gene structures in human genomic DNA. PhD thesis. Stanford University, Stanford, CA.

2. Burge C., Karlin S. (1997), Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268: 78--94.

3. http://genes.mit.edu/README.

4. Costantini M., Clay O., Auletta F., Bernardi G. (2006), An isochore map of human chromosomes. Genome Res. 16: 536-541.

5. Fearnhead P., Vasileiou D. (2006), Bayesian Analysis of Human Isochores. Poster. 8th Valencia International Meeting on Bayesian Statistics and the 2006 World Meeting of the International Society for Bayesian Analysis, June 2-6, 2006, Valencia.

6. Wen S.-Y., Zhang Ch.-T. (2003), Identification of isochore boundaries in the human genome using the technique of wavelet multiresolution analysis. Biochemical and Biophysical Research Communications 311: 215-222.

7. Bernaola-Galván P., Grosse I., Carpena P., Oliver J.L., Román-Roldán R., Stanley H.E. (2000), Finding Borders between Coding and Noncoding DNA Regions by an Entropic Segmentation Method. Physical Review Letters 85, 6.

8. ftp://ftp.ncbi.nih.gov/genomes/H_sapiens/Assembled_chromosomes/.

9. http://ruby-doc.org/core/classes/IO.html#M002294.

10. Cormen T.H., Leiserson Ch.E., Rivest R.L., Stein C. (2005), Wprowadzenie do algorytmów. Polish trans. by K. Diks, Wydawnictwa Naukowo-Techniczne, Warszawa.

11. Karlin S. (1994), Statistical studies of biomolecular sequences: scorebased methods. Phil. Trans. R. Soc. Lond. Biol. 344: 391-402.

12. Xu Y., Uberbacher E.C. (1997), Automated gene identification in large-scale genomic sequences. J. Comput. Biol. 4: 325-338.

BIT PATHWAYS – A TOOL FOR TEACHING AND LEARNING CLINICAL PATHWAYS

ANDRZEJ A. KONONOWICZ[a], THOMAS HOLLER[b]

[a] *Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, ul. Św. Łazarza 16, PL 31-530 Kraków, Poland*
[b] *Canton Hospital Aarau, Kantonsspital Aarau AG, Tellstrasse, CH-5001 Aarau, Switzerland*

Corresponding author: Andrzej A. Kononowicz, Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, Kopernika 7e, 31-034 Kraków, Poland, phone/fax: +48 12 4227764, email: mykonono@cyf-kr.edu.pl

E-learning

Abstract: The primary objective of clinical pathways is to improve the quality of healthcare and to reduce variations and costs, however they also have the potential to be used as an efficient learning resource. Their introduction can help medical professionals to keep their knowledge up-to-date and to better learn their roles and responsibilities in a hospital. >>mipp> is a clinical pathway project undertaken at the Aarau Canton Hospital in Switzerland. Existing >>mipp> tools have been extended to support better dissemination of clinical pathways in the hospital and to serve also as an educational tool. Newly implemented authoring and visualization tool – BIT Pathways – has made it possible to develop pathways easily. The tool may be used outside the Aarau Canton Hospital for educational purposes. A group of 47 multidisciplinary postgraduate students has preliminarily tested the software with encouraging results.
Keywords: applications in subject areas, architectures for educational technology systems, authoring tools and methods

Introduction

Pressures on quality, time and financial resources in the health care system have resulted in the need to organize medical processes more efficiently. It should always be remembered that financial and organizational improvements cannot reduce treatment quality. One of the attempts to maximize quality and cost efficiency is the introduction of clinical pathways. Clinical pathways (also known as Integrated Care Pathways [Campbell, Hotchkiss, Bradshaw & Porteous, 1998], CaseMaps [Bollmann & Beck, 2002] or Critical Pathways [Pearson, Goulart-Fischer & Lee, 1995]) are documents which describe the activities involved in the treatment of patient groups with specific diagnoses or conditions. These documents are created by interdisciplinary teams of professionals based on evidence found in medical standards and verified scientific publications. In contrast to general clinical guidelines, clinical pathways are constructed with the consideration of conditions and resources of local health care institutions. The activities determined by clinical pathways involve all the professional groups in a hospital (not only the physicians but also the nursing and technical staff). The standardization process reduces variations in clinical processes, which are often the source of errors [Panella, Marchisio & Di Stanislao, 2003]. Pathways may decrease the patient's length of stay in the hospital, and reduce complications and costs [Weiland, 1997]. The transparency of treatments carried out according to clinical pathways has the potential to improve the quality of medical processes, for example redundant activities can now be easily detected and eliminated. The use of clinical pathways can reduce medical documentation time by only recording deviations from the pathway ("charting by exception") [Pearson et al., 1995] [Roeder & Küttner, 2006]. Clinical pathways are an invaluable source of information about hospital procedures. Their construction can integrate fragmented knowledge from different divisions of health care institutions. If regularly updated, they can provide medical staff with current, evidence-based data on hospital procedures.

The publication of clinical pathways can help professionals to better understand their roles and responsibilities [Panella et al., 2003]. Pathways have the potential to improve cohesion in teams by clarifying relationships between different medical staff groups (physicians, nurses and technicians). Furthermore, pathways can assist new employees greatly by introducing

them to the details of hospital procedures. Since clinical pathways are financially optimized they are an effective method for teaching students about cost-effective practice.

A potential risk of using clinical pathways for educational purposes cited by many authors is that they undermine training by discouraging experimentation, and stifle innovation and independent thinking [Woolf, Grol, Hutchinson, Eccles & Grimshaw, 1999][Campbell et al, 1998][Pearson et al., 1995]. Some physicians fear that physicians trained on clinical pathways will not be capable of treating patients autonomously. However, it should be noted that clinical pathways are constructed only for the most common and well known procedures, therefore physicians will still have to deal with rare cases independently. In addition, it is still necessary to select the correct pathway and react immediately in a situation where the treatment results do not fit the values anticipated by the standard.

Clinical pathways which are specially adapted for lay readers can be used to inform patients about forthcoming medical activities. They can answer the most frequent questions which arise during each treatment phase. Derived from guidelines, clinical pathways may be used to present the pros and cons of treatments and empower patients to make better informed health care choices [Woolf et al., 1999]. They have the potential to build better relationship between patients and physicians.

Materials and methods

The legacy system

The >>mipp> project has been started at the Aarau Canton Hospital (Switzerland) in 1995 [Holler, Schmid, Müller, Reemts, Bissat & Rieben, 2002] [Rieben, Müller, Holler & Ruflin, 2003]. In 2005 118 clinical pathways were implemented and deployed in the hospital (22 in general medicine, 77 surgery, 12 gynaecology, 7 paediatrics) amounting to more than 40% of all inpatient treatments [Canton Hospital Aarau, 2005]. Clinical pathways are currently used in Aarau to account for procedures which are funded by health care insurance.

Clinical pathways in >>mipp> project are comprised of components. These components consist of descriptions of activities executed within a single medical service by members of particular staff groups (physicians, treatment and nursing staff, and others). Not all activities are mandatory as some of them carry a probability value which indicates the frequency of occurrence of this activity. Each activity also carries a cost value and a cost centre. Many of those activities have their origins in standard activity catalogues (e.g. TARMED [TARMED Suisse, n.d] in Switzerland) or the hospital's internal activity catalogue. Besides these activities components also contain other properties such as quality requirements, time specifications and general descriptions. All components from a path are depicted as a flowchart.

>>mipp> clinical pathways were created using general purpose tools. The >>mipp> pathways were stored in a *Microsoft Access* database and were visualized by flowcharts created using *Microsoft Visio* or *Power Point*. A Microsoft Access application was implemented to utilize and modify the data. Clinical pathways were published on the hospital's intranet as Access reports in PDF format.

Extension of the >>mipp> system

The clinical pathways developed within the >>mipp> project have proven to be very useful in the hospital, however, there are certain aspects which need to be refined or extended. The pathways contain mainly financial data as they were created for billing and costs transparency purposes, therefore only little advantage has been taken of their educational potential. The >>mipp> flowcharts are not uniform because they were created with general purpose tools by different authors, and the links between components are stored only in graphical form; not as relations in the database. This prohibits the possibility of integrating the pathways automatically into the workflow of the hospital information system[1]. The existing flowcharts are difficult to modify and changes made to the flowcharts do not automatically modify the database. Additionally, Microsoft Access which was used for storing pathways' data, is in our opinion not well suited to work in web environment.

The objective of the proposed tool is to maximize the educational potential of the >>mipp> project. The knowledge stored in pathways should be presented in a more accessible and interactive form. Therefore, the extension is expected to provide different views on the data depending on the target audience (medical staff, patients) and the intention of user (e.g. staff group oriented, flowchart oriented). Additionally, a function is required which includes external knowledge sources (e.g. EBM publications, e-learning sites) in the pathway description. It was decided that the extension of the >>mipp> project should not modify the legacy database schema or the >>mipp> Access application. The new system should be flexible enough to support further extensions (e.g. integration with the new hospital information system that will be implemented in the near future) and to function also outside the hospital in Aarau.

Implementation

To fulfil the requirement for flexibility we have decided to use XML to store and transfer data. In fig.1 we present the overall architecture of the system after extension. We developed a new graphical editor called '*BIT Pathways*' to create and modify new pathways. The editor is a Java Swing application. The choice of Java technology has the advantage of not restricting the application's usage to any particular operating system. We have not decided to develop *BIT Pathways* entirely as a web application due to the complexity of the flowchart editor, which is still very difficult to implement efficiently in a generic web browser. Legacy data is imported from the >>mipp> database (JDBC-ODBC bridge) using a web service (Axis WS implementation). Clinical pathways can be stored locally as files or remotely in a native XML database (eXist [Meier, 2003] accessed using the XML:DB interface). Users can view selected pathways with a JavaScript-enabled web browser. The data in the XML database is accessed by XQuery scripts, translated into HTML using XSLT. Sections of the pathways (e.g. component properties) can be downloaded asynchronously using

---

1 For integration of clinical pathways and guidelines into clinical workflow refer to the SAGE model [Campbell, Tu, Boyer & Mansfield, 2003] [Tu, Shankara, Campbell, Hrabak, McClayb & Huffc, et al., 2004] or other approaches e.g. [Dazzi, Fassino, Saracco, Quaglini & Stefanelli, 1997].
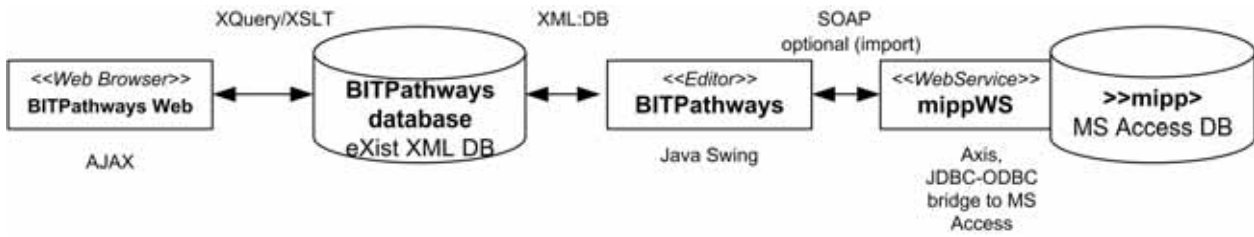
Fig. 1. System architecture of the new clinical pathways environment

AJAX (*Asynchronous JavaScript and XML*) technology [Crane, Pascarello & Darren, 2006], which provides better performance for the user interface.

The clinical pathways created in the extended >>mipp> system are stored locally as XML files or remotely on an XML database server (authentication and authorization functions were implemented). We chose to use a native XML database server rather than an XML enabled RDBMS because the conversion overhead caused by the properties of the stored data would nullify the superior performance of a relational database. A native XML database which accesses the data without conversion is very advantageous when processing large quantities of XML data [Fong, Wong & Fong, 2003]. We selected eXist

as the native XML database server implementation because it is stable, open source, and it supports XQuery processing. In addition to the XML data model – for efficiency reasons – an automatically generated graphical file of the flowchart (in PNG format) is also stored in the database.

Several proposed specifications for encoding clinical guidelines already exist. Among them are e.g. Arden Syntax, GLIF [Ohno-Machado, Gennari, Murphy, Jain, Tu & Oliver, 1998], GEM [Shiffman, Agrawal, Deshpande & Gershkovich, 2001] [Georg, Séroussi & Bouaud, 2005] and CPG-RA [CPG-RA, n.d]. Our goal of designing pathways, which can be used as educational resources has lead our research towards narrative guidelines (like GEM or CPG-RA) rather than methods which

Tab. 1. Summary of an exemplary attributes template created for teaching purposes

| Whole pathway properties	
Attrib Group	*Attributes*
General properties	**Id**:string, **Name**:string, **Diagnosis codes**:icd[], **Procedure codes**:icd[], **Pathway description**:html, **Authors**:string[]
Sources	**National sources**:ref[], **International sources**:ref[], **Web sources**:url[]
Patient	**Patient – Path description**:html, **Lay Link**:url[]
Pathway component properties	
Attrib Group	*Attributes*
General properties	**Name**:string, **Phase**:enum, **Category**:enum, **Description**:html, **Risk factors**:html, **Duration**:time
Activities	**Activities**:activity[]
Sources	**National sources**:ref[], **International sources**:ref[], **Web sources**:url[]
Drugs	**Drugs**:drug[]
Patient	**Patient – Path description**:html, **Lay Link**:url[]
Pathway branch properties	
Attrib Group	*Attributes*
General properties	**Name**:string, **Description**:html
Sources	**National sources**:ref[], **International sources**:ref[], **Web sources**:url[]
Patient	**Decision rational**:html, **Lay Link**:url[]
Pathway edge properties	
Attrib Group	*Attributes*
General properties	**Description**:html, **Probability**:double
Start element properties	
Attrib Group	*Attributes*
General properties	**Name**:string, **Inclusion criteria**:html
End element properties	
Attrib Group	*Attributes*
General properties	**End premises**:html, **Recommendations**:html
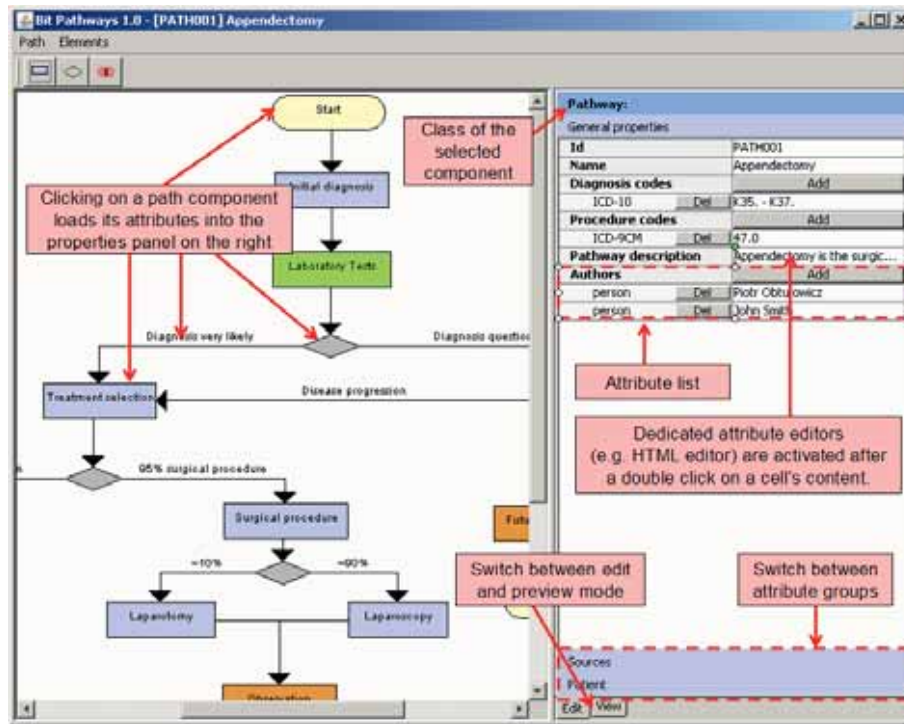
Fig. 2. The BIT Pathways editor – Appendectomy Pathway

specify executable code (Arden Syntax, GLIF). However, our goal is also to preserve maximal flexibility for the teachers in reaching their educational goals. This reason leads us to the opinion that the mapping of the current structure into an established standard is, for the time being, impractical. We have decided to create our own XML schema which is very easily modifiable and which may, if needed, be mapped to an exchangeable format like GEM.

We have decided to divide our XML documents into 3 parts:

- Header – containing general attributes describing the whole pathway (e.g. the pathway name or detailed characteristics of patients which fit into the pathway, information about authors of the pathway, and multiple choice-questions about the pathway).
- Elements – containing properties of all workflow elements. Including properties describing the graphical position of the element in the diagram, and the purpose of the element.
- Connections – this section contains relations between flowchart elements.

Elements are characterized by their class. We distinguish start elements, end elements, components (corresponding to the major task steps of the described procedure), branches and edges (connections). There is also a class that represents the whole pathway. Each class has its own attribute set distributed among one or more attribute groups. For instance, elements of the class '*component*' contain attribute groups: *general properties*; *activities*; *sources, drugs, patients*. The attribute group '*sources*' contains attributes which contain literature references for medical staff willing to extend their knowledge about the given step of the pathway. Attributes may consist of a single value or a list of values. Values have types – beside standard data types (string, boolean and numerical values) we have

specified more complex types such as activity, classification code, person, multiple-choice-question and *enum*. Type '*enum*' means a value from a controlled set of terms.

The list of available attributes and attribute groups is defined by templates in XML format and may be changed easily. This feature allows the usage of BIT Pathways in various institutions and for diverse purposes. Tab. 1 presents a summary of an exemplary template created for the use in teaching of clinical pathways. Presented are attributes for all classes of elements. Each attribute (in bold) is additionally characterized by its type. Types ending with [] denote list of attributes of the same type.

Results

A learning process based on clinical pathways involves, in our opinion, two main groups of users:

- medical educators – i.e. those preparing learning materials regarding pathways
- learners – i.e. those interested in gaining knowledge about topics covered by pathways. Members of this user group can be medical professionals (physicians, nursing staff, and technicians), students or lay readers (patients or patient's family members). The diversity of users in this group results in the requirement for medical educators to prepare differently profiled materials for the same component.

The task of medical educators is facilitated by the newly developed pathway editor called *BIT Pathways* (fig. 2). This editor (currently available in 3 language versions: English, German and Polish) assists the educator in developing new clinical pathways, adding educational materials to pathways and importing existing pathways or their elements from the legacy
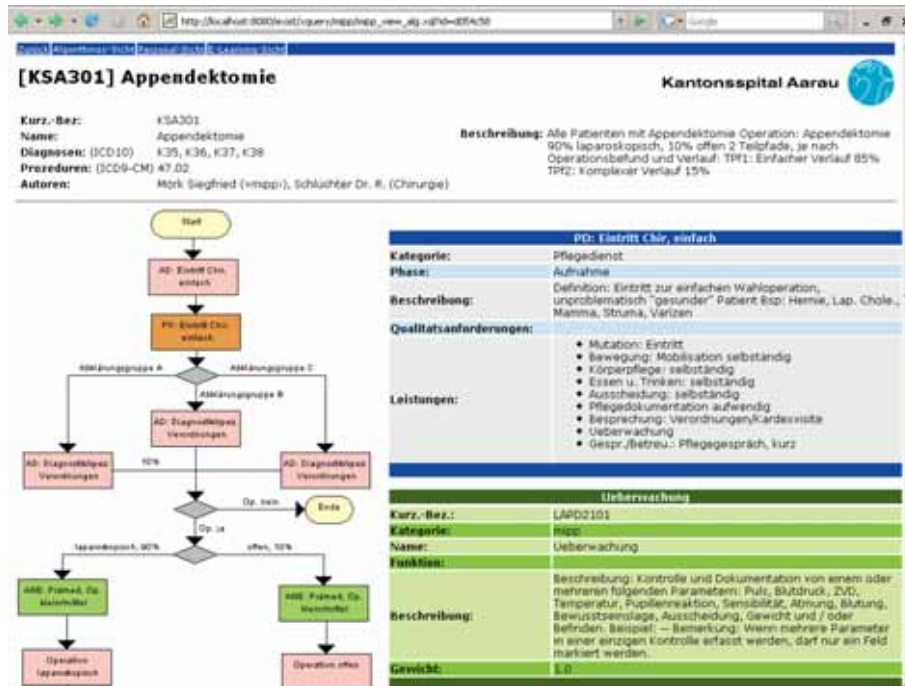
Fig. 3. A clinical pathway imported from the >>mipp> system displayed in the flowchart view of the BIT Pathways system

system. The import process of pathways is enabled by a web service. The user can choose from more than 100 existing pathways and than select the components to import.

Clinical pathway flowcharts are displayed in the main panel of the application. The right-hand panel displays the attributes of the selected flowchart element. Attributes are displayed in read-only mode as an HTML page, or as an editable grid in edit mode. Attributes with complex types have their own editors activated by a double click on the attribute value row. Activities can be imported from an integrated hospital activity catalogue browser (*TARMED* [TARMED Suisse, n.d] for Switzerland, and *Katalog Świadczeń Szpitalnych NFZ* [Polish National Health Fund, n.d.] for Poland). The adaptation of a component's content for different user groups is realised by two means. Firstly, by setting a value in the component's category attribute according to the intended user group (e.g. physicians, treatment staff, nursing staff, etc). Secondly, by the use of attribute pairs in which one attribute contains data for medical staff members (students) and the other for lay readers (e.g. medical activity description/patient's activity description, medical EBM link/patient's educational link). The colour of a component in the graphical view indicates its assignment to a hospital staff group.

It is intended that learners will access information about pathways via the Internet or hospital's intranet site using a web browser. Storing the data in an XML database makes it easy to provide different views of the same data for different user groups.

At present we provide three different views of clinical pathways: two for medical staff (students) and one for patients. The views for medical staff are *flowchart* and *staff group* view. For patients we provide *general patient view*.

The *flowchart view* (fig.3) displays the flowchart at the left-hand side of the window. When a component is clicked on its properties are asynchronously downloaded and displayed to the right of the diagram. Clicking on activities displays the general properties of the selected activity. The *general patient view* is similar to the *flowchart view* in that sense that it also displays a flowchart, however, clicking on a component displays information intended for patients. The *flowchart view* displays very detailed information written in a formal language, whereas the *patient view* presents general information about ongoing activities displayed in easily comprehensible form for the lay reader.

We will illustrate this division using the *Appendectomy* pathway. The pathway's global attributes define the inclusion and exclusion criteria of the pathway, and also give some links to EBM sources about appendectomy in general (pathogenesis, symptoms, diagnosis, differentiation, complications). The same attributes in the *patient's view* explain some basic terms like *appendix, (perforated/non-perforated) appendicitis* or *appendectomy*. The post-operative component informs, for example, medical professionals about recommended means (drugs and their dosage), and patients about post-operative side effects (nausea, lethargy, dizziness, etc.) and, for instance, in the case of a perforated appendix, what a NG tube is and what is it used for. This approach could be extended in pediatrics by providing different materials for children of different ages and, in addition, materials for their parents (this feature is currently not supported).

The *staff group* view displays components ordered by different staff groups. This view is also called the *'swimlane view'* because the data is arranged into columns. The first column contains all tasks to be done by physicians, the next contains activities for nursing staff, the next for technicians', etc. The aim of this view is to compare the activities of different professional groups.

Pilot evaluation

The BIT Pathways system has been preliminary tested on an interdisciplinary group of 47 post-graduate students at the Jagiellonian University Medical College. Most of them graduated two years ago in medicine, pharmacy, physiology or health care management. All of them participated in a course on IT technologies in life sciences. Divided in groups of 2-4 persons course participants prepared in approx. one month time 14 pathways using the BIT Pathways editor. The pathways pertained topics from various medical disciplines including endocrinology(3 pathways), gynaecology(2), infectious diseases(1), gastroenterology(3), genetic disorders(1), cardiology(2) and paediatrics(2). After presentation of the projects students were asked to take part in a short electronic survey concerning the applied course. Twenty six questionnaires have been returned (participation rate 55%). Seventeen students (65%) agreed (or strongly agreed) with the statement that BIT Pathways was useful in preparation of their clinical pathway (in a 1-5 Likert scale, 1 strongly disagree, 5 strongly agree, median = 4, mean = 3.76, SD = 0.88, one answer was not given). In a free text field students could point out problems related to usage of the program. Among the most frequent comments was the difficulty of inserting multimedia files (images, movies) into pathway's description or the inability of selecting more than one pathway components. Some students wished to have more flowchart symbols available in the toolbar. Course participants were also inquired about their general opinion on clinical pathways. Nineteen students (73%) agreed (or strongly agreed) with the statement that clinical pathways have a chance to be accepted in Polish hospitals (Likert scale, median = 4, mean = 3.92, SD = 0.80) what proves the relevance of the touched on subject. Asked about the major benefits of clinical pathways, students pointed out introduction of evidence-based practice, improvement in personnel communication, improved hospital audit and effective teaching resource for novice employees. The major hindrances of introducing clinical pathways in Polish hospitals were in the student opinion the scepticism of the hospital's managements and high costs of clinical pathways' implementation.

Discussion

Analysis of the projects delivered within the course on clinical pathways proved once more that construction of clinical pathways is not easy and should be practiced. Improper usage of branching elements, too detailed modelling of clinical pathways fragments, difficulties in distinguishing between patient's state and actions performed within the clinical pathway were the most common errors. Most of the course participants were also not aware (or unable to express) the tasks undertaken by other professional groups involved in the modelled medical process.

The conclusion based on the experience of Canton Hospital Aarau is that development of clinical pathways in general purpose tools like Microsoft Visio or Microsoft Power Point pose the threat of discrepancies in interpretation of clinical pathways even if they are developed in one single institution. This results in pathways incomparability and difficulties in implementing and updating them in hospital information systems. A good system for authoring and learning clinical pathways

should be a trade-off between giving the practitioners/medical teachers free hand in expressing their knowledge and putting constraints to preserve uniformity of clinical pathways. Unfortunately, awareness of dedicated tools for pathway construction is not common knowledge. The BIT Pathways system attempts to fill that gap by offering a means of practicing clinical pathways in academic settings. Being derived from an existing system imparts the tool with a certain degree of authenticity required for successful learning. The concept of templates setting up the list of permitted attributes makes the system suitable for many different purposes.

We believe that BIT Pathways have the potential to become a platform for learning about medical processes as well as acquiring skills of pathway construction. The pilot evaluation delivered many ideas for extending BIT Pathways by new features as e.g. the possibility of creating hidable subpathways or adding multimedia to the pathway description. More in-depth studies evaluating the educational outcomes are definitely needed and will be presented in the future.

Conclusions

Clinical pathways are a modern tool for ensuring quality and reducing costs and variations in health care. They rely upon evidence based documents and are tailored for a hospital's individual needs, making them to an invaluable source of knowledge about hospital procedures. The process of creating, modifying and publishing clinical pathways is supported and improved by the IT methods outlined in this paper. The BIT Pathways system which we present here is an extension of the >>mipp> clinical pathway project and represents a method for ensuring easier access to clinical pathways for learning purposes. The system was implemented mainly using XML technologies, which gives it the potential to be easily integrated into any hospital's IT infrastructure. Due to the choice of flexible technologies and a clear separation of different system components, selected modules can be used outside the Canton Hospital Aarau. The preliminary evaluation of BIT Pathways at the Jagiellonian University Medical College encourages in further development and evaluation of the system.

Acknowledgements

References

1. Bollmann M. & Beck M. (2002), Geplante Behandlungsabläufe und CaseMaps – Wirkung, Nutzen und Anwendungsfelder im Krankenhaus der Zukunft. In: Hellmann W. (ed.),

Klinische Pfade: Konzepte, Umsetzung, Erfahrungen. Landsberg/Lech: ecomed, pp. 239-248.

2. Campbell H., Hotchkiss R., Bradshaw N., & Porteous M. (1998), Integrated care pathways. *BMJ, 316*, 133-137.

3. Campbell J., Tu S.W., Boyer J., & Mansfield G. (2003), *The sage guideline model: A knowledge representation framework for encoding interoperable clinical practice guidelines.* Stanford: Stanford Medical Informatics; Report No.: SMI-2003-0962.

4. Canton Hospital Aarau (2005), *Annual Report.*

5. Crane D., Pascarello E., & Darren J. (2006), *Ajax in Action.* Greenwich: Manning Publications Co.

6. Dazzi L., Fassino C., Saracco R., Quaglini S., & Stefanelli M. (1997), A patient workflow management system built on guidelines. In: *Proceedings of the AMIA Annual Fall Symposium.* Hanley & Belfus, pp. 146-150.

7. Fong J., Wong H.K., & Fong A. (2003), Performance Analysis between an XML-Enabled Database and a Native XML Database. In: Chaudhri A.B., Rashid A., Zicari R. (ed.), *XML Data Management: Native XML and XML-Enabled Database Systems.* Addison-Wesley, pp. 547-565.

8. Georg G., Séroussi B., & Bouaud J. (2005), Extending the GEM model to support knowledge extraction from textual guidelines. *International Journal of Medical Informatics, 74*, 79-87.

9. Holler T., Schmid K., Müller H.P., Reemts Ch., Bissat K., & Rieben E. (2002), Praktische Pfadarbeit – Konstruktion, Implementierung, Controlling von Patientenpfaden. In: Hellmann W. (ed.), *Klinische Pfade: Konzepte, Umsetzung, Erfahrungen.* Landsberg/Lech: ecomed, pp. 38-129.

10. Meier W.M. (2003), eXist Native XML Database. In: Chaudhri A.B., Rashid A., Zicari R. (ed.), *XML Data Management: Native XML and XML-Enabled Database Systems.* Addison-Wesley, pp. 43-68.

11. Ohno-Machado L., Gennari J., Murphy S., Jain N., Tu S., Oliver D. et al. (1998), The GuideLine Interchange Format: A Model for Representing Guidelines. *Journal of American Medical Informatics Association, 5*, 357-372.

12. Panella M., Marchisio S., & Di Stanislao F. (2003), Reducing clinical variations with clinical pathways: Do pathways work? *International Journal for Quality in Health Care, 15*(6), 509-521.

13. Pearson S.D., Goulart-Fisher D., & Lee T.H. (1995), Critical Pathways as a Strategy for Improving Care: Problems and Potential. *Annals of Internal Medicine, 123*(12), 941-948.

14. Rieben E., Müller H., Holler T., & Ruflin G. (2003), *Pfadkostenrechnung als Kostenträgerrechnung – Kalkulation und Anwendung von Patientenpfaden.* Landsberg/Lech: ecomed.

15. Roeder N. & Küttner T. (2006), Behandlungspfade im Licht von Kosteneffekten im Rahmen des DRG-Systems. *Internist, 47*(7), 684-689.

16. Shiffman R., Agrawal A., Deshpande A., & Gershkovich P. (2001), An Approach to Guidelines Implementation with GEM. In: Patel V. et al. (ed.), *MEDINFO 2001.* Amsterdam: IOS Press, pp. 271-275.

17. Tu S.W., Shankara R., Campbell J., Hrabak K., McClayb J., Huffc S.M., et al. (2004), Modeling Guidelines for Integration into Clinical Workflow. In: Fieschi M. et al. (ed.), *MEDINFO 2004.* Amsterdam: IOS Press, pp. 174-178.

18. Weiland D.E. (1997), Why use clinical pathways rather than practice guidelines? *American Journal of Surgery, 174*, 592-595.

19. Woolf S.H., Grol R., Hutchinson A., Eccles M., & Grimshaw J. (1999), Potential benefits, limitations, and harms of clinical guidelines. *BMJ, 318*, 527-530.

Web References

1. CPG-RA: The Clinical Practice Guideline – Reference Architecture, http://www.cpg-ra.net (3.06.2008).

2. Polish National Health Fund (Narodowy Fundusz Zdrowia NFZ), http://www.nfz.gov.pl (3.6.2008).

3. TARMED Suisse, http://www.tarmedsuisse.ch (3.6.2008).

E-learning

CREATIVE E-LEARNING

WOJCIECH LASOŃ, PIOTR WALECKI, WIESŁAW PYRCZAK, KRZYSZTOF SARAPATA

Department of Bioinformatics and Telemedicine, Jagiellonian University, Medical College, Kopernika 7, 31-034 Krakow

Abstract: This paper presents questions related to new computer-aided didactic techniques such as e-learning and knowledge management. The main problems are the choice of content and the organization of knowledge and thus the ensuing creation of a unified educational system by employing diverse methods. One example of this could be individualized educational programmes or problem-based and integrated teaching combined with cross-subject presentation of the material being taught. Our proposal is include to e-education techniques individual extraordinary students' homework which show their creativity.
Keywords: computer-aided didactic techniques, e-learning, knowledge management

E-learning

Introduction

The development of internet techniques which offer interactive and multimedia-based didactic programs has forced educators to convert courses which were previously taught by traditional methods into Web-Based (WB) format. The main problems related to the new computer-aided didactic techniques are the choice of content and the organization of knowledge and thus the ensuing creation of a unified educational system by employing diverse methods. One example of this could be individualized educational programmes or problem-based and integrated teaching combined with cross-subject presentation of the material being taught. All of these efforts have the common aim of coordinating and unifying the effects of the basic elements of the teaching-learning system including the teacher, the students, the subject content and the teaching environment. Furthermore, the goal of such an approach to education is the improvement of the quality of the teaching process and its adjustment to dynamically changing educational needs, individual conditions and opportunities [1].

E-learning

E-learning is a new form of education which consists in the use of electronic tools and appropriately chosen methods to aid in the didactic process. Interactive and multimedia databases and specially profiled courses, most of all Information and Communication Technologies (ICT), are transforming the educational process and giving it a new shape which is frequently not present in the traditional, direct approach to teaching. On the one hand vast resources of knowledge are almost immediately available through the Internet, but on the other hand the shear amount of incomplete and even distorted information and the difficulties inherent in finding sought-for and credible data cause both students and teachers to sometimes feel lost and overwhelmed. Moreover, the constant development of the technological infrastructure forces one to keep up-to-date with these changes and to constantly learn new computer techniques. Therefore, the elaboration of standards and the unification of methods is so important. However, before a concrete e-learning solution can be accepted as a standard model, it is called a specification and undergoes verification and assessment by accredited educational organizations (such as the IEEE). Practically all of the elements and methods of e-learning require standardization in order to make communication with other elements of the system possible and in order to ensure optimal productivity and compatibility of individual implementations [2-5].

E-learning-based education can take place in two time modes: asynchronic and synchronic. Asynchronic solutions are geared toward an individualized learning program. These methods presume that there is no communication between the student and the teacher during the learning process. The student sets his own times and pace for learning and on his own uses available sources of information (databases) which are frequently only accessible locally and offline (Computer Based Training – CBT), for example CD or DVD disks containing the course programme. Synchronic mode learning supposes real-time student-teacher communication. Therefore, the course or training have precisely set times and dates and a specified time framework. The communication between the course participants and the teacher may not be direct as in the traditional method of teaching and instead can occur through electronic means of remote communication of which the most commonly used is the internet (Web Based Training – WBT) [6].

Distance Education

One of the key e-learning methods is Distance Education (DE) which uses the internet network. One could say that tele-education refers to a mode of knowledge transfer in which the teacher and the student are separated from each other by a geographical and/or temporal distance. DE is also referred to by several other related names: Distance Learning (DL), Internet-Based Training (IBT), Internet-Based Instruction (IBI), Computer-Based Distance Education (CBDE), Web-Based Learning (WBL), Web-Based Distance Education (WBDE), or Web-Based Instruction (WBI). DE can be described using a series of elements, one of which assumes the uses of various media and communication tools to connect the teacher and the student and the others describe the method for transferring information as well as its degree of structure [7]. Media technologies used in DE can be accurately described using a model called a "4-element map of group software options" [8]. This model contemplates 2 parameters, the time and the place, which must be taken into account in the case of group collaboration in a course taught using the DE method. Group collaboration support technologies have been classed into four groups based on these parameters: 1) same time / different place, 2) same place / different time, 3) same time / same place, 4) different time / different place [9].

Web-Based Instruction

Web-Based Instruction (WBI) is a type of technology which fits into the fourth category (different time / different place), and consequently offers a high degree of flexibility which permits learning in any place and at any time. This allows for academic programmes to be offered to a wider group of potential students who do not have time to study in a traditional university programme. The WBI method is just as effective as traditional teaching. Experimental studies have been made in order to compare the results obtained by students learning with the traditional method and with the WBI method. It was concluded that students who learn via WBI courses either obtain higher results or that there is no significant difference between the two groups [10]. Besides establishing the effectiveness of WBI, it was also shown that students enjoy these courses and speak positively about them.

Computer-moderated communication helps fulfil a new model paradigm based on the construction of knowledge in the learning process. The most important and distinguishing characteristic of WBI is that emphasis is placed on active learning and not just on providing information. If providing information would be taken as the fundamental goal then WBI would to a great degree resemble passive methods of computer based training. Therefore WBI should be designed based on the theory of cognitive psychology, according to which the students fully collaborate with the environment and take an active part in the learning process thus fulfilling the knowledge construction paradigm [11-12].

Learning Management System

Learning Management System (LMS) is the computer system which provides access to the e-learning environment. It allows one to manage user accounts, courses, lessons and also to monitor and report learning progress (results). A similar system is Electronic Performance Support System (EPSS) although it is designed more for professional online support for increasing work productivity (Vocational Education and Training – VET).

The tools used to preparing interactively lessons are Authoring Tools and Authoring Systems that is a part of Learning Content Management System (LCMS) [13-14].

