Analysis of Heterogeneous Biomedical Data

Habilitation Thesis

Czech Technical University in Prague

Daniel Novák

23rd of February 2016

Acknowledgements

This thesis would not have been written without the help of number of people. Sincere thanks to:

Olga Štěpánková, from the Czech Technical University who contributed with many valuable remarks and comment observations in the area of biomedical engineering and who also helped to correct and improve the thesis.

Vladimír Eck and Lenka Lhotská from the Czech Technical University who initiated my interests at biomedical engineering and medical data analysis.

David Cuesta from the Technical University of Valencia, for his contribution, many constructive inputs, invaluable help in the field of signal processing and long lasting friendship.

Tomáš Sieger, Eduard Bakštein and Václav Křemen for their valuable recommendations and helpful suggestions.

Filip Španiel, Marek Kuzma, Martin Haluzík and Robert Jech for giving me a chance to intensively cooperate on challenging biomedical problems.

Andres Orozco-Duque, Juan Pablo Ugarte and Matthew Harris for beneficial discussions.

My doctoral students for helping me in biomedical data processing and analysis.

My family for bearing up with my long working hours.

Contents

Preface

I am presenting this thesis to qualify for habilitation at the Czech Technical University in Prague. The thesis is a review of my biomedical engineering research since 2005. The research work has been driven mainly by cooperation with medical partners in the field of neurology, internal medicine, cardiology and psychiatry.

My interest in biomedical engineering research has been influenced more than 15 years ago mainly by the three activities: i)long term stay at the Universitat Politecnica de Valencia and work with Dr. David Cuesta Frau ii) collaboration with group of prof. Bustamente and his Ph.D. student Andres Orozco from Medellin University of Colombia and iii)post-doc stay in Philips Research Laboratories in Aachen. Initially, with Dr. David Cuesta Frau we explored possibilities of speech recognition techniques for electrocardiogram processing. This direction was thereafter followed by processing of surface and internal electrocardiogram signal in application of atrial fibrillation analysis.

The outcomes contributed to my master thesis, PhD thesis and are summarized in several papers and in a world patent.

My recent research work has stemmed from mutual cooperation with researchers at several Czech medical institutes and labs. In 2005, after my post-doc studies, I co-initiated cooperation with the Department of Neurology and Center of Clinical Neuroscience, the 1st Faculty of Medicine and the 3rd Department of Medicine, both at Charles University, and Mental Institute of Health (previous Prague Psychiatric Center). Together we worked on several local and international projects. I am positive that the results of my research have a great practical outcome and increase the knowledge of designing appropriate protocol for medical data and clinical trials management.

In neurology field, I established a small research group of Computational Neuroscience. We have intensively cooperated with Prof. Robert Jech on microelectrode data originating from deep brain stimulation procedure.

In internal medicine, my activities were focused on processing of glycemic profile data and metabolomics analysis in close cooperation with Prof. Martin Haluzik.

In psychiatry field, I contributed in data processing in Itareps system which was the first telemedicine system for relapse prediction. Here, I have been closely working with Dr. Filip Spaniel.

I am also one of the founders and lecturers of a new Neuroinformatics course taught in the newly established study programme in Biomedical engineering and informatics.

The thesis is organized as an annotated collection of 15 papers assorted from among my publications since the year 2005. Nine papers were published in impacted scientific journals (the impact factors of 9.8, 6.3, 4.1, 3x3.8, 3.4, 2.8 and

1.8), one paper made a chapter in an international book, one contribution is a world patent and the other four articles appeared in conference proceedings. As I preferred consistency in topics to completeness of my research profile, I did not include my earlier or parallel works on other than biomedical research topics as well as the conference articles that reasonably overlap with their later journal extensions. The full list of my publications can be accessed at https://www.researchgate.net/profile/Daniel_Novak2.

The thesis is split in five chapters.. It starts with a self-contained overview of the remaining chapters containing the individual papers. Chapter 2 summarizes the research on complexity measures in atrial fibrillation. Chapter 3 gives an overview of processing of single-neuron recordings. Chapter 4 concerns with metabolomics and statistical data analysis in diabetes mellitus. Chapter 5 provides the book chapter and the journal paper that reviews the research on time event series analysis in schizophrenia and bipolar disorder.

Chapter 1

Habilitation Thesis Overview

Analysis and processing of biomedical data is placed at the intersection of engineering, the life sciences and healthcare. While the man is a more complex system than even the most elaborated machine, many of the same approaches that focus on programming a machine or algorithm can be applied to therapeutic gadgets, biological structures and diagnostic protocols. This thesis deals with several case studies of biomedical data processing elaborated on both clinical basic research and daily clinical practice. My particular contribution is summarized below each paper in italic font.

Analysis of nonlinear measures and multifractal tools in atrial fibrillation (Chapter 1)

Atrial fibrillation (AF) is the most commonly clinically-encountered arrhythmia. Catheter ablation of AF is mainly based on trigger elimination and modification of the AF substrate. Substrate mapping ablation of complex fractionated atrial electrograms (CFAEs) has emerged to be a promising technique. To improve substrate mapping based on CFAE analysis, automatic detection algorithms need to be developed in order to simplify and accelerate the ablation procedures. The nature of CFAE is generally nonlinear and nonstationary, so the use of complexity measures is considered to be the appropriate technique for the analysis of AF records.

In the four papers selected for Chapter 1, measures from the theory of nonlinear dynamics as sample entropy, approximate entropy and tools based on fractal theory were applied for the characterization of the level of fractionation. I and my co-authors demonstrated that complexity measures and its optimized parameters can be proposed as a tool for grade fractionation associated with the detection of target sites for ablation in AF. Furthermore, we have defined a new measure that improve the extraction of information about the shape of the multifractal spectrum. Our results indicate that the new parameter is capable of grading fractionation better than other fractal and multifractal indexes. Last contribution is a world pattern describing an approach consisting of three particular methods for atrial fibrillation detection.

The results presented in this chapter were reached in close collaboration with Dr. David Cuesta Frau of Politecnic University of Valencia. David has focused at time health series analysis using nonlinear dynamic measures. The second closest collaborator was Andreas Orozco from Universidad Pontificia Bolivariana, Medellín, Columbia, former Ph.D. student of Prof. John Bustamente, now lecturer in Instituto Tecnologico Metropolitano, Medellin. He developed multifractal framework, and extracted new parameters from multifractal spectrum. Further collaborator was Eva Cirugeda, Ph.D. student of Dr. David Cuesta, who worked on optimization of complexity measures parameters. The data were registered in two clinics: Institute for Clinical and Experimental Medicine, Czech Republic (provided by Dr. Dan Wichterle) and Staedtisches

Klinikum Karlsruhe, Germany (provided by Dr. Claus Schmitt). Last but not least, all time I was tightly working with my colleague Dr. Václav Křemen from Department of Cybernetics, Czech Technical University in Prague. Finally, my research in atrial fibrillation started during post-docs stay at Philips Research laboratories in Aachen, Germany. I was closely cooperating with Dr. Matthew Harris, Dr. Ralf Schmidt and Dr. Michael Perkuhn. The data were registered at University Hospital, RWTH Aachen, Germany and managed by Moritz Arndt.

In particular, the chapter contains the following papers:

- Eva M. Cirugeda–Roldan, D. Novak, V. Kremen, D. Cuesta–Frau, M.W. Keller, C. Schilling, O. Doessel, C. Schmitt, A. Luik, Characterization of Complex Fractionated Atrial Electrograms by Sample Entropy: An International Multi–Center Study, Entropy, 17(11), p.7493-7509, 2015
  *Contribution: **25%,** involvement of parameterization scheme, multi-centre study approach*

- Andrés Orozco-Duque, Daniel Novak, Vaclav Kremen and John Bustamante, Multifractal analysis for grading complex fractionated electrograms in atrial fibrillation, Physiological Measurement, Physiological Measurement, 36(11), p.2269-84, 2015
  *Contribution: **50%,** paper design, DFA and wavelet method analysis, multifractal spectra parameterization*

- Juan Pablo Ugarte Macías; Andrés Orozco-Duque; Catalina Tobón; Vaclav Kremen; Daniel Novak; Javier Saiz; Tobias Oesterlein; Clauss Schmitt; Armin Luik; John Bustamante, Dynamic approximate entropy electroanatomic maps detect rotors in a simulated atrial fibrillation model, PLOS One, 9(12), 2014
  *Contribution: **25%,** mapping complexity measures to model of atrial fibrillation, comparison to standard fractal indexes*

- D. Novak, V. Kremen, D. Cuesta, K. Schmidt, V. Chudacek, L. Lhotska, Discrimination of endocardial electrogram disorganization using a signal regularity, 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'09), p. 1812 – 1815, 2009
  *Contribution: **70%**, paper design, developing idea of complexity measures for analysis of atrial fibrillation fragmentation*

- Schmidt R., Harris M., Novak D., Perkuhn M., Atrial Fibrillation Detection, Philips Research Aachen, Germany, world patent no. WO 2008/007236 A2,2008
  *Contribution: **60%**, atrial fibrillation detection using QRTS cancellation and P wave detection*

Analysis of single-neuron recordings in Parkinson desease (Chapter 2)

Parkinson's disease (PD) is the second most common neurodegenerative disease. Besides motor symptoms, which include movement slowness, muscle rigidity, and resting tremor, PD patients manifest also cognitive, affective, sleep, and autonomous disturbances. Deep brain stimulation (DBS) is a surgical treatment

of diseases of the central nervous system, which include mainly movement disorders (such as Parkinson's disease, tremor, and dystonia), chronic pain, and more recently also some psychiatric disorders. The impairment in motor function is due to the progressive loss of dopaminergic neurons in the substantia nigra, which affects the function of the basal ganglia, composed of multiple subcortical nuclei situated deeply inside the forebrain.

This neuroscientific direction focuses at defining the role of the basal ganglia in the human brain, particularly with focus on understanding the function of the basal ganglia in Parkinson's disease patients treated with deep brain stimulation. New insights explored would help not only to optimize the treatment of neurodegenerative diseases, but in general to contribute to the understanding of the human brain. Analysis was performed on single-neuron microelectrode recordings acquired during stereotactic surgery when the stimulation electrodes of DBS system were implanted.

Specifically, the goal was to find answers to important neuroscientific questions regarding the functioning of the human brain in general, and the mechanism of the deep brain stimulation in particular. While our team has been the first to report the finding of basal ganglia neurons being connected to the human oculomotor system, we have also found basal ganglia neurons related to emotional and visuoattentional processing.

The results can be summarized as follows. Considering eye movement-related neurons, about 20% of the basal ganglia neurons studied were found to be related to planning, execution and/or control of eye movements. Regarding Emotion-Related Neurons, 15% of the basal ganglia neurons were identified to be related to processing or responding to distinct types of emotional stimuli**.** One half of the basal ganglia neurons studied were revealed to be activated during a visuo-attentional task.

Additionally, we coped with significant problem of proper classification of action potentials from extracellular recordings which is essential for making an accurate study of neuronal behavior and its important step before proceeding to any further high-level analysis. Even if the activity of several neurons is recorded with only a single electrode, spike sorting allows the researcher to measure the activity of the individual neurons separately. Three widely-used publicly-available spike sorting algorithms based on unsupervised learning were compared with regard to their parameter settings. An optimization technique based on Adjusted Mutual Information was employed to find near-optimal parameter settings for a given artificial signal and algorithm. As a conclusion, all three algorithms performed significantly better with optimized parameters than with the default ones. An important part of the sorting is determining the number of constituent clusters which best describe the data. I have contributed to a method for automatic unsupervised determination of the number of neurons in extra-cellular recording. The proposed methodology has an advantage of setting the minimum number of parameters and is very robust to background noise.

I established a small research group of Computational Neuroscience (http://neuro.felk.cvut.cz/) which covers efforts of the above mention research

activities. There is one more significant output related to this topic. Our neuron group released two software packages: (i) sigInspect, a graphical user interface application for Matlab, developed for inspection and annotation of extracellular microelectrode recordings and (ii) idendro, a new R package enabling to inspect dendrograms, resulting from hierarchical cluster analysis. This software tools are publicly available to neuroscience community and are suitable for data visualization.