Creative e-learning – our experience

A special portal created using Content Management System (CMS) technology has been prepared as part of the practical exercises for students taking course in "Biomedical signals
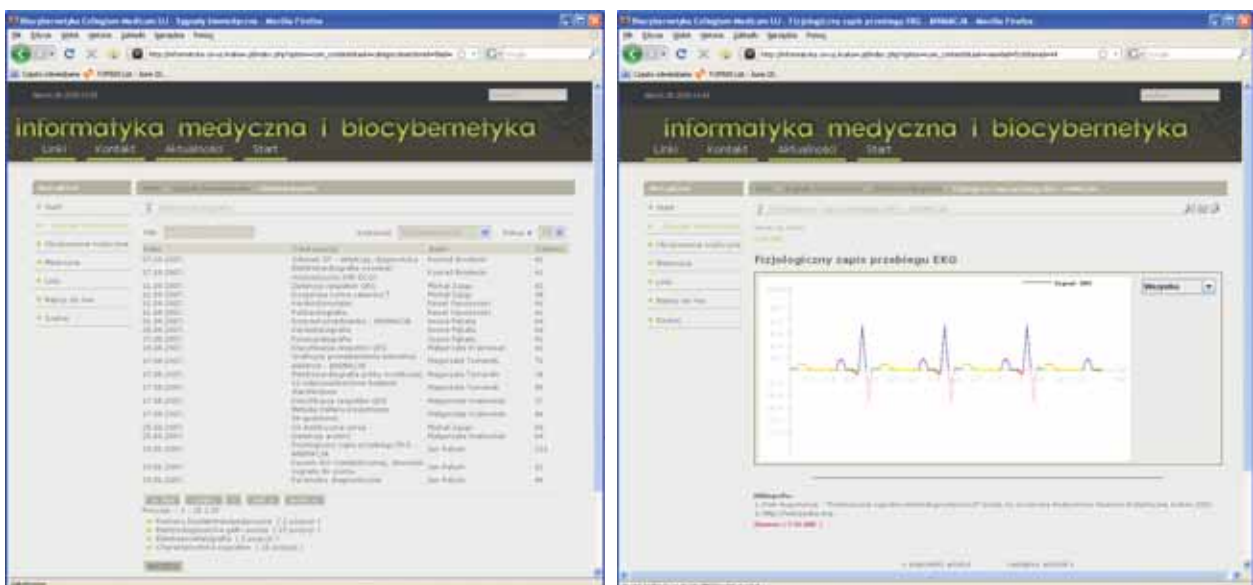

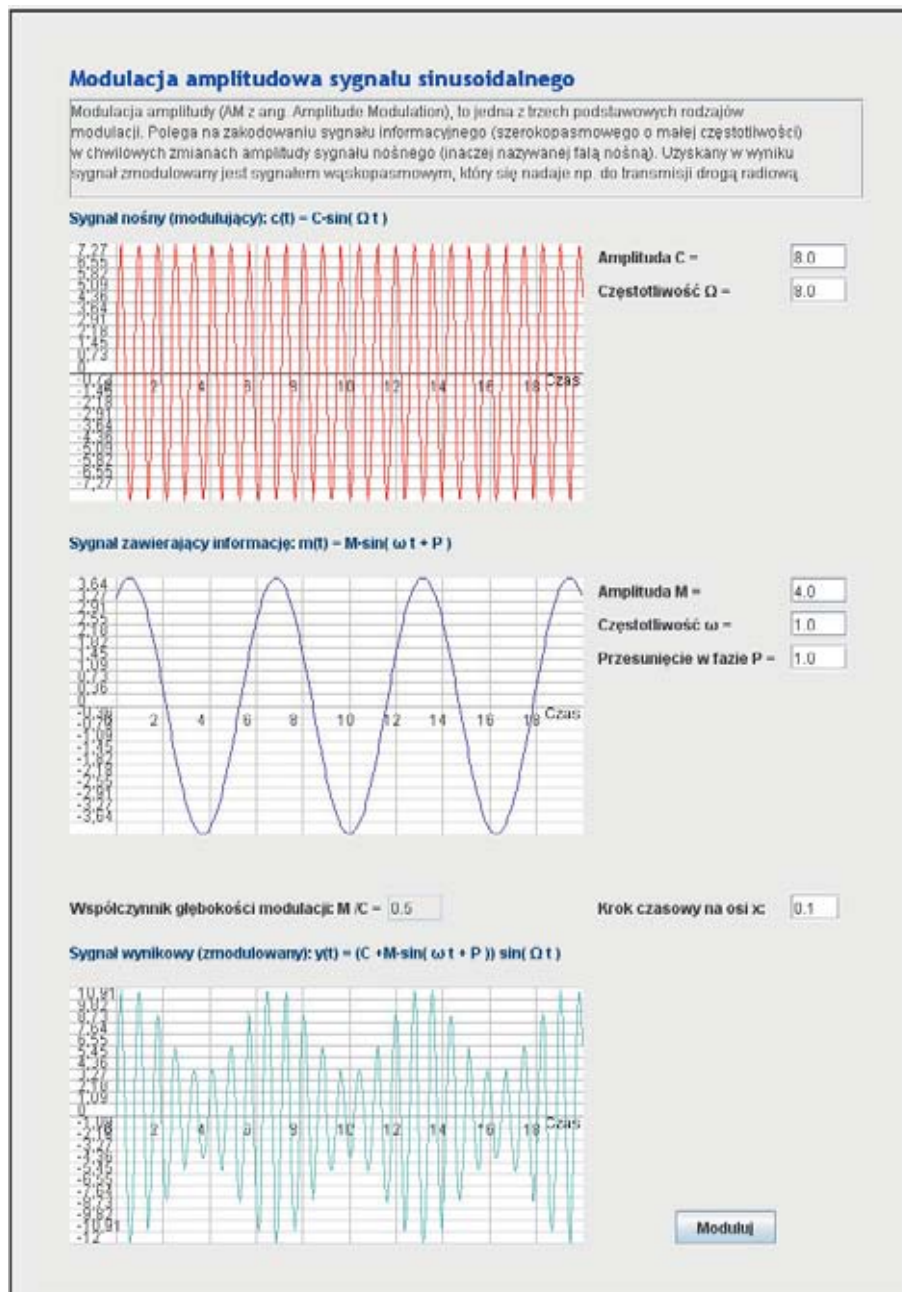Fig. 1. Screenshots from our portal web page. Source: www.informatyka.cm-uj.krakow.pl

Fig. 2. An interactive simulator for amplitude modulation of sinusoidal signal.
Author: Magorzata Tomanek. Source: www.informatyka.cm-uj.krakow.pl

analysis". The course was conducted in the Jagiellonian University, Medical College, Department of Bioinformatics and Telemedicine. E-education materials created by the students have been placed on the portal (Fig. 1).

During the first stage, the collected materials described two areas of medical diagnosis: biomedical signal analysis and medical imaging methods. Students independently carried out e-learning projects dealing with topics from the application of modern technologies in medical diagnostics. The students loaded the prepared projects containing the educational material onto the portal pages by themselves.

The materials which have been placed on the portal include the following topics:

- – *Biomedical signals*
- o Characteristic of signals
- ▪ Modulated deterministic signals
- ▪ Amplitude modulation of sinusoidal signal (see Fig. 2)
- ▪ Sinusoidal signal generator with linear frequency modulation
- ▪ Quantization (see Fig. 3)
- o Electrocardiography (see Fig. 4)
- o Electroencephalography (see Fig. 5)
- o Oculography and nystamography (see Fig. 6)
- o Bioimpedance measurements (see Fig. 7)
- o Magnetoencefalography (see Fig. 9)
- o Chromatography
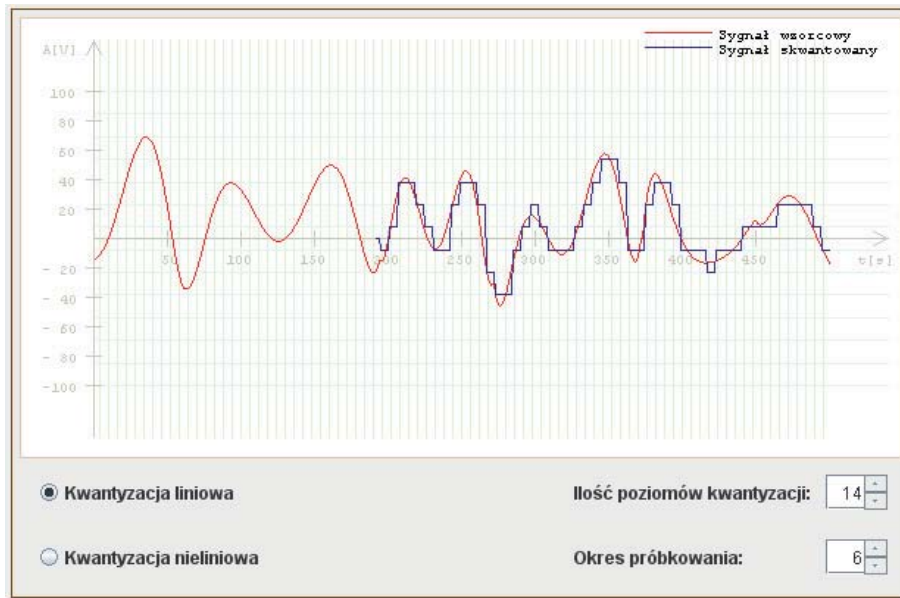
Fig. 3. Quantization – an interactive simulator.
Author: Jan Paluch. Source: www.informatyka.cm-uj.krakow.pl
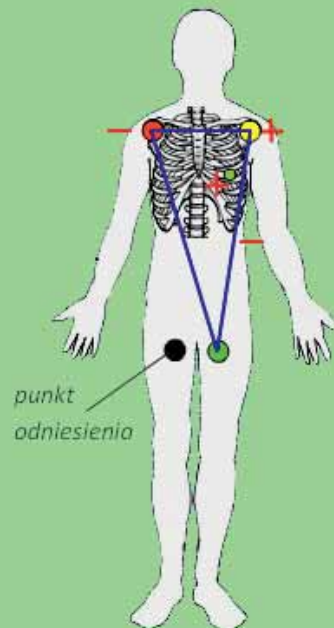
Fig. 4. The placement of EKG electrodes – an interactive simulator.
Author: Magorzata Tomanek. Source: www.informatyka.cm-uj.krakow.pl

Fig. 6. Measurement of eye movements – an animation.
Author: Michał Zając.
Source: www.informatyka.cm-uj.krakow.pl

Fig. 5. The placement of EEG electrodes –
an interactive simulator.
Author: Magorzata Tomanek.
Source: www.informatyka.cm-uj.krakow.pl
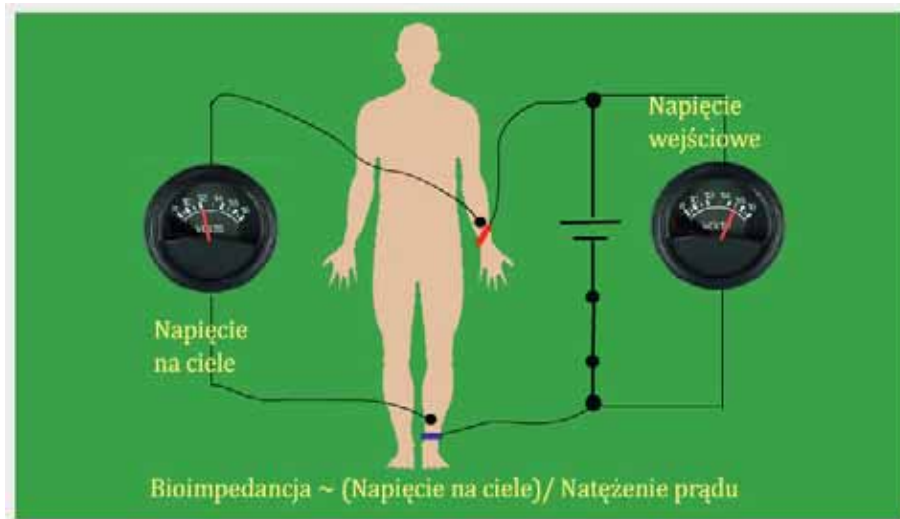
Fig. 7. Bioimpedance measurements – an animation.
Author: Michał Zając. Source: www.informatyka.cm-uj.krakow.pl
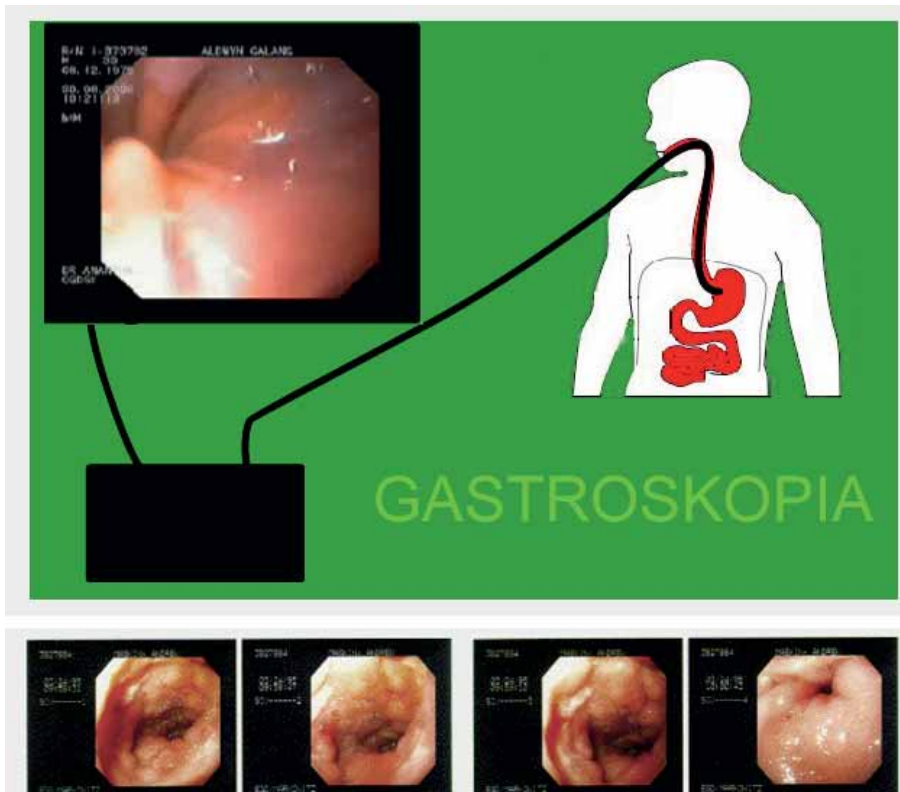
Fig. 8. Gastroscopy (Endoscopy) – an animation.
Author: Przemysław Kozioł. Source: www.informatyka.cm-uj.krakow.pl
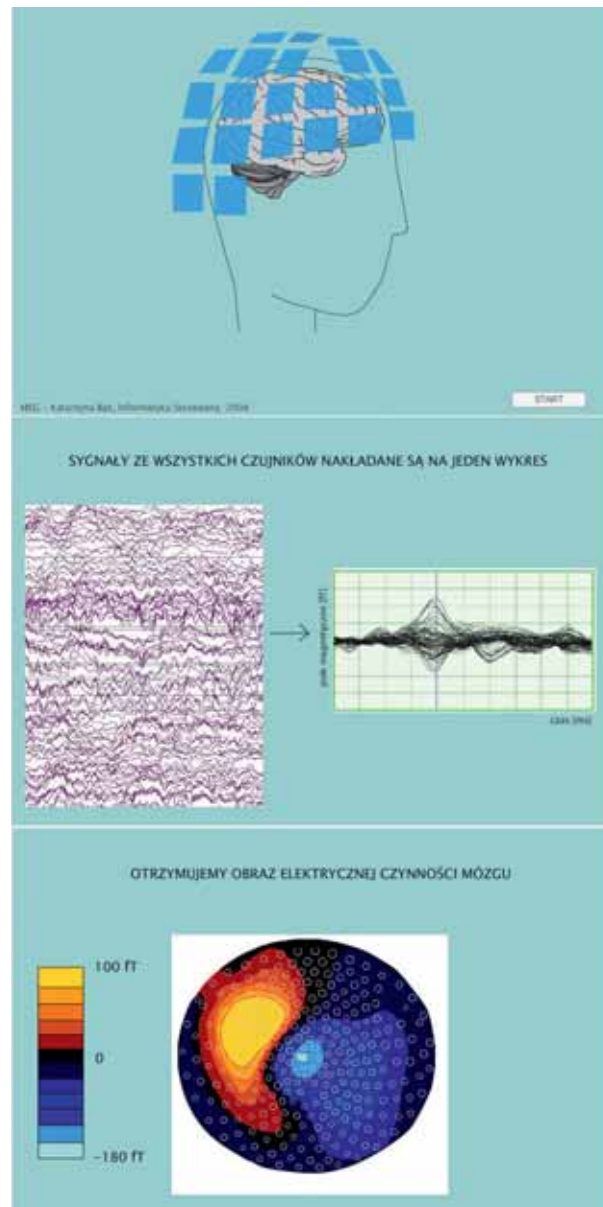
Fig. 9. Magnetoencefalography – an animation.
Author: Katarzyna Bąk. Source: www.informatyka.cm-uj.krakow.pl

- *Medical imaging*
o Endoscopy (see Fig. 8)
▪ Endoscopic measurement techniques
o Ultrasonography
▪ Ultrasound-based diagnostic imaging technique
▪ Ultrasound images of arteries in the neck
▪ Vertebral Artery Ultrasonography
▪ Musculoskeletal ultrasonography
▪ Laser Doppler flowmetry (see Fig. 11)
o Angiography
▪ Coronography (see Fig. 10)
o Tomography
▪ Computed tomography (see Fig. 12)
▪ Electrical impedance tomography
o Near infrared spectroscopy (NIRS)

o Image Recognition
o Microscope image processing
o Optoelectronics in medicine
o Morphometric parameters in biomedical measurements

Conclusion

Knowledge Management (KM) is becoming one of the fundamental goals of developed societies in which e-learning constitutes the standard educational method of an Flexible Learning Framework (FLF) which permits learning *just-in-case* and *just-in-time* [14-15]. E-learning can be viewed as an independent educational method or as an aide to classical didactic approaches. It can also constitute a set of techniques used for
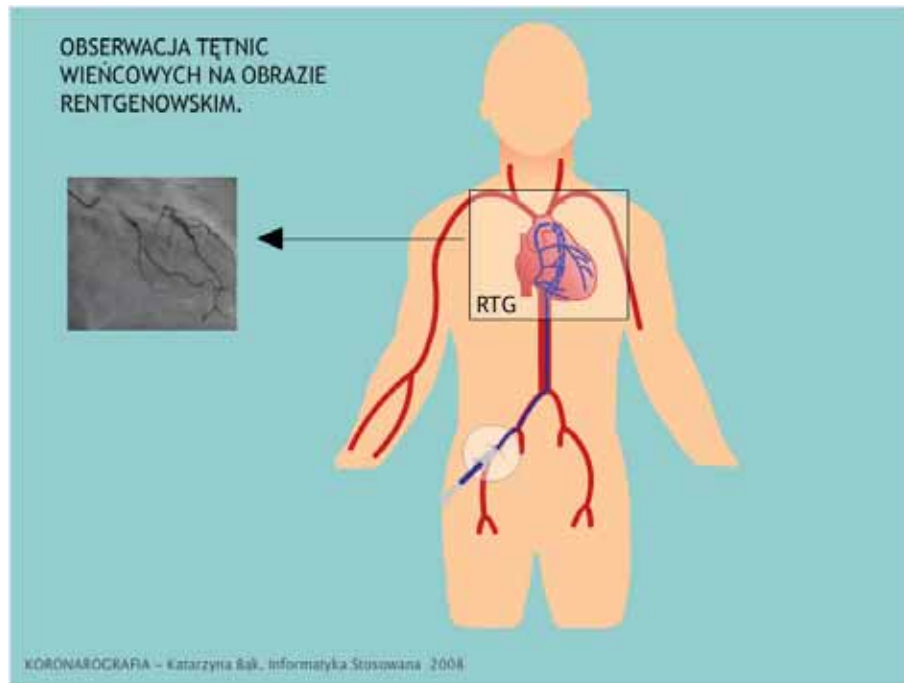
Fig. 10. Coronography – an animation.
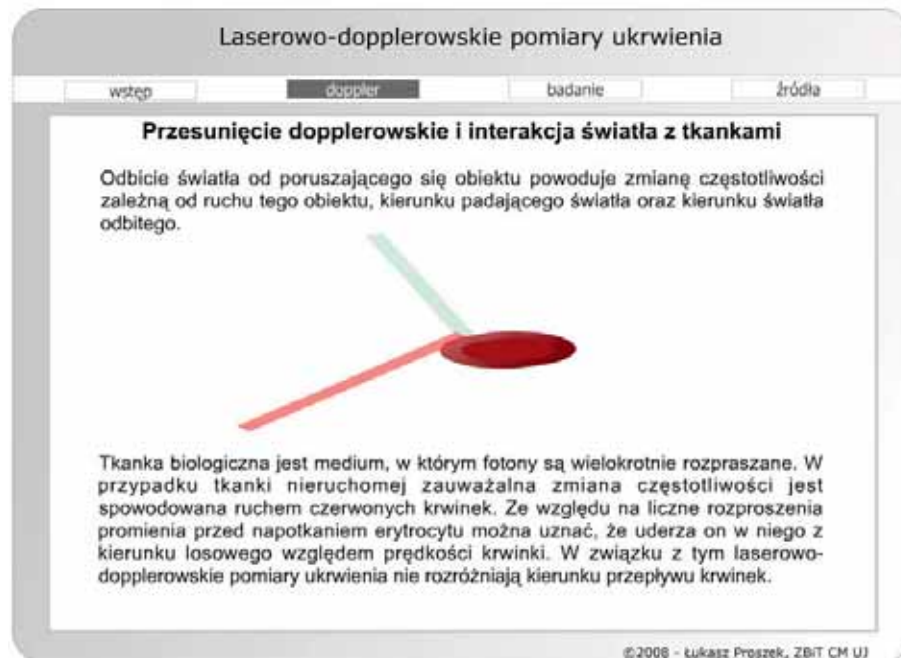Author: Katarzyna Bąk. Source: www.informatyka.cm-uj.krakow.pl

Fig. 11. Laser Doppler flowmetry – a flash presentation.
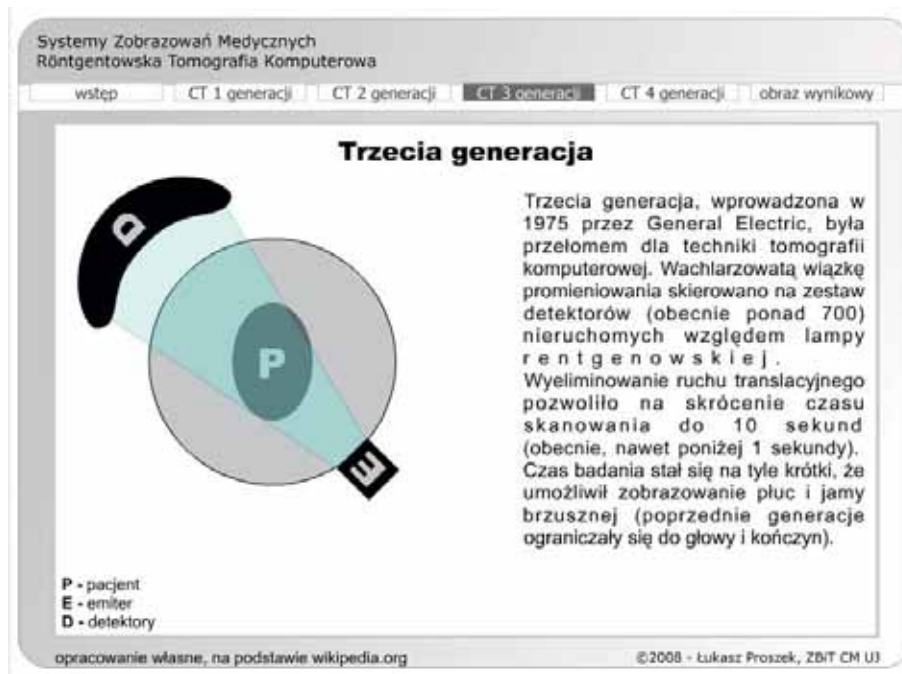Author: Łuksz Proszek. Source: www.informatyka.cm-uj.krakow.pl

Fig. 12. Computed tomography (CT) – a flash presentation.
Author: Łuksz Proszek. Source: www.informatyka.cm-uj.krakow.pl

keeping knowledge up-to-date and organized (for example, as an element of language courses) [16]. KM is above all a collection of important elements of informative and procedural (know-how) knowledge whose description and operationalization is very complicated [17]. Databases built using KM permit an optimal organization of access to data and ensure confidentiality and security [18-19]. The combination of KM and e-learning creates a new and effective platform which serves the transmission of knowledge and know-how in accordance with the needs of both corporate and individual users. Training can be based on internet databases supported by expert systems which incorporate different methods of artificial intelligence [20-21]. A series of telemedical accomplishments have already been achieved based on the use of advanced multi-level databases [22-31].

New educational techniques, as e-learning, which are used in instructing medicine in Department of Bioinformatics and Telemedicine UJ are based on Evidence-Based Medicine methodology [32-36].

Creative e-learning is a new form of organizing e-educational lessons. The most of educational tasks is executing by individual extraordinary students' homework. Our experience in teaching according to this paradigm has granted our expectations.

References

1. Friedrich K.R., Armer L. (1999), *The Instructional and Technological Challenges of a Web Based Course in Educational Statistics and Measurement,* ERIC Document: ED432262.

2. Pyrczak W., Lasoń W., Walecki P., Sarapata K., Trąbka J. (2003), *Teleinformatyczne systemy w edukacji medycznej,* w: *Komputer w edukacji,* red. J. Morbitzer, Wyd. Naukowe Akademii Pedagogicznej, Kraków, s. 226-230.

3. Walecki P., Lasoń W., Porębski G., Obtułowicz K. (2005), *Skórna postać nadwrażliwości na aspirynę – pilotażowy telemedyczny system informacyjno-diagnostyczny,* MSCB, p. 37.

4. Porębski G., Obtułowicz K., Obtułowicz P., Walecki P., Roterman-Konieczna I., Czarnobilska E., Krzanowska K., Kwapińska A., Targosz K. (2005), *Internet as a potential source of information about allergy for patients,* The XIXth World Allergy Organisation Congress, June 26 – July 1, 2005, Munich, Germany.

5. Porębski G., Walecki P., Obtułowicz K., Czarnobilska E., Krzanowska K., Kwapińska A., Targosz K. (2005), *Częstość wykorzystywania różnych źródeł wiedzy o alergii przez chorych ze schorzeniami atopowymi,* Allergy Asthma Immunology, 7th Central and Eastern European Conference, June 8-11, 2005, Łódź, Poland.

6. Jones E.R. (1999), *A Comparison of an All Web-Based Class to a Traditional Class,* ERIC Document: ED432286.

7. Williams M.L., Paprock K., et al. (1999), *Distance Learning: The Essential Guide,* Sage, Thousand Oaks, CA.

8. Keegan D.J. (1980), *On Defining Distance Education,* Distance Education, 1(1): 13-36.

9. Johansen R., Martin A., Mittman R., Saffo P. (1991), *Leading Business Teams: How Teams Can Use Technology and Group Process Tools to Enhance Performance,* Addison-Wesley, Reading, MA.

10. Tuzun H. (2003), *Przekształcanie kursów tradycyjnych na format WBI,* e-mentor, 2: 4-12.

11. Ingebritsen T.S., Flickinger K. (1998), *Development and Assessment of Web Courses That Use Streaming Audio and Video Technologies,* ERIC Document: ED422859.

12. Ingebritsen T.S., Brown G.G., Pleasants J.M. (1997), *Teaching Biology on the Internet,* ERIC Document: ED429536.

13. Pyrczak W., Roterman-Konieczna I., Łasocha B., Matuszewski P. (2004), *E-learning w naukach medycznych*, Krakowskie Studia Małopolskie, 8: 403-412.

14. Imel S. (1997), *Web-Based Training: Trends and Issues Alerts,* ERIC Document: ED414446.

15. Alavi M., Leidner D.E. (1999), *Knowledge management systems: issues, challenges, and benefits,* Communications of the AIS, 1(2).

16. Addicott R., McGivern G., Ferlie E. (2006), *Networks, Organizational Learning and Knowledge Management: NHS Cancer Networks,* Public Money & Management, 26(2): 87-94.

17. McInerney C. (2002), *Knowledge Management and the Dynamic Nature of Knowledge*, Journal of the American Society for Information Science and Technology, 53(12): 1009-1018.

18. Serenko A., Bontis N. (2004), *Meta-review of knowledge management and intellectual capital literature: citation impact and research productivity rankings*, Knowledge and Process Management, 11(3): 185-198.

19. Sarapata K., Usarz M., Trąbka J., Walecki P., Pyrczak W., Roterman-Konieczna I. (2004), *Security of net servers in biological approach,* Journal of Artificial Intelligence, Vol. 1, 1(23): 31-38.

20. Sarapata K., Bilo G., Pyrczak W., Kononowicz A., Roterman-Konieczna I., Kawecka-Jaszcz K. (2005), *Zastosowanie kryptografii w telekonsultacji kardiologicznej,* Bio-Algorithms and Med-Systems, Vol. 1, 1/2: 365-368.

21. Walecki P., Lasoń W., Wiśniowski Z., Roterman-Konieczna I. (2004), *Abductive logic programming in medical diagnosis*, Journal of Artificial Intelligence, Vol. 1, 2(24): 107-115.

22. Walecki P., Trąbka J. (2004), *Wnioskowanie abdukcyjne we wspomaganiu diagnozy lekarskiej*, Krakowskie Studia Małopolskie, 8: 389-401.

23. Walecki P., Pyrczak W., Lasoń W., Sarapata K., Roterman-Konieczna I., Klekawka T., Korab-Chrzanowska E., Balwierz W., *Skoordynowana baza danych w chorobie Hodgkina (HD) zorientowana na diagnostykę w systemie rozproszonym*, Przegląd Lekarski, t. 60, supl. 6, s. 78.

24. Walecki P., Sarapata K., Lasoń W., Pyrczak W., Roterman-Konieczna I., Balwierz W. (2004), *Telemedical Database of Hodgkin Disease*, Transformation of Healthcare with Information Technologies, Studies in Health Technology and Informatics, Amsterdam – Berlin – Oxford – Tokyo – Washington, DC, pp. 51-60.

25. Walecki P., Sarapata K., Lasoń W., Pyrczak W., Roterman-Konieczna I., Balwierz W. (2004), *Telemedical Database of Hodgkin Disease*, in: *E-health in Common Europe, Critical Conditions for Development of E-health Environment in Transforming Economies*, Kraków, s. 27.

26. Walecki P., Sarapata K., Lasoń W. (2004), *TDCHD – Zaawansowane rozwiązanie telebazy medycznej. Sztuczna inteligencja w inżynierii biomedycznej*, SIIB 2004, Kraków, s. 1-4.

27. Walecki P., Lasoń W., Porębski G., Obtułowicz K. (2005), *Skórna postać nadwrażliwości na aspirynę – pilotażowy telemedyczny system informacyjno-diagnostyczny*, MSCB, p. 37.

28. Walecki P., Lasoń W., Wiśniowski Z., Roterman-Konieczna I. (2004), *Abductive logic programming in medical diagnosis*, Journal of Artificial Intelligence, Vol. 1, 2(24): 107-115.

29. Lasoń W., Walecki P., Pyrczak W., Sarapata K., Trąbka J. (2003), *Zintegrowane telemedyczne bazy danych pacjentów z chorobami nowotworowymi*, w: *Sztuczna inteligencja – Organizacje wirtualne*, red. M. Kłopotek, J. Tchórzewski, Wydawnictwo Akademii Podlaskiej, Siedlce, s. 35-40.

30. Lasoń W., Walecki P., Trąbka J. (2004), *Telemedyczne bazy danych – teoria i praktyka*, w: *Społeczeństwo informacyjne. Wizja czy rzeczywistość?*, red. L. Haber, AGH Uczelniane Wydawnictwa Naukowo-Techniczne, Kraków, s. 349-360.

31. Sarapata K., Walecki P., Pyrczak W., Lasoń W., Roterman I., Szybiński Z. (2003), *Medical Database of thyroid cancer*, w: *International Conference on E-he@lth in Common Europe, June 5-6*, red. K. Zieliński, M. Duplaga, Kraków, p. 347-354.

32. Walecki P., Lasoń W., Pyrczak W., Roterman-Konieczna I., Sarapata K. (2006), *Szpitalny System Informacyjny (HIS) oparty na metodologii EBM*, w: *Pozyskiwanie wiedzy i zarządzanie wiedzą*, red. M. Nycz, M. Owoc, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, Wrocław, s. 179-188.

33. Lasoń W., Walecki P., Sarapata K., Roterman-Konieczna I. (2006), *Nauczanie medycyny w oparciu o paradygmat Medycyny Opartej na Dowodach Naukowych (Evidence Based Medicine)*, w: *Komputer w edukacji*, red. J. Morbitzer, Wyd. Naukowe Akademii Pedagogicznej, Kraków, s. 125--130.

34. Walecki P., Pyrczak W., Lasoń W., Sarapata K. (2006), *E-learning i telemedycyna – problemy strukturyzacji wiedzy*, w: *Komputer w edukacji*, red. J. Morbitzer, Wyd. Naukowe Akademii Pedagogicznej, Kraków, s. 243-248.

35. Żochowski A.P., Lasoń W. (2006), *Advanced system for organization of work and scientific data exchange*, Bio-Algorithms and Med-Systems, Vol. 2, 4: 65-77.

36. Lasoń W., Walecki P., Pyrczak W., Sarapata K., Roterman-Konieczna I. (2005), *Medycyna Oparta na Dowodach (EBM) – nowe podejście optymalizacyjne wykorzystujące systemy komputerowe*, w: *Computer Methods and Systems*, red. R. Tadeusiewicz, A. Ligęza, M. Szymkat, s. 455-462.

THE HOSPITAL SYSTEM OF PATIENT TREATMENT MANAGEMENT

MARCIN GODYŃ[1,2]

[1] *Department of Bioinformatics and Telemedicine, Collegium Medicum, Jagiellonian University, Lazarza 16, Krakow, Poland*
[2] *Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, Reymonta 4, Krakow, Poland*

Abstract: The effective management of gathered medical data is of utmost importance as accurate diagnosis entirely relies on the data; management cannot improve the data but by all means poor management can deteriorate any conclusion drawn from them.

Introduction

Nowadays health care has raised high expectations, not only about effective treatment of patients, but also about proper documentation and fast access archival information. Currently a lot of information can be stored in electronic form. The computer technology makes possible to store, archive and manage the medical data.

Patient cards, examination results, information about availability visit at specialist, state of materials etc., all of this forms huge amount of data that pile up, need to be processed and managed.

Currently in Poland just starts implementation of computer systems in public health care. There are many different technologies which allow to solve this problem. Proper choice of them will be a problem, but also will be a base to create efficient system.

I would like to focus at creating simple system for hospital which main job is to help access to information about patient, both for medical staff, and for patient.

Description of the project

Primary goal of the system will be to help access to all archives information about patient and history of patient diseases in hospital, medical examination results. Furthermore system will allow to reserve time of visit in the hospital clinic.

Because of continuous changes in rules concerning treatment of patients and changes in requirements of National Health Fund (NFZ) there is a need to design and to develop system in that way, to allow simple extension and reconstruction in adaptation to new requirements and rules.

System should allow to handle full patient support from the moment of registration in the hospital through diagnostic and any kind of treatments up to the moment of discharge of patient from hospital from hospital.

By applying uniform system one gets uniform appearance of patient cards and release cards. Currently on almost every ward they looks different.

Requirement concerning personal data security puts restraints as concerns access to the data according to role which user will be granted.

Data

In the system will be data placed by staff and also some dictionary data (like drugs and ICD10). All data will be stored in relational database ie.: PostgreSQL [1].

Dictionary data will contain information about medication admitted to trading in Poland, also classification diseases according to International Classification Diseases (ICD-10) [2] managed by World Health Orgranization.

Medication dictionary:
Medication dictionary consist of three tables. First of them is *leki* (drugs) table, second one is *leki_podmioty* (drugs-companies) table, which consist list of all companies and importers of medication, third table (*leki_postacie*)(drug-form) consist form of medicine.

Data about medication admitted for trade on territory of Poland are available to gain on web pages of Office for Registration of Medicinal Products [3].

Electronic health record

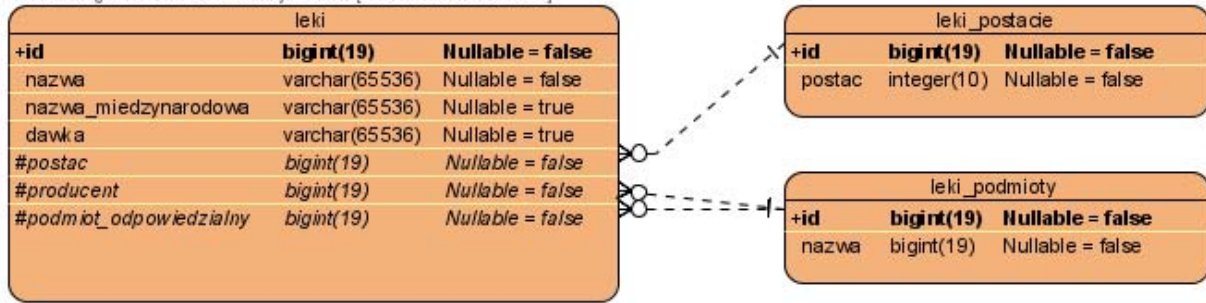Visual Paradigm for UML Community Edition [not for commercial use]

Fig. 1. Medication database schema [8]

Tools

System will be accessible as a web application [4,5], so from user side requirements are moderate.

From the server side is ability to split application on three separate layers:

- the presentation layer – the application server (JBoss AS [6]),
- the business layer (JBoss Seam Framework [7]) managing,
- and processing data and database layer (any relational database, ie. PostgreSQL).

Fig. 2. System layers (The user – www-level – Seam framework – database level) [8]

That architecture of the system allows to its easy scalability, from little system in which all layers (without user layer) will be on the same one server, to large system using by large hospitals, in which every layer may be on separate server or cluster.

In general the system server will not need large computing power (this lowers cost of hardware). All demanding operations can be scheduled for nights and weekends.

The system

For patient data security the system uses roles granted to the user at moment of logging. For correct login user does not need to have any role, but then will be no ability to make any operation. According to granted roles user will have access to proper function.

There are three ways for grant roles in the system :

- By manager, in the process of creating the doctors employed. The manager can grant roles: KIEROWNIK (Manager), LEKARZ (Doctor), LABORANT (Laboratory worker), PIELĘGNIARKA (Nurse), RECEPCJONISTKA (Receptionist).
- In the moment of creating patient there is automatically granted role PACJENT (Patient) if for person exist entry in pacjenci (patients) table.
- In the moment of log-ing according to duty chart is granted role LEKARZ_IP(ER Ward doctor).

If the doctor have ability to browse patient cards from his ward have to have duty on that ward.

Requirement of sensitive data security requires that server of Web application should allow only exclusive connections in encrypted HTTPS mode (using certificate of authenticity).

In fresh installed system is predefined user: *admin* with same password which has granted roles: KIEROWNIK (boss), LABORANT (lab staff), LEKARZ (doctor), PACJENT (patient), PIELEGNIARKA (nurse), RECEPCJONISTKA (registration). It is accessible right after installation, and should be removed after creating account for manager. It is created to avoid adding first user by a hand from database layer.

Patient card is the most important part of the system. It is divided into five tabs: examination, permanent condition, treatment, medical examination and discharge. Over them are placed general patient data like first name, second name, personal identification number (PESEL – ID), age, gender, race, ward, NFZ department, sign in date and extraction date. If pa-

Fig. 3. Login page [8]

Fig. 4. Patient card – empty [8]

tient age is less than one year is given in months, also if age is less than one month is given in days.

Conclusions

Installation and usage of the system is straightforward; system while remaining compact is ready for further upgrades.

First and most important is to change medicine database. My database is a simple extract from Office for Registration of Medicinal Products newsletter, but it has many pollution like tooth paste or cream for skin.

Second one is adaptation. From my observation looks that, different hospitals works in different way, so there is need to adapt application for its specification.

Another improvement can be creating reports for hospital management which may allow make better decision.

The system forms a good starting platform for indispensable nowadays patient data processing at typical contemporary health care centers.

References

1. PostgreSQL. http://www.postgresql.org.
2. ICD-10 Updates Pages. http://www.who.int/classifications/icd/icd10updates/en/.
3. Office for Registration of Medicinal Products. http://www.urpl.gov.pl/.
4. Wikipedia. http://pl.wikipedia.org/wiki/JavaScript.
5. World Wide Web Consortium (W3C). http://www.w3.org/TR/CSS2.
6. JBoss Application Server. http://www.jboss.org/jbossas/.
7. JBoss Seam framework. http://seamframework.org/.
8. Thesis: *System Obslugi Pacjentow* – Marcin Godyn.

Electronic health record

ELECTRONIC HEALTH RECORD

MICHAŁ TOMCZAK, WOJCIECH LASOŃ

*Department of Bioinformatics and Telemedicine, Collegium Medicum, Jagiellonian University,
Kopernika 7, 31-501 Krakow, Poland*

Abstrakt: Information technologies are widely present in modern health care and in teaching medicines students. It is important to create tools that will be universal and will employ newest technologies and medical standards. Creating an electronic health record application is a trial to implement such a tool in order to teach students how real life application works and how to work with medical dictionaries such as ICD-10.
Keywords: electronic health record, patient, PostgreSQL, ICD--10, Jboss Seam

Streszczenie: Rozwiązania z zakresu informatyki są integralną częścią wspomagającą różnorodne aspekty współczesnej służby zdrowia. Powinnością projektantów i wykonawców tych narzędzi jest dbałość o to, aby tworzone były zgodnie z najnowszymi, uznanymi standardami informatycznymi oraz medycznymi. Nowoczesne nauczanie medycyny wymaga równocześnie pełniejszego zapoznawania studentów z tymi rozwiązaniami.
Słowa kluczowe: elektroniczny rekord pacjenta, pacjent, PostgreSQL, ICD-10, Jboss Seam

Introduction

Modern health care faces a lot new requirements. Fulfilling those requirements is essential for its proper functioning; providing patients with comprehensive health services. „Due to vast technical Progress many areas of health care face necessity to introduce new technologies".

Nowadays all patients data, such as medical records, test results, images, bills and bookings, orders and medicines are stored in electronic format. Health care facilities generate was amount of data that needs to be stored, transformed and available at all times, twenty four hours a day, seven days a week.

Managing such a vast number of data is not easy, it requires a lot of effort from architect and software developer around the world to create tools that fulfill the need of modern health care.

The process of creating an application that will run in health care facilities is very complex. It starts with understanding how such facilities work, what is their business model. To understand those need software developers need to become aware of the health care system and the right to govern it. Close cooperation between doctors, nurses, health care administrative staff and software developers need to be established. There is a great need to create tools allowing to translate medical terms used in health service to terms used in modern IT. All this creates a great challenge for information technology specialists around the world.

Along with fast development of ideas and tools used in information technologies goes fast development of techniques and standards used in health care. A modern software developer faces a challenging task in creating and developing application that are used in health services.

One of consequences of that fact is grooving complicity of health services applications. Most of them cover lot of activities that happen in health care facilities. A lot of information is passed and produced every day. Therefore it difficult for a new health services employee to understand how such applications works and how to effectively use them. For that reason there is a need to teach basics of health care application on an academic level. Unfortunately it introduces limitations, like the amount of time a student and teach can spend on exploring health care software. That creates a need to develop simpler version of such software, purely for teaching purposes. Such software should cover some basic ideas of regular health service application. It should demonstrate some basic processes that are present in a real world application.

Thanks to that a student will know that such application exists and are there for him, to help with student's daily work.

In our work we propose such an educational tool. It covers some basic clinical paths so that a student can see how modern software tries to adjust to existing real life heath service situations.Using this software student can register a new patient, diagnose him, send him to surgery, prescribe medicines and finally sign him out.

Additionally while performing tests or diagnosing a patient student deals with ICD-10 dictionaries embedded into the application. Thanks to that he should be familiar with such dictionaries and will know how to deal with them in the future.

Technologies used to create ERP application

The application we proposed is based on Java programming language and Java J2EE (Enterprise Edition) platform. They are one of the most popular technologies nowadays with large community of users which guaranties good support and knowledge exchange. Tools that they provide cover large amount of techniques and technologies.

In the center of application resides JBoss Seam Framework. Seam is a development platform designed for building internet applications written in Java language. It integrates many technologies such as Asynchronous JavaScript and XML (AJAX) JavaServer Faces (JSF), Java Persistence (JPA), Enterprise Java Beans (EJB 3.0) and Business Process Management (BPM).

Seam design eliminates architectural and API complexity. It enables to create web based application in very little time consuming manner, using annotations, sets of UI components and just a bit of XML. Large amount of XML descriptor was a drawback of older technologies such as EJB 2.1 where user needed to create a large amount of XML descriptors to configure and run an application. In Seam such requirements are minimized therefore application development proceeds much easier and clearer.

Seam introduces a new approach to the concept of conversation, passing object instances is easy and eliminates many problems that were common in standard web application. Additionally Seam provides many different state management tools, defining granularity that enables a programmer to control lifespan created object in very precise and sophisticated way.

The data layer is based on PostgreSQL database Management System and Java Persistence API (JPA). Thanks to JPA we were able to easily map between relation database and Java objects. Therefore data manipulation in business logic layer was very easy and consisted only of operations related to JPA entities. There was no need of creating direct SQL calls and queries, all of that is performed by the JPA layer without any additional effort from the programmers side.

Presentation layer is based on Java Server Faces (JSF) and ICE Faces. JSF uses a component based approach to creating user interfaces. It integrates witch Seam in order to store the state of UI components. As an addition to JSF components ICE Faces were used to provide AJAX functionality to the UI.

Clinical path model introduced in the system

The diagram below demonstrates the model of clinical path implemented in the system. It is based on clinical paths present in real hospitals but adjusted in a way that it can be implemented using technologies described earlier. It is also a simplification of clinical path present in real health care facilities.

After successfull log in a new case (situation, patient) is presented to the user. The user must decide whether the new patient is known or unkown. Known patient is a patient whose data can be somehow entered into the system, someone that is either conscious, a relative can provide his data for him or has and id. An unknown patient is a patient that does not fall into criteria described for a known patient. That is a patient who is not conscious and there is no way to enter his data at the moment of signing him in to a hospital. Those to kinds of patients were separated because whole data model for them is very different. Known patient has a lot more data associated with him and is identified by his PESEL number. Unknown patient is indentified by an ID given to him by the system, data model for an unknown patient doesn't store personal or address information.

After this decision clinical path is split into two separate sub paths that have the same element but operate on different data. At this point a user fills forms with patients data. In case of an unknown patient that is only a brief state description.

After registering one or more patients a user goes to patients view. There he can see all his patients and decide who to diagnose first. Here an unknown patient can become a known patient if his data is found.

After choosing a patient to diagnose another step in the path begins. First a user performs a mandatory blood test and then performs test and apply medicines.

After that the user goes to final diagnosis stage where using ICD-10 dictionary one can set diagnosis and decide whether patient requires a surgery. And decide whether patient requires a surgery.

Decision regarding a surgery splits path in two. If a patient is send to surgery his state is changed in the system and a user can select procedures to perform from ICD-10 PCS dictionary.

If surgery is not required or after it is finished a patient is held on the ward for observation. Then some more tests can be performed and some additional medicines can be applied.

Last state that deals with patients is the sing out state where a user can provide some final data for the patient and finally sign him out.

Last state is the log out state where a user leaves the application.

ICD-10 and ICD-10 PCS dictionaries

One of the most important parts of health service are medical procedures and diagnoses. Many authorities are aware of that situation and therefore they have started the initiative to classify and codify all of them. As a result of this initiative two dictionaries were created, one that holds all the diseases and their codes and second one which is responsible for medical procedures.

There is a great need for students to be aware of those dictionaries and their structure. On the other hand it is impossible to properly diagnose a patient without access to some codified set of procedures and disease. Therefore making those dictionaries available in a medical application is very important.

All of the data associated with ICD dictionaries is stored in database records. It can be then retrieved from the database in many different ways. One of them is Java Persistence API (JPA). Thanks to JPA a programmer can access database through simple Java object without any need to create complex SQL queries and JDBC calls.

Below all of the tables and JPA entities related to ICD dictionaries are described. One database table has one JPA assigned to.
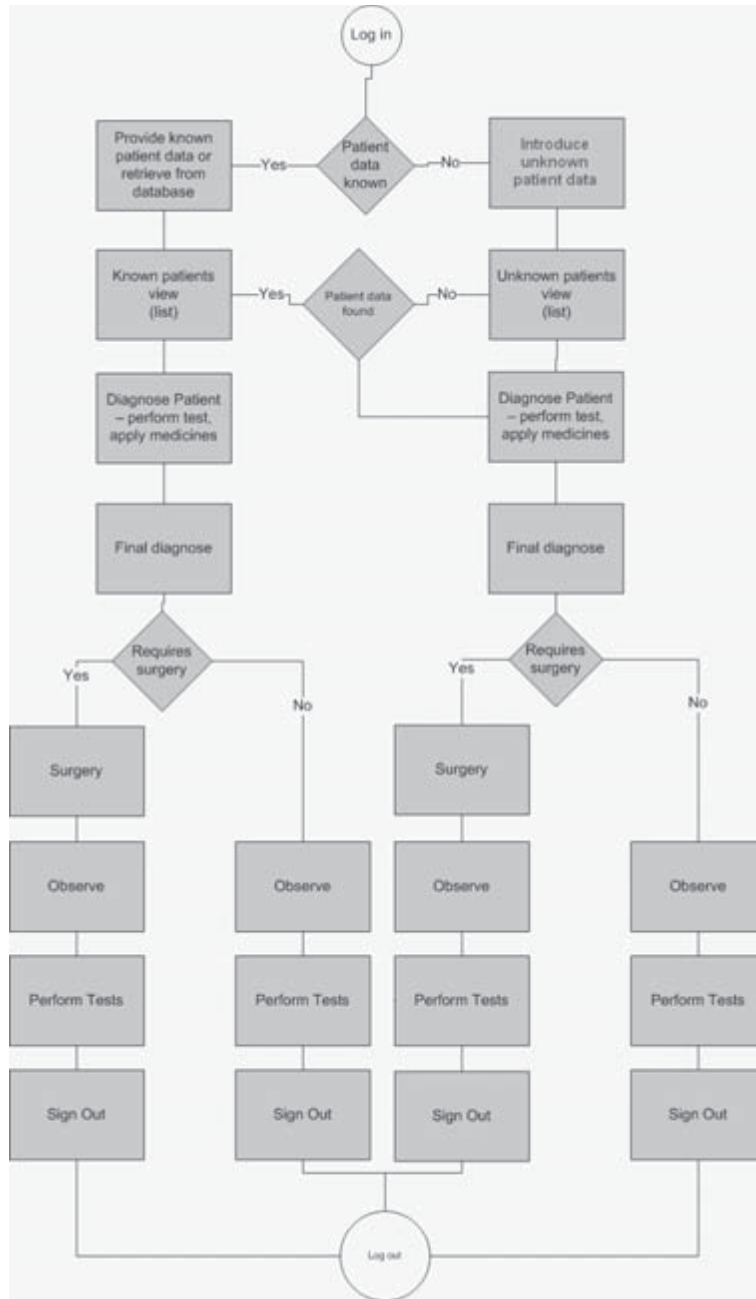
Fig. 1. System Path

Fig. 2. ICD Blocks

```
«Java Class, Entity, Name, Table, Scope»
            Ⓖ ICDChapter
─────────────────────────────────────────
🔧 serialVersionUID : long
 ▫ id : int
 ▫ block : String
 ▫ description : String
 ▫ chapter : String
 ▫ has_subcategory : boolean
 ▫ remarks : String
─────────────────────────────────────────
 ♂ ICDChapter ( )
 ♂ ICDChapter ( )
 ♂ ICDChapter ( )
 ● «Id, NotNull» getId ( )
 ● setId ( )
 ● getBlock ( )
 ● setBlock ( )
 ● getDescription ( )
 ● setDescription ( )
 ● getChapter ( )
 ● setChapter ( )
 ● isHas_subcategory ( )
 ● setHas_subcategory ( )
 ● getRemarks ( )
 ● setRemarks ( )
 ● «Override» toString ( )
```

icd_chapters	
PK	<u>id</u>
U1	**block** **description** **chapter** has_subcategory remarks

Fig. 3. ICD Chapters

```
«Java Class, Entity, Name, Table, Scope»
          Ⓖ ICDSubChapter
─────────────────────────────────────────
🔧 serialVersionUID : long
 ▫ id : int
 ▫ block : String
 ▫ description : String
 ▫ has_subcategory : boolean
 ▫ parent_category : String
─────────────────────────────────────────
 ♂ ICDSubChapter ( )
 ♂ ICDSubChapter ( )
 ♂ ICDSubChapter ( )
 ● «Id, NotNull» getId ( )
 ● setId ( )
 ● getBlock ( )
 ● setBlock ( )
 ● getDescription ( )
 ● setDescription ( )
 ● isHas_subcategory ( )
 ● setHas_subcategory ( )
 ● getParent_category ( )
 ● setParent_category ( )
 ● «Override» toString ( )
```

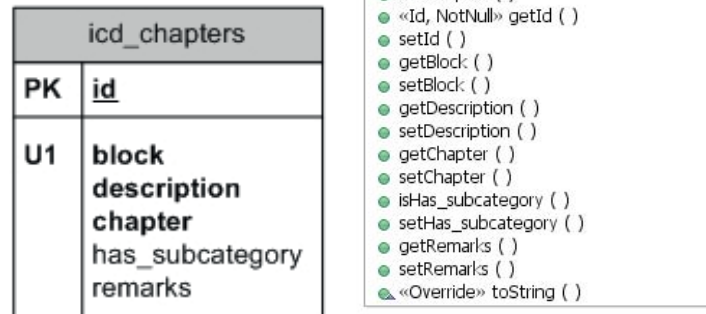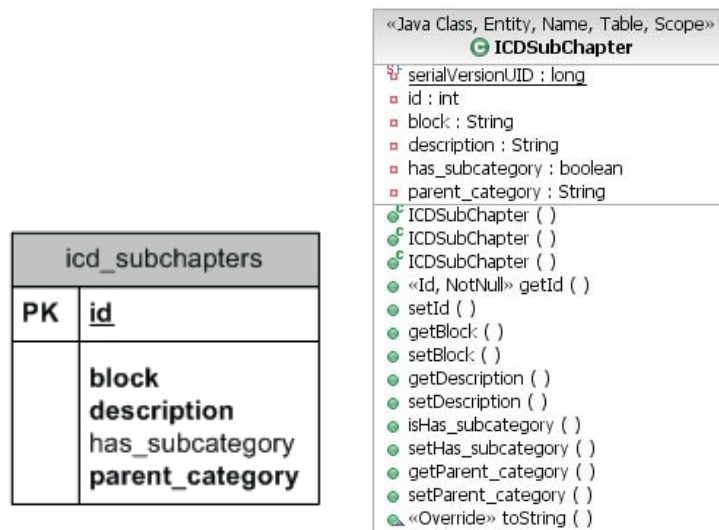| icd_subchapters	
PK	<u>id</u>
	block **description** has_subcategory **parent_category**

Fig. 4. ICD Subchapters

CD Blocks

ICD Blocks are on top of the ICD-10 dictionary hierarchy. They hold the most general information about a disease which is a group of diseases. Foe example: *Diseases of the circulatory system.* User firsts chooses the ICD Block, the process of obtaining a proper ICD Code begins form here.

The table structure is very simple. It contains of an unique identifier. Chapter number, block number and a title.

ICD Chapters

Chapters consist of more specific information. For example one of the chapters from *Diseases of the circulatory system* block *is Cerebrovascular diseases.*

Chapters allow user do find more detailed disease information, therefore narrowing the search criteria for a specific ICD code.
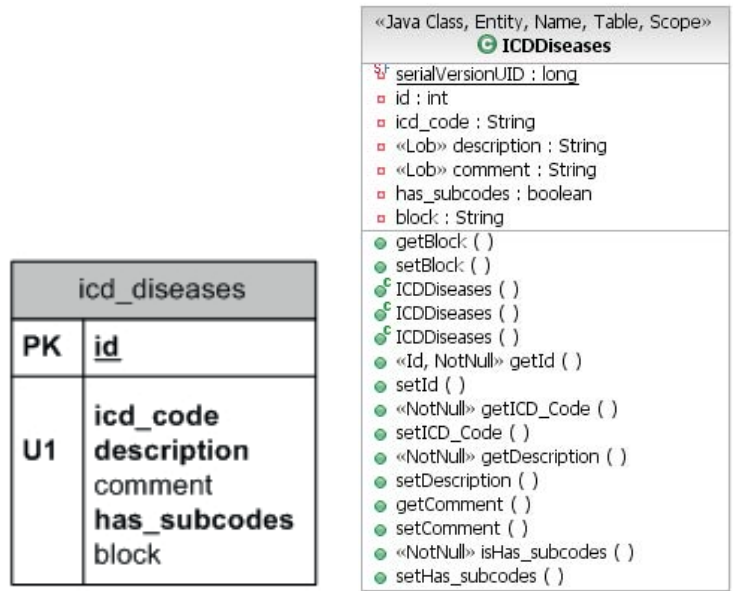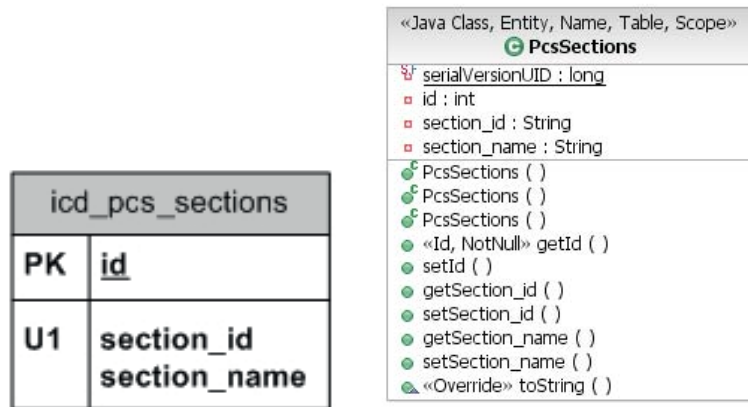
Fig. 5. ICD Diseases

Fig. 6. ICD PCS Sections

Chapters table is also very simple. Contains an unique identifier, information about a block to which a certain chapter belongs; short description of chapter code, some additional information and a logical variable which is used to determine whether a certain chapter contains some subchapters.

ICD Subchapters

Some ICD chapters contain big amount of information, therefore they have been divided into subchapters. For example chapter *Soft tissue disorders* contains three subchapters: *Disorders of muscles, Disorders of synovium and tendon and Other soft tissue disorders.*

Structure of this table is similar to the ICD Chapters table. Additionally it contains information about parent category.

Some of the subchapters contain other subchapters, there are also store in this table.

This table is the biggest table related to ICD dictionary. It has the most complex structure and also it holds a large amount of over 40 thousand records.

This table holds information about all disease stored in the ICD-10 dictionary.

It consist of and unique id field, a filed that holds an ICD Code for a disease, disease description, some comments, number a block to which this disease belongs and a Boolean variable that determines whether a particular disease has a sub diseases. For example: *Open wound of scalp has sub diseases: Unspecified open wound of scalp, Laceration without foreign body of scalp, Laceration with foreign body of scalp, Puncture wound without foreign body of scalp, Puncture wound with foreign body of scalp* and *Open bite of scalp.*
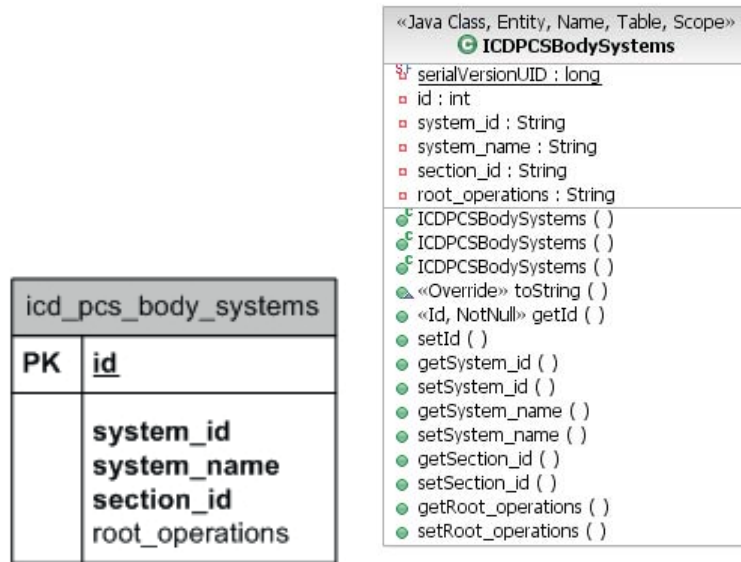
Fig. 7. ICD PCS Body Systems

Fig. 8. ICD PCS Root Operations

ICS PCS Sections

ICD-10 PCS is divided differently that ICD-10. The top most element of its structure are sections. Table structure for sections is very simple, it consists of a unique id, sections id, which is the first symbol in ICD-10 PCS code and a section name. From this point a user start to build an in ICD-10 PCS code for the procedure.

ICD PCS Body Systems

Just below section in ICD-10 structure are body systems. They represent the second level of the structure and body systems id is the second symbol in an ICD-10 PCS code.

The structure of table consists of a unique identifier, system identifier, system name used by user to identify the correct code (it is displayed in the UI); id of the section to which

this body system belongs and root operations that belong to this body system.

ICD PCS Root Operations

Root operations table is next in the ICD-10 structure. It has a simple structure similar to the tables that precede it.

The table consists of an unique identifier, root operation identifier, root operation name that is displayed in the UI and ad identifier of the section that the root operation belongs to.

ICD PCS Medical and Surgical Section

Analysis of the ICD-10 PCS dictionary shows that is has a large amount of data and that it can be divided into two separate sections. One of them are all medical and surgical procedures.
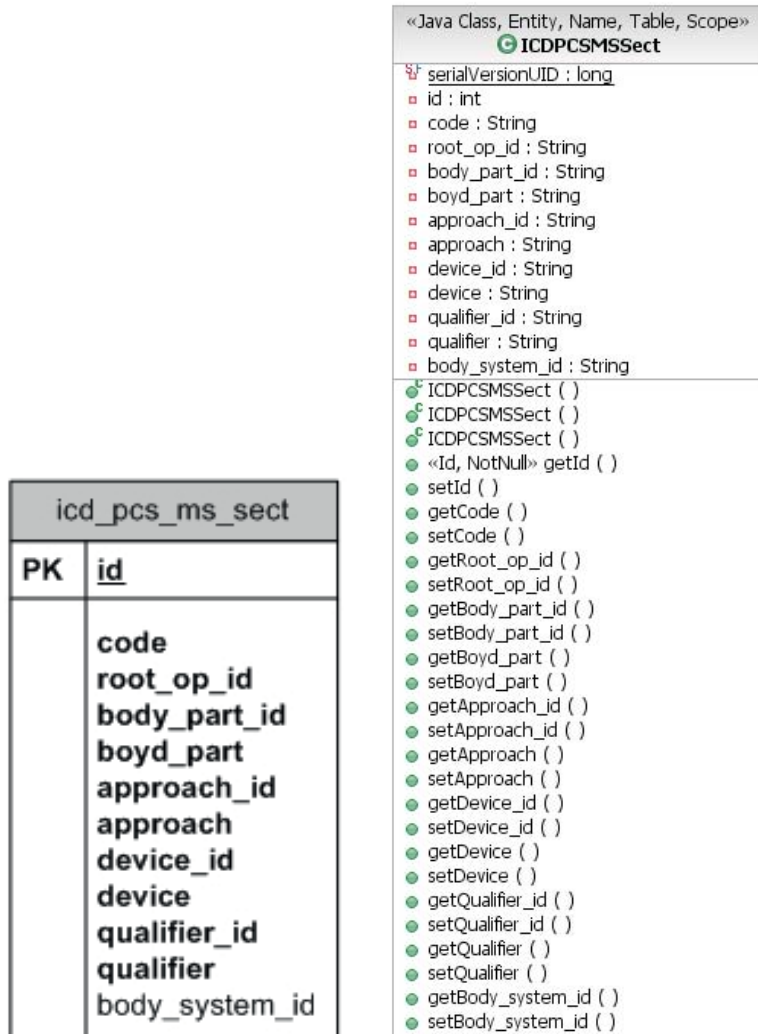
Fig. 9. ICD PCS MSSect

The second one are other ones that are not surgical. This division is done for two reasons, to simplify complicity and improve performance.

ICDPCSMSSect is a very complex table that holds a large amount of over seventy thousands records. It holds all the code from medical and surgical section.

It structure consists of:
– Identifier
– ICD-10 PCS code
– Id of he root operation to which particular procedure refers to
– Body part identifier which is a part of ICD-10 PCS code
– Body part name
– Approach identifier which is a part of ICD-10 PCS code
– Approach name
– Device identifier which is a part of ICD-10 PCS code
– Device name
– Qualifier identifier which is a part of ICD-10 PCS code
– Qualifier name
– Body system identifier to which particular procedure refers to

ICD PCS Rest of the sections

This table and entity has a structure identical to the ICDPC-SMSSect table.The only difference is that it stores different data which is related to every ICD-10 PCS code that doesn't belong to medical and surgical section.

Summary

Creating an Electronic Health Record application is a good way to learn how modern health service facilities operate, how
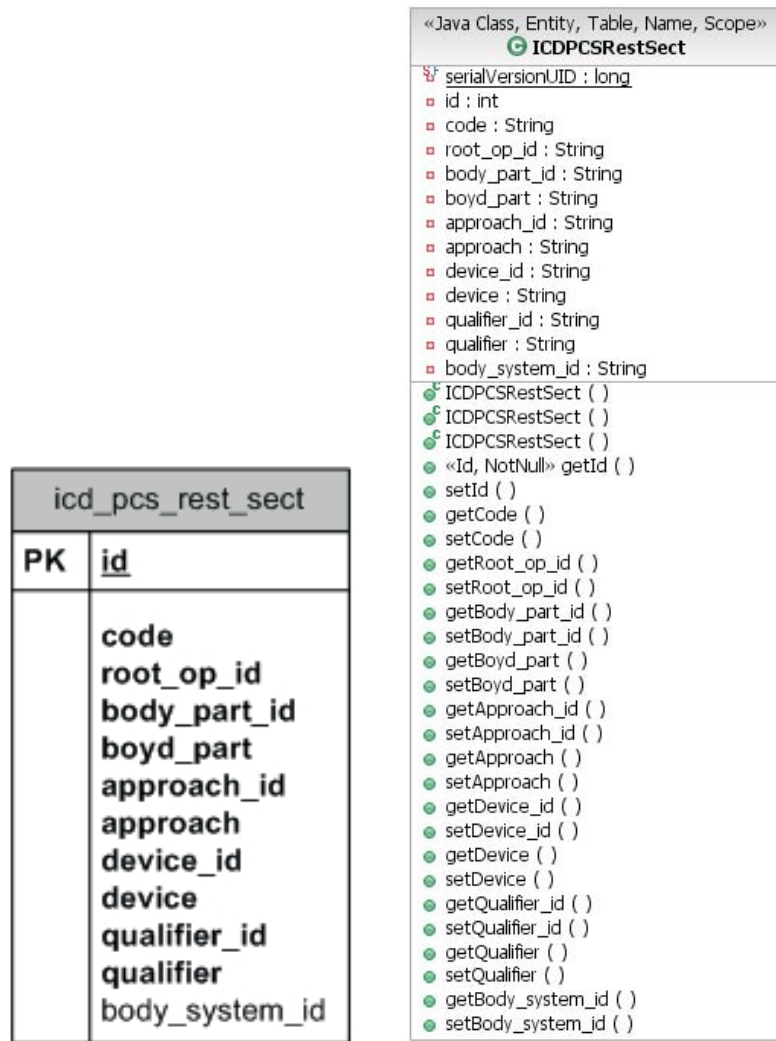
Fig. 10. ICD PCS Rest Sect

does their business model look and how all the people involved interact with the application and each other using it. There are a lot of processes and tasks that are simultaneously occur in such an environment, controlling them is crucial to proper functioning of health services.

New health service employees of hospitals and similar health care facilities have to face an unknown to them complicated world. This creates a need for tool that will teach them basic of this world and its rules. The tool that was created aims exactly that students and tries to teach them what they will encounter in the real world in the future.

References

1. http://www.cms.hhs.gov/icd10/.
2. http://www.seamframework.org.
3. http://www.icefaces.org/main/home/index.jsp.
4. http://www.who.int/classifications/apps/icd/icd10online/.

TELEDICOM ENVIROMENT AS A MODERN TELEMEDICAL TOOL FOR EDUCATION OF MEDICAL STUDENTS

WOJCIECH LASOŃ[1], ANDRZEJ A. KONONOWICZ[1], PIOTR OBTUŁOWICZ[2], TOMASZ KUŁAGA[1]

[1] *Department of Bioinformatics and Telemedicine, Jagiellonian University, Collegium Medicum, Kopernika 7, 31-501 Cracow, Poland*
[2] *Radiology Department, Jagiellonian University, Collegium Medicum, Kopernika 19, 31-501 Cracow, Poland*

Abstract: The article presents the concept of scenario of telemedicine classes for medical faculty students at the Jagiellonian University Medical College. The aim was to present practical exercises using modern tools to conduct remote medical consultations. The idea involves exercises simulate a real medical consulting by using remote collaboration system TeleDICOM. Students working in groups of 2-3 people are preparing a case by case, searching elementary information and then present to the specialist – consultant using telemedical tools. At the same time the remaining group of students listen to and watches the consultation, which is displayed through a multimedia projector. Classes are held in three physically separate locations and are interactive. Proposed scenario largely reflects real medical teleconsultation giving an impression of a real medical consultation session.
The high rating of class issued by the students confirmed the desirability and effectiveness of the proposed classes scenario.
Keywords: telemedicine, teleconsultation, medical e-learning

Streszczenie: W artykule zaprezentowano koncepcję scenariusza zajęć z zakresu telemedycyny dla studentów kierunku lekarskiego UJ Collegium Medicum. Celem ćwiczeń było przedstawienie praktycznego wykorzystania współczesnych narzędzi informatycznych do przeprowadzania zdalnych konsultacji medycznych. Pomysł ćwiczeń zakłada zasymulowanie rzeczywistej konsultacji zdalnej przy wykorzystaniu systemu TeleDICOM. Studenci pracując w grupach 2-3 osobowych przygotowują określony przypadek, wyszukują podtawowe informacje, a następnie prezentują specjaliście – konsultantowi wykorzystując narzędzie teleinformatyczne. Równocześnie pozostała część grupy przysłuchuje się i obserwuje konsultację, która wyświetlana jest poprzez rzutnik multimedialny. Zajęcia odbywają się w trzech, fizycznie odrębnych lokalizacjach i mają charakter interaktywny. Proponowany scenariusz w dużym stopniu odzwierciedla rzeczywistą telekonsultację medyczną. Wysoka ocena ćwiczeń wystawiona przez studentów potwierdza celowość i skuteczność zaproponowanego scenariusza.
Słowa kluczowe: telemedycyna, telekonsultacja, e-learning w medycynie

Introduction

Knowledge of the use of modern telemedical tools becomes an important element in the medical education of 21st century. As an example of such a tool the TeleDICOM system for medical teleconsultation is presented at the Department of Bioinformatics and Telemedicine, Jagiellonian University Medical Collage (Zakład Bioinformatyki i Telemedycyny – ZBiT).

The simplest definition of teleconsulting says that it means a use of computers and telecommunications technology to share electronic medical records and to deliver medical consulting from a distance.

Interactive class using TeleDICOM system is a part of coursework to illustrate how tools in telemedicine are used. The clinical version of this system has been implemented in several medical centres including the John Paul II Hospital in Cracow. An educational version of the system has been prepared for use at the ZBiT.

The program of practical exercises for students foresees that several consultations will be carried out in accordance with the principles of clinical practice. Each team of two to three students will be assigned to perform a consultation about a specific case. These consultations which are recorded in the database have been previously passed onto the server. Each case description contains several (usually four) images taken by different diagnostic imaging devices. Thus the students also have this opportunity to compare imaging methods and to learn the differences between them. A written description of the case, symptoms, blood count results, etc. is also attached. Based on the case description, each team gathers fundamental information about the case from the Internet.

TeleDICOM enviroment

TeleDICOM is an environment for collaborative and interactive work on medical documents. Application integrates an interac-

tive teleconsultation system with a medical database which contains medical images and documents. The idea of the system is to provide an easy but powerful way to exchange expert knowledge from any place where computer network is available. System was created by Distributed Systems Research Group Department of Computer Science at AGH – University of Science and Technology [1].

System is characterized by user friendly interface, is easy to use and fulfill requirements of security and privacy protections. Is capable of present information for an optimal work flow, route and manage medical cases and integrating into Hospital / Clinical Information System (HIS/CIS), Radiology Information System (RIS), Picture Archiving and Communication System (PACS) and other medically related databases.

System supports **DICOM** standard, using advantage of information supplied by DICOM. According to National Electrical Manufactures Association (NEMA) definition DICOM (Digital Imaging and Communication in Medicine) is a document which defines a method of communication for the various types or equipment of digital medical imaging devices or softwares (modalities). This standard for handling, storing, transmitting, displaying and printing information in medical imaging is now widely used by the majority of medical imagering hardware manufacturers. DICOM represents a progress in the interconnection of various imaging modalities, through a *Document of Conformity* or *Conformance Statement* emitted for each device or software following this standard. DICOM enables the integration of printers, scanners, workstations, servers and network hardware from multiple manufacturers into a picture archiving and communication system (PACS). DICOM standard includes a file format definition and a network communications protocol. The file format definition contains both all image data and header. Image data can be compressed to reduce the image size. Header contains textual information about the patient's name, the type of scan, image dimensions, etc. Standard is developed by the *DICOM Standards Committee*, and the NEMA

holds the copyright to DICOM standard, known as NEMA Standard PS3, and as ISO Standard 12052 [18].

Measurements can be made on the images during the TeleDICOM consultation. The real distance between two points, surface area or angles can be measured. Tools for volume estimation have also been provided. To make a measurement the image should first be calibrated (that is, one should indicate what real distance corresponds to a selected segment of the image). DICOM-type files are automatically calibrated. For images in other formats the calibration should be carried out manually using the appropriate tool. To make the calibration, one should chose two points on the image (this is the reference segment) and then type in the corresponding real distance. Some images have a special printed scale for manual calibration.

System offers Hounsfield Window adjustment. The **Hounsfield Scale** quantitatively indicates how much x-ray radiation is absorbed by different types of tissue and Hounsfield Unit is a measure of the weakening of the radiation in a volume element of tissue. By convention it takes on values from – 1024 to 3071 (a – 1024 value signifies vacuum, values close to – 1000 air, 0 – water, soft tissues: 20 – 80, calcification: 80-400 and values range 400-1000 corresponds to bone).

The main window of TeleDICOM system is presented below.