The results presented in this chapter were reached in collaboration with medical partners and other members of Computational Neuroscience group. The research work was driven by prof. Robert Jech from Dept. of Neurology and Center of Clinical Neuroscience, Charles University in Prague. Data were provided by Dr. Milan Urgošík from Dept. of Stereotactic and Radiation Neurosurgery, Na Homolce Hospital, Prague, Czech Republic. Dr. Tomáš Sieger elaborated the topic of eye-movement and emotional neurons. My Ph.D. student Jiří Wild explored spike sorting comparison. My Ph.D. student Eduard Bakštein contributed to annotation tool sigInspect. Finally, my Ph.D. student Pavel Vostatek worked on single-neuron visualization techniques.

Specifically, the chapter consists of the following papers:

- Tomas Sieger, Tereza Serranova, Pavel Vostatek, Jiri Wild, Daniela Stastna, Cecilia Bonnet, Daniel Novak, Evzen Ruzicka, Dusan Urgosik, Robert Jech, Distinct Populations of Neurons Respond to Emotional Valence and Arousal in the Human Subthalamic Nucleus, Proceedings of the National Academy of Sciences (PNAS), 112(10), p.3116-3121, 2015
  *Contribution: **15%**, visualization of the STN model*

- Sieger T, Bonnet C, Serranová T, Wild J, Novák D, Růžička F, Urgošík D, Růžička E, Gaymard B, Jech R., Basal Ganglia Neuronal Activity during Scanning Eye Movements in Parkinson's Disease, PlosOne, 8(11), 2013
  *Contribution: **10%**, time relation between electrooculography and neuronal activity*

- J.Wild, Z.Prekopcsak, T.Sieger, D.Novak, R.Jech, Performance comparison of spike sorting algorithms for single-channel recordings, Journal of Neuroscience Methods, 203(2), p.369-76, 2012
  *Contribution: **30%,** state of the art, citation histograms, paper design, unsupervised algorithm analysis*

- D. Novak, J. Wild, T.Sieger and R.Jech, Identifying Number of Neurons in Extracellular Recording, 4th International IEEE EMBS Conference on Neural Engineering, Antalya, Turkey, p.742-745, 2009
  *Contribution: **80%,** paper design, penalty measures, comparison with sorting algorithms*

Metabolomic and statistical data analysis in diabetes mellitus (Chapter 3)

Diabetes mellitus is a chronic metabolic disorder characterized by increased glucose levels due to insulin secretion deficiency and decreased sensitivity of the peripheral tissues to insulin effects, which can ultimately lead to the

development of acute (hypoglycemia, hyperglycemia) and longterm complications (cardiovascular, cerebrovascular, renal, ophthalmic).

This chapter offers several approaches to analysis of longterm complications in diabetes mellitus. Our early work joint with prof. Martin Haluzík and Dr. Miloš Mráz resulted in design of a telemedicine system for Diabetes Mellitus compensation under EU FP7 project OLDES "Older people's e-services at home". A small pilot focusing on elderly patients suffering from Type 2 was carried out. Time series as three point glycemic profile, blood pressure measurement and caloric in-take and out-take were recorded and consequently analysed. We successfully demonstrated the feasibility of daily use of a low-cost home access point that makes a range of services available, supports social interaction and efficiently man-ages some problems related to diabetes mellitus.

Subsequent work with the same team and Jan Balham from Department of Anesthesia, Resuscitation and Intensive Medicine concentrated on tight glucose control during elective cardiac surgery. A total of 2383 subjects were randomized into the trial between January 2007 and December 2010, 1134 in the perioperative group and 1249 in the postoperative group. To sum up, perioperative initiation of intensive insulin therapy during cardiac surgery reduced postoperative morbidity in nondiabetic patients while having a minimal effect in diabetic subjects. During mutual cooperation, I gained valuable experience at dealing with major clinical trial and data work-flow. Ph.D. student Jiří Anýž, supervised by prof. Olga Štěpánková, elaborated statistical analysis of patients data cohort.

Recent works with Helena Pelantová and Marek Kuzma from Institute of Microbiology, Academy of Sciences of the Czech Republic, deals with metabolic studies in Type 2 diabetes mellitus. Since metabolic pathways throughout the body are mutually interconnected, the course of diabetes is frequently coupled with visceral fat accumulation and obesity resulting in an increased risk of simultaneous cardiovascular diseases. Metabolomics can not only help to understand the biochemical signature of diabetes and concomitant problems but can also contribute to early diagnostics development. We used nuclear magnetic resonance-based metabolomics of urine of mice-a model of obesity and insulin resistance-to identify novel metabolites associated with obesity phenotype. I cooperated again with Jiří Anýž on statistical analysis of selected metabolites.

This topic is covered by the following papers:

- Jan Balham, Miloš Mráz, Petr Kopecký, Martin Stříteský, Michal Lipš, Michal Matias, Jan Kunstýř, Michal Pořízka, Tomáš Kotulák, Ivana Kolníková, Barbara Šimanovská, Mykhaylo Zakharchenko, Jan Rulíšek, Jiří Anýž, Daniel Novák, Jaroslav Lindner, Roman Hovorka, Štěpán Svačina, Martin Haluzík, Perioperative tight glucose control reduces postoperative adverse events in non-diabetic cardiac surgery patients, The Journal of Clinical Endocrinology & Metabolism, 100(8), p.3081-9, 2015 *Contribution:* ***10%,*** *extraction of features from glycemic profile data, statistical analysis*

- Pelantová, Helena; Bártová, Simona; Anýž, Jiří; Holubová, Martina; Železna, Blanka; Maletínská, Lenka; Novák, Daniel; Lacinová, Zdenka; Šulc, Miroslav; Haluzík, Martin; Kuzma, Marek, Metabolomic profiling of urinary changes in mice with monosodium glutamate-induced Obesity, Analytical and Bioanalytical Chemistry, p.1-12, 2015
  *Contribution: **15%**, NMR spectra processing, statistical analysis*

- Helena Pelantová, Martina Bugáňová, Jiří Anýž, Blanka Železná, Lenka Maletínská, Daniel Novák, Martin Haluzík, Marek Kuzma, Strategy for NMR metabolomic analysis of urine in obesity mouse models – from sample collection to interpretation of acquired data, Journal of Pharmaceutical and Biomedical Analysis, 115, p.225-35, 2015
  *Contribution: **15%**, NMR spectra processing, statistical analysis*

- Daniel Novák, Olga Štepánková, et al, Does IT Bring Hope for Wellbeing?, Handbook of Research on ICTs for Healthcare and Social Services: Developments and Applications, IGI Global Inc., p.270-302, 2013
  *Contribution: **70%**, paper design, glycemic data visualization and processing, user-centered design*

- D. Novak, M. Uller, S. Rousseaux, M. Mraz, J. Smrz, O. Stepankova, M. Haluzik, M. Busuoli, Diabetes management in OLDES project, 31st Annual International Conference of the IEEE 0.12 Engineering in Medicine and Biology Society, p 7228 – 7231, 2009
  *Contribution: **70%**, paper design, glycemic data visualization and processing, diabetes pilot-proof of concept design*

Analysis of event and actigraph time series in schizophrenia and bipolar disorder (Chapter 4)

Psychiatric disorders, including depression, schizophrenia, and bipolar disorder, affect millions of people around the world. Schizophrenia is a major psychotic disorder that has devastating effects on the lives of patients and their caregivers. The illness is accompanied by high rate of relapses and readmissions. Bipolar disorder is an episodic and recurrent condition that is frequently disabling daily life. Bipolar Disorder is associated with an increased suicide risk.

This chapter deals with analysis of data acquired from two telemedicine system for relapse prediction in psychiatric disorders.

The first version of the system ITAREPS (Information Technology Aided Relapse Prevention Programme in Schizophrenia) was developed previously at the Prague Psychiatric Centre by Dr. Filip Španiel, who is now in the National Institute of Mental Health. It represents a step towards a highly customizable prodromal signs monitoring M-Health service platform. The components of the ITAREPS programme include monitoring of non-specific and specific prodromal symptoms, frequent assessments and an involvement of both patients and their caregivers. Crisis interventions such as increasing the dose of antipsychotic medication, based on early detection within a relapse prevention programme, reduced relapse and readmission rates compared to a treatment-as-usual group. My role was processing of event time series data. There was a statistically

significant 60% decrease in the number of hospitalizations in the ITAREPS, compared to the same time period before the ITAREPS entry.

The second programme focuses on the early identification of prodromal symptoms of mania or depression. Two types of data were monitored: objective data represented by movement activity recordings by use of actigraphy and subjective data acquired by questionnaire. I analysed both sources of data; in actigraphy recording I extracted several sleep parameters from actigraphy recording and performed trend analysis. I showed that the framework helps to predict a relaps early enough to suggest a change in medication and improve patient quality of life achieving sensitivity and specificity of 70%. This project is also driven by Dr. Filip Španiel.

The chapter includes one application paper and one conference paper:

- D. Novak, F.Spaniel, F.Albert, Analysis of Actigraph Parameters for Relapse Prediction in Bipolar Disorder: A Feasibility Study, 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,p.4972–4975,2014
  *Contribution: **70%**, paper design, features extraction, time event-series data analysis*

- Spaniel F, Vohlídka P, Hrdlicka J, Kožený J, Novák T, Motlová L, Cermák J, Bednarík J, Novák D, Höschl C., ITAREPS: Information Technology Aided Relapse Prevention Programme in Schizophrenia, Schizophrenia Research,98(1-3),2008
  *Contribution: **20%**, paper design, features extraction, time event-series data analysis*

Conclusions

The major contribution of the thesis to the actual state of biomedical data processing are applications to real-world clinical cases. Particularly, the outcomes of my work in each topic were summarized below each paper in italic font. To conclude, I contributed to the detection, analysis and modelling of atrial fibrillation, to spike sorting and statistical processing of single-neuron unit, to metabolomic analysis and statistical processing of blood glycemia time series and to time-event series analysis in bipolar and schizophrenia disorder.

Chapter 2

Analysis of nonlinear measures and multifractal tools in atrial fibrillation

Entropy **2015**, *17*, 7493-7509; doi:10.3390/e17117493

OPEN ACCESS
entropy
ISSN 1099-4300
www.mdpi.com/journal/entropy

Article

Characterization of Complex Fractionated Atrial Electrograms by Sample Entropy: An International Multi-Center Study

Eva Cirugeda–Roldán [1], Daniel Novak [2], Vaclav Kremen [3], David Cuesta–Frau [1,*], Matthias Keller [4], Armin Luik [5] and Martina Srutova [2]

[1] Technological Institute of Informatics, Polytechnic University of Valencia, Alcoi Campus, Plaza Ferrándiz y Carbonell, 2, Alcoi 03801, Spain; E-Mail: evaci075@gmail.com
[2] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague 121 35, Czech Republic; E-Mails: xnovakd1@labe.felk.cvut.cz (D.N.); srutova.martina@seznam.cz (M.S.)
[3] Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Prague 166 36, Czech Republic; E-Mail: vaclav.kremen@gmail.com
[4] Institute of Biomedical Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe 76128, Germany; E-Mail: matthias.w.keller@gmail.com
[5] Medizinische Klinik IV, Staedtisches Klinikum Karlsruhe, Karlsruhe 76133, Germany; E-Mail: armin.luik@klinikum-karlsruhe.de

* Author to whom correspondence should be addressed; E-Mail: dcuesta@disca.upv.es; Tel.:+34-651028565; Fax:+34-963877579.

Academic Editors: J. A. Tenreiro Machado and António M. Lopes

Received: 11 September 2015 / Accepted: 22 October 2015 / Published: 28 October 2015

Abstract: Atrial fibrillation (AF) is the most commonly clinically-encountered arrhythmia. Catheter ablation of AF is mainly based on trigger elimination and modification of the AF substrate. Substrate mapping ablation of complex fractionated atrial electrograms (CFAEs) has emerged to be a promising technique. To improve substrate mapping based on CFAE analysis, automatic detection algorithms need to be developed in order to simplify and accelerate the ablation procedures. According to the latest studies, the level of fractionation has been shown to be promisingly well estimated from CFAE measured during radio frequency (RF) ablation of AF. The nature of CFAE is generally nonlinear and nonstationary, so the use of complexity measures is considered to be the appropriate technique for the analysis of AF records. This work proposes the use of sample entropy (SampEn), not only as a way to discern between non-fractionated and fractionated atrial electrograms (A-EGM),

but also as a tool for characterizing the degree of A-EGM regularity, which is linked to changes in the AF substrate and to heart tissue damage. The use of SampEn combined with a blind parameter estimation optimization process enables the classification between CFAE and non-CFAE with statistical significance ($p < 0.001$), 0.89 area under the ROC, 86% specificity and 77% sensitivity over a mixed database of A-EGM combined from two independent CFAE signal databases, recorded during RF ablation of AF in two EU countries (542 signals in total). On the basis of the results obtained in this study, it can be suggested that the use of SampEn is suitable for real-time support during navigation of RF ablation of AF, as only 1.5 seconds of signal segments need to be analyzed.