The TeleDICOM program also makes it possible to modify the Hounsfield Window (HW). The window defines the upper and lower limits of this range – WW (window width) is the width (range) of the window. The level is the Hounsfield number at the centre of the window – WL (window level) is the central (median) value of the HW. Each point in the diagnostic image (pixel) is assigned a value in Hounsfield Units (HU). Because the human eye is not capable of distinguishing 4000 shades of grey, a greyscale of 256 shades covering a specific range of HU values is applied in practice. Points with HU values outside of the WW range are either black or white.
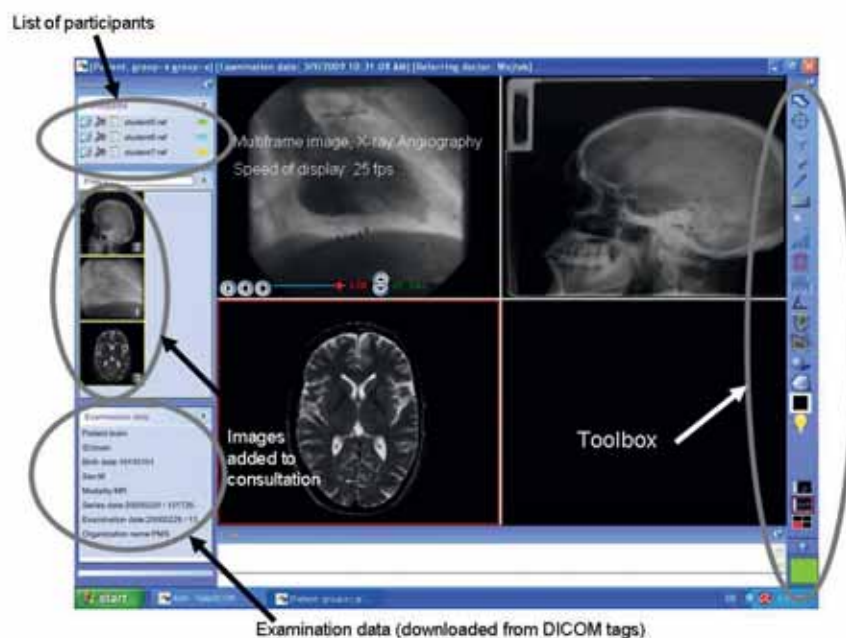

Fig. 1. The main window of TeleDICOM system

Medical diagnostics

The available grey scale is spread over the chosen range of Hounsfield numbers. To produce an image which shows up most major structures a large window is used. For more detailed information about tissues with very similar density a small window is used. The smaller window – more detailed image but the range of tissue density that is seen is reduced. This is chosen so that the window covers the type of tissue you are interested in. To image dense tissues a high level is used and to image low density tissues a low level is used.

TeleDICOM offers also specialized instruments operating on single and multiframe medical images, integrated conferencing tools such as interactive voice, text and pointer exchange, chat giving an impression of a real medical consultation session performed without forcing doctors to move from their workplace.

Teleconsultation

The learning scenario involves three computers situated in remote locations connected to a single TeleDICOM session. Consultation takes part between physician and a pair of students, which are equipped with headphones and microphones. Students based on obtained data present the medical case to the specialist – radiologist. Physician remotely explains diagnosis and treatment methods to the students.

Preparing the consultation

The steps of preparing a consultation are:
- Inform all consultation participants about the date and subject of the meeting
- Log on to the system (user's name, password)
- Find the assigned consultation in the server
- Download the consultation on to the user's local computer

Similar steps are taken by all of the consultation participants. These actions should be completed before the session, so that the system is fully ready when the session begins. This solution is very effective from the point of view of the system's functionality. During the session there is no more need for transferring images because everything has already been downloaded locally to the users' workstations. Only the pointers added by the users and the results of measurements carried out on the images will need to be transmitted.

Conducting the consultation

As depicted in the diagram in Fig. 2, the participants in the consultation are the team of students which presents the specific case and a physician specialist who is a radiologist-consultant and has experience in the analysis of diagnostic medical images.

Both of these two active participants and one passive participant, the auditorium of observers, are physically located in three different locations. After connecting the chosen consultation a list of participants appears and a different pointer colour is assigned to each participant. The set of radiographic images is available and additional information about the medical tests can be accessed by reading the DICOM-tagged information.

The students' team presents the case and the most pertinent information based on the provided case description. The students also share the information they have found in the Internet. Next, they formulate the medical problem by posing a question to the medical consultant. Now the principal part of the conference begins: the joint analysis of the attached images. The participants use attachable or interactive pointers to indicate the precise parts of the images that they are referring to.
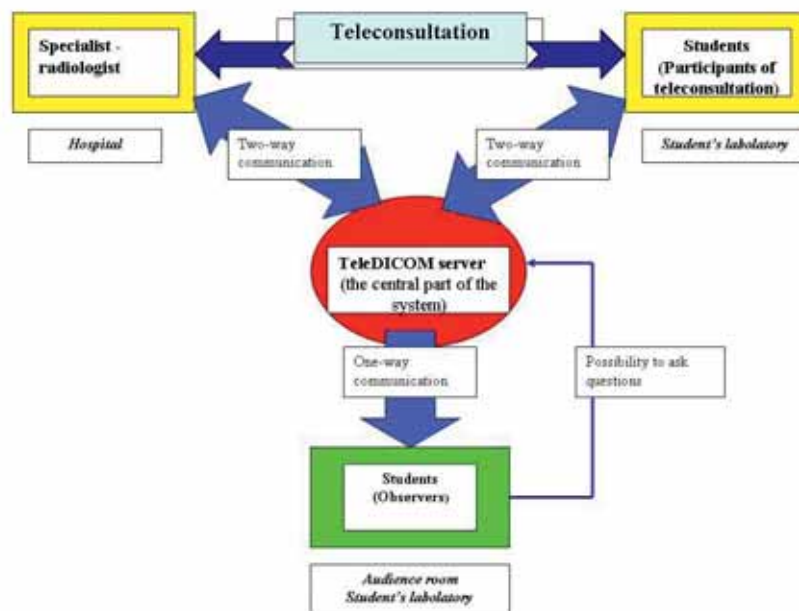

Fig. 2. The concept of teleconsultation scenario

Medical diagnostics

Measurements of distances between two chosen points, surface areas of closed regions and angles between three chosen points can also be made. The consulting physician (radiologist) points out the characteristic features of the images and at the same time analyses the case. He or she can also ask the students questions regarding how the measurements are made or about specific fragments of the image. For the sake of education several images are analysed, each of them having been made by a different diagnostic imaging technique. Multi-frame images can be played as a sequence. The consultation ends with the radiologist presenting the medical diagnosis based on the analysis of the attached images. Therapeutic methods are also discussed. Two-way verbal communication, text communication (chat) and communication via interactive pointers are employed during the session.

During this entire time the teleconsultation is observed by the remaining students in the projection room (auditorium or lab). Projection by means of a multimedia projector provides the possibility of a detailed view of the images. The students are passive observers in the consultation and do not take an active part in it. However, they can constantly exchange comments among themselves. This scenario can also be enriched by adding the possibility for the observing students to ask the specialist physician questions. However, one should be aware that in this case audio feedback can occur.

Exercises for the students

Below is a brief description of the practical exercises for students. It illustrates how a medical teleconsultation using the TD system is conducted. The first part of the exercises entails becoming acquainted with the system's structure, interface, basic operations and the toolbox. In the second part, the students participate actively in a session by conducting a teleconsultation regarding a specific case.

Presentation of the TeleDICOM medical teleconsultation tool

The following steps should be made during the class:
- Becoming familiar with the conditions, benefits and limitations of teleconsultations
- Logging onto the system
  - User's name and password
- Setting up a new consultation
  - Selecting the consultation's priority level
  - Introducing the patient's personal data
  - Adding images generated by various diagnostic imaging devices
- Preparing for a consultation
  - Finding the consultation on the server,
  - Importing the consultation onto the local computer
  - Activating the consultation

Conducting the consultation

The following steps should be made by students during the class:

- Identifying the topic of the teleconsultation:
  - Becoming familiar with the case, gathering basic information
  - Informing the participants about the topic of the consultation
  - Setting a date and time for the meeting
- Preparing the TeleDICOM environment:
  - Logging onto the system
  - Adding all of the participants to the consultation
  - Finding the case on the server
  - Downloading the consultation onto the local computer
- Adjusting the microphone sensitivity and headset volume
- Verbal presentation of the problem by the presenting team
- Loading the images which are to be analysed into the working window
- Identifying and locating organs in the image
- Outlining essential parts of the image
- Attaching pointers
- Making measurements:
  - Distances between two points (for example, to measure the diameter of the oesophagus), surface areas and angles
  - Measuring surface areas
  - Measuring chosen angles
- Setting the H-window parameters in order to highlight specific tissues
- Formulating and posing the question by means of the text chat tool
- Receiving exhaustive information from the consultant about the case being analysed, especially regarding the diagnosis and the methods of therapy
- Ending the teleconsultation session
- Ending work in the system, logging out

After the teleconsultation the students are required to write a report regarding both the medical aspects of the specific case as well as a brief description of the actions taken and measurements made during the session.

Classes assesment

The teachers highly appreciated TeleDICOM as a tool in teaching of medicine. The course conducted in the accepted scenario was interesting and engaging students in solving by themselves the medical cases. Additionally students had the opportunity to get acquainted with real operating system to conduct remote medical consultations. Slightly worse ratings were issued by system administrators. They pointed to some technical problems mostly related to the voice transmission.

Practical exercises using the TeleDICOM system according to the scenario presented here were conducted in the regular medical studies programme. The exercises for the latter group of students were carried out using the English language version of the system and in English.

A short survey was conducted among the English division students asking them to assess the system and its usefulness in the educational process. The students used a scale of 1 to 6 (completely useless, not very useful, moderately useful, useful, very useful, excellent) to express their view about the system as an aide in the formation of medical students according to the scenario presented here. Thirty seven students responded
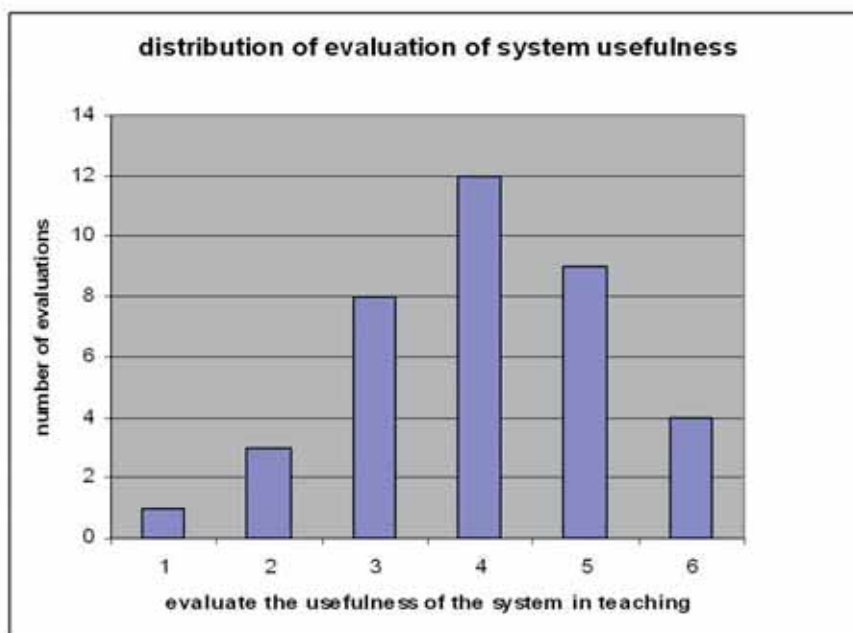
Fig. 3. Evaluation of the usefulness of TeleDICOM system in teaching medicine

in the survey. Histogram Fig. 3 presents a histogram of the responses.

Experience and appreciation classes held for students confirmed the value of the system and the proposed learning scenario of medical education classes, which simulates the real medical teleconsultation.

References

1. Cała J., Czekierda Ł. (2003), TeleDICOM – environment for collaborative medical consultations, w: International Conference on E-he@lth in Common Europe, June 5-6, 2003, Krakow, pp. 307-322.
2. Gackowski A., Cała J., Czekierda Ł., Zieliński K., Nowak M., Powroźnik P., Wolański M., Kruszyński M., Podolec P. (2006), Cyfrowa archiwizacja i telekonsultacja badań echokardiograficznych, w: Echokardiografia praktyczna, t. 4, red. P. Podolec [et al.], Kraków: Medycyna Praktyczna, s. 587-599.
3. Jaworek J., Kańtoch E. (2009), Telemedycyna – nauka przydatna dziś, niezbędna jutro, w: Podstawy inżynierii biomedycznej, t. 2, red. nauk. R. Tadeusiewicz, P. Augustyniak, Kraków: Wydawnictwa AGH, s. 393-398.
4. Kononowicz A.A., Roterman-Konieczna I. (2006), E-Learning as a New Tool in Medicine Teaching on the Example of an Introductory Statistics Course, Bio-Algorithms and Med-Systems, 2(3), pp. 71-75.
5. Koprowski A., Laurentowski A., Radziszowski D. (2003), Telemedycyna w praktyce – modele telekonsultacji medycznych oraz ich wykorzystywanie w ramach Krakowskiego Centrum Telemedycyny, Zdrowie i Zarządzanie, t. 5, nr 1, s. 71-76.
6. Kosiński J., Nawrocki P. (2003), Technical Aspects of a Teleconsultation Organization, w: International Conference on E-he@lth in Common Europe, June 5-6, 2003, Krakow, pp. 187-196.
7. Lasoń W., Pyrczak W., Roterman-Konieczna I. (2006), Edukacja przyszłości – e-learning i telemedycyna, Krakowskie Studia Małopolskie, 10, s. 275-282.
8. Lasoń W., Walecki P., Sarapata K., Roterman-Konieczna I. (2006), Nauczanie medycyny w oparciu o paradygmat Medycyny Opartej na Dowodach Naukowych (Evidence Based Medicine), w: Komputer w edukacji, red. J. Morbitzer, Kraków: Wydawnictwo Naukowe Akademii Pedagogicznej, s. 125-130.
9. Lasoń W., Walecki P., Pyrczak W., Sarapata K., Roterman-Konieczna I. (2005), Medycyna Oparta na Dowodach (EBM) – nowe podejście optymalizacyjne wykorzystujące systemy komputerowe, w: Computer Methods and Systems, red. R. Tadeusiewicz, A. Ligęza, M. Szymkat, Kraków, s. 455-462.
10. Martyniak J. (red.) (2005), Podstawy informatyki z elementami telemedycyny: ćwiczenia dla studentów medycyny, Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
11. Pyrczak W., Lasoń W., Walecki P., Sarapata K., Trąbka J. (2003), Teleinformatyczne systemy w edukacji medycznej, w: Komputer w edukacji, red. J. Morbitzer, Kraków: Wydawnictwo Naukowe Akademii Pedagogicznej, s. 226-230.
12. Sarapata K., Bilo G., Pyrczak W., Kononowicz A., Roterman-Konieczna I., Kawecka-Jaszcz K. (2005), Zastosowanie kryptografii w telekonsultacji kardiologicznej, Bio-Algorithms and Med-Systems, Vol. 1, No. 1/2, pp. 365-368.
13. Tadeusiewicz R. (2007), Systemowe podejście do wybranych zagadnień telemedycyny, w: Techniki informacyjne w badaniach systemowych, red. P. Kulczycki, O. Hryniewicz, J. Kacprzyk, Warszawa: Wydawnictwa Naukowo-Techniczne, s. 341-359.
14. Tadeusiewicz R. (2006), Telemedycyna jako ważny, ale trudny składnik społeczeństwa informacyjnego, w: Spo-

łeczeństwo informacyjne: doświadczenie i przyszłość, red. nauk. G. Bliźniuk, J.S. Nowak, Katowice: Polskie Towarzystwo Informatyczne, Oddział Górnośląski, s. 99-123.

15. Walecki P., Lasoń W., Pyrczak W., Roterman-Konieczna I., Sarapata K. (2006), Szpitalny System Informacyjny (HIS) oparty na metodologii EBM, w: Pozyskiwanie wiedzy i zarządzanie wiedzą, red. M. Nycz, M. Owoc, Prace naukowe Akademii Ekonomicznej we Wrocławiu, s. 179-188.

16. Walecki P., Pyrczak W., Lasoń W., Sarapata K. (2006), E--learning i telemedycyna – problemy strukturyzacji wiedzy, w: Komputer w edukacji, red. J. Morbitzer, Kraków: Wydawnictwo Naukowe Akademii Pedagogicznej, s. 243-248.

17. Żochowski A.P., Lasoń W. (2006), Advanced system for organization of work and scientific data exchange, Bio--Algorithms and Med-Systems, Vol. 2, No. 4, pp. 65-77.

18. http://medical.nema.org/.

Medical diagnostics

THE INFLUENCE OF COFFEE ON AN EYEBALL MOVEMENT

PIOTR WALECKI*, WOJCIECH LASOŃ*, EDWARD J. GORZELAŃCZYK#, IRENA ROTERMAN*

** Department of Bioinformatics and Telemedicine, Jagiellonian University, Medical College,
Kopernika 7 st., 31-034 Krakow, e-mail: pwalecki@cm-uj.krakow.pl*
*# Department of Human Physiology, Chair of Physiology, Nicolaus Copernicus University, Ludwik Rydygier
Collegium Medicum, Jagiellońska 13-15, 85-067 Bydgoszcz*

Abstract: This paper presents a research of influences of coffee on dynamic parameters of eyeballs movements. In this study we made use of a head-mounted saccadometer which measures parameters related to eyeball movement (duration, latency, amplitude, peak velocity). 19 people, selected by age, took part in the study. Coffee-consuming group subjects were chosen at random for the experimental group and proper subjects were assigned to the control group without taking coffee. The results of the study were divided into two groups: the first related to saccade dynamics and the second related to reaction time (latency) and accuracy. The statistical significance has not been found in the saccadic duration, amplitude and peak velocity. A significant difference in latency was observed in the Coffee-consuming group, though an increase in the number of mistakes (errors of saccade accuracy) was not observed. In reverse, in the Coffee-consuming group was observed a decrease in the errors of saccade accuracy.
Keywords: eyeball movement, saccadic latency, coffee, caffeine, neuropsychopharmacology

Introduction

The influence of caffeine on motor and cognitive fitness has been the object of numerous studies [1-5]. Caffeine accelerates task completion but also increases the number of mistakes [6-7]. Research has shown that caffeine stimulates the nervous system by facilitating and accelerating thought processes, raises and strengthens reception of external stimuli, and also suppresses the sensation of fatigue and increases the capacity for mental and physical work [8].

However, most of these studies have been carried out exclusively on caffeine administered in pure form [9]. Yet caffeine is most frequently consumed in beverages such as coffee or tea where it is present together with other compounds which influence the nervous system function. In tea it exists together with theine, an amino acid which is a neurotransmitter and an antagonist of caffeine. (It lowers blood pressure and serotonin levels which caffeine increases.) On the other hand coffee contains a morphine-like substance whose effects are not yet fully known. However, the important function which opioid receptors play in the brain leads us to suppose that coffee's influence on nervous system function does not depend solely on caffeine.

The study of the effects of a series of chemical compounds present in coffee on cognitive and motor processes therefore appears justified. The psychostymulating effect of caffeine is modulated by other substances.

Materials and Methods

In this study we made use of a head-mounted saccadometer (Ober Consulting Poland) which measures parameters related to eyeball movement (duration, latency, amplitude, peak velocity) in response to stimuli projected by special lasers with which the device is also fitted. The saccadometer measures eye movements by utilizing the method of Direct Infra Red Oculography, which is embedded in the Cyclop ODS sensor [10]. The measurements of the left and right eye rotations have been averaged. The combination of a very sensitive device for the detection of eyeball position change with a precise projection of stimuli made it possible to accurately assess changes of the analysed parameters (sampling with 12 bit resolution at 1 kHz) [11].

The study was divided into two stages: before consumption of coffee (200ml) and 45 minutes after consumption of coffee. The subjects always carried out the same task which consisted in focusing vision on a laser-generated point of light which randomly shifted its position along the horizontal axis in 10 angular

The influence of coffee on an eyeball movement

degree increments. Horizontal eye movement in a 20 angular degree range was analysed (±10º from the central vision fixation position). Each study consisted of 20 trials for calibration and 50 actual study measurements. In total, each person made 140 responses to stimuli of which 100 were analysed.

19 people, selected by age, took part in the study. 8 subjects were chosen at random for the experimental group (called *Coffee*-consuming) and 11 were assigned to the control group without taking coffee (called *Control*). The experiment was carried out at 6 p.m., and the subjects did not ingest any stimulatory substances for at least 3 hours prior to the test. The time for conducting the study was specifically chosen so that the subjects would be tired after an entire day of work (which was confirmed by a questionnaire). The experimental conditions were meant to simulate the most common situation in which coffee is consumed: when a person is tired but has to carry out a task which requires concentration.

Results

The results of the study were divided into two groups: the first related to saccade dynamics (Fig. 1-3) and the second related to reaction time (latency) and accuracy (Fig. 4-7).

Basic saccade dynamics parameters such as duration time, amplitude, peak velocity, latency and error of saccade accuracy were analysed. Graphs 1-5 present a comparison for both groups of the results of these parameter measurements before and after consumption of coffee.

Fig. 2. A comparison for both groups (*Coffee*-consuming and *Control*) of the results (average and standard deviation) of saccadic amplitude in angular degrees

Fig. 3. A comparison for both groups (*Coffee*-consuming and *Control*) of the results (average and standard deviation) of saccadic peak velocity in angular degrees per seconds
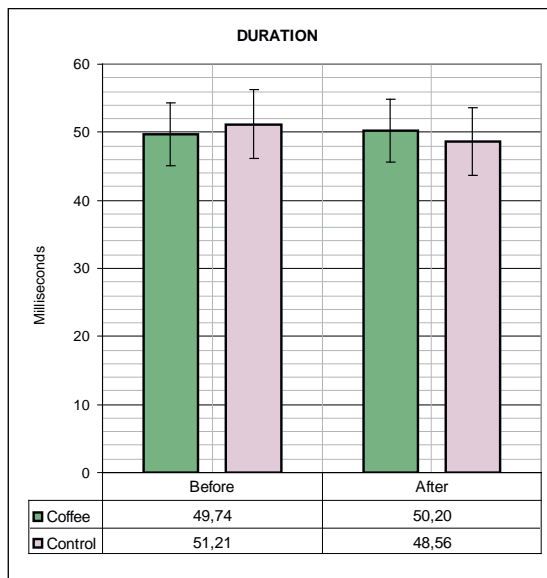

Fig. 1. A comparison for both groups (*Coffee*-consuming and *Control*) of the results (average and standard deviation) of saccadic duration
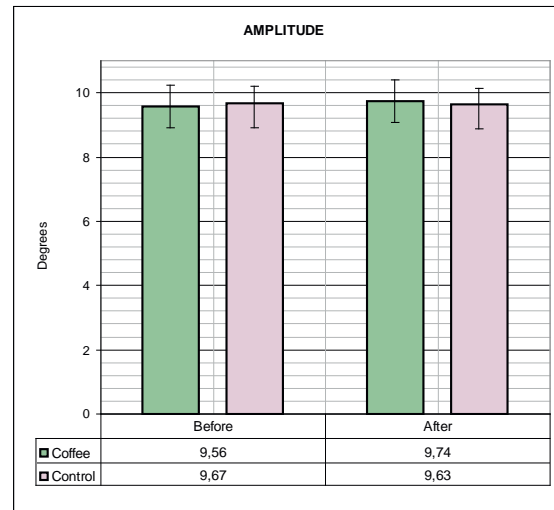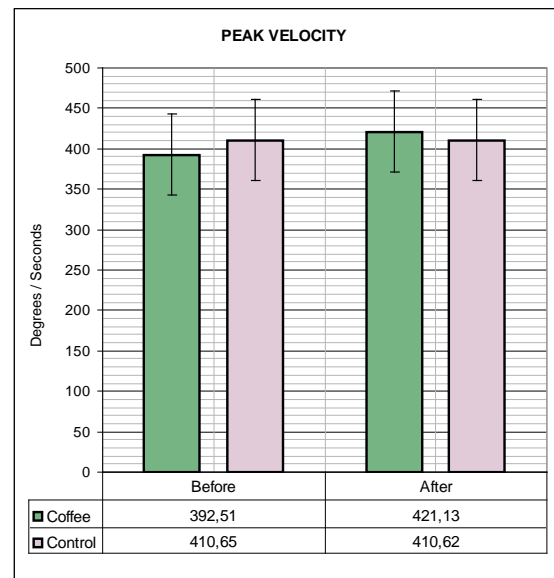
LATENCY

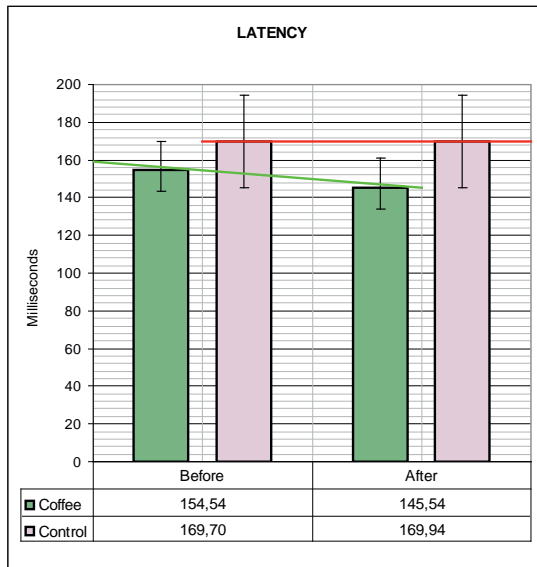|	Before	After
Coffee	154,54	145,54
Control	169,70	169,94

Fig. 4. A comparison for both groups (*Coffee*-consuming and *Control*) of the results (average, standard deviation and trend lines) of saccadic latency in milliseconds

ERROR OF ACCURACY

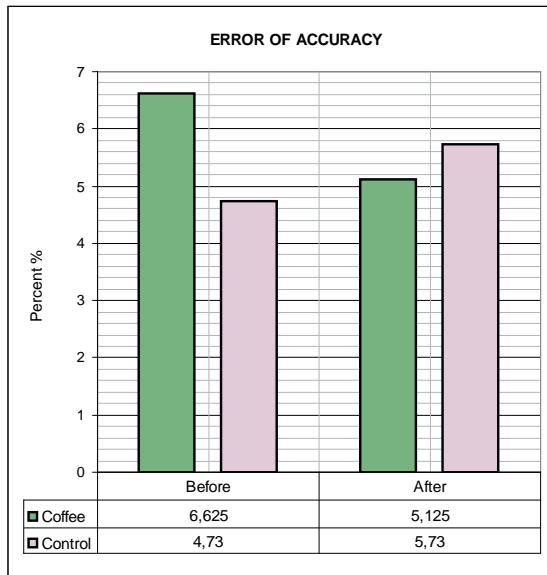|	Before	After
Coffee	6,625	5,125
Control	4,73	5,73

Fig. 5. A comparison for both groups (*Coffee*-consuming and *Control*) of the results (average) of errors of saccade accuracy in percents per seconds

A more detailed analysis was made for movement latency (reaction time to stimulus). The analysis of the movement latency measurements before and after consumption of coffee is shown for both groups in Fig. 6 and 7.

The statistical significance of the observed differences in the *Coffee*-consuming group was verified using the non-parametric Wilcoxon Matched Pairs Test and Sign Test (Fig. 8 and 9).

A significant difference in latency was observed in the *Coffee*-consuming group (Fig. 8), though an increase in the number of mistakes (errors of saccade accuracy) was not observed (Fig. 5). In reverse, in the *Coffee*-consuming group was observed a decrease in the errors of saccade accuracy (Fig. 5).
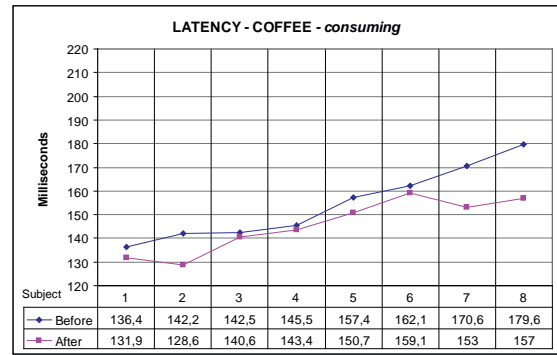

LATENCY - COFFEE - *consuming*

Subject	1	2	3	4	5	6	7	8
Before	136,4	142,2	142,5	145,5	157,4	162,1	170,6	179,6
After	131,9	128,6	140,6	143,4	150,7	159,1	153	157

Fig. 6. Sorted results of saccadic latency in the *Coffee*-consuming group before and after consumption of coffee

LATENCY - CONTROL

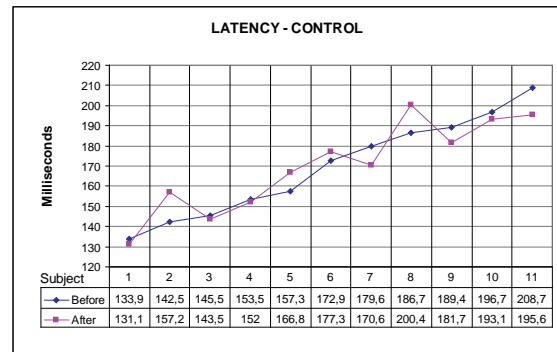| Subject	1	2	3	4	5	6	7	8	9	10	11
Before	133,9	142,5	145,5	153,5	157,3	172,9	179,6	186,7	189,4	196,7	208,7
After	131,1	157,2	143,5	152	166,8	177,3	170,6	200,4	181,7	193,1	195,6

Fig. 7. Sorted results of saccadic latency in the *Control* group before and after consumption of coffee

	Wilcoxon Matched Pairs Test Marked tests are significant at p <,0500			
Pair of Variables	Valid N	T	Z	p-level
Lat Before&After	8	0,00	2,520504	0,011719

	Sign Test (Latency Coffee) Marked tests are significant at p <,0500			
Pair of Variables	No. of Non-ties	Percent v < V	Z	p-level
Lat Before&After	8	0,00	2,474874	0,013328

Fig. 8. The verification for the *Coffee*-consuming group of the statistical significance of the observed differences in latency using the non-parametric Wilcoxon Matched Pairs Test and Sign Test

	Wilcoxon Matched Pairs Test Marked tests are significant at p <,05			
Pair of Variables	Valid N	T	Z	p-level
Lat [ms] & Lat [ms]	11	32,00000	0,088911	0,929153

	Sign Test Marked tests are significant at p <,05000			
Pair of Variables	No. of Non-ties	Percent v < V	Z	p-level
Lat [ms] & Lat [ms]	11	36,36364	0,603023	0,546494

Fig. 9. The verification for the *Control* group of the statistical significance of the observed differences in latency using the non-parametric Wilcoxon Matched Pairs Test and Sign Test
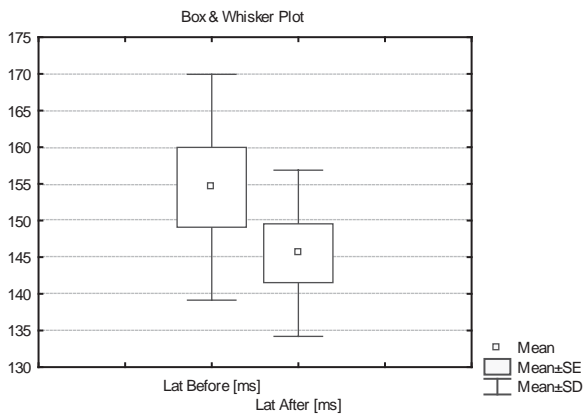
Medical diagnostics

Fig. 10. Box and Whisker Plot of saccadic latency in the *Coffee-consuming* group before and after consumption of coffee

Conclusion

In summary, this study presents an important contribution to the planning of experiments which will reflect the real conditions in which psychostimulating substances are ingested. Successive studies which will be carried out on a larger sample include plans for monitoring the duration and extinction of the observed effect. A series of identical experiments at different time intervals after ingestion of the substance will be carried out for this purpose. The device used in this research is very appropriate for these follow-up studies, because saccade dynamics and latency are not trained with multiple repetition of similar tasks. The saccadometer measures physiological reactions which occur on a millisecond timescale and which are beyond voluntary control. Therefore the use of a saccadometer contributes to increase measurement accuracy and repeatability.

Visual attention, which allows eyesight to be fixed on a chosen object, is connected to cognitive processes. By using appropriately chosen visual stimuli, it is possible to construct more complicated tasks which engage different types of cognitive functions. Thus, more complex experiments can be expected in the future which make use of eye movement measurements in cognitive tasks and study the influence of psychostimulating substances on the performance of such tasks. The use of eyeball movement measurements is a promising tool for the dynamically growing science of neuropsychopharmacology.

References

1. Kalmar J.M. (2005), The influence of caffeine on voluntary muscle activation. Med. Sci. Sports Exerc., Dec; 37(12): 2113-9.
2. Cauli O., Morelli M. (2005), Caffeine and the dopaminergic system. Behav. Pharmacol., Mar; 16(2): 63-77. Review.
3. Specterman M., Bhuiya A., Kuppuswamy A., Strutton P.H., Catley M., Davey N.J. (2005), The effect of an energy drink containing glucose and caffeine on human corticospinal excitability. Physiol. Behav., Jan 17; 83(5): 723-8.
4. Cauli O., Morelli M. (2002), Subchronic caffeine administration sensitizes rats to the motor-activating effects of dopamine D(1) and D(2) receptor agonists. Psychopharmacology (Berl), Jul; 162(3): 246-54. E-pub. 2002, May 14.
5. Pelissier A.L., Gantenbein M., Bruguerolle B. (1999), Caffeine-induced modifications of heart rate, temperature, and motor activity circadian rhythms in rats. Physiol. Behav., Aug 1; 67(1): 81-8.
6. Cerqueira V., de Mendonça A., Minez A., Dias A.R., de Carvalho M. (2006), Does caffeine modify corticomotor excitability? Neurophysiol. Clin., Jul-Aug; 36(4): 219-26. E-pub. 2006, Aug 22.
7. Lorist M.M., Snel J., Kok A. (1994), Influence of caffeine on information processing stages in well rested and fatigued subjects. Psychopharmacology (Berl)., Jan; 113(3-4): 411-21.
8. Bovim G., Naess P., Helle J., Sand T. (1995), Caffeine influence on the motor steadiness battery in neuropsychological tests. J. Clin. Exp. Neuropsychol., May; 17(3): 472-6.
9. Jacobson B.H., Winter-Roberts K., Gemmell H.A. (1991), Influence of caffeine on selected manual manipulation skills. Percept. Mot. Skills, Jun; 72(3 Pt 2): 1175-81.
10. http://www.ober-consulting.com/page/spec_saccadometer/.
11. Ober J.K., Przedpelska-Ober E., Gryncewicz W., Dylak J., Carpenter R.S., Ober J.J. (2003), Hand-held system for ambulatory measurement of saccadic durations of neurological patients. In: Gajda J. (ed.), Modelling and Measurement in Medicine. Warsaw: Komitet Biocybernetyki i Inżynierii Biomedycznej PAN, pp. 187-198.

Medical diagnostics

DIRECTION-DEPENDENT SACCADIC WAVEFORM ASYMMETRY

PIOTR WALECKI[*], WOJCIECH LASOŃ[*], EDWARD J. GORZELAŃCZYK[#]

*Department of Bioinformatics and Telemedicine, Jagiellonian University, Medical College, Kopernika 7,
31-034 Krakow, e-mail: pwalecki@cm-uj.krakow.pl*
*# Department of Human Physiology, Chair of Physiology, Nicolaus Copernicus University, Ludwik Rydygier
Collegium Medicum, Jagiellońska 13-15, 85-067 Bydgoszcz*

Abstract: This paper presents response direction-dependent saccadic velocity waveform asymmetry and direction-dependent difference between means of saccadic parameters of eyeball movement such as duration, latency, amplitude, peak velocity. The research was a preparatory study which aimed to determine repetitive patterns in saccadic velocity waveform (SW). The study also aimed to distinguish different types of SW and eyeball movement direction-dependent waveform asymmetry (SWA).
Keywords: eyeball movement, saccadic waveform, saccade velocity, eyeball movement direction-dependent asymmetry

Streszczenie: W artykule przedstawiono zależną od kierunku ruchu oczu asymetrię w profilu prędkości sakad i zależną od kierunku ruchu oczu różnicę pomiędzy średnimi wybranych parametrów skokowych ruchów gałek ocznych tj.: czasem trwania, latencją, amplitudą, prędkością szczytową. Badanie ma charakter wstępnego studium, którego celem jest opisanie powtarzających się wzorców profilu prędkości sakad. Celem badania jest również wyodrębnienie różnych przebiegów profili prędkości sakad oraz ich asymetrii, która związana jest z kierunkiem ruchu oczu.
Słowa kluczowe: ruch skokowy oczu, profil sakady, prędkość sakady, asymetria zależna od kierunku ruchu oczu

Background

The eyeball is securely fastened in the orbit by the orbital fascia and held in place by the retrobulbar fat. Consequently, its range of movement is small. The eye's movement is limited by the suspensory ligaments which stop the rotating eyeball from stretching the muscles to their limit. The eyeball can move along an arch of about 50° in every direction though the area of fixation of each eye is additionally medially restricted by the nose. Usually, however, eye movements do not cover the entire field of fixation but only a small field of about 30° around the so-called initial position (eyeball position when looking straight ahead). If an observed object would require eye movement of more than 30°, then the entire head also turns in that direction. The exception to this is a movement of the eyes directly down along the central line. In this case, the eyes can rotate even up to 60° without movement of the head.

The eyeball rotates around the X, Y and Z Fick axes. The X axis is the horizontal axis which allows the eyeball to move up and down, and the Z axis is the vertical axis around which the eyes make horizontal left and right movements. Both the X and Z axes lie on the Listing plane which passes through the equator of the eyeball. The Y axis is the front-back axis around which the eyeball can rotate clockwise or counter-clockwise. This type of movement is called torsion. The Y axis is perpendicular to the Listing plane.

During horizontal movement, the point around which the eyeball rotates moves in a semicircular movement on the plane of rotation along the so-called spatial centroid. The centre of rotation of the eyeball is the point of intersection of the spatial centroid with the body's centroid. Because shifting movements can occur during rotation, it is accepted for convenience that an adult person's centre of rotation of the eye lies 13.5 mm behind the cornea and 1.6 mm away from the geometric centre of the eyeball from the nasal side.

For orientation, the most important voluntary eye movement is the one which fixes vision in such a way that the field of vision is projected onto the central part of the retina (macula lutea) where best focus (maximum resolution) and colour perception occur. As a consequence of the fact that only a small area of the retina has optimal focus, vision fixation has to occur in such a way that the entire scene is projected for at least a moment onto the central area of the retina. This is facilitated by quick, intermittent movements of the eyes called saccades. Saccades are used to observe the surroundings. They provide information about the surroundings by directing a chosen scene to the centre of vision (retinal macula). In humans saccades occur almost constantly, every several hundred millisec-

Medical diagnostics

onds, so that on average a person makes about 170000 saccade movements everyday.

The velocity of saccades is beyond voluntary control. Therefore it is not possible to execute a slow saccade. This characteristic of saccade movement which is conditioned by the way the neuronal system is constructed is used in medical diagnostics. Saccade velocity dynamics change in various diseases or when certain structures of the nervous system are damaged.

Eye measurements using precise devices which check eye position every millisecond make it possible to analyse the dynamics of these movements. Thanks to these precise measurements, it is possible to analyse individual elements of the movement. A further development of this method is the analysis of the changes of individual parameters in time (saccadic waveform – SW).

Material and methods

This study was conducted over a period of 3 months on one healthy subject aged 30. The research was a preparatory and a distinctive study which aimed to determine repetitive patterns in saccadic velocity waveform (SW). The study also aimed to distinguish different types of SW and eye movement direction-dependent waveform asymmetry (SWA).

A head-mounted saccadometer (Ober Consulting Poland) was used in the study (sampling with 12 bit resolution at 1 kHz). For the eye movement measurement the saccadometer system utilizes the Cyclops-ODS (Oculus Dexter Sinister) technology optimized for easy set-up and minimal intrusiveness. Saccadometer is the medical research instrument, allowing evoking the strictly controlled saccadic responses, measuring it and automatically analysing their latencies, durations, amplitudes, and peak velocities [14].

The experimental procedure consisted in presenting stimuli along a horizontal line. The subject was asked to focus his or her vision on the appearing stimulus. The response to stimuli projected by special lasers with which the device is also fitted.

Only the horizontal movement of the eyes in both directions was measured. The stimuli appeared either randomly or according to a planned sequence in one of several fixed positions. The focusing of vision on the stimulus induced a movement of the eye of about 10°, and 20°. Two types of experimental tasks were assigned in the study (Latency and Dynamic). The Latency task consisted in focusing vision on a laser-generated point of light which randomly shifted its position along the horizontal axis in 10 angular degree increments. The Latency task is a simple 10° where duration and direction is randomised. In the Dynamic task the laser-generated point of light shifted its position along the horizontal axis in 20 angular degrees. In the Dynamic task the saccadic tracking of a target jumping between the two peripheral targets at regular intervals. The task is intended for measuring saccadic velocity.

Each task consisted of 20 calibration trials for and 50 experimental trials (movements of the eye caused by the appearance of the stimulus).

In total, the subject made 21 Latency tasks (1470 responses to stimuli of which 1050 were analysed) and 16 Dynamic tasks (1120 responses to stimuli of which 800 were analysed).

Results

Saccadic velocity waveform asymmetry (SWA) was observed in the study. Several types of SWA were distinguished in different tasks (Fig. 3 and 5). A strong correlation was established between the all trials (eye movements) in the left direction responses and the all trials in the right direction responses which were observed in the one task (Fig. 3 and 5).

As a result of several dozen trial repetitions conducted on the same subject at different times, it was determined that SWA changes dynamically and sometimes does not occur at all. Based on the study it was determined that SWA depend on such factors as a day time, a type of activity executing before the examination or a current psychophysical condition.

Table 1 and Table 3 present a comparison for both response directions (left and right) of the results of saccades duration (mean and standard deviation – SD), saccades latency (mean and SD), saccades amplitude (mean and SD), and saccades peak velocity (mean and SD).

Tab. 1. Latency tasks – a comparison of values of left and right response

	Left Mean	Left SD	Right Mean	Right SD
Duration [ms]	54,50	4,00	55,24	3,29
Latency [ms]	174,78	17,34	199,17	21,94
Amplitude [deg]	10,19	1,21	10,20	1,18
Peak Velocity [deg/s]	376,46	59,59	363,44	54,72

Tab. 2. Latency tasks – a statistical verification of direction-dependent difference between means of saccadic parameters

Wilcoxon Matched Pairs Test
Latency Task – amplitude 10º
Marked tests are significant at p < ,05

	Valid	T	Z	p-level
Duration [ms] Left vs Right	21	98,5	0,242662	0,808267
Latency [ms] Left vs Right	21	1,0	3,979751	0,000069
Amplitude [deg] Left vs Right	21	114,5	0,034758	0,972273
Peak Velocity [deg/s] Left vs Right	21	91,0	0,851562	0,394458

Table 2 and Table 4 present a statistical verification of direction-dependent difference between means of saccadic parameters. In the Latency tasks only latency is a significant at $p < 0,05$. In the Dynamic task latency and amplitude is a significant at $p < 0,05$.

Box & Whisker Plot
Duration [ms]

A.

□ Mean
▫ Mean±SE
⊥ Mean±SD

Box & Whisker Plot
Amplitude [deg]

C.

□ Mean
▫ Mean±SE
⊥ Mean±SD

Box & Whisker Plot
Latency [ms]

B.

□ Mean
▫ Mean±SE
⊥ Mean±SD

Box & Whisker Plot
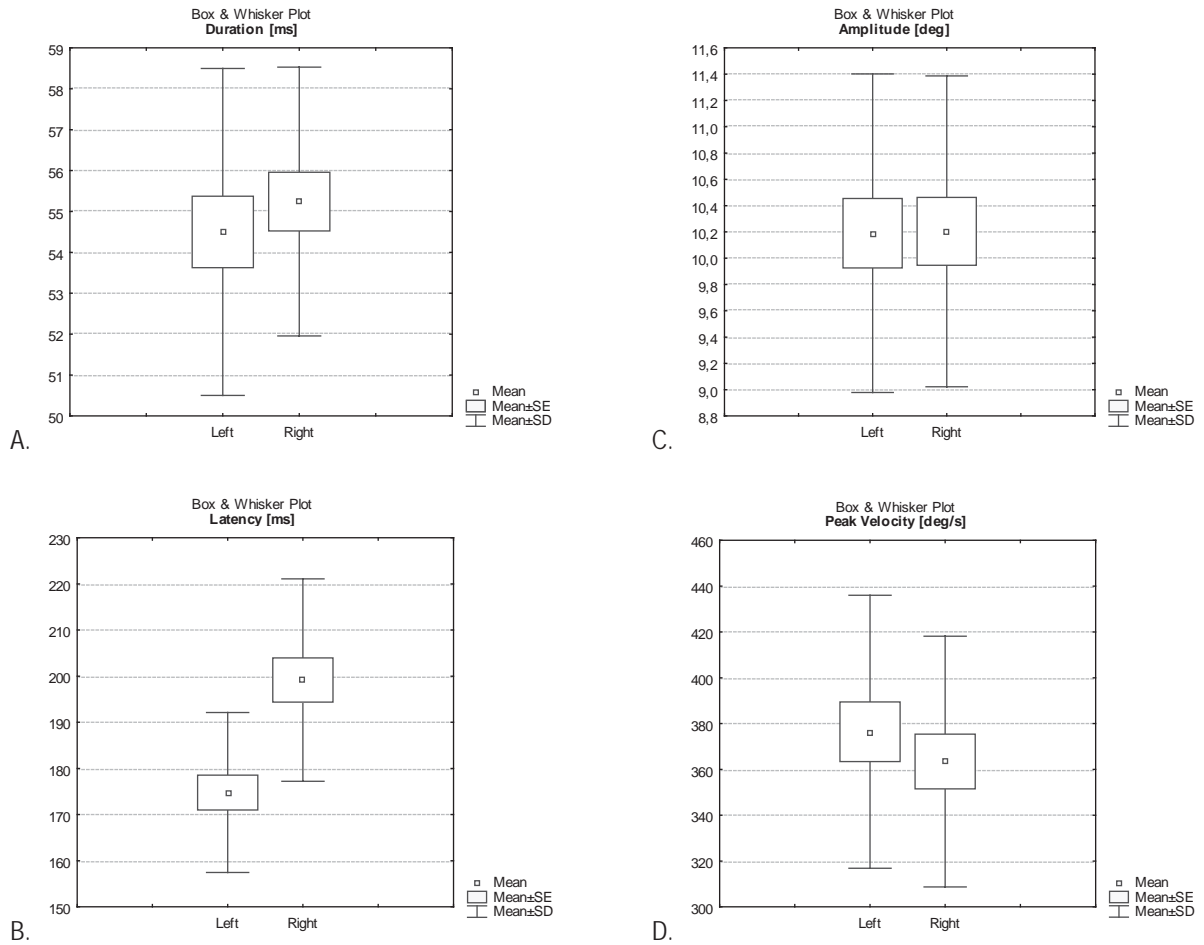**Peak Velocity [deg/s]**

D.

□ Mean
▫ Mean±SE
⊥ Mean±SD

Fig. 1. Latency tasks – a comparison for left and right response of all tasks means, standard deviations and standard errors (A – Duration, B – Latency, C – Amplitude, D – Peak velocity)

Tab. 3. Dynamic tasks – a comparison of values of left and right response

	Left Mean	Left SD	Right Mean	Right SD
Duration [ms]	91,79	18,36	87,69	7,99
Latency [ms]	154,33	11,23	168,13	14,04
Amplitude [deg]	19,78	1,39	21,63	1,67
Peak Velocity [deg/s]	452,05	78,93	484,59	82,76

Tab. 4. Dynamic tasks – a statistical verification of direction-dependent difference between means of saccadic parameters

Wilcoxon Matched Pairs Test
Dynamic Task – amplitude 20º
Marked tests are significant at $p < {,}05$

	Valid	T	Z	p-level
Duration [ms] Left vs Right	16	60,0	0,413670	0,679116
Latency [ms] Left vs Right	16	0,0	3,516196	0,000438
Amplitude [deg] Left vs Right	16	10,0	2,999109	0,002708
Peak Velocity [deg/s] Left vs Right	16	35,0	1,706389	0,087937

Medical diagnostics

A.

DURATION

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Left response	53	42	56	56	52	53	52	58	56	57	54	54	58	51	56	57	56	51	54	59	61
Right response	50	55	56	59	61	55	53	59	60	56	52	53	59	53	54	55	56	49	55	59	55

B.

LATENCY

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Left response	186	172	215	159	186	194	191	172	155	180	188	172	179	171	182	182	182	159	152	149	143
Right response	219	202	236	213	216	223	205	181	185	240	215	182	208	194	188	192	177	177	168	207	156

C.

AMPLITUDE

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Left response	11	7,6	12	8,7	11	13	10	11	11	9,1	11	11	10	9,7	9,6	9,9	9	9,9	9,6	11	11
Right response	8,8	9,3	8,8	8,2	10	10	9,2	9,6	9,5	13	9,1	9,2	12	11	12	11	11	10	11	11	11

D.

PEAK VELOCITY

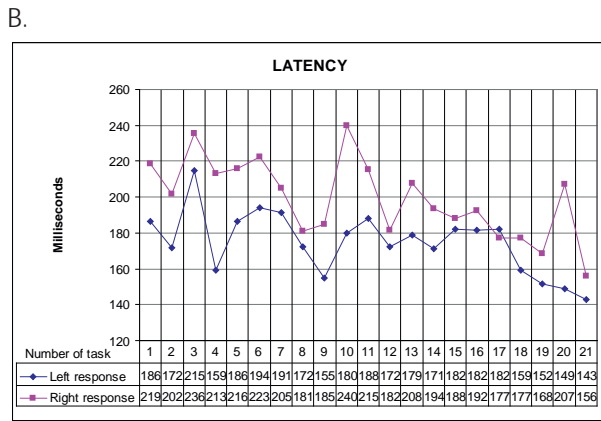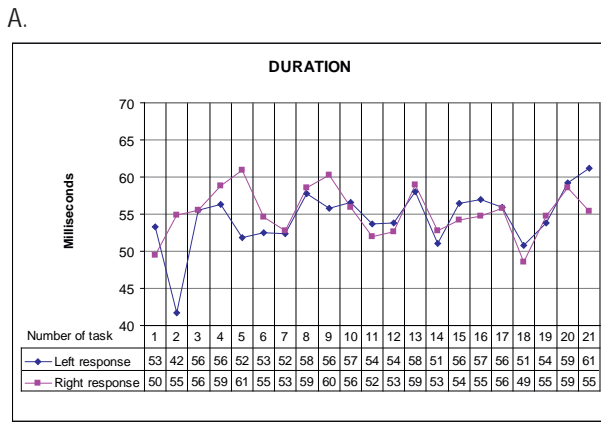| Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Left response	373	420	423	327	369	533	372	366	348	322	454	452	307	395	296	347	383	440	336	324	320
Right response	333	324	276	264	320	417	320	288	362	427	441	443	350	373	440	375	422	343	394	365	358

Fig. 2. Latency tasks – a comparison of task means of left and right response for each individual task separately (A – Duration, B – Latency, C – Amplitude, D – Peak velocity)

Fig. 3. Latency tasks – a comparison for left and right response of velocity waveform

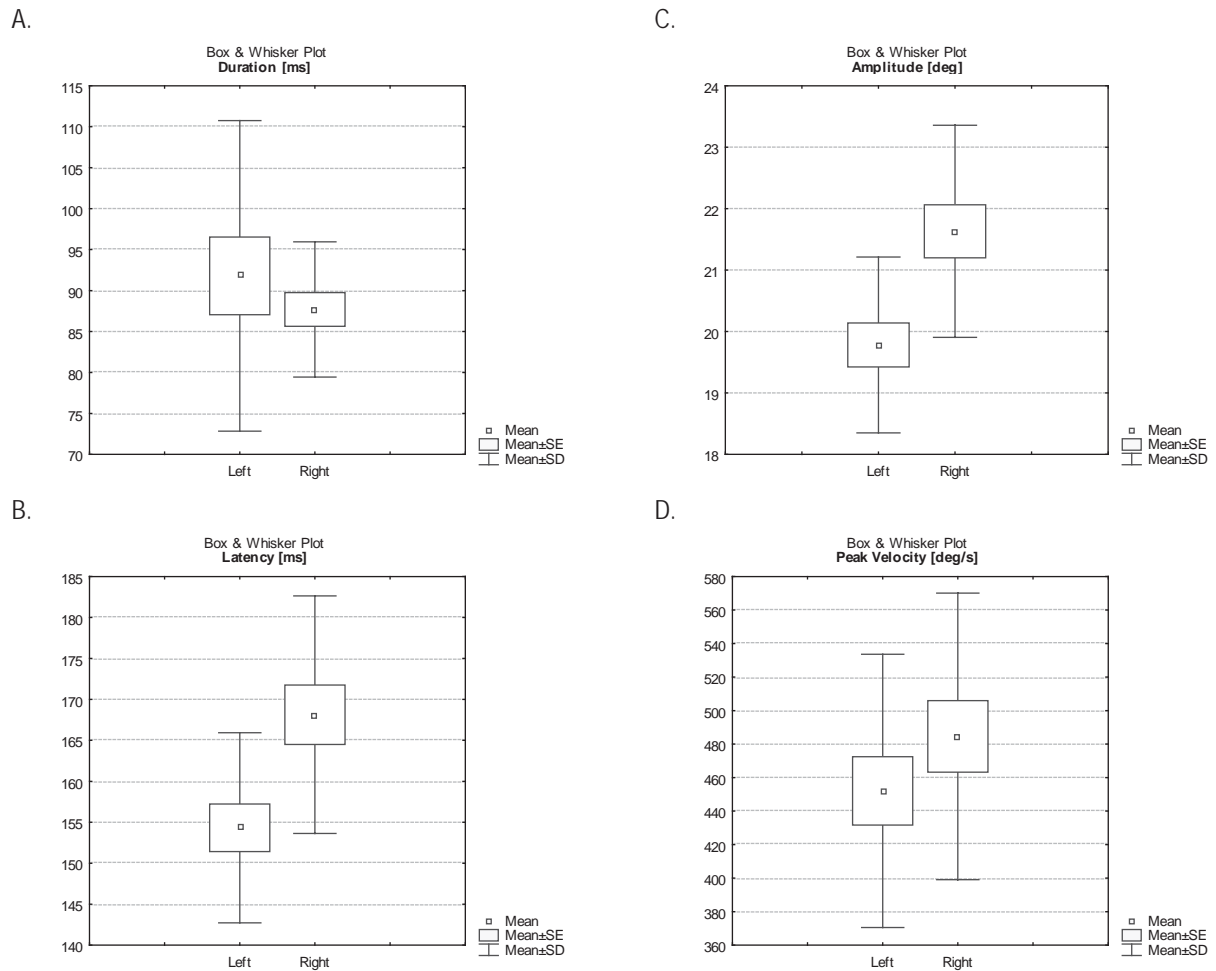Medical diagnostics

A.

C.

B.

D.

Fig. 4. Dynamic tasks – a comparison for left and right response of all tasks means, standard deviations and standard errors (A – Duration, B – Latency, C – Amplitude, D – Peak velocity)

As a result of several dozen trial repetitions conducted on the same subject at different times, it was determined that SWA changes dynamically and sometimes does not occur at all (Fig. 3 and Fig. 5).

The search for a factor which conditions the appearance of asymmetry was unsuccessful.

It was noticed, however, that certain factors (such as sleep, rest or physical activity) can be conducive to the appearance of SWA.

Conversely, other actions or substances (for example coffee) can have the opposite effect of suppressing SWA.

Conclusion

The explanation of SWA appearance requires the elucidation of the neuronal control mechanisms of the eyeballs. Most probably SWA is related to the phenomenon of domination of one of the hemispheres of the brain. Classical theories supposed that this domination is invariant and that it is consolidated during personality development. However many new studies as well as the studies presented here indicate that the concept of domination of one of the cerebral hemispheres is not accu-

rate. Some faculties such as speech are predominantly located in one cerebral hemisphere. Thus, we can speak of a functional asymmetry of the cerebral hemispheres. Hemispherical asymmetry has an effect on the domination of a specific hemisphere, but it does not determine it. The corresponding cerebral systems are activated depending on the task that is being performed.

However the domination of a certain hemisphere in relation to the task being performed at the moment is not a constant characteristic such as, for example, right – or left-handedness. Hemisphere domination changes depending on the task being performed. Of course, often in a person one hemisphere dominates most of the time just as one eye usually dominates. However, at different times this situation can temporarily change.
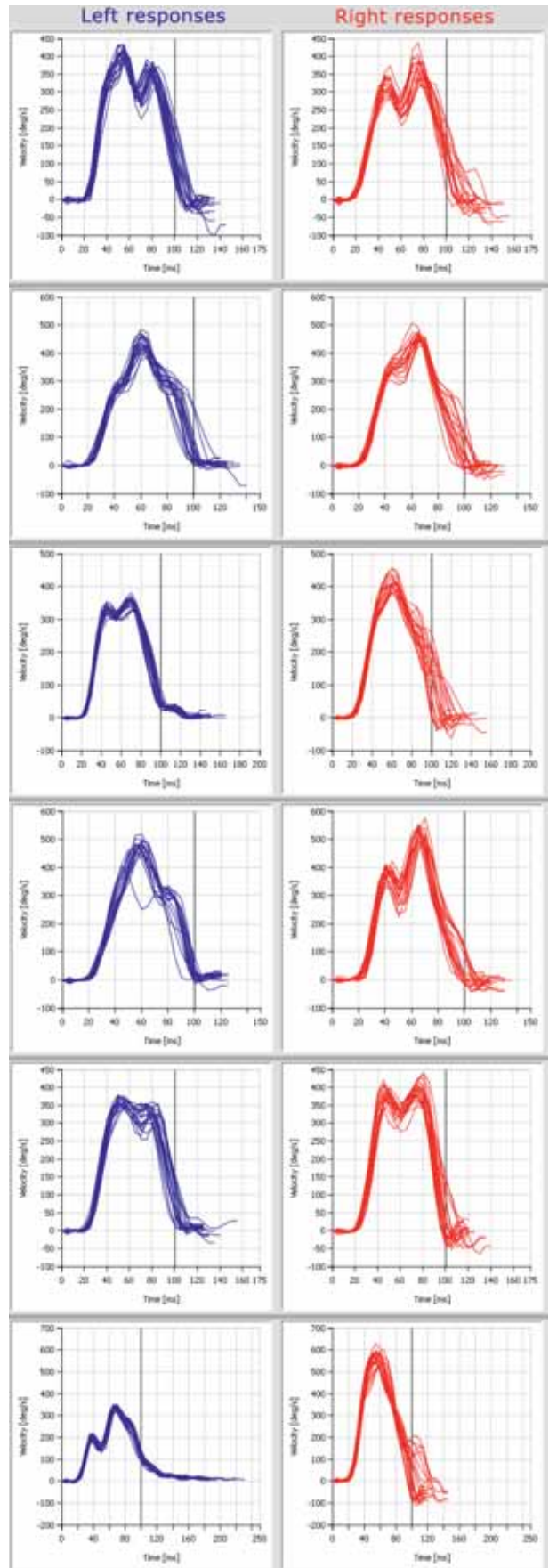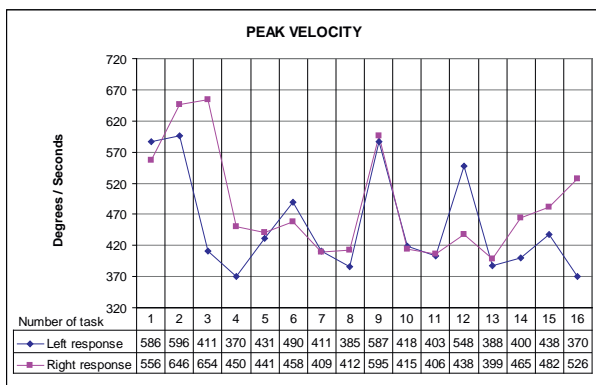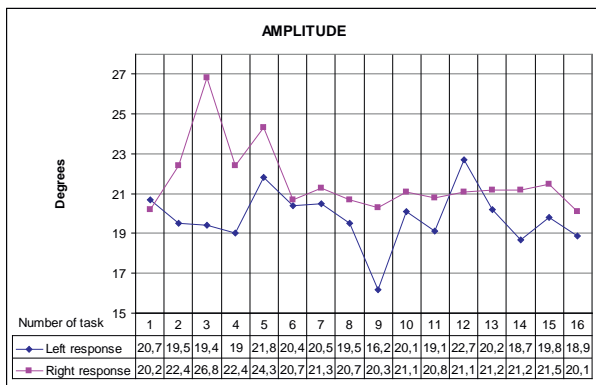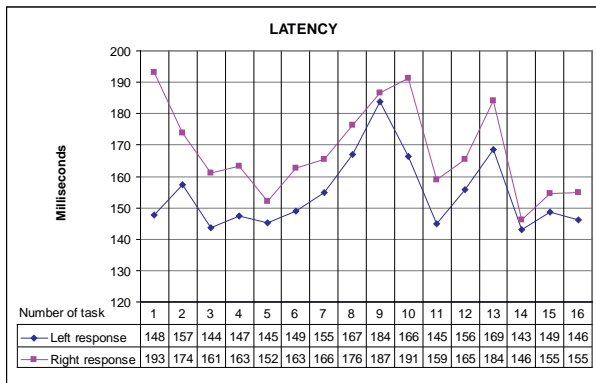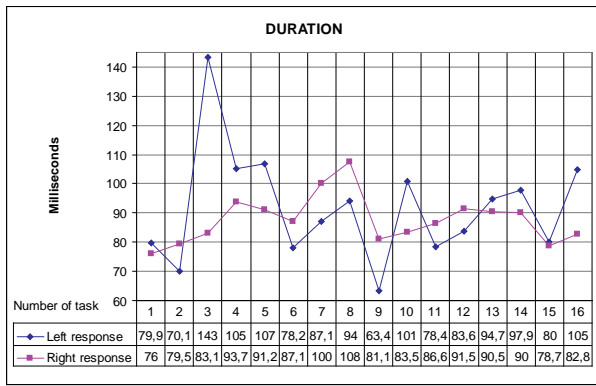
Medical diagnostics

DURATION

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Left response	79,9	70,1	143	105	107	78,2	87,1	94	63,4	101	78,4	83,6	94,7	97,9	80	105
Right response	76	79,5	83,1	93,7	91,2	87,1	100	108	81,1	83,5	86,6	91,5	90,5	90	78,7	82,8

LATENCY

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Left response	148	157	144	147	145	149	155	167	184	166	145	156	169	143	149	146
Right response	193	174	161	163	152	163	166	176	187	191	159	165	184	146	155	155

AMPLITUDE

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Left response	20,7	19,5	19,4	19	21,8	20,4	20,5	19,5	16,2	20,1	19,1	22,7	20,2	18,7	19,8	18,9
Right response	20,2	22,4	26,8	22,4	24,3	20,7	21,3	20,7	20,3	21,1	20,8	21,1	21,2	21,2	21,5	20,1

PEAK VELOCITY

Number of task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Left response	586	596	411	370	431	490	411	385	587	418	403	548	388	400	438	370
Right response	556	646	654	450	441	458	409	412	595	415	406	438	399	465	482	526

Fig. 5. Dynamic tasks – a comparison of means of left and right response for each individual task separately (A – Duration, B – Latency, C – Amplitude, D – Peak velocity)

Fig. 6. Dynamic tasks – a comparison for left and right response of velocity waveform

Left responses   Right responses

Medical diagnostics

References

1. Schmidt D., Dell'Osso L.F., Abel L.A., Daroff R.B. (1980), Myasthenia gravis: dynamic changes in saccadic waveform, gain, and velocity, Exp. Neurol., May; 68(2): 365-77.
2. Schmidt D., Dell'Osso L.F., Abel L.A., Daroff R.B. (1980), Myasthenia gravis: saccadic eye movement waveforms, Exp. Neurol., May; 68(2): 346-64.
3. Chioran G.M., Yee R.D. (1991), Analysis of electro-oculographic artifact during vertical saccadic eye movements, Graefe's Arch. Clin. Exp. Ophthalmol., 229(3): 237-41.
4. Thornton W.E., Uri J.J. (1991), Oculomotor function during space flight and susceptibility to space motion sickness, Acta Astronaut., 23: 53-61.
5. Brooks-Eidelberg B.A., Adler G. (1992), A frontal cortical potential associated with saccades in humans, Exp. Brain Res., 89(2): 441-6.
6. Abadi R.V., Scallan C.J. (2000), Waveform characteristics of manifest latent nystagmus, Invest. Ophthalmol. Vis. Sci., Nov; 41(12): 3805-17.
7. Optican L.M., Robinson D.A. (1980), Cerebellar-dependent adaptive control of primate saccadic system, J. Neurophysiol., 44: 1058-76.
8. Liao K., Kumar A.N., Han Y.H., Grammer V.A., Gedeon B.T., Leigh R.J. (2006), Comparison of Velocity Waveforms of Eye and Head Saccades, Annals of the New York Academy of Sciences, Vol. 1039, pp. 477-479.
9. Wutz R.H., Goldberg M.E. (eds.) (1989), The Neurobiology of Saccadic Eye Movements, Elsevier, Amsterdam.
10. Harwood M.R., Mezey L.E., Harris C.M. (1999), The spectral main sequence of human saccades, J. Neuroscience, 15: 9098-106.
11. Inchingolo P., Spanio M., Bianchi M. (1987), The characteristic peak velocity – mean velocity of saccadic eye movements in man. In: Eye Movements: From Physiology to Cognition, J.K. Regan, A. Levy-Schoen (eds.), Elsevier, Amsterdam, pp. 17-26.
12. Van Opstal A.J., Van Gisbergen J.A.M. (1987), Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades, Vision Res., 27: 731-45.
13. Fioravanti F., Inchingolo P., Pensiero S., Spanio M. (1995), Saccadic eye movement conjugation in children, Vision Research, 35(23-24): 3217-28.
14. Ober J.K., Przedpelska-Ober E., Gryncewicz W., Dylak J., Carpenter R.S., Ober J.J. (2003), Hand-held system for ambulatory measurement of saccadic durations of neurological patients. In: Modelling and Measurement in Medicine, J. Gajda (ed.), Komitet Biocybernetyki i Inżynierii Biomedycznej PAN, Warsaw, pp. 187-198.

Medical diagnostics

ADVANCED MEASUREMENT OF SACCADIC EYE MOVEMENTS – STUDY OF INVARIABLE AND SIGNIFICANT PHYSIOLOGICAL FEATURES

PIOTR WALECKI, WOJCIECH LASOŃ, IRENA ROTERMAN-KONIECZNA

Department of Bioinformatics and Telemedicine, Jagiellonian University, Medical College, Kopernika 7, 31-034 Krakow

Abstract: Precise measurements of rapid eye movements (saccades) are becoming more widespread. Because is important to answer the question whether how is a measurement error. This paper presents the measurement of about 3000 saccades in one person. It was performed in 14 times the same set of tests. For each of the 14 studies calculated mean and standard deviations of each parameter of the dynamics of eyeball movement, i.e. duration, peak velocity, average velocity, slope, sharpness, Q, and amplitude. These parameters were calculated depending on the amplitude, the type of the task and direction of eyeballs movement (left or right). The results show that the saccadic parameters are a good indicator to assess psychophysical functions.

Keywords: eye movements, assessment, occulometry

Introduction

The visual system is human's special, favorite, sensory system, because most of the information about the outside world reaches the brain just through it. Human functioning appears to be largely determined by this system. Highly developed – it has become a fundamental element of the evolutionary advantage of Homo sapiens. The sense of sight is one of the most important Homo sapiens brain functions. This system is involved both in functioning of the basic cognitive mechanisms such as attention or memory, as well as in functioning of those most complex, as consciousness and self-consciousness.

Examination of the visual system always involves analysis of the function of the oculomotor system, because it is an integral part of it, and significantly affects its characteristics. The motor system of the eyes, which consists of muscles, is closely related with data transmitted by the sensory receptors of the visual system.

The eye movement abnormalities were observed in neurological disorders associated with damage done to the central nervous system (CNS) or cranial nerves, as well as in diseases such as schizophrenia, autism and Parkinson's disease and disorders such as dyslexia or ADHD. Measurement of eye movement may provide important information on the functions and structure of the CNS. On the example, we can analyze the mechanism of sensory-motor integration.

Although we've created more precise measuring devices such as oculometers and saccadometers which are recording the eye movement with a sample rate of 1000Hz, still these measurements are complicated and do not give unambiguous results.

The aim of this study was to establish whether the dynamic properties of the eyeball movement are relatively stable characteristics that may serve as a diagnostic evaluation. The study focused on the analysis of fast, rapid eye movements (saccades), whose parameters such as duration and velocity are beyond the control of will.

In previous studies a high interindividual and intraindividual variability of saccadic velocity, duration, or latency was observed. The values of the saccade parameters change with the age and other reasons. We can also see that the values of the saccade parameters are varying in one person examined in the same period of time The saccades with different amplitudes have various peak velocities and durations and may vary depending on the direction of eye movements. Therefore, to obtain reliable data we have to measure a large number of saccades in one person.

Materials and methods

The tested person was a 42 years old man, with impaired attention span and working memory disorders. The subject volunteered for the study. The patient was examined using the long version of the self-report form Conners' Adult ADHD Rating Scales (CAARS:S-L). CAARS results indicate that the person exhibits symptoms of ADHD in adults.

In this paper a test of the stability and dynamics of saccadic eye movement parameters such as latency, duration, peak velocity and amplitude is presented. Tests of eyeball movement on one person using oculometer device (Jazz Ober) was carried out 14 times to assess the stability of the parameters. Single test lasted for about 5 minutes. In each particular test, two tasks were performed: Saccadic Eye Movement Task (SEM) and Smooth Pursuit Eye Movement Task (SPEM). Based on the results of the study the standard error and standard uncertainty of measurement were calculated. An additional objective was to compare the dynamic parameters occurring during the test designed to generate saccadic eye movements (SEM) and saccades which occurred at the time when the trigger stimulus was in continuously smooth moving (SPEM). Saccades in the SPEM task were secondary (corrective, intrusive) against the smooth pursuit eye movements.

Figure 1 shows the results changes in the position of the eyes along the horizontal (eyex) and vertical line (eyey) of a single SPEM task, and plots the angular velocity of the eye movement (Vel_eyex; Vel_eyey). Also shown in Figure 1: changes of head position (Head_X, Head_Y) and velocity, acceleration of horizontal and vertical head rotation (Head_Vel_X; Head_Vel_Y, accx; accy). Velocity and acceleration signals of the head rotation allow the detection of head movements of the tested person and the automatic calibration in degrees of eye movement in the horizontal axis. The SEM task (see Figure 2) consists of a visual fixation on a point-light stimuli displayed on the computer screen that rapidly changed their location along the horizontal line. The fixation stimuli were changing

position of every 5, 10, 15 and 20 degrees. The task included two parts, first served generation of variable amplitude saccades, and the second contained the sequence of stimuli with a constant amplitude of 20 degrees.

The research tool used to measure eye and head movement was a JAZZ oculometer (producer Ober Consulting Poland). The system was optimized to obtain the simplest and shortest preparation for work (the setup time) and minimal invasiveness, what is crucial for monitoring behavior in non-laboratory environment. The inner canthi of left and right eyes are lighted up with low intensity infrared light (0.6 mW/cm2). The difference between the amounts of infrared light reflected back from the eye surface provides information about eye position changes.

Measurement of eye movement in the JAZZ system shall be made by Cyclops-ODS technology (Oculus Sinister Dexter) based on InfraRed Oculography (IRO). The majority of measuring devices based on the technique of IRO require the deployment of emitters/detectors in the close vicinity around the eye. A set of sensors in Cyclops-ODS optoelectronic technology is located between the eyes. Thanks to this positioning a minimum reduction of the field of vision was obtained, thus reducing the possibility of the measurement system's interference with the visual exploration of the subject's environment.