Keywords: atrial fibrillation; catheter ablation; complex fractionated atrial electrograms; sample entropy; signal classification

1. Introduction

Atrial fibrillation (AF) is the most commonly clinically-encountered arrhythmia. For paroxysmal AF, on the one hand, a focal trigger inside the pulmonary veins (PV) has been shown to be the dominant mechanism for its initiation [1]. The technique of isolating PV by catheter ablation has therefore been established [2]. On the other hand, persistent AF is more related to the maintenance of the mechanisms rather than to the trigger [3–5]. It can be shown that the AF substrate of patients with persistent AF has a shorter cycle length and a higher degree of disorganized activity than that of patients with paroxysmal AF [6]. Nademanee [7] was the first to propose these areas with high frequency disorganized activity as ablation target sites and was able to achieve a respectable success rate. However, these results could not be reproduced by any other group, even in a combined approach of antral PV isolation and complex fractionated atrial electrogram (CFAE) ablation [8]. Still, there are several limitations in CFAE ablation. First, the interpretation of CFAE varies due to the absence of objective detection algorithms. Second, the role of CFAE in the initiation and continuation of persistent AF is still unclear. In addition, the genesis of CFAE may be affected by anatomical conditions or by other temporary effects. Algorithms to identify the dominant frequency and the mean cycle length of the local activity (CFEmean, St. Jude Medical, Minneapolis, MN, USA) have been developed in order to improve the categorization of the AF substrate.

The process for identifying these sites has been studied in recent years, and several methods, measures and approaches have been described. Nevertheless, they do not satisfy the sensitivity and specificity requirements. Kumagai [9] described different patterns of activation, and others have shown and quantified different characteristics of atrial electrograms (A-EGM) in both the time and frequency domains [10–13]. Some of the cited algorithms are implemented into commercially available mapping systems, so they can be used for daily AF ablation routines. However, their performance is not optimal, and they often require initial parameter manual setting. This manual configuration makes them user dependent and, thus, subjective. To overcome this problem, several new approaches and methodologies have been developed so as to measure and quantify the complexity of A-EGM. Kremen *et al.* [14] introduced a novel methodology based on several A-EGM measures extracted from the raw A-EGM

signal. Combining the information provided by these measures, it enabled a better and robust enough classification, which was not possible with the use of one single A-EGM measure. Ciaccio *et al.* [15] used extrinsic and intrinsic features of the A-EGM shape to characterize the level of fractionation in CFAE sequences. Lin *et al.* [16] showed that a linear analysis (mean fractionation interval and dominant frequency) and a nonlinear-based waveform similarity analysis of the local A-EGM could determine the area in which CFAE sites, which are important for AF maintenance, are localized.

According to the latest studies, the level of fractionation has been shown to be promisingly well estimated by the use of complexity measures. Ng *et al.* [17] showed that Shannon entropy (ShEn) can be used to quantify the complexity of A-EGM and therefore provides an objective and automated method for the identification of CFAE sites that are indicated for AF ablation. Ganesan *et al.* [18] introduced ShEn as a tool that could assist the mapping of the AF rotor. Jacquemet *et al.* showed that micro-scale obstacles cause significant changes to electrogram waveforms and that conduction velocity, electrogram amplitude and also the degree of fractionation can be used to discriminate the nature of the substrate and the characteristics of fibrosis, giving rise to slower conduction [19]. Many previous works, e.g., [19,20] suspect that when the changes in the substrate become more acute, the A-EGMs become more irregular and, thus, more complex, and for this reason, it can be analyzed in terms of complexity. We suspect that SampEn is expected to be increased when the AF substrate is changing or if AF is triggered or sustained at a site under the measurement catheter. In order to evaluate the changes in the substrate, we introduced artificial individual classes of A-EGM fractionation for the purposes of this study. Current expert opinion is that CFAE can be either substrate based or caused by temporary effects (changes in cell repolarization time).

To the best of our knowledge, a direct relationship between the subjective perception of A-EGM fractionation and quantification in a continuous or semi-continuous scale has never been performed. Additionally, the data used for the analysis in these studies are often given and classified by a single source, even though the evaluation is performed by more than one independent physician (clinician or the same working group).

In this work, we propose the use of sample entropy (SampEn) for the quantification of the CFAE fractionation degree in two different A-EGM databases. The A-EGM recorded in the Czech Republic (CZ database) have been classified by at least three different clinicians, while the A-EGM database recorded in Germany (GE database) was classified by two clinicians. SampEn has been chosen because it is a well-characterized complexity measure with robust performance in other cardiac applications, e.g., in beat to beat (RR) intervals, heart rate variability (HRV) signals and ECG signals [21–24], and it is one of the few complexity measures that can deal with the analysis and complexity of CFAEs.

Moreover, SampEn can be used not only to segment the main CFAE, but also to quantify the substrate changes in a continuous way by discerning among four different classes of CFAE fractionation levels. The analysis described is based only on A-EGM records instead of surface ECGs [25–28].

2. Materials and Methods

The core of the method proposed is the computation of the SampEn metric for A-EGM records. To improve the standard SampEn results, a novel parameter optimization stage is included. This new step calculates the results for a range of m and r values in order to find the best combination. The optimization is performed in terms of maximal segmentation between the main CFAE classes and minimum SampEn variability in a database training subset. Specifically, a K-fold bootstrap algorithm for variability reduction over the ROC area (AUC) curves is used so as to find the optimal (m, r) parameter set. Then, the best parameter combination (m, r) obtained can be used for SampEn computation over the entire input database to obtain the final classification.

A statistical test was used to assess the accuracy of the partition. A Mann–Whitney test and ROC computation as a function of sensitivity and specificity were used to evaluate the segmentation capabilities of the method proposed. The classification goal was to discern among the main A-EGM record types and also between individual levels of fractionation. The general scheme of the method proposed is shown in Figure 1, and a more detailed description of each stage is developed in the following subsections, including the validation experimental database.

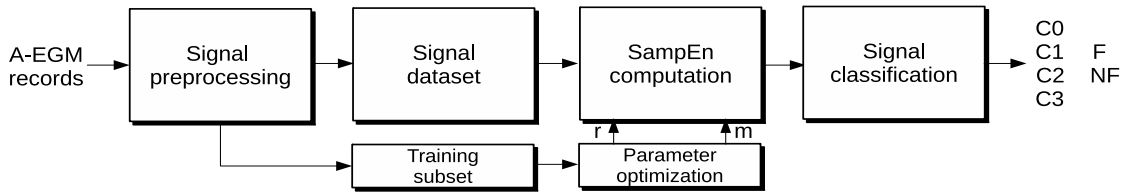


Figure 1. Block diagram of the new methodology proposed, from the raw atrial electrograms (A-EGM) input signal preprocessing stage, up to the classification results.

2.1. Sample Entropy

Sample entropy (SampEn) was introduced to counteract the well-known Approximate entropy (ApEn) weaknesses [29]. SampEn is less sensitive to data series length and input parameters; it does not compute self-matches and yields better relative consistency. SampEn was defined as the negative natural logarithm for conditional properties that a series of data points of length m would repeat at $m + 1$.

Given a time series $x(n)$ of length N, sequences of m-length vectors are formed and compared. Two of these vectors x_i, x_j commencing at $n = l$ and of length m are considered alike if the distance between them $(d_m(i, j))$ is below a threshold r. Then, the matching ratios A and B can be computed as:

$$B_i^m(r) = \frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} (d_m(i,j) \leq r) \tag{1}$$

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \tag{2}$$

$$A_i^m(r) = \frac{1}{N-m-1} \sum_{\substack{j=1 \\ j \neq i}}^{N-m} (d_{m+1}(i,j) \leq r) \tag{3}$$

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \tag{4}$$

Specifically, the dissimilarity measure used for SampEn metrics can then be defined as:

$$d_m(i,j) = d(x_m(i), x_m(j)) = \max_l \{|x(i+l-1) - x(j+l-1)|\} \tag{5}$$

Finally, SampEn is computed as the negative natural logarithm of the conditional probabilities defined in Equations (1) and (3):

$$\text{SampEn}(m,r,N) = \ln(A^m(r)) - \ln(B^m(r)) \tag{6}$$

SampEn m and r Parameter Optimization

SampEn parameters m and r configure how matches are addressed. The accuracy and the confidence of the SampEn estimates depend on the number of matches. It can be increased by choosing small m or large r values, but since entropy is defined under the assumptions $m \to \infty$ and $r \to 0$, choosing extreme values can lead to incorrect results. A large r value can yield a SampEn estimate smaller than it should be. Choosing a small m value can hide physical processes that are relevant at larger scales [24].

The parameter optimization method proposed searches for the best m and r configuration in the vicinity of usual values ($r = 0.25, m = 2$ or $m = 3$). The m parameter initially spans from 1–10. The r parameter ranges from 0.1–0.7 in 0.05 steps. SampEn values are computed for each possible resulting combination. Using a labeled training subset, ROC curves are respectively computed for each combination between classes of interest (C0 $\cup$ C1 *vs.* C2 $\cup$ C3 in this case). To further improve the robustness of the results, a K-fold bootstrap stage is included, with K=10. The optimal pair of parameters is chosen to be the pair that additionally satisfies maximum area under the curve (AUC), which can be interpreted as the probability that a randomly-chosen CFAE is correctly classified as fractionated or non-fractionated A-EGM and minimum mean variability at most of the different K folds. The optimal combination is searched according to the selection criteria value (SCV) described in (7). Analytically, the parameters are obtained from the following expression:

$$(m,r) = \text{argmax} \left\{ \frac{\text{AUC}}{\frac{1}{N_s} \sum_{q=1}^{N_s} |\text{SampEn}_q - \text{median}(\text{SampEn})|} \right\} \tag{7}$$

the SampEn value being computed according to (6) for signal q in fold s. In this way, the variability of the measure for each group is minimized, whereas AUC is maximized.

2.2. Validation

The proposed algorithm was validated following the steps described for SampEn and ROC computation. The specificity and sensitivity values for the optimal pair of parameters (m, r) were obtained from records of the experimental databases. Segmentation capabilities were assessed using a Mann–Whitney U-test, the corresponding sensitivity and specificity ROC parameters and the area under ROC (AUC). This optimal parameter set was then validated using the rest of the records from the experimental database not used for training.

In addition, the results obtained were also compared to those yielded by the standard methodology, complex fractionated (CFE) mean [30], which is a standard mapping procedure. The CFEmean algorithm (St. Jude Medical) measures the mean cycle length during atrial fibrillation. The baseline signal noise level is usually set between 0.02 and 0.04 mV in order to avoid noise detection. Each dV/dT is detected with a deflection width of 10 ms after a refractory period of 30–50 ms. The mean cycle length (CL) of a 5-s electrogram is calculated and displayed color-coded on the left atrium (LA) anatomical shell [31]. According to previous studies [2,31], regions with mean CL < 120 ms can be defined as CFAEs.

2.3. Experimental Dataset

Two different A-EGM databases recorded at two independent cardiology departments in two different EU countries were used separately. These databases were also combined to form a mixed third database. Each database contains at least four classes of CFAE, each class representing a different fractionation level. Signals in both databases were preselected manually by skilled electrophysiologists (EP) to select high quality CFAE signals with no ventricular far field influence, no artifacts caused by obvious catheter movements and visually-detectable presence of signal noise [14,32]. These classes can also be roughly separated at a higher conceptual level into two main classes, non-fractionated (NF) A-EGM and fractionated (F) A-EGMs. Recordings with a strong far field were excluded before annotation because they were out of the scope of the study.

The first database was recorded and independently annotated by three different EP that perform radio frequency ablation (RFA) of AF on a regular basis in the Czech Republic [14]. The second one was recorded and annotated by two different EP working at a German unit [32]. Both databases were post-processed with the same algorithms for baseline wander and high frequency noise removal. They were also resampled to an equal sampling frequency ($F_s = 1KHz$). Epochs of 1500 ms were analyzed in order to keep consistency on the parameters that can influence SampEn computation [24,29,33].