The measurement of eye movements was performed with high temporal (1 kHz) and spatial (0.1 degree) resolution, thus allowing a precise detection of saccades. An important part of the analysis of eye movement was the separation of three types of signal changes in eye movement, namely saccades,
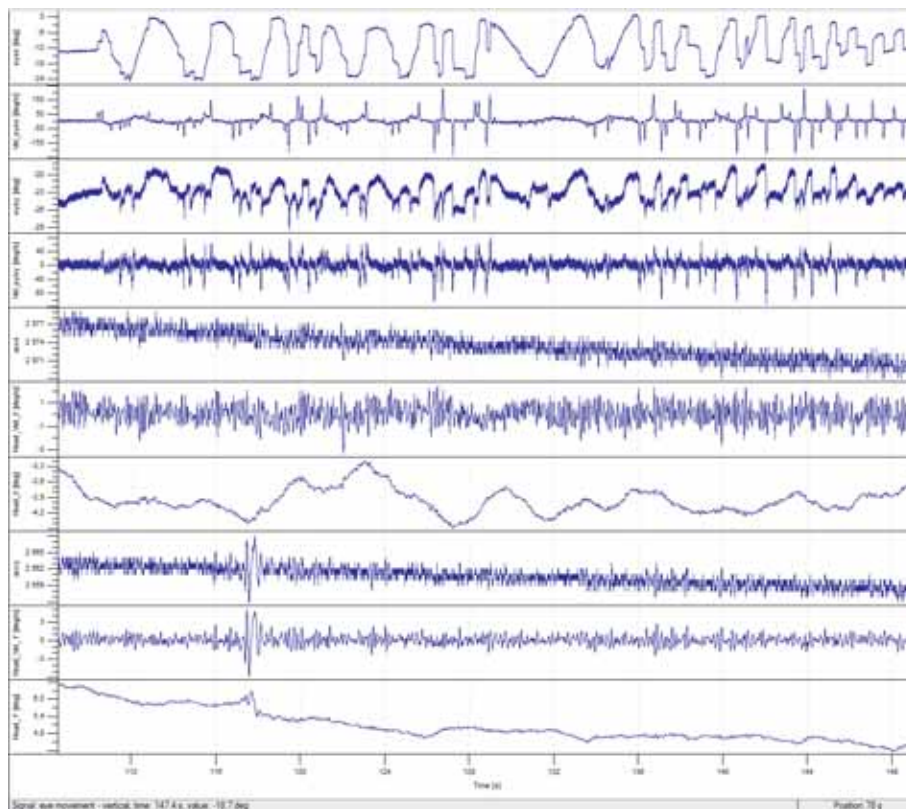

Fig. 1. Graphs of the Smooth Pursuit Eye Movement Task (SPEM). Changes of the horizontal eyes position (eyex) and vertical eyes position (eyey), the angular velocity eyeball movement (Vel_eyex; Vel_eyey). Changes of horizontal head position (Head_X) and vertical head position (Head_Y) and velocity, acceleration of horizontal (Head_Vel_X, accx) and vertical head rotation (Head_Vel_Y, accy)
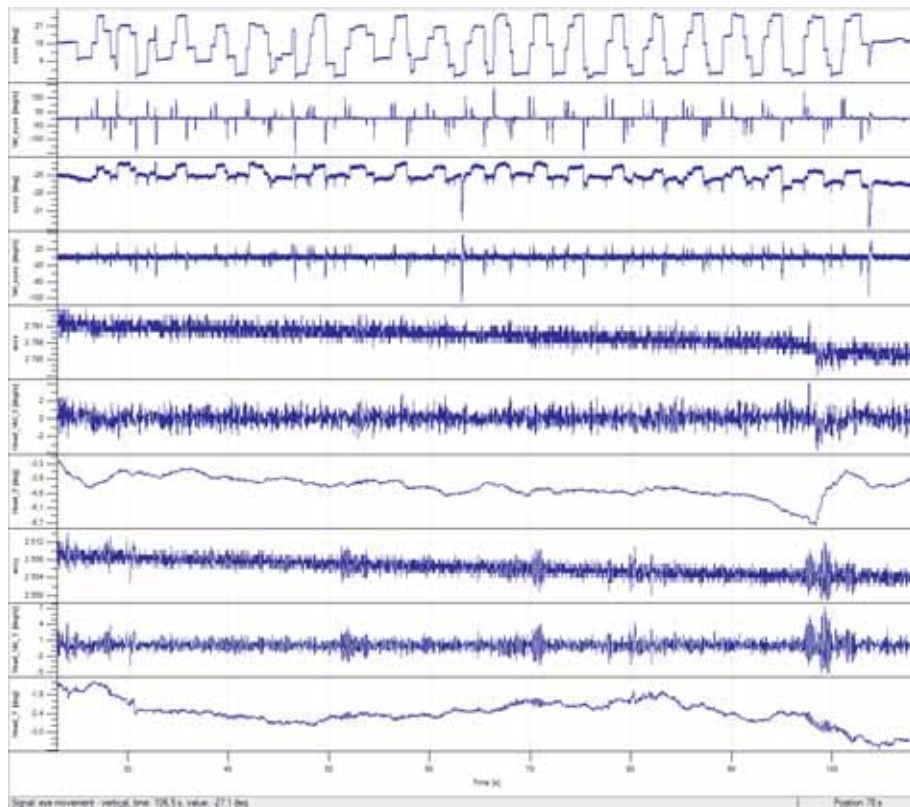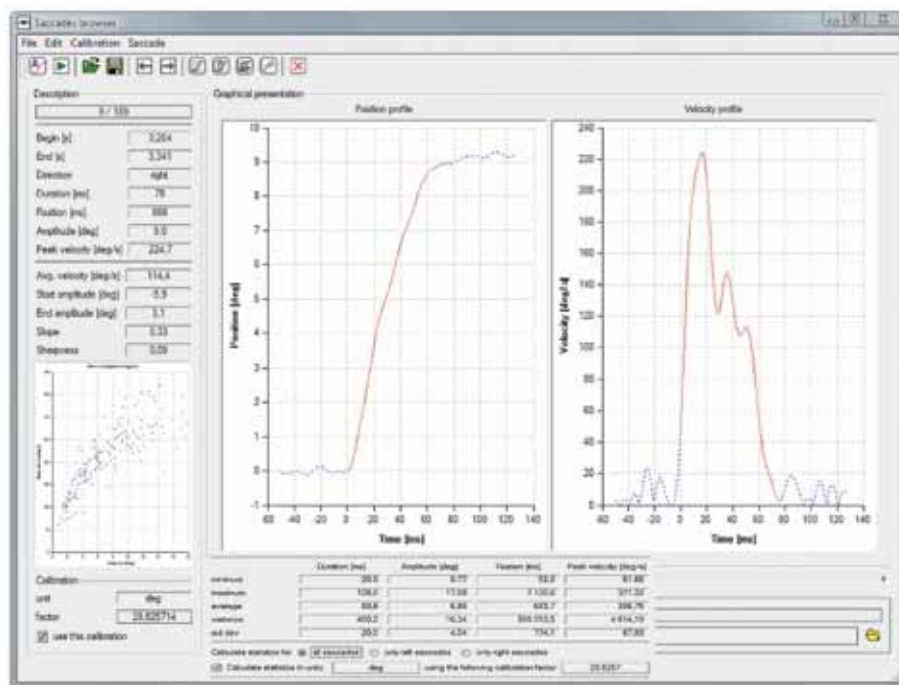
Fig. 2. Graphs of the Saccadic Eye Movement Task (SEM). Changes of the horizontal eyes position (eyex) and vertical eyes position (eyey), the angular velocity eyeball movement (Vel_eyex; Vel_eyey). Changes of horizontal head position (Head_X) and vertical head position (Head_Y) and velocity, acceleration of horizontal (Head_Vel_X, accx) and vertical head rotation (Head_Vel_Y, accy)

Fig. 3. The main window of the Saccades Browser application. At the right side the description panel is placed with information about currently viewed saccade. In the graphical presentation panel there are two saccade graphical profiles (position profile, and velocity profile)

continuous "slow" eye movements and artifacts. Acquisition of measurement data was performed using JazzRecorder applications, which enabled the reception, recording and visualization of data from the measuring system. Analysis and processing of data conveying the dynamics of change in the position of the eyes was performed using JazzManager applications. Detection and browsing of saccades (fast eye position changes) in eye movement signals was performed using Saccades Browser application (see Figure 3).

Results

The study was designed to estimate the measurement error, i.e. the difference between the measurement result and the true value measured quantity. To estimate the measurement error is selected the measurement uncertainty, a parameter characterizing the dispersion of results (the dispersion of the values attributed to a measured quantity). Were statistically analyzed a series of 14 measurements to evaluation of type A uncertainty. As a best estimate of measurement uncertainty was taken as an international ISO standard uncertainty $u(\bar{x})$ calculated from the formula:

a) $\quad u(\bar{x}) = \sqrt{\dfrac{1}{n(n-1)}\sum_{i-1}^{n}(x_i - \bar{x})^2}$

After calculating the uncertainty of measurement determined the standard error (x) for all parameters analyzed according to the formula:

b) $\quad x = \bar{x} \pm u(\bar{x})$

Tab. 1. Standard errors in SEM task

standard error = mean ± standard uncertainity ± standard uncertainity %	n	Duration [ms]	Peak Velocity [deg/s]	Average Velocity [deg/s]	Slope	Sharpness	Q	Average Amplitude [deg]
Amplitude 0.4º – 2.9º	497	29,68 ±0,47 ±1,60%	97,36 ±3,41 ±3,51%	57,40 ±2,12 ±3,69%	32,73 ±2,02 ±6,17%	17,00 ±0,95 ±5,60%	1,70 ±0,016 ±0,96%	1,74 ±0,04 ±2,28%
Amplitude 3º – 5.9º	510	44,53 ±0,62 ±1,39%	174,20 ±3,70 ±2,12%	101,88 ±1,93 ±1,90%	46,75 ±3,21 ±6,87%	20,69 ±1,44 ±6,95%	1,71 ±0,016 ±0,92%	4,50 ±0,04 ±0,95%
Amplitude 6º – 9.9º	539	58,85 ±0,79 ±1,35%	222,04 ±3,68 ±1,66%	133,73 ±1,92 ±1,44%	50,35 ±4,73 ±9,40%	20,27 ±1,72 ±8,47%	1,67 ±0,019 ±1,12%	7,75 ±0,03 ±0,44%
Amplitude 10º – 14.9º	257	77,64 ±1,27 ±1,63%	258,75 ±6,04 ±2,33%	157,95 ±2,53 ±1,60%	46,78 ±4,30 ±9,20%	17,78 ±1,43 ±8,02%	1,65 ±0,019 ±1,14%	12,01 ±0,12 ±0,99%
Amplitude 15º – 20.9º	97	99,60 ±1,84 ±1,84%	286,21 ±7,05 ±2,46%	172,32 ±3,01 ±1,74%	41,75 ±3,44 ±8,25%	15,34 ±1,48 ±9,66%	1,66 ±0,018 ±1,08%	16,89 ±0,22 ±1,30%

Tab. 2. Standard errors in SPEM task

standard error = mean ± standard uncertainity ± standard uncertainity %	n	Duration [ms]	Peak Velocity [deg/s]	Average Velocity [deg/s]	Slope	Sharpness	Q	Average Amplitude [deg]
Amplitude 0.4º – 2.9º	323	31,86 ±0,67 ±2,10	98,06 ±3,38 ±3,45	58,41 ±2,16 ±3,70	31,27 ±2,46 ±7,88	15,24 ±1,03 ±6,76	1,69 ±0,01 ±0,82	1,87 ±0,05 ±2,89
Amplitude 3º – 5.9º	668	49,30 ±0,88 ±1,79	203,79 ±5,58 ±2,74	116,81 ±2,90 ±2,49	47,13 ±3,60 ±7,64	20,81 ±1,64 ±7,88	1,76 ±0,04 ±2,19	5,79 ±0,11 ±1,84
Amplitude 6º – 9.9º	143	73,39 ±1,84 ±2,51	281,32 ±7,84 ±2,79	163,39 ±3,41 ±2,09	49,08 ±3,96 ±8,07	18,76 ±1,61 ±8,60	1,74 ±0,04 ±2,26	11,81 ±0,14 ±1,22
Amplitude 10º – 14.9º	17	99,98 ±5,70 ±5,70	309,84 ±14,12 ±4,56	172,93 ±7,47 ±4,32	38,58 ±5,65 ±14,66	15,16 ±2,08 ±13,70	1,80 ±0,05 ±2,55	16,83 ±0,30 ±1,80
Amplitude 15º – 20.9º	11	102,60 ±8,23 ±8,02	319,03 ±26,71 ±8,37	178,55 ±17,05 ±9,55	33,4 ±23,63 ±10,85	14,30 ±2,52 ±17,65	1,81 ±0,06 ±3,09	17,42 ±0,32 ±1,82

The SEM and SPEM tasks results of the standard errors in 14 trials are shown in Tables 1 and 2.

Presents the means of saccadic parameters (duration, peak velocity, average velocity, slope, sharpness, Q, amplitude) and the standard uncertainity and the percentage of the standard uncertainty.

The values of saccadic parameters are changeable with increasing amplitude, therefore the results of each individual task is divided into five categories based on the scope of the amplitude (0.4°–2.9°, 3°–5.9°, 6°–9.9°, 10°–14.9°, 15°–20.9º).

In addition to the typical eye movement parameters as duration, speed and amplitude, was calculated sharpness, skewness and Q .

* **Direction** – the saccades direction (right or left)
* **Duration [ms]** – the mean saccades duration time in milliseconds
* **Amplitude [deg]**– the mean amplitude of the saccades
* **Peak velocity [deg/s]** – the mean peak velocity of the eye movement during saccades
* **Average velocity [deg/s]** – the mean average velocity of the eye movement during saccades
* **Slope** – the mean tilt of the increasing slope of the saccades velocity profile
* **Sharpness** – the mean peak velocity to duration time ratio
* **Q –** the mean peak velocity to the mean average velocity ratio

The results are shown in Figures 4-9. Bar charts have been separated by the amplitude, direction of eye movements (left or right) and the type of task (SEM or SPEM). In Figures 4-6 and 18-19 shows the increasing duration and velocity depending on the growing amplitude. On all the charts you can see the asymmetry in the values of various parameters depending on the direction of eye movement. The subject had a reduced velocity and increased duration of right movements in relation to the left movements. This relationship was confirmed at the velocity profiles (Figure 10) and the main-sequence diagrams (Figure 18 and 19).

Figures 11-16 show the means o saccadic parameters in individual trials separated by the amplitude. The largest differences were recorded in the first and the eighth trial. It was also noted that the higher the amplitude the higher the standard deviation and greater differences between the trials.
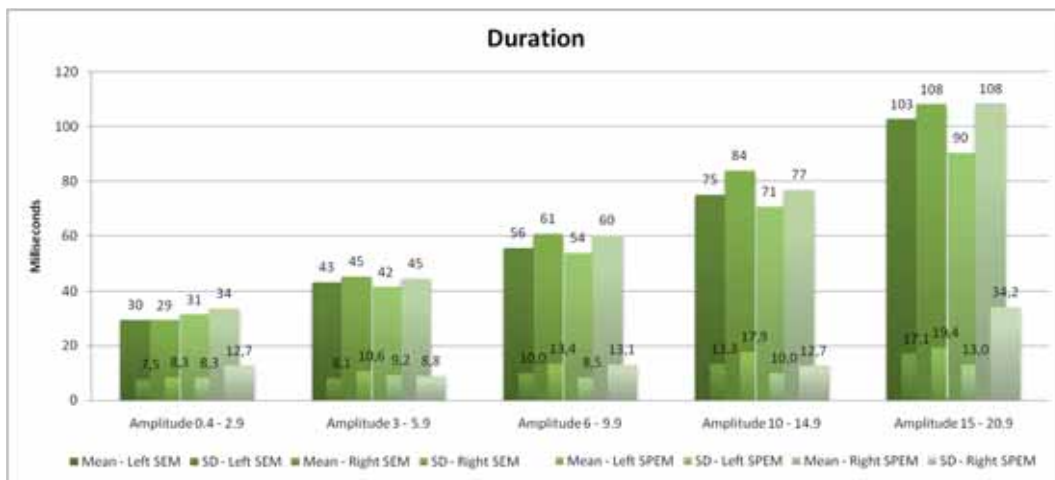
Fig. 4. Duration [ms] – the saccade duration time in milliseconds (means and SD) separated by the amplitude and the direction
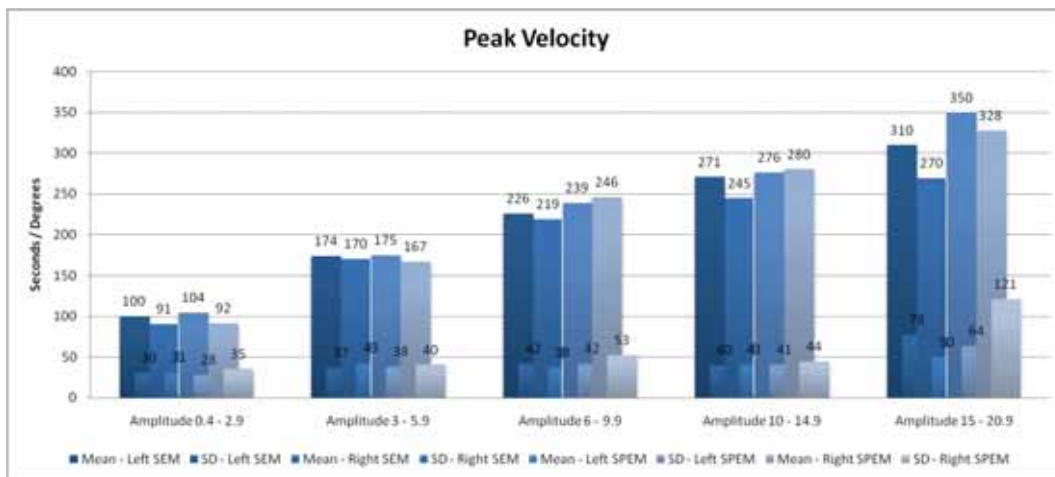
Fig. 5. Peak velocity [deg/s] – the peak velocity of the eye movement during saccade in angular degrees per seconds (means and SD) separated by the amplitude and the direction
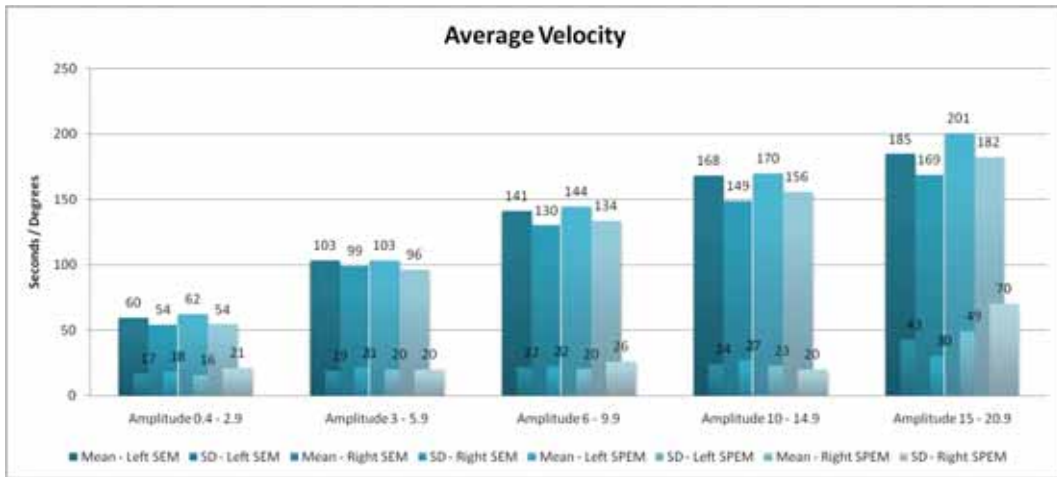
Fig. 6. Average velocity [deg/s] – the average velocity of the eye movement during saccade in angular degrees per seconds (means and SD) separated by the amplitude and the direction
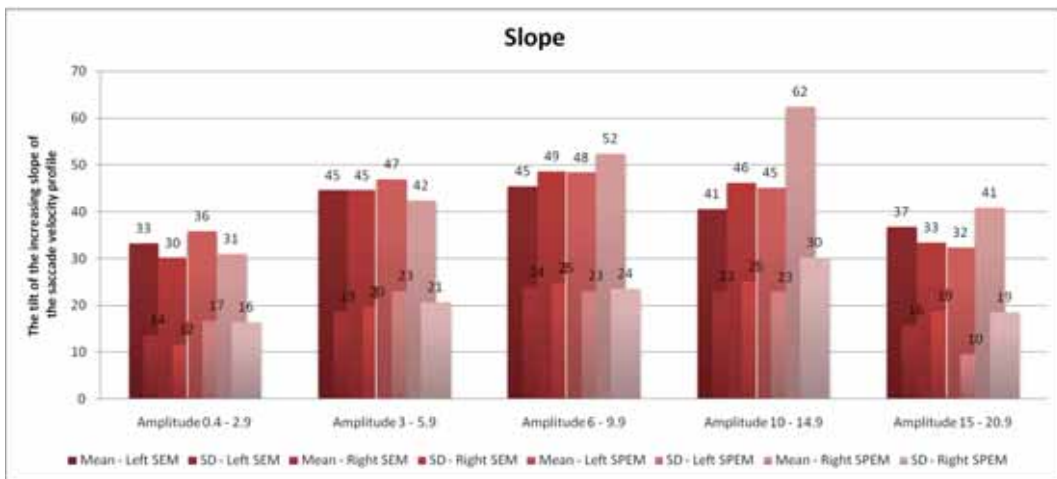

Fig. 7. Slope – the tilt of the increasing slope of the saccade velocity profile (means and SD) separated by the amplitude and the direction
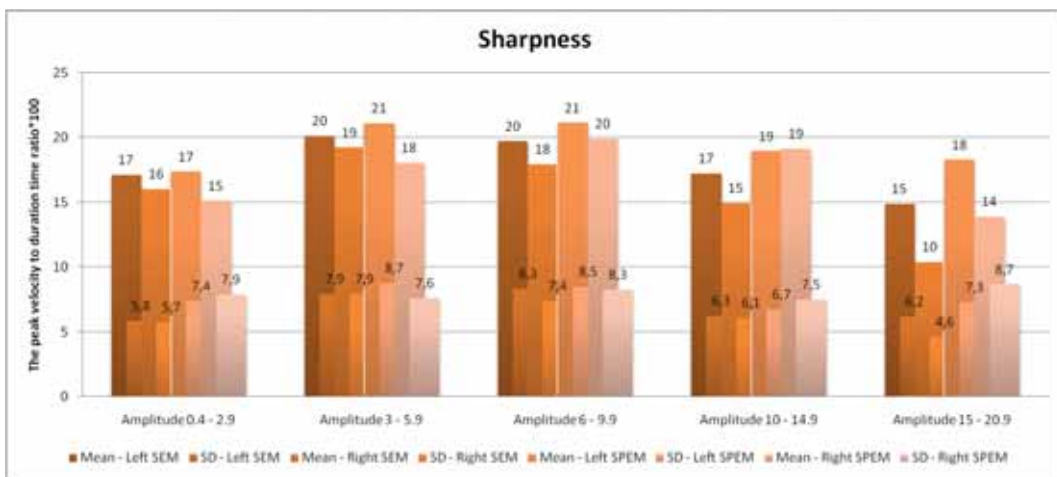

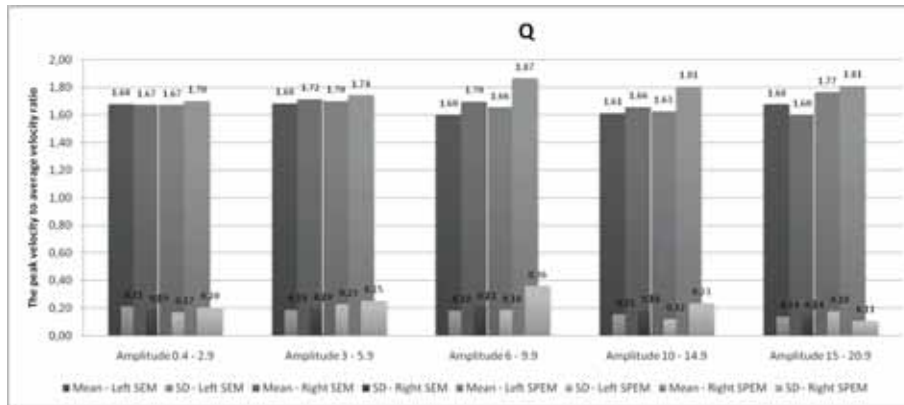**Fig. 8**. Sharpness – The peak velocity to duration time ratio (means and SD) separated by the amplitude and the direction

Fig. 9. Q – The peak velocity to average velocity ratio (means and SD) separated by the amplitude and the direction
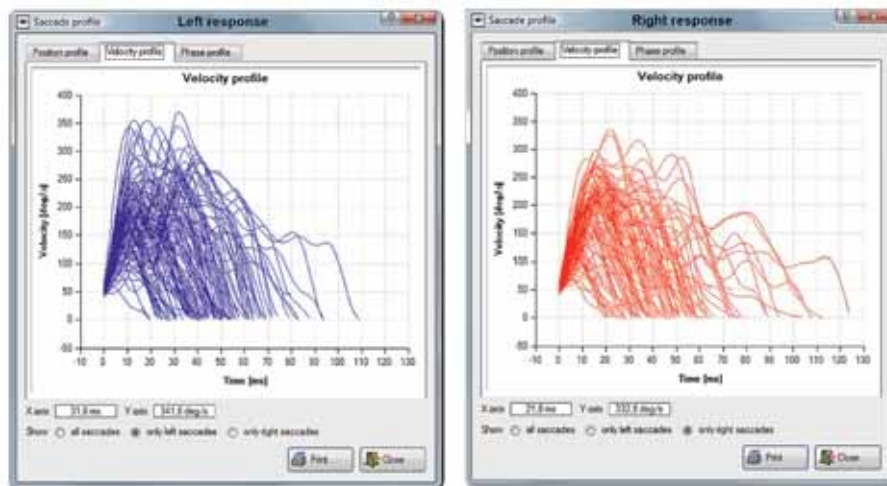
Fig. 10. The windows of the Saccades Browser application. There are two saccade graphical velocity profiles for left saccades and right saccades
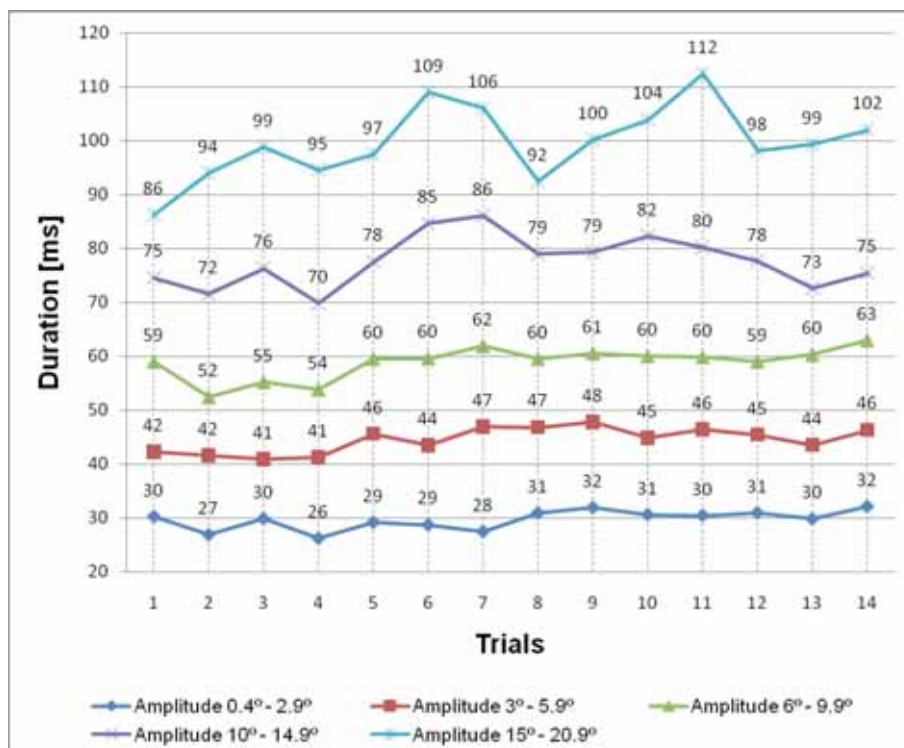
Fig. 11. The means of saccade duration time [ms] in each trial separated by the amplitude
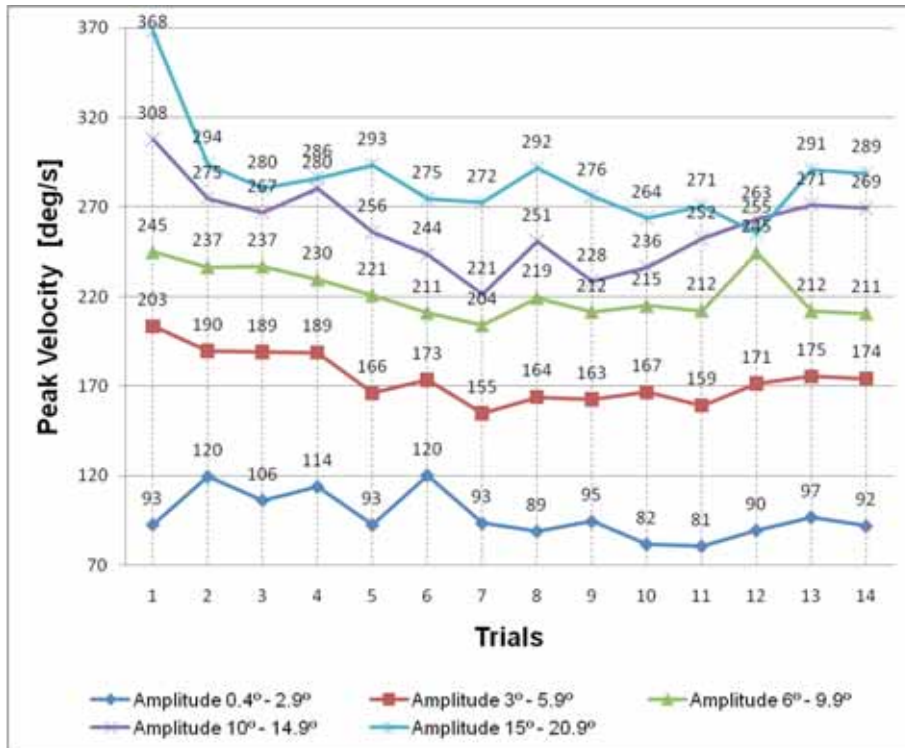
Fig. 12. The means of saccade peak velocity [deg/s] in each trial separated by the amplitude
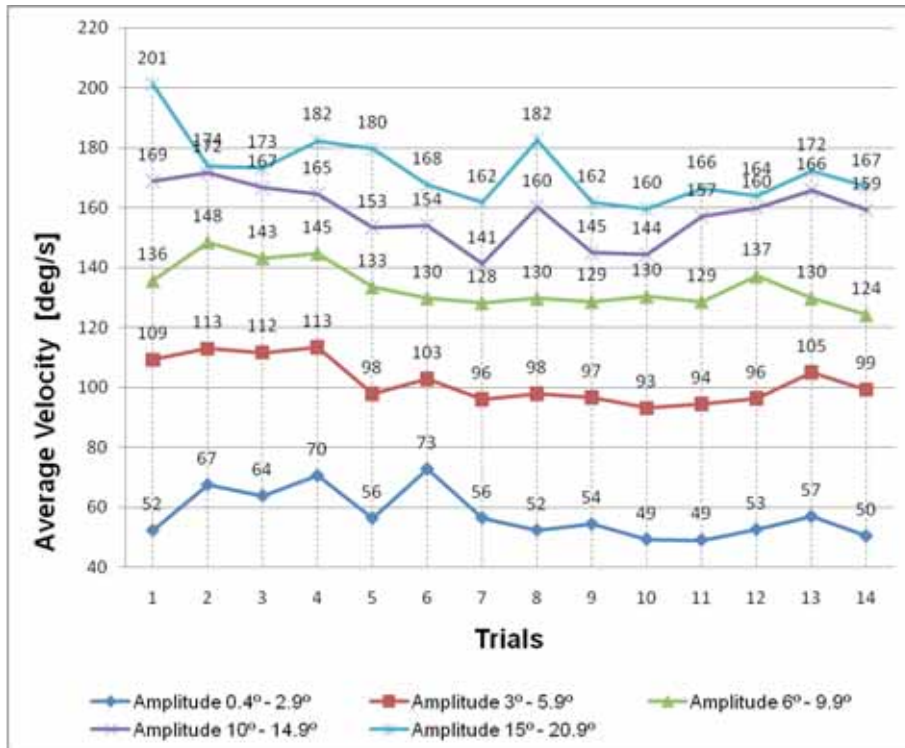
Fig. 13. The means of saccade average velocity [deg/s] in each trial separated by the amplitude

Medical diagnostics
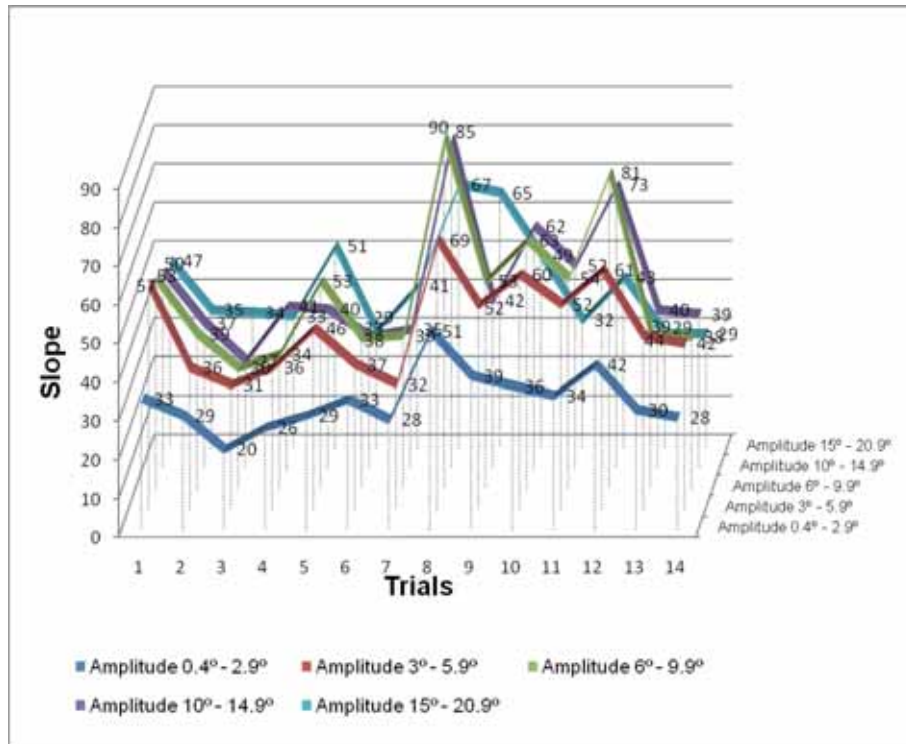
Fig. 14. The means of slope in each trial separated by the amplitude
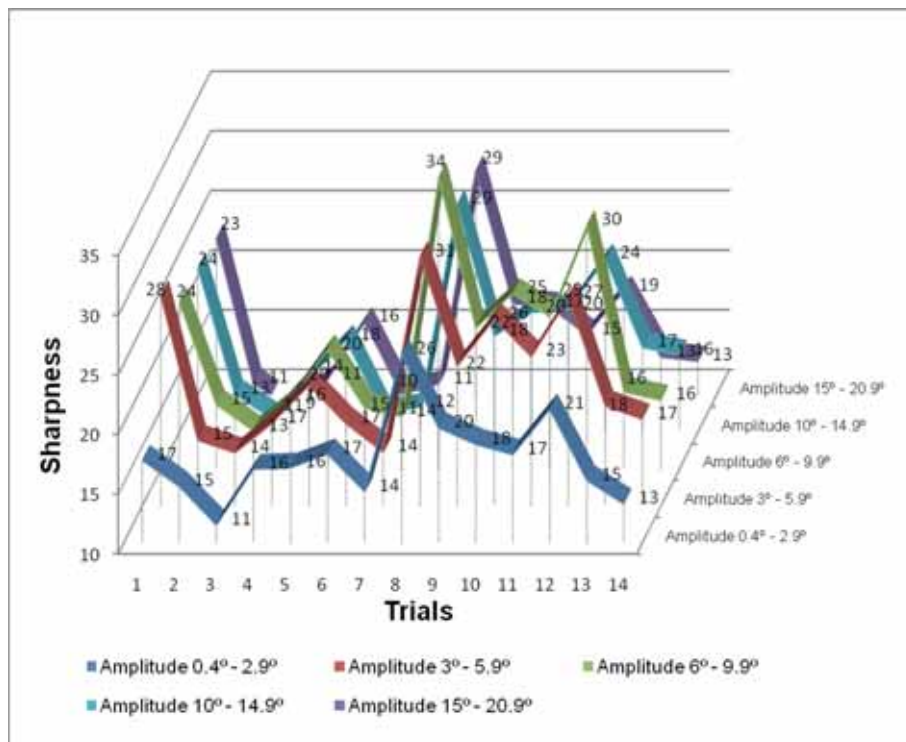
Fig. 15. The means of sharpness in each trial separated by the amplitude
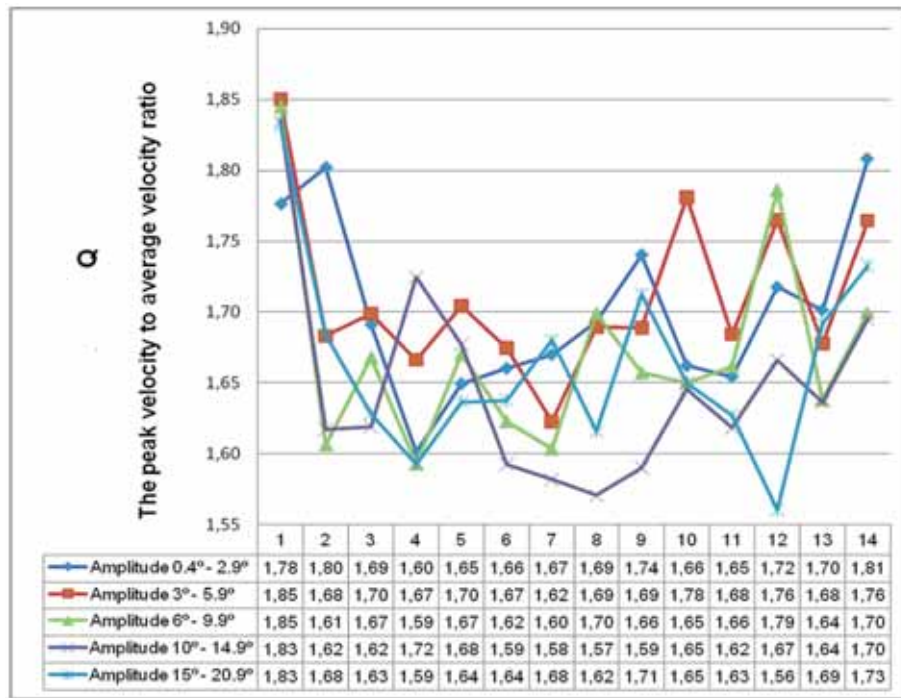
Fig. 16. The means of Q value (the peak velocity to average velocity ratio) in each trial separated by the amplitude

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Amplitude 0.4°- 2.9°	1,78	1,80	1,69	1,60	1,65	1,66	1,67	1,69	1,74	1,66	1,65	1,72	1,70	1,81
Amplitude 3°- 5.9°	1,85	1,68	1,70	1,67	1,70	1,67	1,62	1,69	1,69	1,78	1,68	1,76	1,68	1,76
Amplitude 6°- 9.9°	1,85	1,61	1,67	1,59	1,67	1,62	1,60	1,70	1,66	1,65	1,66	1,79	1,64	1,70
Amplitude 10°- 14.9°	1,83	1,62	1,62	1,72	1,68	1,59	1,58	1,57	1,59	1,65	1,62	1,67	1,64	1,70
Amplitude 15°- 20.9°	1,83	1,68	1,63	1,59	1,64	1,64	1,68	1,62	1,71	1,65	1,63	1,56	1,69	1,73
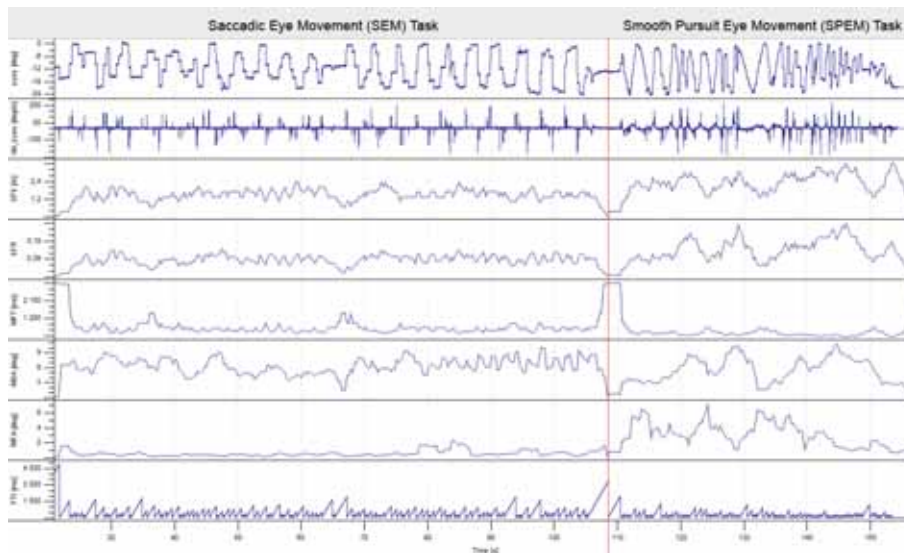
Fig. 17. Graphs of the Saccadic Eye Movement Task (SEM) and the Smooth Pursuit Eye Movement Task (SPEM). The comparison of number of saccades per second (SPS), saccadic time to fixation time ratio (SFR), mean fixation time (MFT), mean saccadic amplitude (MSA), mean fixation amplitude (MFA), fixation time integrator (FTI) in the SEM and SPEM tasks

Figure 17 presents graphs which show how different tasks. It has been calculated and presented graphically, the following indicators of saccadic activity.

Figure 17 presents graphs that show how varied the SEM and SPEM tasks due to the saccadic activity. The results were calculated and presented in graphical form, taking into account the basic indicators of saccadic activity:

- **Number of Saccades Per Second (SPS)** – the average number of saccades per second performed in the given time window.

- **Saccadic time to Fixation time Ratio (SFR)** – the total duration time of saccades to total duration time of fixation in the given time window ratio.
- **Mean Fixation Time (MFT)** – mean fixation time in the given time window.
- **Mean Saccadic Amplitude (MSA)** – mean amplitude of saccades in the given time window.
- **Mean Fixation Amplitude (MFA)** – mean amplitude of fixations in the given time window.
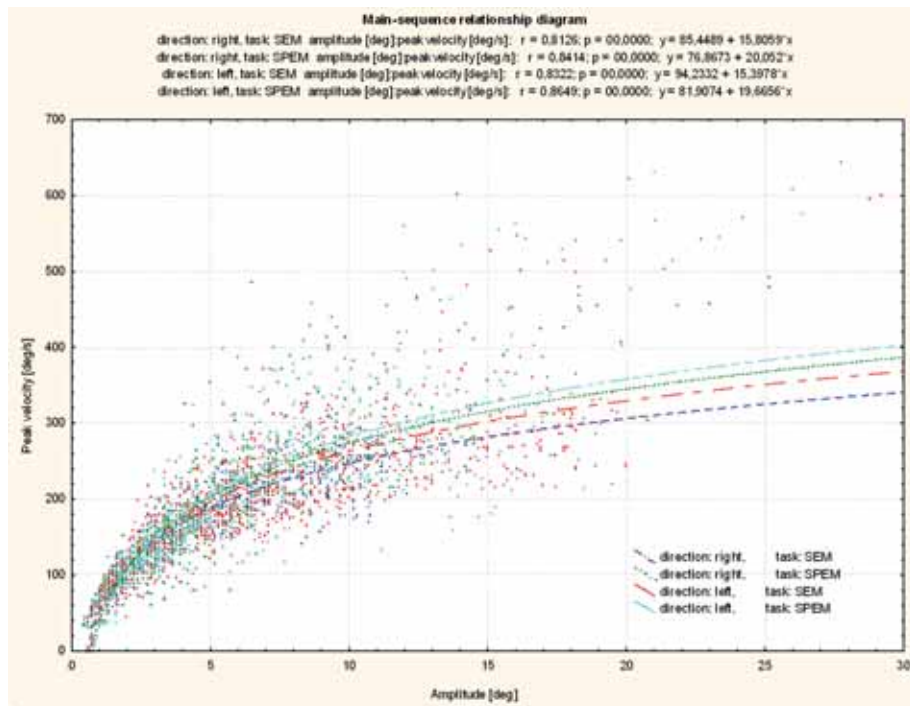
Fig. 18. The main-sequence diagram – the amplitude and peak velocity relationships of multiple saccades separated by the direction and the task
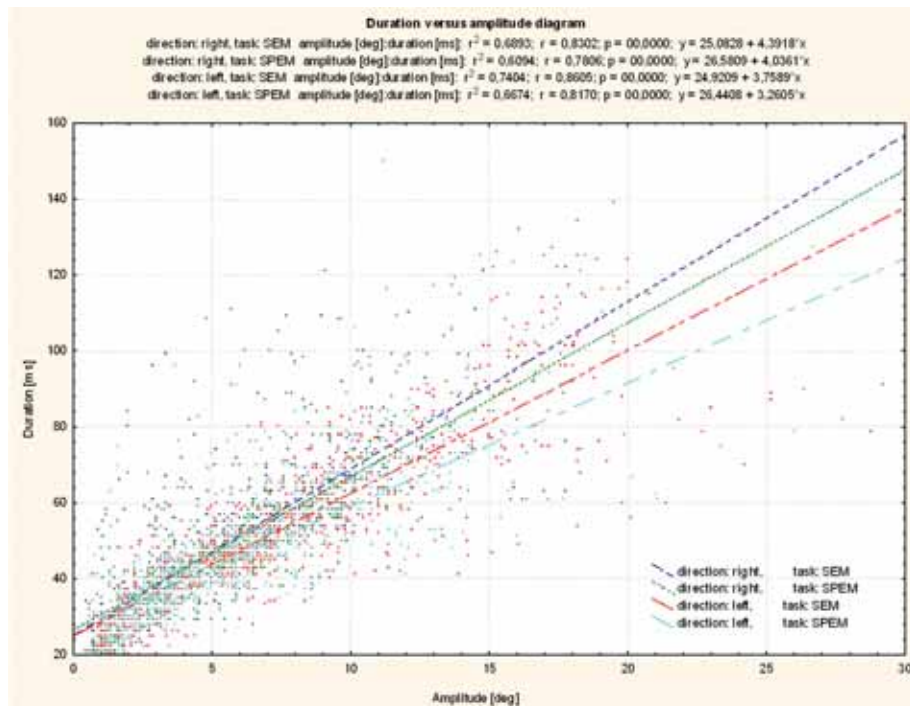

Fig. 19. The main-sequence diagram – the amplitude and duration relationships of multiple saccades separated by the direction and the task

- **Fixation Time Integrator (FTI)** – the integrator's level which growths up during fixation and drop to zero at the end of the saccade.

In Figures 18 and 19 were presented the main-sequence diagrams illustrating the relationships between the amplitude and peak velocity and the amplitude and duration. The relationships of multiple saccades were separated by the direction and the task. In Figure 18 to the dispersion plot are matched logarithmic regression line with r and p calculated. In Figure 19 to the dispersion plot are matched linear regression line with r, r^2 and p calculated.

Conclusion

The standard error of all the parameters analyzed was small and dependent of amplitude. The smallest errors was found in duration and Q ($x < 2\%$), while the highest error was found for slope and sharpness ($5\% < x < 10\%$).

We conclude from the results of tests that measure the dynamics of eye movements using Jazz oculometer may be a good diagnostic method characterized by their actual physiological level. Based on a comparison of the saccadic parameters was noted asymmetry of the left and right directions of eyeball movement in saccades duration, velocities and sharpness.

Although human behavior is characterized by enormous complexity, incomparable to other organisms, it is often carried out through sensory-motor systems with a relatively simple construction. Analysis of sensory and motor systems contributes in this way to a better understanding of the functioning of more complex structures of the CNS. That knowledge gives us the understanding of cooperation agreements in the sensory-motor processing of sensory stimuli and behavioral expression that is used in neuropsychology, and medical diagnostics.

Knowledge of the neurophysiological basis of the eyes' movement is essential for the physician. It also provides important information for the neuropsychologist showing the basis for the mechanism of visual attention and for exploring the evolutionary neuroscience organizational structure of the various sensory-motor systems. Visual and oculomotor system is an excellent example of sensory-motor integration through which sensory signals are transformed to the output motor. The difficulty of the test of eye movements lies in their dynamics, because the phenomena lasts for a very short period of time – to the tune of milliseconds. However, the use of modern diagnostic equipment such as oculometers or saccadometers and the use of computers for processing and analyzing the results, contributes to the dissemination of such research. Diagnosis of eye movements is used not only in medicine, but more often in the industry and the military, where certain characteristics of the motion carry key information on the effectiveness of specific activities.

References

1. Accardo A.P., Pensiero S., Da Pozzo S., & Perissutti P. (1995), Characteristics of horizontal smooth pursuit eye movements to sinusoidal stimulation in children of primary school age. Vision Research, 35, 539-548.
2. Becker W. (1989), The neurobiology of saccadic eye movements (eds. Wurtz and Goldberg). Amsterdam: Elsevier.
3. Carpenter R.H.S. (1988), Movements of the eyes. London: Pion.
4. Fischer B., Biscaldi M., & Gezeck S. (1997), On the development of voluntary and reflexive components in human saccade generation. Brain Research, 754, 285-297.
5. Hoffman J.E. & Subramaniam B. (1995), The role of visual attention in saccadic eye movements. Perception and Psychophysics, 57, 787-795.
6. Jacobsen L.K., Hong W.L., Hommer D.W., Hamburger S.D., Castellanos F.X., Frazier J.A., et al. (1996), Smooth pursuit eye movements in childhood-onset schizophrenia: Comparison with attention-deficit hyperactivity disorder and normal controls. Biological Psychiatry, 40, 1144-1154.
7. Jaśkowski P., Sobieralska K. (2004), Effect of stimulus intensity on manual and saccadic reaction time. Perception and Psychophysics, 66, 535-544.
8. Karatekin C. (2007), Eye tracking studies of normative and atypical development. Developmental Review, 27, 283--348.
9. Krauzlis R.J. (2005), The Control of Voluntary Eye Movements: New Perspectives. Neuroscientist, 11(2), 124-137.
10. Leigh R.J. & Zee D.S. (1999), The neurology of eye movements (3ed.). Philadelphia: F.A. Davis.
11. Levy D., Holzman R., Matthyse S., Mendeil N. (1993), Eye tracking dysfunction and schizophrenia: a critical perspective. Schizophr. Buli., 19, 3, 461-536.
12. Liversedge S.P. & Findlay J.M. (2000), Saccadic eye movements and cognition. Trends in Cognitive Sciences, 4, 6-14.
13. Missal M., Keller E.L. (2002), Common inhibitory mechanism for saccades and smooth-pursuit eye movements. J. Neurophysiol., 88, 1880-1892.
14. Sheliga B.M., Riggio L., Rizzolatti G. (1995), Spatial attention and eye movements. Exp. Brain Res., 105, 261-275.
15. Tanaka M., Lisberger S.G. (2001), Regulation of the gain of visually guided smooth-pursuit eye movements by frontal cortex. Nature, 409, 191-194.
16. Walecki P. (2007), Neurofizjologia ruchu gałek ocznych (Neurophysiology of eye movement), Episteme, 4, 159--176.
17. Yan Y.J., Cui D.M., Lynch J.C. (2001), Overlap of saccadic and pursuit eye movement systems in the brain stem reticular formation. J. Neurophysiol., 86, 3056-3060.

Medical diagnostics

DIGITALIZED DRAWING TEST (DDT) – THE TOOL FOR AN EARLY SUBCLINICAL MOTOR SYMPTOMS DIAGNOSIS

EDWARD J. GORZELAŃCZYK[1, 2, 3,], KRZYSZTOF WŁODARCZYK[1], EWA LASKOWSKA[2],

MAREK KUNC[4], JACEK JELONEK[5]

[1] *Sue Ryder Home, Bydgoszcz, Poland*
[2] *Laboratory of Neuropsychology and Psychophysiology, Institute of Psychology, Kazimierz Wielki University, Bydgoszcz, Poland*
[3] *Regional Mental Hospital in Świecie, Poland*
[4] *University of Leeds, UK*
[5] *Institute of Computing Science, Poznań University of Technology, Poznań, Poland*

Abstract: Motor functioning is linked with cortico-subcortical circuits and can be the biological marker of some disorders. Motor and tension disorders of specific muscles can be monitored during drawing test completion. The advantage of computer aided tests is that they allow to perform mathematical and statistical analysis. By analyzing primary movements parameters and spectrograms of hand tremor during a drawing test, the researcher obtains a diagnostic data set.

A digitalized drawing test consists of a pattern (designed by Edward Jacek Gorzelańczyk) on a piece of paper with special figures. Each shape is of diagnostic value for certain motor abilities (precise movement, muscle tone, kinetics of whole hand/arm). A PC digital tablet with an electro-magnetic ballpen allows to register digital data during the test. The application MedTablet, co-designed by the author of this article, is used to store and analyze the obtained data.

Alcohol or other psychoactive substance addiction gives specific tremors that can be observed in a momentary velocity profile. The profile can also contribute to Parkinson's disease or intention tremor distinction. What is more, motor disorders can be diagnosed in the preclinic phase of the disease. This digitalized drawing test was also used to evaluate the level of thalamotomy or pallidotomy. A tremor affecting any area of the body is transmitted to hand during the drawing task.

Digitalized drawing test can become a subclinical motor symptoms diagnostic tool in PD, addiction and motor dysfunctions in other diseases.

Keywords: early diagnostic, digitalized drawing test, Parkinson's disease, kinetic abnormalities, movement parameters, motor circuit

Introduction

Early diagnosis in medicine gives a chance for an appropriate treatment. Some disorders are difficult to diagnose, especially in early stages. The control of motor, emotional and cognitive functions is linked with cortico-subcortical circuits. Motor circuit is responsible for automatic movements, static position of human body and muscle tone [1]. This circuit starts, controls and ends movement that depends on will. Abnormalities in its components can be observed as bradykinesia and muscle stiffness (hypertonic disorder), akinesia and hypokinesia [2, 3]. In Parkinson's disease dysfunctions of motor circuit are one of the main mechanisms causing tremor symptoms.

Materials and methods

Different tests have been performed (e.g. digitalized drawing test) during past 5 years in some groups of patients including subjects with alcohol dependency, other psychoactive substances (drugs) dependency, Parkinson's disease (PD) – in different stages of disorder or on different treatment, with Alzheimer's disease and healthy controls. Some of the results are presented in this article. PD causes hypokinesia that can be easily registered and diagnosed during drawing tests, especially digitalized drawing tasks.

Digitalized drawing test (original digitalized drawing task)

This test and the pattern was designed by Edward Jacek Gorzelańczyk in 2000 in order to evaluate patients with PD after pallidotomy or thalamotomy [4, 5]. Additionally the tool enables to diagnose patients with Parkinson's disease in early stages, when a tremor is subtle and difficult to detect clinically.

Medical diagnostics

Digital tablet with drawing pattern

Special drawing patterns have been designed (fig. 1.). They consist of different shapes – 13 separate continuous lines grouped into 8 figures. Each group of shapes represents different kinetic properties.

The test allows to evaluate motor activity of a diagnosed patient. The examination can be conducted in a casual office room or hospital room without any particular external conditions. The equipment that is essential to carry out the examination comprises only of the following: sheets of paper with patterns of figures and a digital tablet linked to a PC with software collecting and analyzing signals emitted during task completion.

A patient is given a drawing pattern on a tablet (as on a holding board) and an ink or non-ink ballpen – connected to the tablet. The kind of electro-magnetic pens can modify examined person's movement parameters, such as pressure, distance of drawing, etc. The researcher explains how to use the writing tool to the patient and shows the pattern. Nearly routine activity like drawing or writing decreases the stress during the task.

Graphical interface (MedTablet Application)

The software has been designed for the purpose of using certain peripherals and special drawing patterns. It has also wide variety of possible applications. New drawing tasks or even spontaneous drawings can be registered and analyzed. Graphic interface (fig. 2.) is intuitive and easy to use, even without explanation.

The movement of the electro-magnetic ballpen is registered by the digitalized tablet connected to a computer. The parameters of each position in the 3-dimensional scale are stored.

Horizontal coordinates (X, Y) are strictly linked with the pattern and the vertical coordinate (Z) is connected with the pressure of the drawing. It's important to use tablets with at least 1024 grades of pressure.

During drawing the movement over the pattern is registered simultaneously as a pair of points on the horizontal surface (X, Y) and the height parameter (Z). The position is shown in a four-dimensional arrangement p (x, y, z, t). Data collected in this way (on the basis of the figures grouped in single shapes) allow to use the FFT (Fast Fourier Transform) with relation to inclusion conditions. It shows not only the dynamic parameters of the movement but also their spectrum.

While performing a drawing task (from the beginning to the end) a wide variety of parameters is calculated. Most of them are acquired only if the actual pressure of the drawing is different than "0", which means that a non-drawing is observed.

Dynamic parameters of the movement are acquired during the examination:

a) distance of drawing (DIST) [cm] – represents the length of all drawn shapes
b) time of drawing (T) [s]
c) mean speed (MN_SP) [cm/s] – obtained by division of DIST by T
d) mean momentary velocity (MN_LSP) [cm/s] – the average of all instantaneous velocities
e) maximum momentary velocity (MX_LSP) [cm/s]
f) standard deviation of momentary velocity (SD_LSP) [cm/s] – represents the dispersion of instantaneous velocities and it is connected with a movement fluency
g) mean pressure of drawing (MN_PRS) [a.u., 0-1023] – which is also shown in the drawing (task visualization) by different colors
h) standard deviation from mean pressure [a.u.]
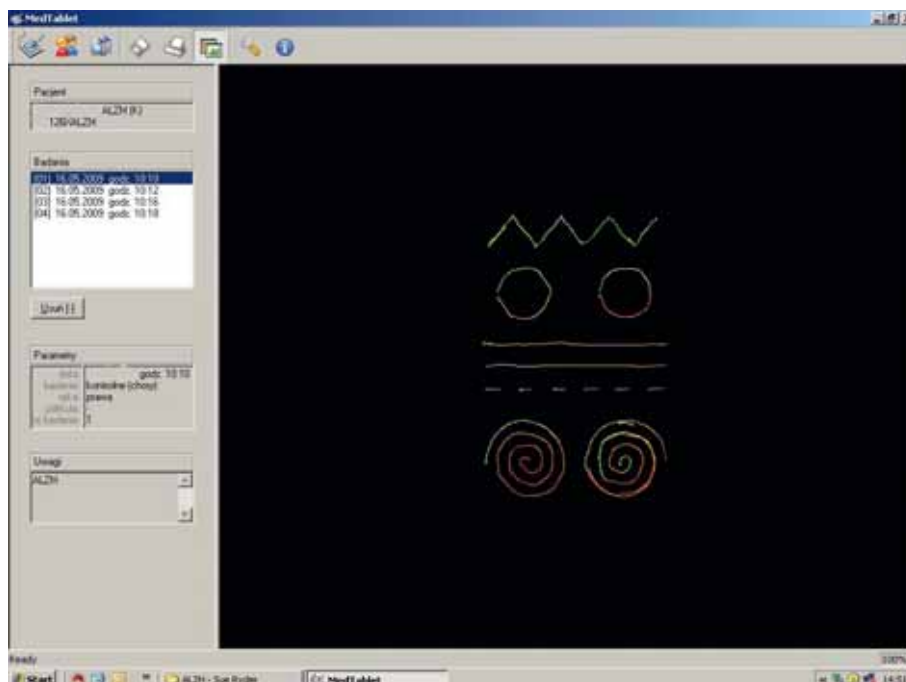i) maximum pressure of drawing [a.u.].
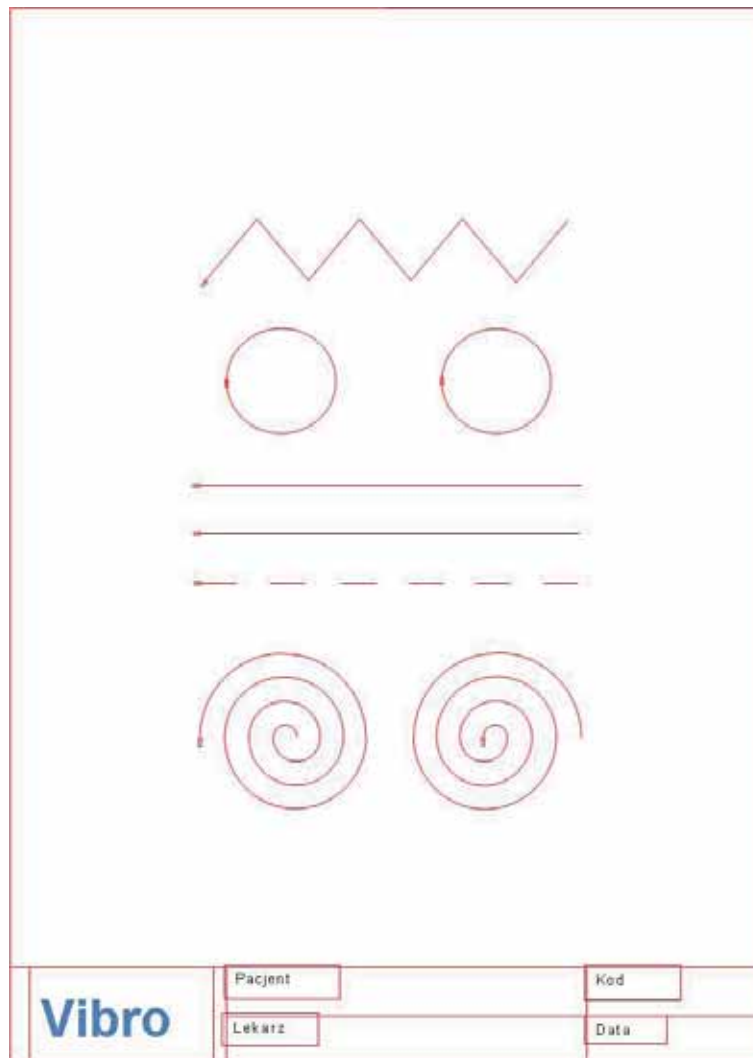

Fig. 1. Drawing task pattern

Fig. 2. MedTablet interface – during the examination of the patient with Alzheimer's disease

The pattern used in the test is constant, so the distance of drawing should be similar for all performers. The patients with better motor control function obtain results (relating to distance) that are approximate to the pattern. A tremor of hands or other areas of the body extend the distance of drawing.

Time is related with distance of drawing, but only to some extent. It is possible to perform a long distance drawing in a short period of time, especially with a non-ink pen. Through analyzing mean velocity the examiner is able to compare different groups of patients. Changes of instantaneous velocity represent fluency of the movements.

The pressure of drawing is a diagnostic parameter, especially for patients with hypo – or hyperkinetic dysfunctions [6]. Current resolution of pressure scale enables to register 1024 positions, which is enough to diagnose Parkinson's disease.

More robust analysis of the drawing task provides such variables as:
a) length of each (separate) shape [cm]
b) shape drawing time [s]
c) time of switching between shapes [s]
d) spectrum of each parameter.

Each figure on the pattern can be analyzed as a separate drawing. This allows analyzing of 8 different figures, which can be very useful when searching for a definite abnormality in some motor functions. The analysis of (…) specific drawing elements: straight lines, circles and curved shapes gives other diagnostic clues. Some diseases cause hyperkinetic (Huntington's disease), other hypokinetic (Parkinson's disease) dysfunctions or fluctuating tonic abnormalities.

Obtaining the spectrum of dynamic parameters of movement enables deriving many conclusions. Alcohol addiction and Parkinson's disease causes tremors of hands, but with different profile [5].

In the MedTablet application the tremor analysis is made in the range from 1 to 16 Hz.

Results

The results formulation allows stressing the diagnostic strength of the DDT. One pattern can be used in a variety of tests. The tests of 24-year-old man with intention tremor of right hand before and after stereotactic thalamotomy were performed on the Wacom Intuos 2 Tablet (Fig. 3.).

Even without a pressure visualization, a tremor can be observed. Additional scale and the spectrogram (Fig. 4.) of devia-
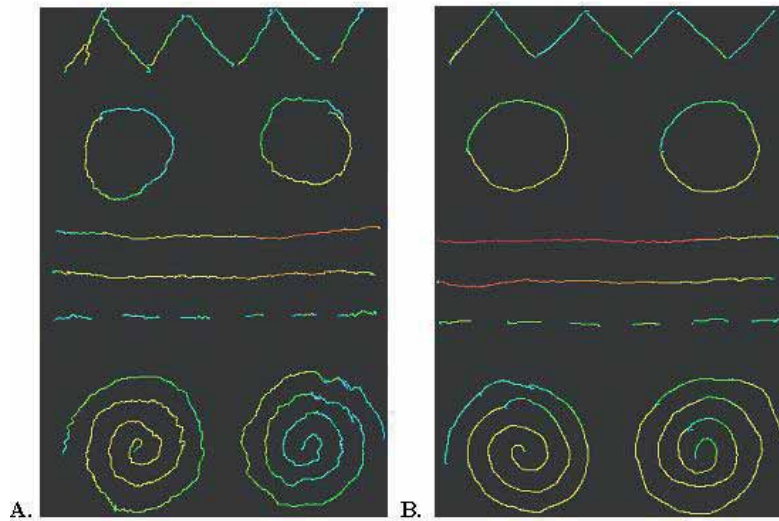
Medical diagnostics

Fig. 3. Visualization of drawing test of the patient with intention tremor (right-handed performance): A – before left-sided thalamotomy; B – after left-sided thalamotomy
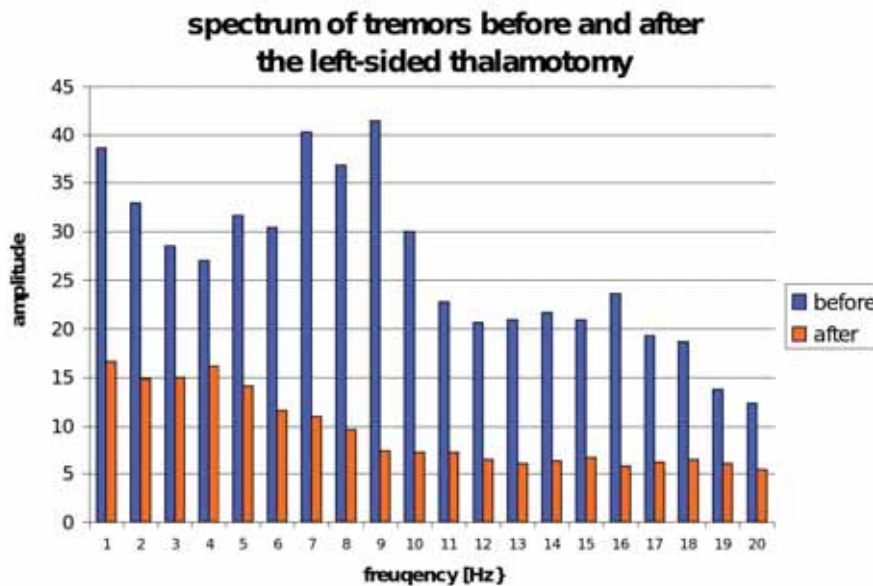

Fig. 4. Spectrogram of tremor in the patient with intention tremor (right-handed performance): A – before left-sided thalamotomy; B – after left-sided thalamotomy

tion from the pattern indicates that some specific tremors have drastically changed. The frequency of tremors is consistent with PD [4, 5].

Tremor frequency ranges between 5 Hz and 10 Hz, but the highest amplitude of it is in the range of 7 and 9 Hz. Before thalamotomy a tremor in this range is over 1 Hz (value in healthy patients). After thalamotomy, although tremors in this specific range in the PD patient are not observed, they are present. In healthy patients exponential tremor atrophia is very characteristic. Primary parameters of the movement can be of diagnostic value. Time and distance of the drawing can be the indicator of a hand-eye coordination (motor and oculomotor circuit). Comparisons of movement parameters of 50 healthy

controls of different age (25 young controls – 18-25 years old and 25 controls – 40-55 years old) and 25 patients hospitalized for alcohol dependency have been made (fig. 5., fig. 6.).

The momentary velocity of each group is statistically significantly different (Anova test, $p<0.01$). Patients addicted to alcohol perform drawing task slower than the healthy ones. Motor control function is not as good as in healthy controls in the same age or older. Young controls perform the drawing test with more than 25% higher instantaneous velocity than addicted patients. Velocity depends on age, but also on motor circuit functioning. The conclusion from the patient examination is made that motor circuit is damaged by alcohol abuse.
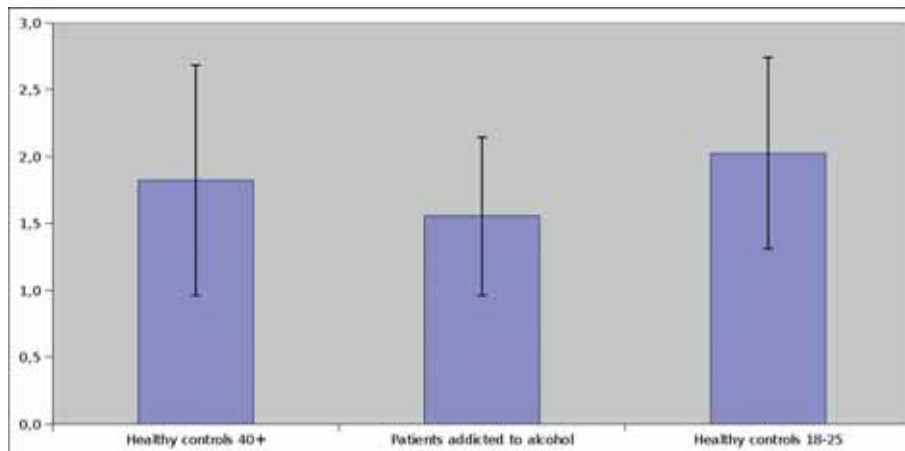
Fig. 5. Mean momentary velocity of healthy controls (in 2 groups) and alcohol addicted patients during performance of digitalized drawing test
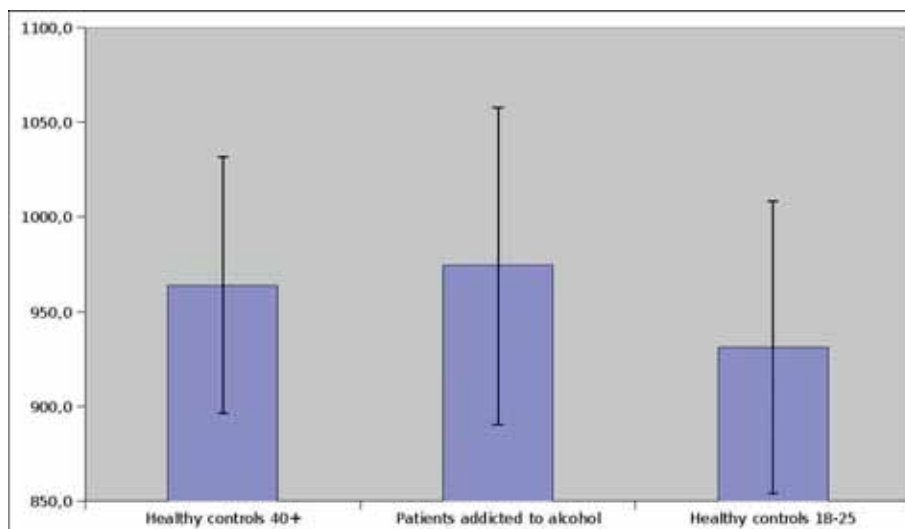

Fig. 6. Maximum pressure of drawing of healthy controls (2 groups) and alcohol addicted patients during performance of digitalized drawing test

The maximum pressure of drawing is statistically different in all 3 groups. Patients addicted to alcohol have the highest level of pressure among these groups. The examination also confirms that the motor functioning is disturbed by the lack of alcohol in addicted patients or by organic brain damage. The lowest value of pressure was noticed in the group of young healthy controls. It indicates that better hand-eye coordination with a good motor functioning is proportional (with minus coefficient) to the pressure of drawing. High pressure of drawing or the changes of pressure indicate some psychosomatic or somatic abnormalities.

Also the spectrograms of momentary velocity (fig. 7, fig. 8.) are diagnostic for some disorders in motor dysfunctions. In two men:

– middle-aged alcohol addicted

– 25-year-old addicted to alcohol and psychoactive substances organic damages were confirmed.

Both men had hand tremor and other body areas tremor (lesser than hands) on a daily basis. The tremors exacerbated after alcohol withdrawal or the withdrawal of other psychoac-

tive substances. Other exacerbating factors noted were tiredness and stress.

The highest values of amplitudes of instantaneous velocity are shown on the spectrogram from 6 Hz to 8 Hz and 13 Hz to 15 Hz. The tendency can also be observed in deviations and spectrogram of the pressure of drawing. The frequency of tremor has lower value in alcohol dependent patients than in patients with PD, but probably it is connected with hypertonic disorders. The results achieved may depend on co-existing diseases, which should become known to the specialist in a medical interview.

The tremor in this case occurs in a narrower range between 7-8 Hz, but the profile of momentary velocity is slightly less physiologic (less exponential). Tremors of both patients can be observed and are connected with high level of muscle tonus and improper neurotransmission. The profiles of the spectrograms can indicate causes of disorders and are specific for each of them.
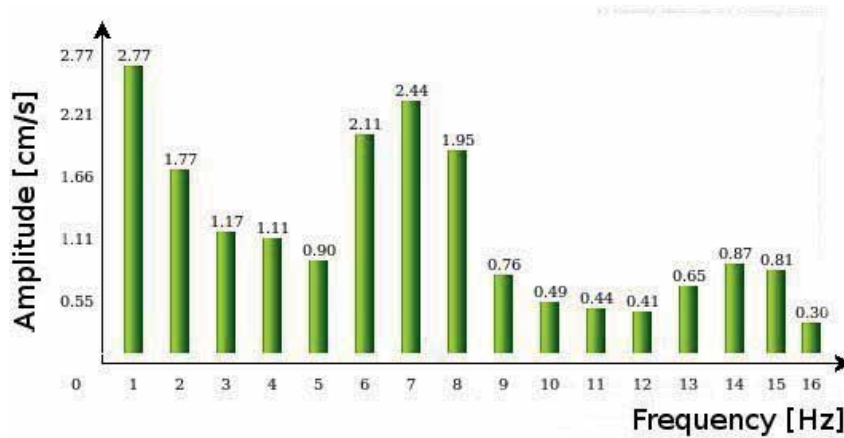
Medical diagnostics

Fig. 7. Spectrogram of momentary velocity of alcohol addicted patient
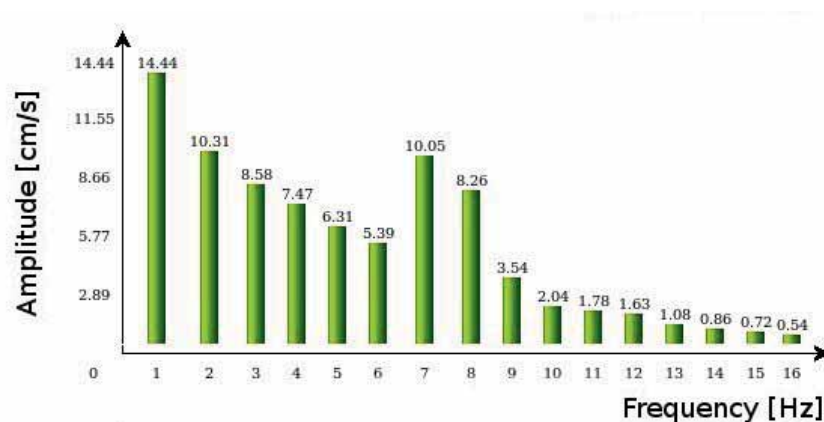

Fig. 8. Spectrogram of momentary velocity of alcohol and drug addicted patient

Discussion

The diagnostic value of a digitalized drawing test have been confirmed during the clinical test with patients with PD [6], addicted to alcohol or other psychoactive substances. Some of the motor dysfunctions can be observed [7] and evaluated without using any tool, but during clinical research prediction is insufficient. If tremors of hands or other parts of body can be observed, it means that more advanced stage of a disease or a dysfunction has developed.

Tremors can be observed in the spectrograms of deviations, momentary (instantaneous) velocity or pressure of drawing or through deeper analysis of each figure in the drawing test with passing time.

Velocity and maximum or momentary pressure of drawing are the parameters that indicate the abnormalities in co-operation of occulomotor and motor circuit. Early diagnostic procedures can incorporate all or some of the obtained parameters. When looking for specific dysfunctions (for example tremors caused by the lack of alcohol in alcohol dependence) the researcher can observe only a few, the most important variables and parameters.

The usefulness of this method should be proved during a clinical test in larger experimental and control groups (200-400 patients). Larger number of the patients tested would make the obtained results more reliable.

Since the first use of DDT, results confirming its effectiveness and diagnostic advantages have been acquired. The experimetal groups were tested in hospitals all over Poland to extend the examination area and observe the tendency in the groups evaluated.

The main disadvantage of the DDT is the lack of standarization. The examinators are still developing norms (with strict range) for different groups of patients. Gender, age, education and dexterity are the examination classifiers.

The main advantage of the DDT is simplicity in performance and wide variety of possible usage contexts. In the data analysis extended modules have been used that are going to be replaced with Digital Signal Processing Modules during subsequent stages of the tests. No restriction is imposed on the patients' age and history. The same drawing pattern can be used for the patients with various diseases and abnormalities as well as for the healthy ones. People with highly damaged arms or with motor dysfunctions can use an electromagnetic pointer, fastened to their arms or other parts of body, that enables obtaining of the motor parameters.

DDT is evolving and expanding its application range.

Medical diagnostics

References

1. Groeneweg H.J. (2007), CNS Spectr., 12(12), 887-892.
2. Delong M.R., Wichmann T. (2007), Circuits and circuit disorders of the Basal Ganglia. Arch. Neurol., 64, 20-24.
3. Ring H.A., Serra-Mestres J. (2002), Neuropsychiatry of the basal ganglia. J. Neurol. Neurosurg. Psychiatry, 72, 12-21.
4. Gorzelańczyk E.J., Jelonek J., Harat M., Lenart-Jankowska D., Olzak M. (2003), Objective test of tremors in patients with Parkinson's disease. European Journal of Neurology, 10 (Suppl. 1), 45-124, 85.
5. Gorzelańczyk E.J., Harat M., Konnek M., Laskowska I., Michalak M., Olzak M. (2004), Cognitive outcome following pallidotomy in patients with Parkinson's disease. In: Korczyn A.S. (ed.), Mental Dysfunctions in Parkinson's Disease. Bologna, Medimond – International Proceedings Division, 131-135.
6. Olzak M., Laskowska I., Jelonek J., Michalak M., Szolna A., Gryz J., Harat M., Gorzelańczyk E.J. (2006), Psychomotor and executive functioning after unilateral posteroventral pallidotomy in patients with Parkinson's disease. Journal of the Neurological Sciences, 248, 97-103.
7. Verdejo-García A., Bechara A., Recknor E.C., Pérez-García M. (2006), Executive dysfunction in substance dependent individuals during drug use and abstinence: an examination of the behavioral, cognitive and emotional correlates of addiction. J. Int. Neuropsychol. Soc., 12(3), 405-415.