There are different types of A-EGM during AF, as described in [34]. The signals in both databases were classified into four levels of fractionation [35]. On the one hand, fractionated signals were classified by experts into three levels of fractionation (C1, C2 and C3). On the other hand, non-fractionated A-EGM signals were considered as Level 0 (C0). Classes C0 and C1 were assumed not to trigger or maintain AF, as EP would consider C0 and C1 as sites not recommended for RFA. They were termed as non-fractionated (NF). Signals in C2 and C3 could indicate sites of substrate changes that could either trigger or sustain AF. This characterization is similar to those described in other works, such as [14,32].

These four classes were introduced to enable objective quantification for the study purpose and are summarized as follows:

- C0: Non-fractionated A-EGM.
- C1: Fractionated A-EGM with periodic activity.
- C2: A mixture of periodic fractionated and periodic non-fractionated A-EGM.
- C3: High frequency A-EGM with continuous activity. No regular activation can be seen.

2.3.1. German Database

The German (GE) A-EGM database contains signals from 11 patients who suffered from AF. Eight patients suffered from persistent AF and three patients from paroxysmal AF. All patients were indicated for catheter ablation. Intracardiac A-EGM recordings from a multipolar circular catheter, such as Lasso (10 polar, Biosense Webster, Diamond Bar, USA), Optima (14 polar, St. Jude Medical, St. Paul, USA) or Orbiter (14 polar, Bard Electrophysiology, Lowell, USA), were performed after PV isolation.

This database consists of 429 signals. Each signal has a duration of 5 s (only the first 1500 ms were used for consistency with the other database) and was recorded at a sampling rate of 1.2 KHz. The data were filtered by the allurement system with a band pass filtering at [30, 250] Hz, and the remaining baseline wander and high noise were reduced by wavelet decomposition.

The 429 signals were divided into the four classes as follows: 153 C0 signals, 75 C1 signals, 148 C2 signals and 53 C3 signals. The inter-observer agreement was 100%, because only records with such an agreement level were included in the dataset. An example of one signal belonging to each class is given in Figure 2A.

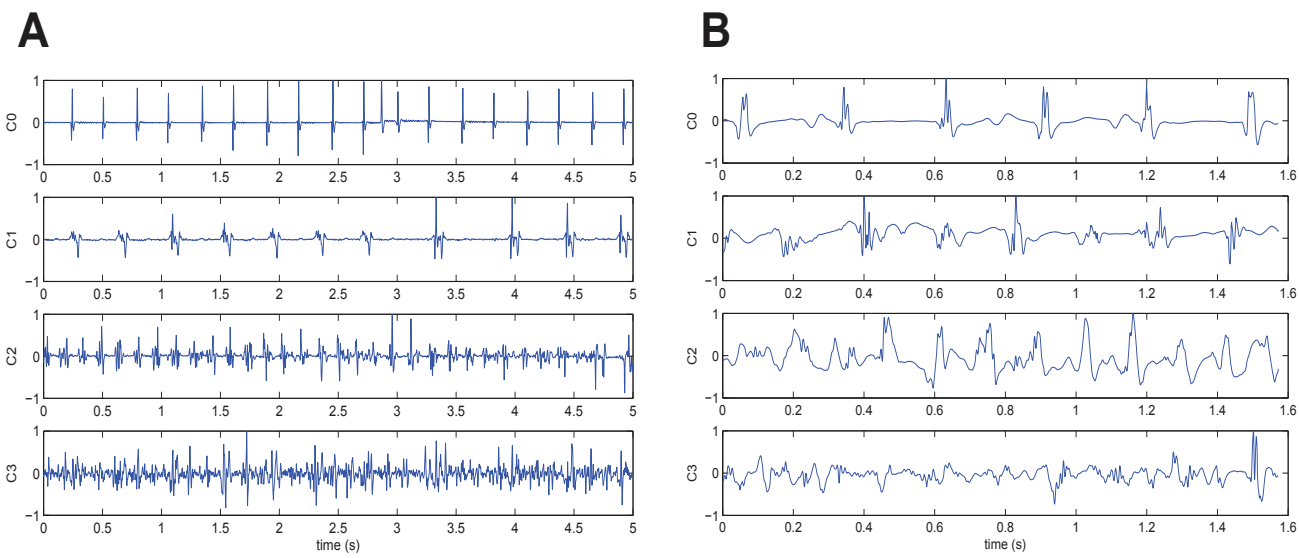


Figure 2. One signal from each fractionation level, ranging from C0 (top) to C3 (bottom). Original raw signals after baseline wander removal. (**A**) German (GE) database and (**B**) Czech (CZ) database.

2.3.2. Czech Database

The Czech database (CZ) consists of 113 A-EGM, preselected by experts from a larger database of A-EGM from the Czech Republic. This database was recorded during AF mapping procedures performed on patients that were indicated for RFA of AF [14].

The signals were recorded by CardioLab7000, Prucka Inc. during AF procedures at a sampling rate of 997 Hz and resampled to 1 KHz. Each preselected A-EGM signal in this dataset has a duration of 1500 ms.

The EP expert signal selection was driven by the goal to get a good signal-to-noise ratio for later evaluation of the degree of A-EGM fractionation by an expert.

The 113 signals were divided into the four classes as follows: 22 C0 signals, 42 C1 signals, 36 C2 signals and 13 C3 signals. The inter-observer agreement of ranking to classes was higher than 90% for Classes C0, C2 and C3 and 79% for Class C1. An example of one signal belonging to each class is given in Figure 2B.

2.3.3. Mixed Database

A mixed database (termed BT) was created for training purposes in the SampEn parameter optimization process. This database contains 75% of the signals in each class from the CZ database (as it has the lowest number of signals per class) and the same number of signals from each class of the GE database, so as to provide a weighted database. The BT database was made up of 158 signals. Class C0 contained 30 signals, C1 contained 64 signals, C2 contained 54 signals and C3 contained 20 signals.

3. Experiments and Results

The method proposed was applied to the experimental dataset described in Section 2.3.3. First, a labeled training subset was randomly chosen for SampEn parameter optimization. The resulting (m, r) values were estimated in terms of the best segmentation obtained between non-fractionated (C0 $\cup$ C1) and fractionated (C2 $\cup$ C3) A-EGM groups and minimum SampEn variability. Table 1 shows the (m, r) values corresponding to the maximum SCV at each fold. For most of the folds, 88% of the AUC was achieved when $m = 4$ and $r = 0.65$. Only three folds (K = 2, 6, 8) do not correspond to these parameter values. In any case, the results are very robust, since the minimum performance achieved was an AUC of at least 89%, 81% specificity and 75% sensitivity.

Table 1. Selection criteria value (SCV), AUC and sample entropy (SampEn) parameters (m, r) obtained for the training subset using the optimization scheme described. Sensitivity (Se) and specificity (Sp) values are shown for the corresponding (m, r) parameters in each row. The K-fold column refers to each of the folds considered from 1–10.

K-fold	SCV	AUC (%)	m	r	Sp (%)	Se (%)
1	1.278	88.6	4	0.65	85.1	76.4
2	1.299	90.9	8	0.65	80.6	78.7
3	1.331	89.2	4	0.65	88.1	76.4
4	1.380	90.9	4	0.65	91.0	77.5
5	1.195	89.1	4	0.65	86.6	78.7
6	1.161	91.2	8	0.65	88.1	76.4
7	1.140	88.3	4	0.65	88.1	75.3
8	1.266	84.7	2	0.65	89.6	77.5
9	1.238	90.0	4	0.65	83.6	79.8
10	1.295	88.5	4	0.65	85.7	77.8

The segmentation capabilities of the method proposed were then assessed using these $(m, r) = (4, 0.65)$ parameters over the entire remaining dataset (the one containing all CZ and GE signals not used for training). The statistical test applied was a Mann–Whitney test. Non-Fractionated A-EGM (NF) included C0 and C1 levels, whereas fractionated (F) A-EGM corresponded to C3 and C4 levels. The results of this step are shown in Table 2. This table enlists results about the F and NF classes, such as mean, median, standard deviation (SD) and 95% confidence intervals (CI). It can be seen that mean/median values of each group at each database are located far away and none of their CI overlap. The Mann–Whitney statistical test provided a significance level lower than 0.001 for each comparison.

Table 2. SampEn statistics for the fractionated/non-fractionated (F/NF) A-EGM for $m = 4, r = 0.65$. DB stands for database. Results are summarized using mean and median SampEn values, standard deviation (SD), confidence intervals (CI), AUC, sensitivity (Se), specificity (Sp) and the p-value of the Mann–Whitney statistical test.

DB	class	mean	median	SD	95% CI	AUC (%)	Se (%)	Sp (%)	p
BT	NF	0.054	0.049	0.003	[0.047,0.061]	89.3	86.9	77.5	0.001
	F	0.115	0.108	0.005	[0.105,0.125]	89.3	86.9	77.5	0.001
CZ	NF	0.050	0.050	0.003	[0.044,0.055]	88.7	79.5	89.1	0.001
	F	0.090	0.086	0.003	[0.083,0.097]	88.7	79.5	89.1	0.001
GE	NF	0.043	0.023	0.003	[0.038,0.048]	93.4	94.5	82.9	0.001
	F	0.137	0.134	0.003	[0.130,0.143]	93.4	94.5	82.9	0.001

A number of different training subsets were used for comparative purposes. Although the optimal m and r values obtained were not equal, the segmentation capability of the SampEn remained unchanged. Two examples of these results are shown in Table 3.

Table 3. SampEn statistics for the F/NF A-EGM for individual parameters (m, r) optimized individually on the CZ and GE databases. DB stands for database and params for parameters (m, r). Results are summarized using mean and median values, standard deviation (SD), confidence intervals (CI), area under ROC (AUC), sensitivity (Se), specificity (Sp) and the Mann–Whitney statistical probability (p).

DB	Class	params	mean	median	SD	95% CI	AUC (%)	Se (%)	Sp (%)	p
CZ	NF	(5,0.15)	0.120	0.131	0.007	[0.106,0.133]	90.9	82.0	79.2	0.001
	F		0.224	0.227	0.008	[0.208,0.240]	90.9	82.0	79.2	0.001
GE	NF	(4,0.15)	0.202	0.162	0.011	[0.180,0.225]	88.5	87.8	78.5	0.001
	F		0.452	0.454	0.008	[0.437,0.468]	88.5	87.8	78.5	0.001

The three databases were analyzed with the initially-optimized parameters, but considering the individual Classes C0, C1, C2 and C3 instead of the global F and NF groups. Graphically, Figures 3 and 4 show the results for each database and each class. From these boxplots, it can be stated that the entropy of the signals estimated by the SampEn metric significantly increases. These trends can also be observed with different parameter configurations, as depicted in Figure 4. This is in agreement with expectations, since manual expert ranking of the A-EGM signals coincides with these results qualitatively. The numerical results corresponding to these figures are shown in Table 4.

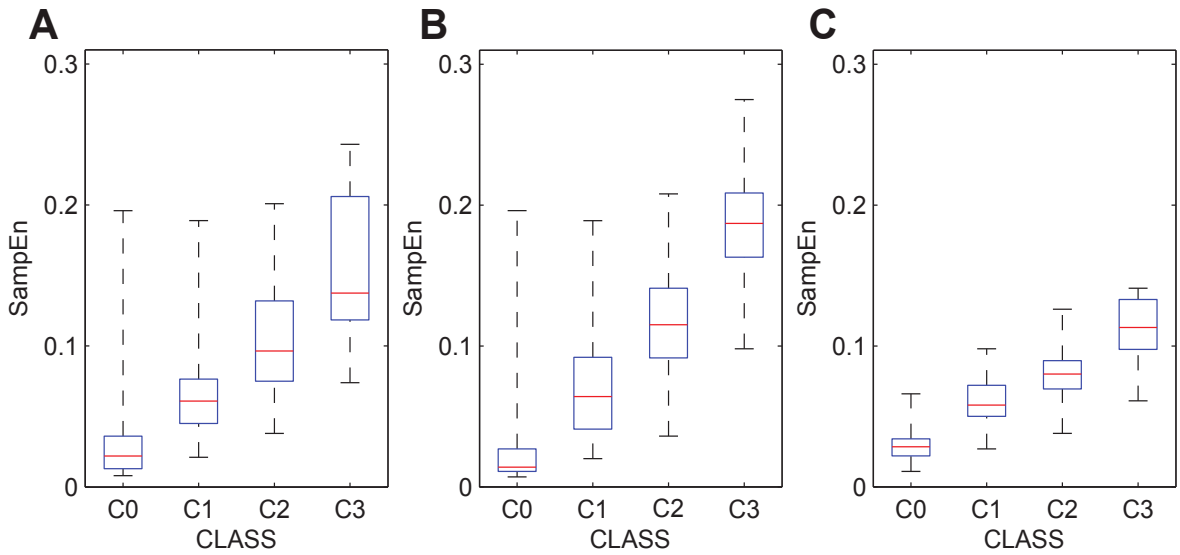


Figure 3. Boxplot distribution of the SampEn values computed to the initially-optimized parameters for individual levels of fractionation, (m, r) =(4,0.65). (**A**) The BT database. (**B**) The GE database. (**C**) The CZ database.