Medical diagnostics

ELECTROMAGNETIC DIGITIZING SYSTEM FOR PRECISE LOCATION OF EEG ELECTRODES

Tomasz Zyss

Department of Psychiatry, University Hospital, Cracow, e-mail: mzzyss@cyf-kr.edu.pl

Abstract: The paper describes the method of precise location of electrodes used in EEG technique. The modus of electromagnetic digitizer is presented. The author presents visualization possibilities of two software packages.

Keywords: EEG electrodes, localization, digitization

Streszczenie: Praca opisuje metodę dokładnego wyznaczania położenia elektrod stosowanych w technice EEG. Zaprezentowano działanie digitizera elektromagnetycznego. Przedstawiono możliwości wizualizacyjne dwóch pakietów oprogramowania.

Introduction

The electrophysical phenomena occurring in the brain neurons, especially on their cell membrane, can be recorded from the head skin in the form of the electroencephalogram (EEG). A standard EEG recording uses 19-21 electrodes positioned at definite points on the head surface – in accord with the international 10-20 system. Traditionally, the EEG technician who installs the cap with electrodes usually positions the electrodes by approximation. To increase precision of the method, use of measuring tape is advised [1].

When using computer assisted analytical methods for the evaluation of EEG data, accuracy of the positional measurement of the EEG electrodes on the patient's head has a direct influence of the results [2-5].

Some special systems have been developed for highly precise determining of EEG electrode positions. There are three main applications for those digitization systems:
– determining the 3D coordinates of electrodes already applied to the patient's head,
– positioning the electrodes after specification of the exact coordinates of the measuring point,
– sensing the cranial shape of the patient in 3D coordinates for correlation with data from imaging methods.

Digitizer CMS20 by *Brain Vision* (*Zebris*) can be an example of a system for electrode placement (fig. 1). It collects the data regarding the placement of electrodes by analysis and processing of ultrasonic signal generated by a special antenna. Investigations have shown that this kind of ultrasonographic digitizers is less accurate that the equipment using electromagnetic emission [6].

However, the digitizers whose work is based on electromagnetic field emission are more commonly used.

Fig. 1. Digitizer CMS20 working on the basis of ultrasonic emission

Fastrak Digitizer

The Fastrak tracking system (*Polhemus* Firm) uses electromagnetic fields to determine the position and orientation of a remote object (fig. 2) [7]. The technology is based on generating of near field, low frequency magnetic field vectors from a single assembly of three concentric, stationary antennas called a transmitter, and detecting of the field vectors with a single assembly of three concentric, remote sensing antennas called a receiver. The sensed signals are input to a mathematical algorithm that computes the receiver's position and orientation relative to the transmitter.

The FASTRAK consists of a System Electronics Unit (SEU), one to four receivers and a single transmitter.

SEU contains the hardware and software necessary to generate and sense the magnetic fields, compute position and orientation, and interface with the host computer.

The transmitter contains electromagnetic coils enclosed in a molded plastic shell that emit the magnetic fields. The transmitter is the system's reference frame for receiver measurements.

The receiver contains electromagnetic coils enclosed in a molded plastic shell that detect the magnetic fields emitted by the transmitter. A lightweight cube, the receiver's position and orientation are precisely measured as it is moved. The receiver is completely passive, having no active voltage applied to it.

The system is capable of operating at any of four discrete carrier frequencies. Different carrier frequencies allow for operation of up to four FASTRAKS simultaneously and in close proximity to one another. The FASTRAK interfaces to the host computer via RS-232 serial communication or USB port. Any single receiver may be operated at the fastest update rate (120 Hz); any two receivers at one half of this rate; any three at one-third of this rate; or all four at one fourth of the fastest rate. Obviously, the unit must be set to output at a high enough baud rate to receive these updates. Mixed rates are not permitted meaning that all active receivers operate at the same update rate, i.e. one cannot be operated faster than another. Active receivers are selected by physical receiver cable connections and software configuration commands.

Since FASTRAK uses patented low-frequency magnetic transducing technology, there's no need to maintain a clear line-of-sight between the receivers and transmitters.

Additionally, FASTRAK may be used with a stylus or a 3ball device instead of a standard package receiver. Tip offsets are automatically calculated for the stylus and no special commands are required for this mode of operation. Switch functionality is provided with both the stylus and the 3ball device. The stylus and the 3ball may be used in any of the receiver ports.

FASTRAK is a 3D digitizer and a quad receiver motion tracker, making it perfect for a wide range of applications requiring high resolution, accuracy, and range. By computing the position and orientation of a small receiver as it moves through space, it provides dynamic, real-time measurements of position (X, Y, and Z Cartesian coordinates) and orientation (azimuth, elevation, and roll).

The azimuth, elevation, and roll angles that define the current orientation of the receiver coordinate frame with respect to the designated reference frame. The Euler angle coordinates that are output by the 3space as one measure of receiver orientation are graphically defined in figure 3. The x,y,z and X, Y, Z tri-axis arrays represent independent, three-dimensional orthogonal coordinate frames. The x,y,z triad represents the receiver frame in its current orientation state. The X,Y,Z triad represents the reference frame against which the relative orientation of the receiver frame is measured. By definition then, the X,Y,Z frame also represents the zero-orientation reference state of the receiver frame.
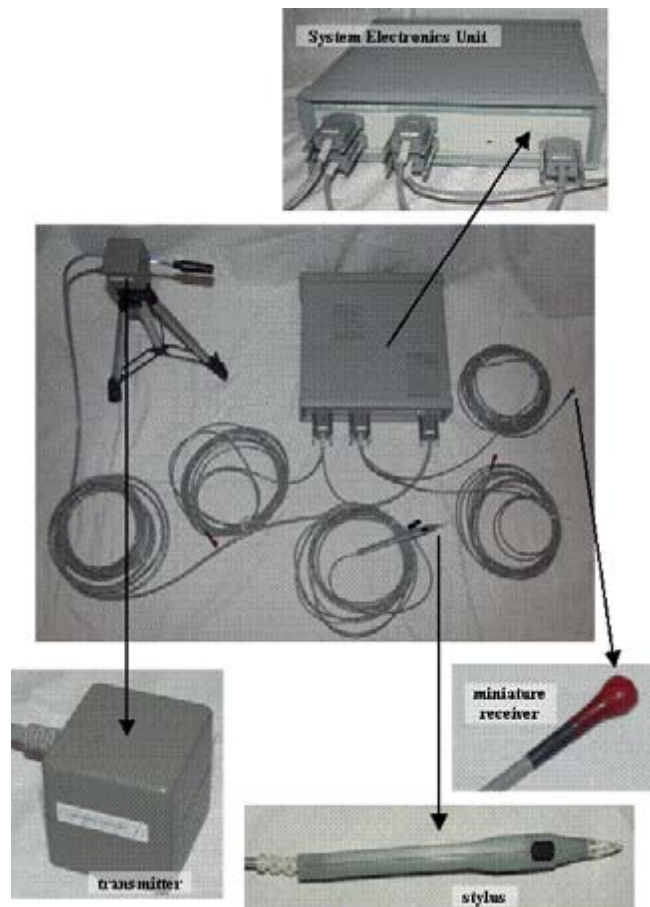

Fig. 2. Electromagnetic FASTRAK Digitizer form Polhemus

The 3D-space Euler angles, azimuth, elevation and roll, are designated y, q, and f in figure 3. These angles represent an azimuth-primary sequence of frame rotations that define the current orientation of the receiver with respect to its zero-orientation state. The defining rotation sequence is an azimuth rotation followed by an elevation rotation followed by a roll rotation.

The azimuth angle y is defined in figure 3 as a rotation of the X and Y reference axes about the Z reference axis. Note that the transition axes labeled X′ and Y′ represent the orientation of the X and Y axes after the azimuth rotation.

The elevation angle q is defined as a rotation of the Z reference axis and the X′ transition axis about the Y′ transition axis. Note that the transition axis labeled Z′ represents the orientation of the Z reference axis after the elevation rotation. Note also that the current x-axis of the current receiver frame represents the orientation of the X′ transition axis after the elevation rotation (fig. 3).

Lastly, the roll angle f is defined as a rotation of the Y′ and Z′ transition axes about the x-axis of the receiver frame. Note that the y and z-axes of the current receiver frame represent the orientation of the Y′ and Z′ transition axes after the roll rotation. The azimuth, elevation and roll rotations are positive, negative and positive respectively.

The resolution of the system is 0.0002 inches per inch of transmitter and receiver separation (0.025^0 orientation). Static accuracy is 0.03 inches RMS for the X, Y, or Z position (0.15^0 RMS for receiver orientation).

A four-parameter quantity representing a vector and a scalar.

The quaternion q = q0 + i q1 + j q2 + k q3 can be used to represent the receiver's orientation without the need for trigonometric functions. The attitude matrix output from the 3space can be equivalently represented by the following matrix using quaternion:

X Directional Cosines	Y Directional Cosines	Z Directional Cosines

$$\begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_3q_0+q_1q_2) & q_0^2-q_1^2+q_2^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_1q_0+q_3q_2) & q_0^2-q_1^2-q_2^2+q_3^2 \end{bmatrix}$$

Electrode locations may be determined by touching with a stylus. Head movement is automatically compensated with 1-3 head mounted sensors (fig. 4).
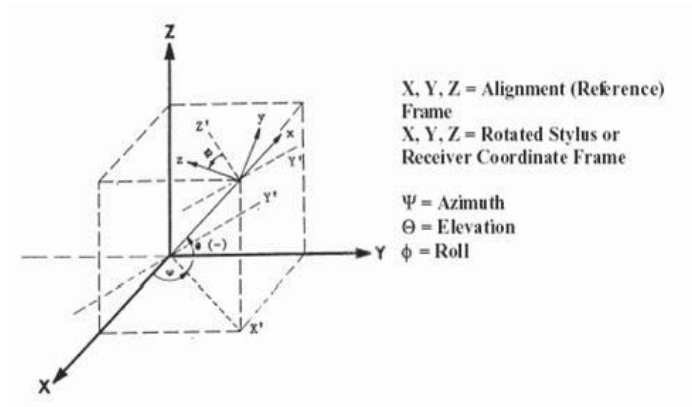

X, Y, Z = Alignment (Reference) Frame
X, Y, Z = Rotated Stylus or Receiver Coordinate Frame

Ψ = Azimuth
Θ = Elevation
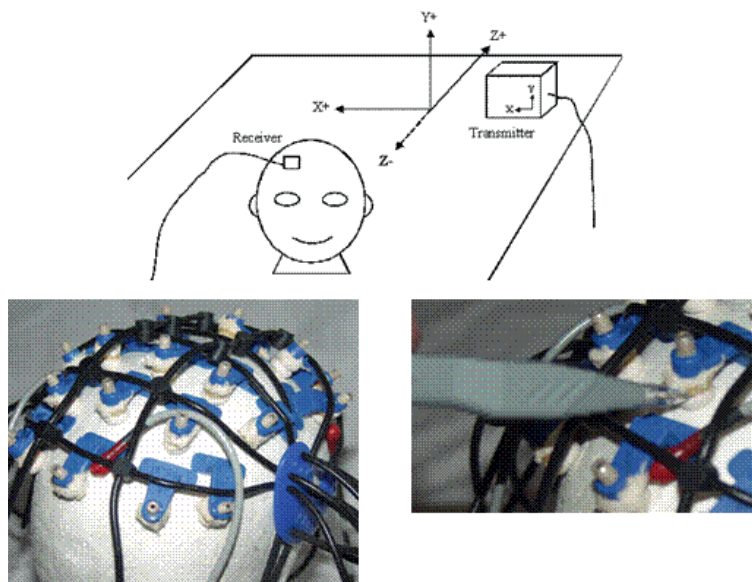φ = Roll

Fig. 3. 3D-space Euler angles

Fig. 4. General diagram for measurement of head point by digitization of EEG electrode (top) and real measurement condition (bottom)

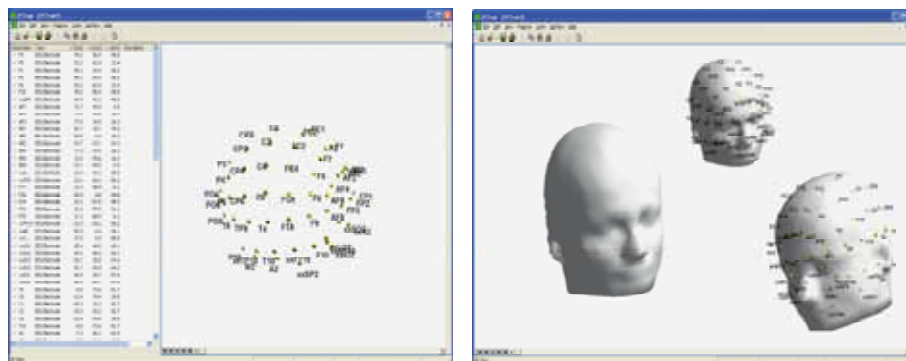**Fig. 5.** Receiving data from the FASTRAK using the GUI software

Fig. 6. Electrodes digitization in *EETrak* software

External head shape locations may be determined by "drawing" over the head surface with a stylus. Head movement is automatically compensated with 1-3 head mounted sensors. Location of the named points (such as nasion, inion, etc.) may be obtained and stored.

Setup functions for *Polhemus* digitizer are available through the software menu.

Interactive 3D display permits viewing of digitized points as they are added or offline. The manual editing function allows for error rejection.

The first three electrodes to be measured are the fiducial points (Left Ear, Right Ear, Nasion). These positions are required to transform the electrodes into the head coordinate system (x to the right ear, y to the nose, z to the top of the head).

Visualizing software

The software package attached to the FASTRAK digitizer is hardly user-friendly. Location parameters of the receivers – calculated in matrix equation – are presented summarily in the software window (fig. 5).

There exist, however, specialized software packages that allow for spatial visualization of the measurement results of EEG electrodes for neurophysiological applications.

Software packages compatible with FASTRAK digitizer include, among others, EETRAK by the Dutch company *ANT* and Locator by the American *SOURCE SIGNAL IMAGING*.

Electrode location in both these packages is calculated in Cartesian coordinates XYZ.

EETrack Software

In this software [8] digitization is executed on the electrodes placed in formerly prepared electrode files (*.elc). An electrode file can be connected with the data included in the head model file (*.vol). In EETRAK software visual presentation is much better than in Locator package. It is, however, much more difficult in it to prepare electrode files comprising any given number of electrodes. Electrode file is an ASCII file and can be edited manually.

Numerical values of XYZ coordinates of the digitized electrodes can be transferred by copy-paste function to appropriate software for digital analysis of EEG signal (fig. 6).

Locator Software

With *Locator* [9] a user can accomplish the following:
*   digitize the position of a set of electrodes in the Cartesian coordinate system defined by three user-selected fiducial points;
*   digitize a shape by registering the position of a series of points representing that shape in the Cartesian coordinate system defined by three user-selected fiducial points;
*   view the x, y, z coordinates of the fiducials and the digitized points (electrodes or shape points);
*   view the digitized electrode or shape points as a simulated 3D display on the screen, which can be zoomed in and out or rotated;
*   save the numerical values in an ASCII text file for further reference.

Locator offers the user a choice between EMSE or BESA coordinate systems (standard systems of EEG electrode location). With EMSE – both the electrodes and the head shape points can be digitized; in BESA – only electrodes.

The display cube provides a consistent frame of reference for the display of the data, as shown in Figure 7a. The cube is displayed with the front edge removed to allow for better viewing of the points inside it. Inside the cube, with origin at the center of the cube, are the three axes x, y, and z and the three fiducial points: Nasion (NZ), Left Preauricular (LPA), and Right Preauricular (RPA). The size, orientation, and location of the display can be modified with the Display Taskbar.

Figure 7b shows small, scattered dots representing the electrodes in the frame of the three fiducials which are displayed with larger dots on the X+, Y+, and Y – axes. Figure 7c shows a portion of the display cube considerably magnified with the Zoom In button to separate the labels for each electrode. The top left corner of the electrode label is always attached to the point showing the electrode location. A shape data display is shown by the figure 7d. A physical head model was digitized with digitization sequences of vertical and circular horizontal motions, and the resulting data are displayed above.
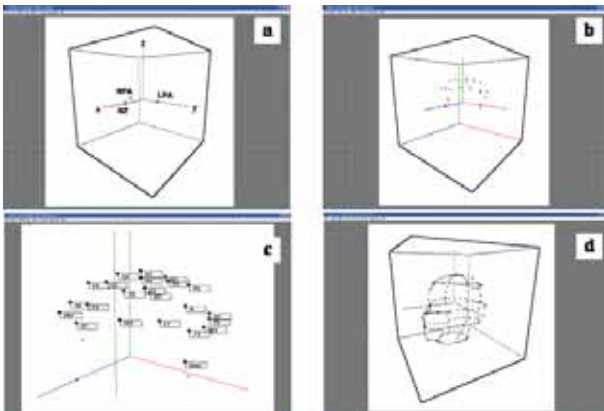

Fig. 7. Electrodes digitization in *Locator* software; explanation in text

In *Locator*, the method of preparing files with a non-standard set of electrodes is much better. A special editor of this kind of files is included in the software package (fig. 8).
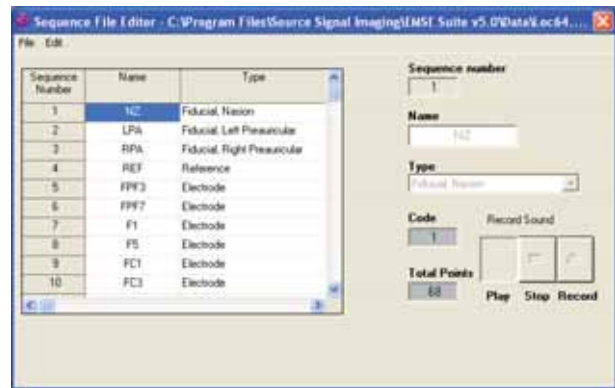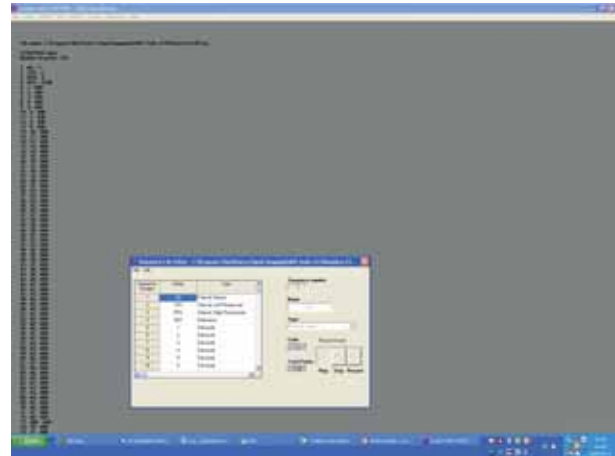

Fig. 8. Sequence file editor inside of the *Locator* software

Moreover, *Locator* software provides an additional function, unavailable in *EETrak* package. It allows for collecting the data on location of the whole series of electrodes, which facilitates drawing the outline of the head (<shape> function). The data concerning the outline of the object are stored in *.hsp files (fig. 9).

Another advantage of *Locator* software is the easy selection of the standards of coordinate systems. I the EMSE standard the X+ axis in directed towards the patient's nose and the Y+ axis – towards the left ear. In turn, in the BESA standard the X+ axis points at the right ear and the Y+ axis – at the nose. In both systems the Z+ axis points at the top of the head.

Example

The results of digitization of a matrix consisting of 64 electrodes (8x8) used in corticography and placed in the left temporal area are shown below (fig. 10) [10].

In *EETrak* software – a text processor (e.g. notepad) serves for preparation of a special file containing information on the number of electrodes (64), the selected unit of distance measurement (mm), electrode locations and their names.

Predefining of electrode location does not require any use of tabular or approximated values. Location of each electrode can be defined as [0, 0, 0]. In the digitization process each electrode will be assigned an adequate real value in accord with its actual location in the Cartesian space XYZ.
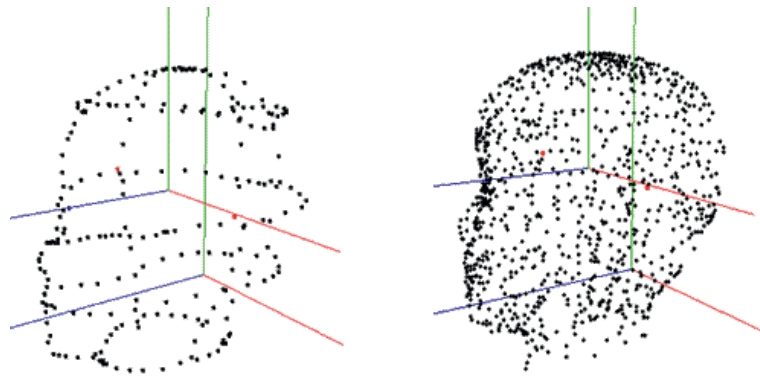
Fig. 9. Visualization of the outline of the head from *.hsp files constructed of, e.g., 235 (left) or 972 dots (right)
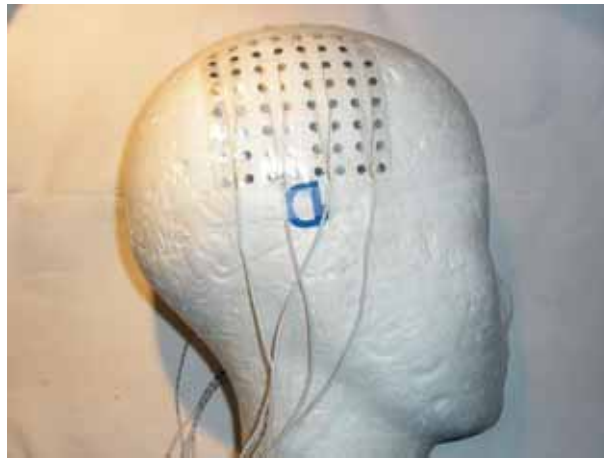

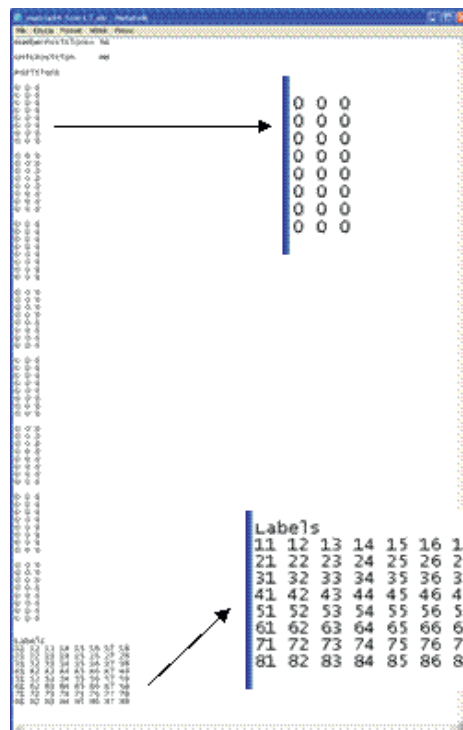**Fig. 10.** 64-point electrode matrix (8x8, space between the electrodes – 10 mm)

Fig. 11. Electrode file *.elc for *EETrak* software with the predefined electrodes of an 8x8 matrix

Since the electrodes of the digitized electrode matrix do not correspond with the electrodes of 10-20 system, they are not assigned letter names, but numerical ones, according to the rule (fig. 11):

in a row: 11, 12, …, 17, 18,

in a column: 11, 21, …, 71, 81.

The window of *EETrak* software shows the results in both the numerical form and the visualized one (fig. 12).

A special processor included in *Locator* software allows for preparation of a list of electrodes for digitization in a much simpler way than that in *EETraku* package. The prepared list of electrodes is stored in a *.seq file (fig. 13).

The procedure of digitization of electrodes itself is similar to that in *EETraku*. At the beginning the software requires marking for the „fixed" fiducial points, like nasion (one of the two major critical points to draw the fibular line, and both preauricular points (critical points for the coronal line).

Having digitized the fiducial points, the software "inquires" about the subsequent electrodes marked in the electrode file. Their digitization is executed after pointing at them with the end of the stylus and pressing its button.

Visualization of the digitized electrodes is presented below (fig. 14):

However, *Locator* window has slightly fewer options than that of *EETrak* software. The electrodes may be represented non-simultaneously in only one projection (in *EETrak* – up to 4 projections). The electrodes can also be rotated at three degrees of freedom, with additional zooming in/out options (fig. 15).

In *Locator* software, the access to the location values of the digitized electrodes is much more difficult (fig. 16).

Conclusions

An electromagnetic digitizer with software executing an analysis and visualization of the coordinates of electrode location is a valuable instrument for experimental neurophysiological studies. It allows for determining of a very precise location of the electrodes recording the EEG signal, which is particularly important for a precise analysis of the dynamics of the activity of bioelectric function of the brain.
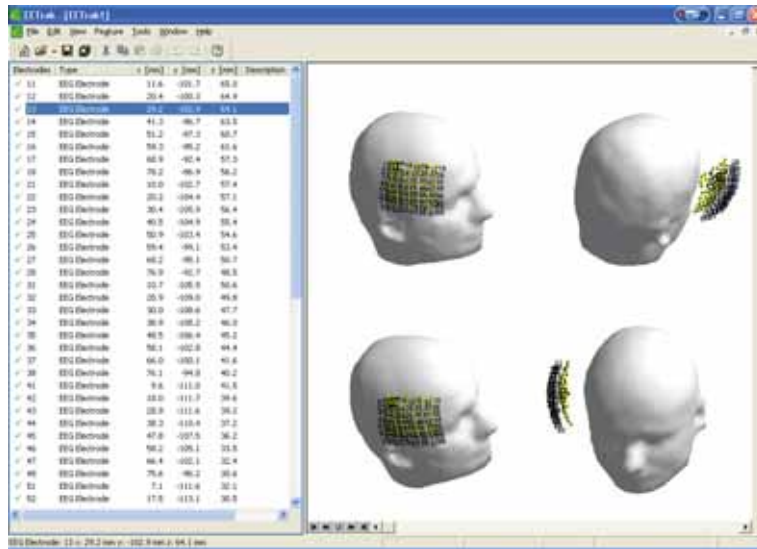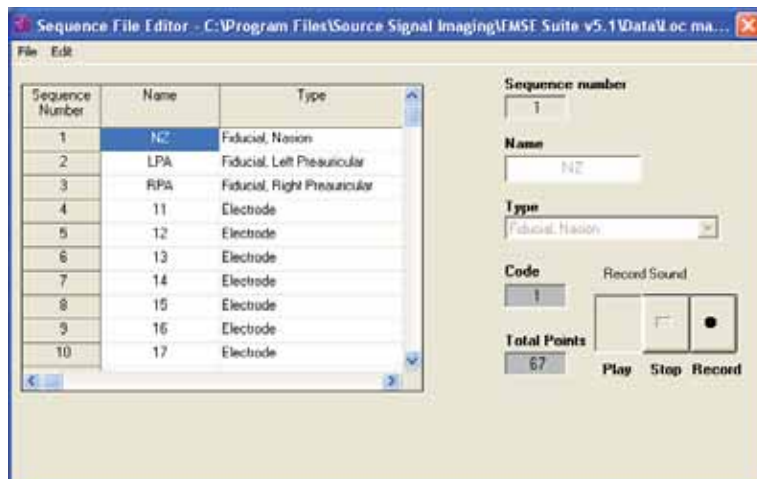

Fig. 12. Digitized 64-electrode matrix

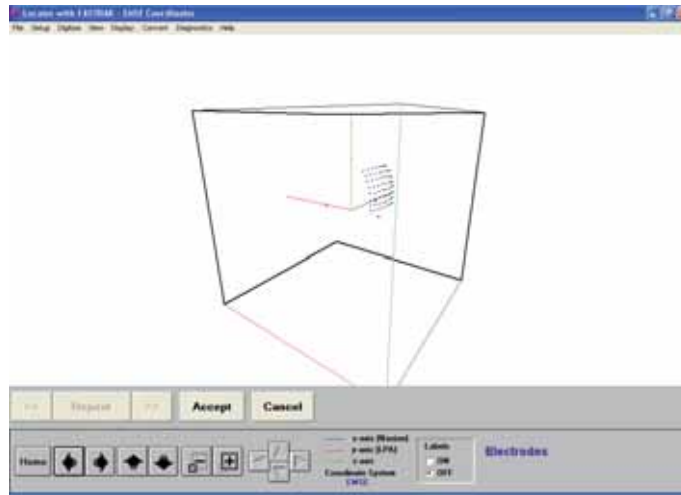Fig. 13. Electrode *.seq file processor for *Locator* software

Medical diagnostics

Electromagnetic digitizing system for precise location of EEG electrodes

Fig. 14. Digitized 64-electrode matrix in *Locator* software
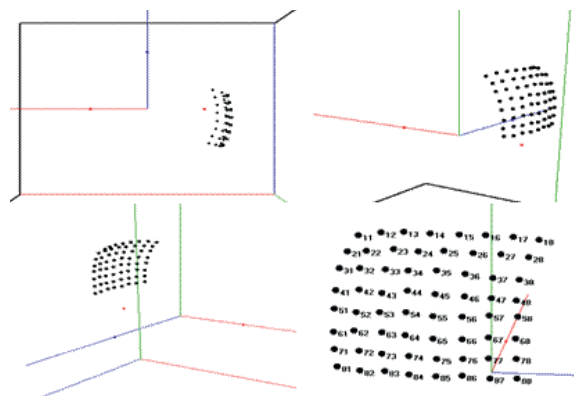

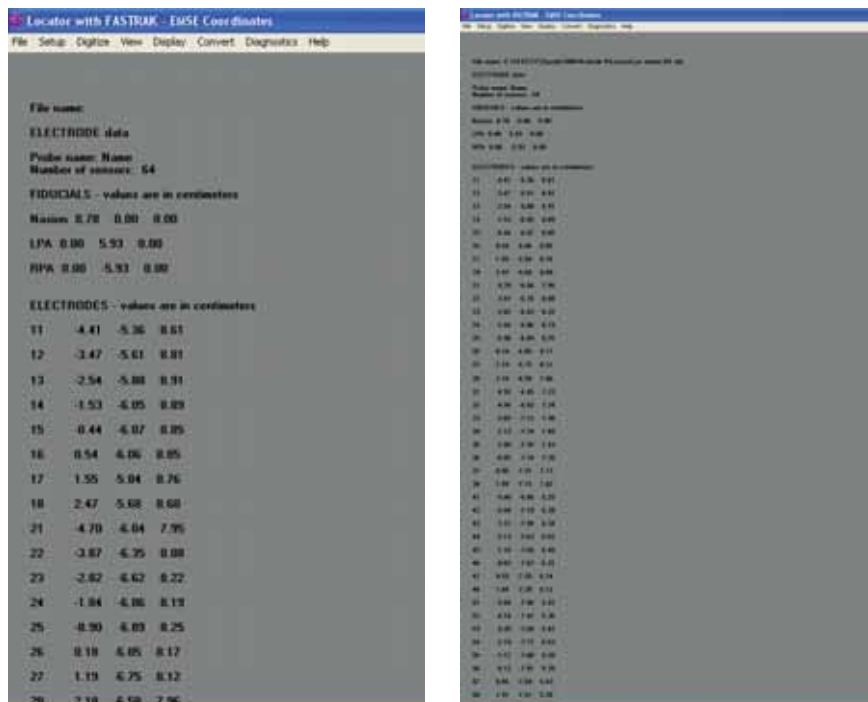**Fig. 15.** Presentation of a digitized 64-electrode matrix in different projections

Fig. 16. An *.elp file with the coordinates of the digitized electrodes

References

1. Zyss T. (red.) (2007), Zastosowanie układu 10-20 do rozmieszczania elektrod do EEG [Applying of 10-20 system to arranging the EEG electrodes]. Biblioteka ELMIKO. Krakowskie Wydawnictwo Medyczne. Kraków.
2. Zyss T. (2004), Wprowadzenie do techniki EEG dużych rozdzielczości [Introduction to high-resolution EEG technology]. Przegląd Elektrotechniczny, 12: 1205-1209.
3. Zyss T. (2004), Definiowanie położenia elektrod dla techniki EEG w standardowym układzie 10-20 [Defining of electrode location for EEG technique in the standard 10-20 system]. XI Krajowa Konferencja „Komputerowe wspomaganie badań naukowych". XI KK KOWBAN '2004 – Materiały. Wrocław, Polanica Zdrój, 81-86.
4. Zyss T. (2004), Rozszerzony układ 5-5 lokalizacji elektrod do EEG – podstawą techniki HR-EEG [Enhanced 5-5 system of electrode location as a basis for HR-EEG technique]. XI Krajowa Konferencja „Komputerowe wspomaganie badań naukowych". XI KK KOWBAN '2004 – Materiały. Wrocław, Polanica Zdrój, 87-92.
5. Zyss T. (2005), Problemy związane z topografią elektrod dla badań EEG – system ich digitalizacji [Problems connected with topography of electrodes for EEG examination – a system of their digitization]. Bio-Algorithms and Med-Systems, 1, 1/2: 325-332.
6. Zebris Medical GmbH: Lokalisation und Repositionierung von EEG-Elektroden 3D Digitalisierung der Kopfform. http://www.zebris.de/deutsch/pdf/d_Technische_Daten_ELGuide.pdf.
7. Polhemus Incorporated: 3space Fastrak user manual. http://www.polhemus.com/fastrak.html.
8. A.N.T. Software BV: EETrak Electrode Digitizer. Getting Started Guide. http://www.ant-neuro.com/products/eetrak.
9. Source Signal Imaging, Inc.: EMSE® Suite. User Manual version 5.0 Locator. http://www.sourcesignal.com/locator.html.
10. Ad-Tech Medical Instrument Corporation: Epilepsy & Neurosurgery Product Guide – Epilepsy / long term monitoring (LTM) subdural grid electrodes. http://www.adtechmedical.com/MKTG-3006-Rev-J.pdf.

Medical diagnostics

INSTRUCTION FOR AUTHORS

1. The original and one Photostat copy of the manuscript should be mailed to: Managing Editor Zdzisław Wisniowski. Authors are strongly urged to submit a CD containing the manuscript in Word for PCs format along with the hard copies. Articles already published or those under consideration for publication in other journals or periodicals should not be submitted.
2. Manuscripts should be in English. Usage of correct language is the responsibility of the author.
3. Submission of manuscripts to us implies that the copyright of the entire material except those already copyrighted by other publications like tables, figures etc and those permission the author(s) have already obtained, belongs to the publisher.

   PRESENTATION
4. The text may be prepared from a good quality laser printer or electronic typewriter. It is to be typed (one side only) on a regular sheet of paper A4 size preferred with single line spacing. New Times Roman font of 12 point size is recommended. Authors are requested to paste figures and tables (if possible) at appropriate places in the text. If this is not possible, they may be given separately, in which case their exact locations in the text may be marked by pencil.
5. The article should begin with a short ABSTRACT followed by a short INTRODUCTION. The rest of the article ma be titled and arranged as per the wishes of the author.
6. REFERENCES should continue the last section of the article. In the text, references to other papers or books should be cited using consecutive numbers in parenthesis (e.g. [1,2]) and they should be listed numerically in the last section.
   Examples:
   I. Pyrczak W., Sarapata K.: Instructions for authors: references. Bio-Alghoritms and Med-Systems, 1, 386, 2005
   II. Tadeusiewicz R., Ogiela M.R.: Medical Image Understanding Technology. Springer Verlag, Heidelberg 2004
   Authors are to make sure that each reference appearing in the text appears in the list of references at the end and vice versa.
7. TITLE PAGE AND FIRST PAGE – The title page of the manuscript, the name(s) of author(s) and institution(s) where the work has been carried out should be typed out on a separate sheet of paper.
8. Section headings (ABSTRACT, INTRODUCTION etc.) are to be typed in upper case letters and placed on a separate line
1. Scientific names should be typed in italics or underlined
2. Tables should be numbered consecutively using Arabic numerals.
3. Illustrations should be limited to materials essential for the text. All figures should be numbered consecutively and should be submitted as sharp, high-quality prints. Figures that are to appear as a group should be photographed and mounted together. All figures and groups of figures should be trimmed at right angles and pasted directly on the typescript in the appropriate places. They should be of a size permitting photographic reduction by 25% together with text (not more than 15 cm across for page width and not higher than 20 cm). Legends to figures should appear directly below the respective figures.
4. In line drawings, all lines should be of uniform thickness; letters and numbers should be of professional quality and proper dimensions, approximately 1mm high when reproduced.
5. RUNNING TITLE – Please make sure to mention the running title (short title) of the article separately in the first page where the title, names of authors, their affiliations, etc. are given.
6. PREFFERED LENGTH – The length of the manuscript may be limited to 20-25 A4 size pages (20 cm width – 26 cms height) through this is not binding.
7. PERMISSION – Authors must obtain permission from other publishers if they use already published and copywrighted materials.
8. COLOUR FIGURES – Cost of printing color figures will be charged to the author at the rate of U.S. $ 125 per plate of four color pages.
9. REPRINTS – The order for reprints should be placed at the time of submission of the article itself, so that it may be out for printing without loss of time, if found suitable for order for not less than 100 reprints is to be made.

MANUSCRIPT SUBMISSION:

1. Electronic – the www.bams.cm-uj.krakow.pl address is available for all authors to submit manuscript. The instruction is available at each step of submission process.
2. The pdf or doc file with figures and tables incorporated into the manuscript body can be sent by email – address: bams@cm-uj.krakow.pl