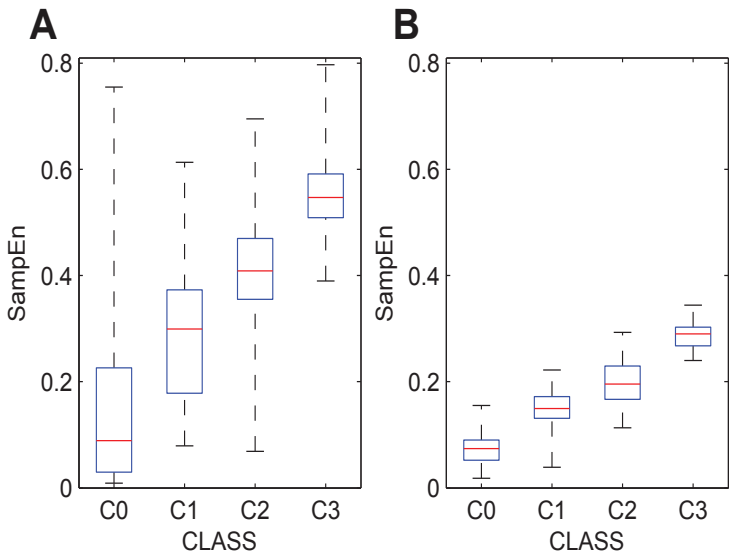


Figure 4. Boxplot distribution of the SampEn values computed with the individual optimized parameters for each of the levels of fractionation present in each database. (**A**) The GE database (4,0.15). (**B**) The CZ database (5,0.15).

Table 4. SampEn statistics for individual classes (4,0.65).

DB	Class	mean	median	95% CI
BT	C0	0.032	0.021	[0.020, 0.045]
	C1	0.065	0.061	[0.058, 0.072]
	C2	0.102	0.097	[0.093, 0.111]
	C3	0.149	0.138	[0.127, 0.172]
CZ	C0	0.029	0.028	[0.024, 0.035]
	C1	0.061	0.058	[0.055, 0.066]
	C2	0.082	0.083	[0.075, 0.088]
	C3	0.113	0.114	[0.100, 0.125]
GE	C0	0.030	0.014	[0.024, 0.036]
	C1	0.070	0.064	[0.062, 0.077]
	C2	0.119	0.115	[0.113, 0.125]
	C3	0.186	0.187	[0.176, 0.196]

Finally, the results were compared to those obtained using a standard CFEmean measure. Figure 5 shows topographical atrial maps of a selected patient. It compares two different indices calculated from the raw signals recorded at sites shown in the figure in 3D coordinates. Figure 5 shows that the SampEn measures reveal broader and more specific areas with a higher level of complexity of CFAEs, whereas the CFEMean indices show single spots with lower complexity around them (see the details of the pulmonary vein in Figure 5B).

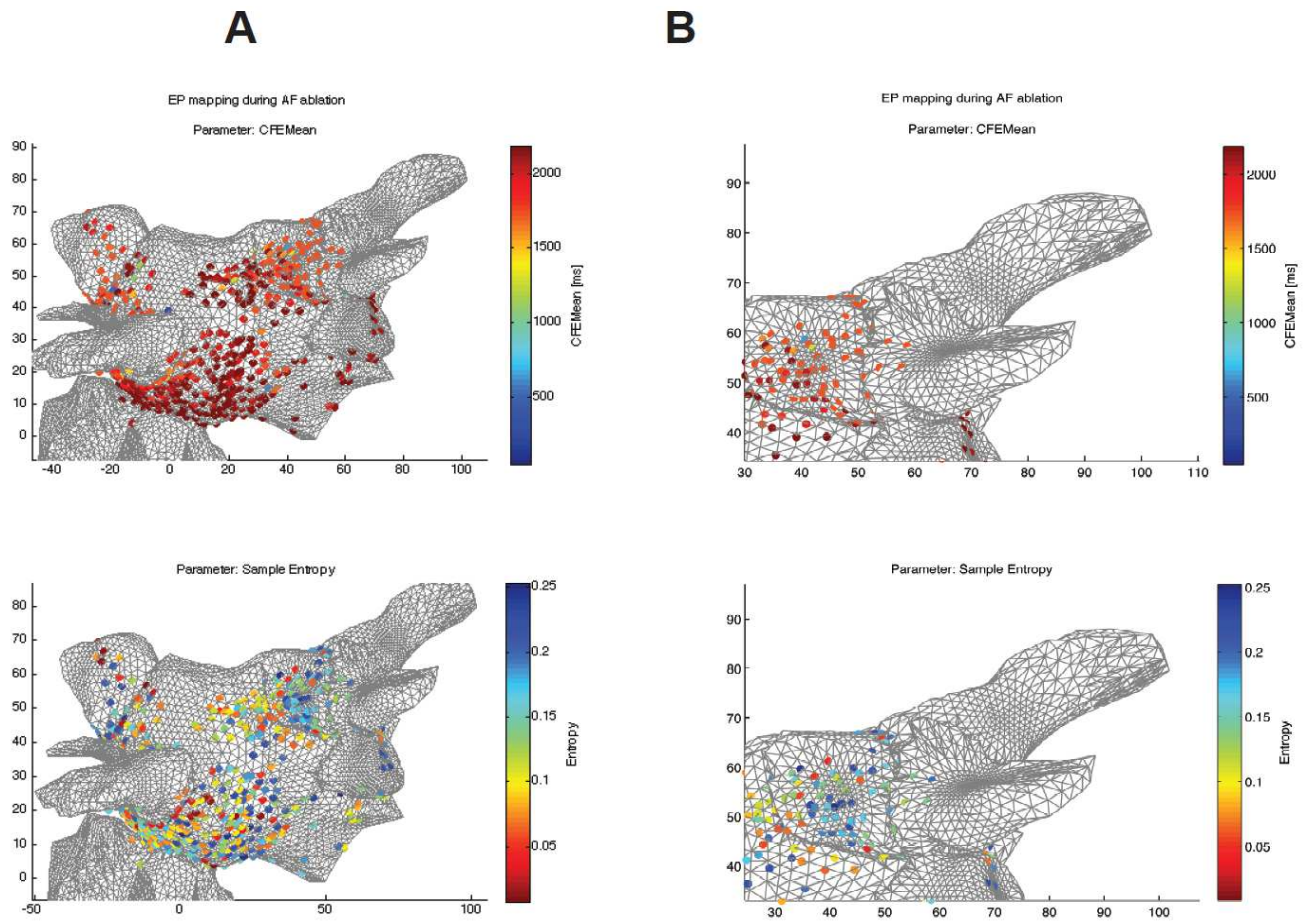


Figure 5. Comparison between both algorithms used for a specific patient undergoing radio frequency ablation (RFA) of atrial fibrillation (AF). SampEn with optimized parameters ($m = 4, r = 0.65$) (bottom) and CFEMean (St. Jude Medical) measurements mapped on a 3D model of heart tissue. The blue color in both measures reveals areas with a higher level of complexity of complex fractionated atrial electrograms (CFAEs). (**A**) 3D atrial topographical map, whole mapped area, frontal view. (**B**) A detail of the area around the pulmonary vein.

4. Discussion

The results obtained confirm the goodness of the SampEn-based method to segment classes in A-EGM records. With the parameters computed from a small training subset, it is possible to distinguish between F and NF groups and even among C0, C1, C2 and C3 classes. In addition, the specific parameter configuration is not very important, since different values yield almost identical performance. This fact also confirms the robustness of the approach.

The findings are in consonance with the work of Ng [17], where CFAEs were found to have higher entropy than non-CFAEs. This work deals with the complete A-EGM signal, while [17] uses fractional intervals computed over the A-EGM signal and also uses Shannon entropy rather than SampEn entropy. It can be seen that different complexity measures working over different signals produce similar results, thus giving a good evaluation of the system. This work suggests that CFAEs are more complex and

random or irregular than non-CFAEs, which is supported by the results that are obtained here. It is important to note that the results are similar even though the analyzed signals are not the same.

With regard to the analysis of the sensitivity and specificity results in Table 1, the results achieved can be considered as really satisfactory based on current similar studies. Although the results are somewhat lower than those achieved in the work of Alcaraz [26], they were dealing with distinguishing between paroxysmal AF and persistent AF in records of at least 10 s in length of surface ECG. In our work, intracardiac electrograms were recorded over a period of only 5 s during AF and the RFA procedure. As the records become shorter, it becomes more difficult to apply entropy metrics with significant results. Moreover, the recorded electrograms were classified into one of four groups, irrespective of the type of AF.

Concerning the optimization of parameters, the rationale to devise a method to select m and r parameters was the fact that no specific guidelines exist for parameter selection. In this regard, only some general values of m and r have been recommended by Pincus [36] and Lake [37]. A criteria based on searching for maximum AUC and minimum variability was defined (7), but providing evidence that the exact values are not crucial to guarantee a good classification performance, since different databases and different parameter arrangements were successfully tested.

From a multicenter point of view, Levels C2 and C3, fractionated electrograms, show different mean or median values in the CZ and GE databases. This may be due to differences in the way the experts in the two countries carried out their annotations. When the parameters in each database are optimized, it can be seen that the SampEn values increase, which unveils lower scale variations that are masked with the training dataset optimum parameters. However, the CZ database still exhibits an increase in complexity due to changes in the substrate, normally distributed SampEn values and non-overlapping interquartile ranges ($p < 0.001$). The same behavior is observed for the GE database, though the distribution is not normal here, and the C0 and C1 interquartile ranges overlap. The median value of one level is not contained inside the interquartile range of the next level, resulting in clear separation of the levels.

Regarding quantification of substrate changes, Figures 3 and 4 show that the complexity of the signal increases in accordance with experts' ranking of the level of complexity of the A-EGM signal. This can be assumed to be related to a change in the substrate characteristics. These results agree with the literature, as higher complexity levels are produced by more irregular and chaotic systems, and a varying substrate is thought to be a more irregular one. This idea is also supported by the study of Hoekstra [38], where epicardial A-EGMs were analyzed by means of correlation dimension, correlation entropy and their coarse-grained variations. Hoekstra *et al.* showed that the complexity of epicardial A-EGMs increases with the type of fibrillation. Interestingly, the results are similar, although this work deals with endocardial recordings rather than with epicardial recordings, and in addition, different complexity measures were considered.

5. Conclusions

This work describes a way to classify A-EGM signals into two main classes, non-fractionated and fractionated A-EGM, and into individual groups, C0, C1, C2 and C3. This method can support the EP decision process during RF ablation of AF. The quantitative and qualitative results also demonstrate that

SampEn is an appropriate entropy measure to perform this task, with powerful discrimination capabilities to evaluate AF substrate changes.

Although signals from different databases may have the same length, sample frequency and have undergone the same post-processing steps, the physician's way of recording, pre-processing and classifying these A-EGM records strongly influence the annotation process, since the optimal SampEn parameters are not the same for each database. However, when analyzing each database with the optimal parameters for a different training subset, the results are still very robust, namely the method is generalizable and transferable to a large variety of situations in the context of A-EGM signal processing.

In summary, we propose the use of SampEn for real-time support during navigation radio frequency ablation of atrial fibrillation based on CFAE. Only 1.5 s need to be analyzed. This method could become a clinical tool in these applications to improve the outcome of patients that have to undergo these procedures by means of a more accurate procedure.

Acknowledgments

This work has been supported by the Spanish Ministry of Science and Innovation, Research Project TEC 2009–14222, by the Ministry of Education Youth and Sports of the Czech Republic, the Grant Agency of the Czech Technical University in Prague No. SGS13/203/OHK3/3T/13 and by the Czech Science 300 Foundation post-doctoral GACR research project GACR #P103/11/P106.

The authors also would like to express their gratitude to Christopher Schilling and Olaf Doessel from the Institute of Biomedical Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, and Claus Schmitt, Medizinische Klinik IV, Staedtisches Klinikum Karlsruhe, Karlsruhe, Germany, for their contribution in terms of datasets, reviews and guidance.

Author Contributions

Eva Cirugeda performed the reseach, analyzed the data and wrote the first paper draft. Daniel Novák and Vaclav Kremen designed the research, supervised the experimental tasks, contributed to the first version of the paper, helped during the review process, and provided the Czech database. David Cuesta-Frau wrote the next drafts and the final version of the paper, submitted the paper, and prepared the responses to reviewers. He was also the leader of the Spanish project that partially supported this research (TEC2009-14222), and provided the entropy analysis framework. Matthias Keller provided the German database, and supported the rest of researchers in undestanding and interpreting the data and the results, as well as in specific database issues of the review process. Armin Luik provided the medical perspective and background. Martina Srutova contributed with the data collection from Czech Site and organization of the data. She also contributed on data pre-processing and answering the reviewer questions. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Haïssaguerre, M.; Jaïs, P.; Shah, D.; Takahashi, A.; Hocini, M.; Quiniou, G.; Garrigue, S.; Le Mouroux, A.; Le Métayer, P.; Clémenty, J. Spontaneous Initiation of Atrial Fibrillation by Ectopic Beats Originating in the Pulmonary Veins. *New Eng. J. Med.* **1998**, *339*, 659–666.
2. Nademanee, K.; Schwab, M.; Porath, J.; Abbo, A. How to perform electrogram-guided atrial fibrillation ablation. *Heart Rhythm* **2006**, *3*, 981–984.
3. Porter, M.; Spear, W.; Akar, J.; Helms, R.; Brysiewicz, N.; Santucci, P.; Wilber, D. Prospective Study of Atrial Fibrillation Termination During Ablation Guided by Automated Detection of Fractionated Electrograms. *J. Cardiovasc. Electrophysiol.* **2008**, *19*, 613–620.
4. Haïssaguerre, M.; Hocini, M.; Sanders, P.; Takahashi, Y.; Rotter, M.; Sacher, F.; Rostock, T.; Hsu, L.; Jonsson, A.; OŃeill, M.; *et al.* Localized Sources Maintaining Atrial Fibrillation Organized by Prior Ablation. *Circulation* **2006**, *113*, 616–625.
5. Schmitt, C.; Ndrepepa, G.; Weber, S.; Schmieder, S.; Weyerbrock, S.; Schneider, M.; Karch, M.; Deisenhofer, I.; Schreieck, J.; Zrenner, B.; *et al.* Biatrial multisite mapping of atrial premature complexes triggering onset of atrial fibrillation. *Am. J. Cardiol.* **2002**, *89*, 1381–1387.
6. Ndrepepa, G.; Karch, M.; Schneider, M.; Weyerbrock, S.; Schreieck, J.; Deisenhofer, I.; Zrenner, B.; Schomig, A.; Schmitt, C. Characterization of Paroxysmal and Persistent Atrial Fibrillation in the Human Left Atrium During Initiation and Sustained Episodes. *J. Cardiovas. Electrophysiol.* **2002**, *13*, 525–532.
7. Nademanee, K.; McKenzie, J.; Kosar, E.; Schwab, M.; Sunsaneewitayakul, B.; Vasavakul, T.; Khunnawat, C.; Ngarmukos, T. A new approach for catheter ablation of atrial fibrillation: Mapping of the electrophysiologic substrate. *J. Am. Coll. Cardiol.* **2004**, *43*, 2044–2053.
8. Oral, H.; Chugh, A.; Good, E.; Wimmer, A.; Dey, S.; Gadeela, N.; Sankaran, S.; Crawford, T.; Sarrazin, J.; Kuhne, M.; *et al.* Radiofrequency Catheter Ablation of Chronic Atrial Fibrillation Guided by Complex Electrograms. *Circulation* **2007**, *115*, 2606–2612.
9. Kumagai, K. Patterns of activation in human atrial fibrillation. *Heart Rhythm* **2007**, *4*, 7–12.
10. Mainardi, L.; Corino, V.; Lombardi, L.; Tondo, C.; Mantica, M.; Lombardi, F.; Cerutti, S. Assessment of the dynamics of atrial signals and local atrial period series during atrial fibrillation: effects of isoproterenol administration. *BioMed. Eng. Online* **2004**, *3*, 37.
11. Ravelli, F.; Faes, L.; Sandrini, L.; Gaita, F.; Antolini, R.; Scaglione, M.; Nollo, G. Wave Similarity Mapping Shows the Spatiotemporal Distribution of Fibrillatory Wave Complexity in the Human Right Atrium During Paroxysmal and Chronic Atrial Fibrillation. *J. Cardiovasc. Electrophysiol.* **2005**, *16*, 1071–1076.
12. Ng, J.; Goldberger, J. Understanding and Interpreting Dominant Frequency Analysis of AF Electrograms. *J. Cardiovasc. Electrophysiol.* **2007**, *18*, 680–685.
13. Takahashi, Y.; O'Neill, M.; Hocini, M.; Dubois, R.; Matsuo, S.; Knecht, S.; Mahapatra, S.; Lim, K.; Jaïs, P.; Jonsson, A.; *et al.* Characterization of Electrograms Associated With Termination of Chronic Atrial Fibrillation by Catheter Ablation. *J. Am. Coll. Cardiol.* **2008**, *51*, 1003–1010.

14. Křemen, V.; Lhotská, L.; Čihák, R.; Vančura, V.; Kautzner, J.; Wichterle, D. A new approach to automated assessment of endocardial electrograms fractionation in human left atrium during atrial fibrillation. *Physiol. Meas.* **2008**, *29*, 1371–1381.
15. Ciaccio, E.; Biviano, A.; Whang, W.; Gambhir, A.; Garan, H. Different characteristics of complex fractionated atrial electrograms in acute paroxysmal versus long-standing persistent atrial fibrillation. *Heart Rhythm* **2010**, *7*, 1207–1215.
16. Lin, Y.; Lo, M.; Lin, C.; Chang, S.; Lo, L.; Hu, Y.; Chao, T.; Li, C.; Chang, Y.; Hsieh, W.; *et al.* Nonlinear Analysis of Fibrillatory Electrogram Similarity to Optimize the Detection of Complex Fractionated Electrograms During Persistent Atrial Fibrillation. *J. Cardiovasc. Electrophysiol.* **2013**, *24*, 280–289.
17. Ng, J.; Borodyanskiy, A.; Chang, E.; Villuendas, R.; Dibs, S.; Kadish, A.; Goldberger, J. Measuring the complexity of atrial fibrillation electrograms. *J. Cardiovasc. Electrophysiol.* **2010**, *21*, 649–655.
18. Ganesan, A.; Kuklik, P.; Lau, D.; Brooks, A.; Baumert, M.; Lim, W.; Thanigaimani, S.; Nayyar, S.; Mahajan, R.; Kalman, J.; *et al.* Bipolar Electrogram Shannon Entropy at Sites of Rotational Activation: Implications for Ablation of Atrial Fibrillation. *Circ. Arrhythm. Electrophysiol.* **2013**, *6*, 48–57.
19. Jacquemet, V.; Henriquez, C. Genesis of complex fractionated atrial electrograms in zones of slow conduction: A computer model of microfibrosis. *Heart Rhythm* **2009**, *6*, 803–810.
20. Jadidi, A.; Duncan, E.; Miyazaki, S.; Lellouche, N.; Shah, A.; Forclaz, A.; Nault, I.; Wright, M.; Rivard, L.; Liu, X.; *et al.* Functional Nature of Electrogram Fractionation Demonstrated by Left Atrial High-Density Mapping. *Circ. Arrhythm. Electrophysiol.* **2012**, *5*, 32–42.
21. Ferrario, M.; Signorini, M.; Magenes, G.; Cerutti, S. Comparison of entropy-based regularity estimators: Application to the fetal heart rate signal for the identification of fetal distress. *Biomed. Eng. IEEE Trans.* **2006**, *53*, 119–125.
22. Lewis, M.; Short, A. Sample entropy of electrocardiographic RR and QT time-series data during rest and exercise. *Physiol. Meas.* **2007**, *28*, 731–744.
23. Al-Angari, H.; Sahakian, A. Use of Sample Entropy Approach to Study Heart Rate Variability in Obstructive Sleep Apnea Syndrome. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 1900–1904.
24. Lake, D.; Richman, J.; Griffin, M.; Moorman, J. Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2002**, *283*, 789–797.
25. Cervigón, R.; Moreno, J.; Reilly, R.; Millet, J.; Pérez-Villacastín, J.; Castells, F. Entropy measurements in paroxysmal and persistent atrial fibrillation. *Physiol. Meas.* **2010**, *31*, 1011.
26. Alcaraz, R.; Rieta, J. The application of nonlinear metrics to assess organization differences in short recordings of paroxysmal and persistent atrial fibrillation. *Physiol. Meas.* **2010**, *31*, 115–130.
27. Orozco-Duque, A.; Novak, D.; Kremen, V.; Bustamante, J. Multifractal analysis for grading complex fractionated electrograms in atrial fibrillation. *Physiol. Meas.* **2015**, *11*, 2269–2284.
28. Ugarte, J.P.; Orozco-Duque, A.; Tobón, C.; Kremen, V.; Novak, D.; Saiz, J.; Oesterlein, T.; Schmitt, C.; Luik, A.; Bustamante, J. Dynamic Approximate Entropy Electroanatomic Maps Detect Rotors in a Simulated Atrial Fibrillation Model. *PLoS ONE* **2014**, *12*, doi:10.1371/journal.pone.0114577.

29. Richman, J.; Moorman, J. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart. Circ. Physiol.* **2000**, *278*, 2039–2049.
30. Stiles, M.K.; Brooks, A.G.; John, B.; Wilson, L.; Kuklik, P.; Dimitri, H.; Lau, D.H.; Roberts-Thomson, R.L.; Mackenzie, L.; Willoughby, S.; *et al.* The effect of electrogram duration on quantification of complex fractionated atrial electrograms and dominant frequency. *J. Cardiovasc. Electrophysiol.* **2008**, *19*, 252–258.
31. Verma, A.; Novak, P.; Macle, L.; Whaley, B.; Beardsall, M.; Wulffhart, Z.; Khaykin, Y. A prospective, multicenter evaluation of ablating complex fractionated electrograms (CFEs) during atrial fibrillation (AF) identified by an automated mapping algorithm: Acute effects on {AF} and efficacy as an adjuvant strategy. *Heart Rhythm* **2008**, *5*, 198–205.
32. Schilling, C.; Keller, M.; Scherr, D.; Oesterlein, T.; Haissaguerre, M.; Schmitt, C.; Dossel, O.; Luik, A. Fuzzy decision tree to classify complex fractionated atrial electrograms. *Biomed. Tech.* **2015**, *60*, 244–255.
33. García-González, M.; Fernández-Chimeno, M.; Ramos-Castro, J. Errors in the Estimation of Approximate Entropy and Other Recurrence–Plot–Derived Indices Due to the Finite Resolution of RR Time Series. *Biomed. Eng. IEEE Trans.* **2009**, *56*, 345–351.
34. Konings, K.; Kirchhof, C.; Smeets, J.; Wellens, H.; Penn, O.; Allessie, M. High-density mapping of electrically induced atrial fibrillation in humans. *Circulation* **1994**, *89*, 1665–1680.
35. Schilling, C. *Analysis of Atrial Electrograms*; KIT Scientific Publishing: Karlsruhe, Germany, 2012.
36. Pincus, S.M.; Goldberger, A.L. Physiological time-series analysis: What does regularity quantify? *Am. J. Physiol.* **1994**, *266*, 1643–1656.
37. Lake, D.; Moorman, J. Accurate estimation of entropy in very short physiological time series: The problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Physiol. Heart Circ. Physiol.* **2011**, *300*, 319–325.
38. Hoekstra, B.; Diks, C.; Allessie, M.; DeGoede, J. Nonlinear analysis of epicardial atrial electrograms of electrically induced atrial fibrillation in man. *J. Cardiovasc. Electrophysiol.* **1995**, *6*, 419–440.

IOP Publishing | Institute of Physics and Engineering in Medicine Physiological Measurement

Physiol. Meas. **36** (2015) 2269–2284 doi:10.1088/0967-3334/36/11/2269

Multifractal analysis for grading complex fractionated electrograms in atrial fibrillation

A Orozco-Duque[1,2], D Novak[3], V Kremen[4] and J Bustamante[1]

[1] Bioengineering Center, Universidad Pontificia Bolivariana, Medellín, Colombia
[2] Grupo de Investigación e Innovación Biomédica, Instituto Tecnológico Metropolitano, Medellín, Colombia
[3] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, 166 36 Praha 6, Czech Republic
[4] Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Praha, Czech Republic

E-mail: andres.orozco@upb.edu.co and john.bustamante@upb.edu.co

Received 5 November 2014, revised 15 June 2015
Accepted for publication 13 August 2015
Published 9 October 2015

Abstract

Complex fractionated atrial electrograms provide an important tool for identifying arrhythmogenic substrates that can be used to guide catheter ablation for atrial fibrillation (AF). However, fractionation is a phenomenon that remains unclear. This paper aims to evaluate the multifractal properties of electrograms in AF in order to propose a method based on multifractal analysis able to discriminate between different levels of fractionation. We introduce a new method, the *h*-fluctuation index (*hFI*), where *h* is the generalised Hurst exponent, to extract information from the shape of the multifractal spectrum. Two multifractal frameworks are evaluated: multifractal detrended fluctuation analysis and wavelet transform modulus maxima. *hFI* is exemplified through its application in synthetic signals, and it is evaluated in a database of electrograms labeled on the basis of four degrees of fractionation. We compare the performance of *hFI* with other indexes, and find that *hFI* outperforms them. The results of the study provide evidence that multifractal analysis is useful for studying fractionation phenomena in AF electrograms, and indicate that *hFI* can be proposed as a tool for grade fractionation associated with the detection of target sites for ablation in AF.

Keywords: mutrifractal analysis, atrial fibrillation, fractionated electrograms, biomedical signal processing

(Some figures may appear in colour only in the online journal)

1. Introduction

Atrial electrograms (EGMs) recorded inside the heart chambers using catheters, are used for studying the propagation patterns in patients with arrhythmias, such as atrial fibrillation (AF). AF is the most common type of atrial tachyarrhythmia, affecting about 2% of the population, and its incidence is increasing.

In AF treatment, ablation is used to remove fibrillatory substrates using catheters and a radiofrequency technique. The electrical isolation of pulmonary veins (PVs) is the mainstream ablation procedure in AF. However, the arrhythmogenic substrate could be located outside the PVs, especially in cases of persistent and permanent AF. Therefore, ablation guided by electrograms is an important option.

The mechanism behind AF remains unclear, but some studies have shown that the morphology of the electrograms during AF could be related to different conduction patterns, e.g. conduction blocks, slow conduction, the collision of activation waves or microreentries (Konings *et al* 1997). Sites with EGM that present high frequency or with irregular and disorganised patterns have been proposed as regions for maintaining the AF (Konings *et al* 1997, Kalifa *et al* 2006, Zlochiver *et al* 2008, Chang *et al* 2013). Accordingly, several indexes have been developed in order to measure the degree of spatial organisation of the activation waves (Barbaro *et al* 2000, 2001, Everett *et al* 2001, Michelucci *et al* 2001, Faes *et al* 2002, Sanders *et al* 2005, Takahashi *et al* 2008).

The most well-established approach for characterising the organisation of the signal is the analysis of complex fractionated atrial electrograms (CFAEs). The CFEA concept was introduced by Nademanee *et al* (2004), and is defined as a signal that has fractionated electrograms composed of two or more deflections, or that has a perturbation of the baseline with continuous deflection. Atrial electrograms with a cycle length less than 120 ms are also considered to be CFAEs.

CFAE mapping has been used as a technique for locating arrhythmogenic substrates. However, CFAE is a controversial concept because not all CFAE patterns are suitable target sites for ablation (Narayan *et al* 2012). Traditional algorithms for detecting CFAEs are based only on measurements of time intervals and voltage thresholds (Nademanee *et al* 2004, Scherr *et al* 2007). Consequently, different fractionation levels and morphologies are not well described (Berenfeld and Jalife 2011). Recent studies propose that only highly fractionated EGMs are related to sites that maintained the AF (Hunter *et al* 2009, Lin *et al* 2013). The discrimination of various levels of fractionation is convenient for differentiating these critical sites.

In order to extract information about the complexity of the signal, nonlinear tools have been proposed to analyse fractionation phenomena (Ng *et al* 2010, Navoret *et al* 2013). Tools based on nonlinear dynamics and fractal theory applied to EGM signals have been reported to detect CFAE or to quantify the level of fractionation (Hoekstra *et al* 1995, Novak *et al* 2009). However, nonlinear methods are not yet used in clinical practice.

To contribute to the field of the nonlinear analysis of CFAEs, the aim of this study is to evaluate multifractal analysis (MF) for characterising fractionation phenomena in electrograms. Fractal signals present self-similarities and scale invariance properties which can be described by a single quantity, e.g. the Hausdorff dimension (Pavlov and Anishchenko 2007) or the Hurst exponent (Eke *et al* 2002). In physiological signals, the fractal properties are not homogeneous (Ivanov *et al* 1999); for example, the local scaling properties in EGM signals change with time, and these signals are characterised by different local Hurst exponents. Multifractal analysis could capture these changes in the global singularity distributions. We reasoned that if there is a nonlinear dynamic behind the apparently random behaviour of AF

conduction patterns, a multifractal approach would be an appropriate mathematical tool for characterising it.

In the present study, we approach the performance of two multifractal methods to grade CFAEs, multifractal detrended fluctuation analysis (MF-DFA) (Kantelhardt *et al* 2002) and wavelet transform modulus maxima (WTMM) (Muzy *et al* 1993). We propose a new method for extracting features from the multifractal spectrum to quantify the level of fractionation. First, we analyse the behaviour of several methods for extracting MF features using synthetic MF signals. Second, we study the performance of these methods for discriminating between different levels of fractionation. Finally, we compare the MF-based measures with other well-established methods for detecting CFAEs.

2. Methods

2.1. Multifractal analysis

Multifractal signals exhibit various scaling properties or various long-range correlations with small and large fluctuations. Accordingly, statistical scaling is characterised by different Hurst exponents (h) and the fractal dimension can be generalised by $D(h)$ (Ivanov *et al* 1999), where $D(h_0)$ is the fractal dimension of the subset that is characterised by the local Hurst exponent h_0. The multifractal approach involves the concept of a singularity spectrum $f(\alpha)$. To understand $f(\alpha)$, suppose that we cover a set with balls of diameter ε; the singular spectrum $f(\alpha)$ characterises the dependence on ε of the number of covering elements N_α corresponding to the points with a singular exponent equal to some value of α:

$$N_\alpha \sim \varepsilon^{-f(\alpha)} \tag{1}$$

where ε is the diameter of the balls that covered the set, and $f(\alpha)$ corresponds to the Hausdorff dimension. Various approaches have been introduced for calculating $f(\alpha)$. They are usually based on calculating the generalised dimensions through a partition function,

$$Z_{(q,\varepsilon)} \sim \sum_{i=1}^{N(\epsilon)} \mu_i^q(\varepsilon) \tag{2}$$

defined as the sum of q powers of μ_i, where μ_i is the measure of the ith element that covers the set and $N(\varepsilon)$ is the number of cover elements of size ε (Pavlov and Anishchenko 2007). Typically Z satisfies the power law:

$$Z_{(q,\varepsilon)} \sim \varepsilon^{(q-1)D(h)} \tag{3}$$

where $D(h)$ is the generalised fractal dimension, and the scaling exponent is,

$$\tau(q) = (q-1)D(h) \tag{4}$$

If $q = 0$, equation (3) is equal to equation (1). Then, a monofractal dimension is acquired, which is a particular case of the multifractal approach. The partition function (equation (2)) for $q > 0$ reflects the scaling of the large fluctuation and strong singularities, and for $q < 0$ it reflects the scaling of the small fluctuation and weak singularities. In monofractal signals, $D(h)$ is constant and $\tau(q)$ is linear, whereas for multifractal signals $\tau(q)$ is a nonlinear function. The generalised Hurst exponent and the generalised fractal dimension are related to $\tau(q)$ as follows:

$$h(q) = \frac{\mathrm{d}\tau}{\mathrm{d}q} \tag{5}$$

$$D(h) = qh - \tau(q) \tag{6}$$

Two popular approaches for calculating the partition function and the generalised Hurst exponents are WTMM and MF-DFA.

2.2. Wavelet transform modulus maxima

In the WTMM method (Muzy *et al* 1993), lines of local extremes of the wavelet coefficient $W(a, t)$ are extracted. The partition function is built from the modulus maxima of the wavelet transform and is defined by,

$$Z_{(q,a)} = \sum_{l \in L(a)} [\sup_{(t,a') \in l} |W(t_l(a'), a')|]^q \tag{7}$$

Where a' corresponds to each wavelet scale, t stands for the time, t_l is the position of the maximum belonging to line l and $L(a)$ is the set of all the lines l that satisfy:

$$(t, a') \in l \rightarrow a' \leqslant a_0 \tag{8}$$

$$\forall\, a' \leqslant a_0, \exists (t, a') \in l \tag{9}$$

In the limit $a \rightarrow 0$, the power can be expressed as:

$$Z_{(q,a)} \sim a^{\tau(q)} \tag{10}$$

Note that here a, the wavelet scale, is analogous to ε in equation (3). The toolbox from Physionet was used for calculating WTMM (Goldberger *et al* 2002).

2.3. Multifractal detrend fluctuation analysis

MF-DFA (Kantelhardt *et al* 2002) uses the qth order fluctuation function,

$$F_q(s) \equiv \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(s, v)]^{q/2} \right\}^{\frac{1}{q}} \tag{11}$$

as a partition function, where $F^2(s, v)$ is the variance of local trend estimation, $v = 1, 2, \ldots, 2N_s$ are the segments in which the local trend is estimated and N_s is the number of segments with equal length s. Being $Y(i) \equiv \sum_{k=1}^{i} [x_k - \mu_x]$, $i = 1, ..., N$, where x_k is a series of length N, μ_x is the mean of $\boldsymbol{x} \in \mathbb{R}^N$. The variance is determined by

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} \{Y[(v-1)s + i] - y_v(i)\}^2 \tag{12}$$

for each segment, $v = 1, 2, ..., N_s$, and

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} \{Y[N - (v - N_s)s + i] - y_v(i)\}^2 \tag{13}$$

for $v = N_s + 1, ..., 2N_s$, where $y_v(i)$ is the fitting polynomial in segment v.

Taking into account the power-law,

$$F_q(s) \sim s^{h(q)} \tag{14}$$

the fluctuation function is analysed using log–log plots for each value of q,

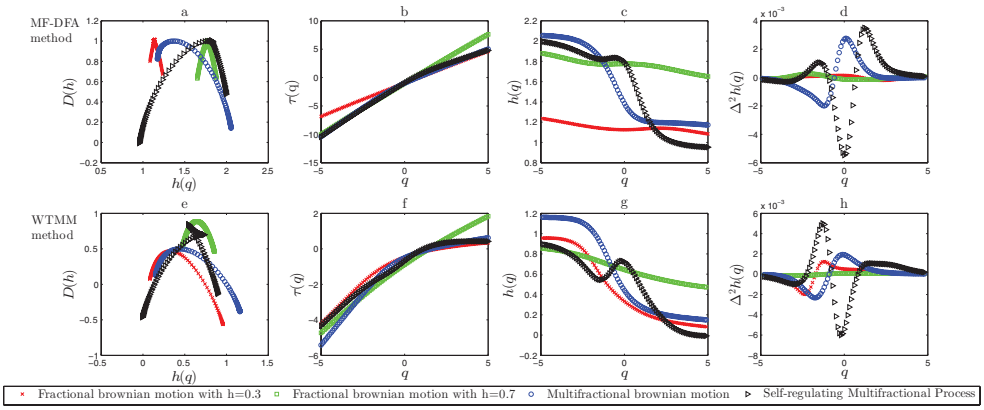


Figure 1. The multifractal spectra of synthetic signals (a) (e), $\tau(q)$ values (b)(f), $h(q)$ (c)(g) and d^2h/dq values (d) (h). (a)–(d) were calculated using MF-DFA, and (e)–(h) were calculated using WTMM. Note that in (a) the spectra are wider than monofractal signals in the case of multifractal signals. The shape of the spectra is related to $\tau(q)$ and to $h(q)$. Fluctuations in $h(q)$ can be analysed using the second derivate. hFI is calculated from the power of $\Delta^2 h(q)$.

$$h(q) = \frac{\log F_q(s)}{\log s} \tag{15}$$

The toolbox developed by Ihlen was used in calculating MF-DFA (Ihlen 2012).

2.4. Features extraction from a multifractal spectrum

Various indexes have been developed for identifying multifractal properties and for extracting useful information from the spectrum (Makowiec *et al* 2011). A test using synthetic signals was performed in order to interpret the properties of the multifractal spectra and the behaviour of the indexes. First, two samples of monofractal signals were built using a function to generate fractional Brownian motion signals from the Wavelet Toolbox in Matlab: the Hurst exponents $h = 0.7$ and $h = 0.3$ were used. Then, two samples of multifractal signals were synthesised using a FracLab toolbox for Matlab developed at INRIA (Vehel and Legrand 2004): (i) a self-regulating multifractional process with midpoint displacement, and (ii) a multifractional Brownian motion using the Wood and Chan circular matrix, with the holder exponent determined by a logistic function with 10 levels of pre-quantification.

Figures 1(a) and (e) show the multifractal spectra calculated from the four synthetic signals using MF-DFA and WTMM. Three indexes can be extracted directly from the spectra: (i) $h_{\max}$, which is defined as $h(q)$ where $D(h)$ is the maximum; (ii) the spectrum width Δ, defined as $\Delta = h_r - h_i$ where $h_r = h(q \to -\infty)$ and $h_i = h(q \to \infty)$. These indexes aim to differentiate between monofractal and multifractal signals based on the observation that the spectrum is narrower on monofractal signals. (iii) the asymmetric ratio (AR):

$$AR = \frac{h_r - h_{\max}}{\Delta}, \tag{16}$$

AR is defined as the ratio between h calculated with negative q and the total width of the spectrum. If the multifractal spectrum is symmetric, AR should be equal to 0.5. This index is based

on the *AR* described by Qiusheng *et al* (2012), and aims to detect the difference between $h(q)$ calculated with negative and positive q. Figure 1(a) shows that the spectra of the two synthetic multifractal signals (multifractional Brownian motion and the self-regulating multifractional process) exhibit differences between the right tails and left tails.

2.5. *h*-fluctuation measure from the multifractal spectrum

The indexes explained above are calculated through an analysis of one or two points from the spectrum, which means that the remaining important information is discarded. The shape of the spectrum depends on the scaling exponent $\tau(q)$. In monofractal signals $\tau(q)$ is linear, which implies that $h(q)$ is a constant, while in multifractal signals $\tau(q)$ shows a nonlinear behaviour. Figures 1(b) and (f) reflect the behaviour of $\tau(q)$, and figures 1(c) and (g) illustrate how $h(q)$ tends to be constant in monofractal signals. In addition, $h(q)$ has more fluctuation when there is a zig-zag shaped spectrum. Note the shape of the spectrum in figure 1(e), corresponding to the signal generated by the self-regulating multifractional process, and compare it with the fluctuation in $h(q)$ in figure 1(g).

The fluctuation can be analysed through the second derivative. Figures 1(d) and (h) show the second order backward difference of $h(q)$ ($\Delta^2 h(q)$); note that the amplitude of $\Delta^2 h(q)$ in the case of multifractal signals is higher than for monofractal signals.

To extract suitable information from $h(q)$ we proposed the following measure, called the h-fluctuation index (hFI), which is defined as the power of the second derivative of $h(q)$,

$$hFI = \frac{1}{2|q_{\max}|+2} \sum_{q=q_{\min}-2}^{q_{\max}} [h(q) - 2h(q-1) + h(q-2)]^2 \tag{17}$$

where $q_{\min} < 0$, $q_{\max} > 0$ and $|q_{\min}| = |q_{\max}|$.

hFI tends to zero in high fractionation signals. *hFI* has no reference point when a set of signals is evaluated, so *hFI* must be normalised, using

$$hFI_n = (hFI - \min(hFI))/(\max(hFI) - \min(hFI)), \tag{18}$$

in order to provide a better understanding of the fractionation levels, where $hFI_n = 1$ is the most organised and the most regular signal in the set.

2.6. Global single measures

To evaluate the differences in performance between multifractal and fractal analysis, we evaluate two fractal-based measures. The correlation dimension with embedded dimension 3 and time delay 4 was evaluated using TSTOOL (Merkwirth *et al* 2009), and the fractal dimension was estimated using the Katz–Sevcik method (Sevcik 1998).

2.7. Established methods for detecting CFAE

To compare the proposed approach with the established method for detecting CFAE, we calculated $\mathrm{CFE_{mean}}$ and dominant frequency (DF). $\mathrm{CFE_{mean}}$ is a measure that is integrated into the 3-D electro-anatomical mapping system EnSite NavX–St. Jude Medical (Hunter *et al* 2009). $\mathrm{CFE_{mean}}$ measures the average of the time intervals between deflections and is based on the detection of negative deflections in EGM signals. Three parameters are used in this algorithm: (i) sensitivity, which is the minimum peak-to-peak voltage amplitude to mark the deflection, (ii) deflection width - the time between the maximum peak and the minimum peak in the

deflection must be lower than this parameter to avoid detecting a broad ventricular event, and (iii) the refractory period, which is the minimum time between two consecutive deflections. In order to have a high performance using CFA_{mean}, all the signals were normalised. Normalisation is important to offset the gaps in the acquisition parameters between signals in the database. Additionally, this process reduces inaccuracy in the measurements due to the variation in amplitude generated by the different orientation of the electrodes. The following parameters were used: sensitivity 0.1, deflection width 20 ms and refractory period 30 ms. The CFE_{mean} calculation was implemented in Matlab according to the description given by Almeida *et al* (2013).

DF is defined as the maximum peak in the power spectrum. Before calculating DF, the signals were preprocessed according to the scheme proposed by Botteron and Smith (1995) as follows: (i) band pass filtering (40–250 Hz), (ii) rectification, (iii) low pass filtering (20 Hz) and the fast Fourier transform with Hamming windows (Everett *et al* 2001).

2.8. Experimental data set

Two different independent EGM databases, already established and reported in previous works (Schilling *et al* 2015, Kremen *et al* 2008), were used. The first database consists of 113 records from the Institute of Clinical and Experimental Medicine, Prague, Czech Republic (CZ), recorded at a sampling rate of 997 Hz and resampled to 1 kHz. The second database consists of 429 records from Stadtisches Klinikum Karlsruhe in Germany (GE) recorded with a sampling rate of 1.2 KHz and resampled to 1 kHz. All the patients were indicated for radiofrequency ablation of AF. The data were filtered with a bandpass filter at 0.3–300 Hz, and the remaining baseline wander and high noise was reduced by wavelet decomposition according to Schilling (2012).

In CZ database EGMs were collected during left-atrial endocardial mapping using a 4 mm irrigated-tip ablation catheter in 12 patients with persistent AF. Data were recorded outside the pulmonary veins and their tubular ostia. The fragments with inadequate endocardial contact, a relatively high signal-to-noise ratio or artifacts were excluded. The EGM signals very close to the mitral annulus were discarded to prevent the interference of relatively sharp ventricular signals with atrial signal analysis.

GE database was constructed using signals from 11 patients. After PVI, bipolar electrograms were recorded using circular multipolar mapping catheters. All the signals were checked visually. Signals that were superimposed with ventricular far field or with ablation artifacts were sorted out. Signals that were not stationary over 5 s were also rejected.

The databases were independently evaluated and ranked by two different teams of electrophysiologists (EPs) from two different countries using the same criteria. Non-fractionated A-EGM signals were considered as level 0 (C0). The fractionated signals were categorised into three levels of fractionation: mild (C1), intermediate (C2) and high (C3), see figure 2. To define the four classes the Wells criterion (Wells *et al* 1978) was used; however, it was modified by swapping the order of types 3 and 4 in order to have a gradual transition between levels of fractionation. Signals from the GE database were labeled by two EPs and signals from the CZ database were labeled by three independent EPs. The EPs were asked to assign each EGM signal to one of four classes. The ranking was done manually by visual inspection of an individual EGM signal. Statistical analysis in this work is compared to this distribution of classes taken as a gold standard for the purpose of the study. A total of 542 signals were distributed into the fractionation classes as follows: 175 signals into C0, 117 signals into C1, 184 signals into C2 and 66 signals into C3.

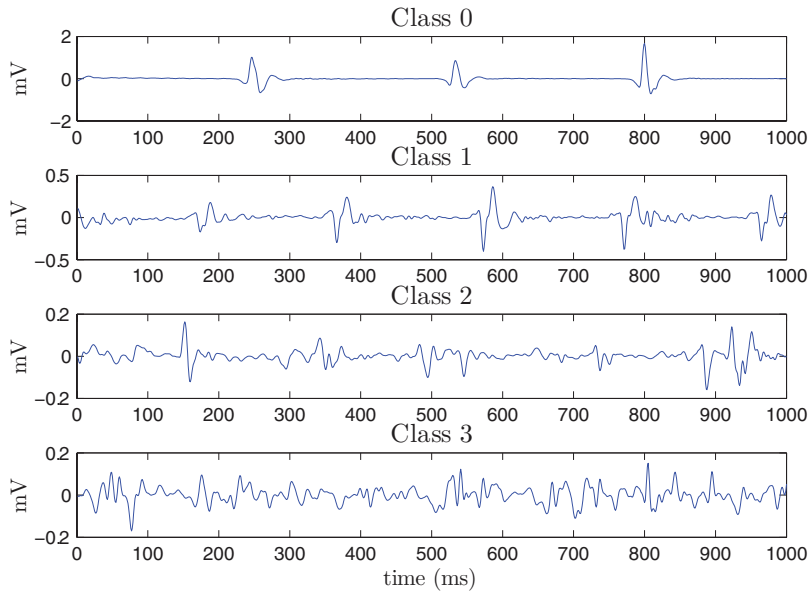


Figure 2. Samples of EGM signals from the GE database. From top to bottom, signals from classes C0, C1, C2 and C3.

2.9. Statistical analysis

ROC measurements were used to evaluate the performance of CFAE discrimination. Multifractal spectrum calculation was applied to signals from both databases. The tests were conducted as follows: first, a ROC curve was created taking into account the separation between non-fractionated signals and fractionated signals; i.e. C0 versus C1 + C2 + C3. The sensitivity and specificity were calculated using the cut-off point for the higher ROC area. Then, a ROC curve was created to separate the lower degree of fractionation from the higher degree, i.e. C0 + C1 versus C2 + C3. Finally, a ROC curve was created to detect the high degree of fractionation C3.

To check the performance of each index, to minimise the dispersion of each class and to maximise the distances between different class measures, we evaluated medians and quartiles 1 and 3 from the distribution of the results for each index.

An important parameter in multifractal analysis is ε, which corresponds to the segment length in MF-DFA and to the scales in WTMM. To find the optimal parameter for MF-DFA, a combination of the following parameters was used: segment length s minimum [8, 16, 32, 64], s maximum [128, 256]. For WTMM, a combination of the following parameters was used: min-scale [1, 2, 4, 8] and, max-scale [16, 32, 64].

2.10. Computing Platform

Calculations were performed using MatLab 2013a in a PC with Windows 8 (64 bits), Core I7 processor and a RAM of 6 GB.

3. Results

The optimal parameters for generating the multifractal spectrum and for calculating the evaluated indexes are shown in table 1. Note that in the case of MF-DFA the optimal parameters

Table 1. Average of sensitivity (Se) and specificity (Sp) in the discrimination of fractionation levels using a cut-off point for the higher ROC area.

		Parameters (%)		Mean	Mean
Feature		$\text{Scale}_{\text{min}}$	$\text{Scale}_{\text{max}}$	Se	Sp
MF-DFA	*hFI*	64	128	86.03	87.9
	AR	8	256	83.63	76.2
	h_{max}	64	256	73.56	84.0
	Δ	64	256	84.13	76.4
WTMM	*hFI*	4	16	82.1	78.0
	AR	8	64	81.1	73.1
	h_{max}	8	64	72.5	67.7
	Δ	8	32	73.16	51.9

Note: optimal parameters for the best Se and Sp were found for each index.

do not include the shorter segments. For WTMM better results are achieved if fine scales (lower than scale 4) are not included. Fine scales and shorter segments can contain mainly high-frequency noise.

Figures 3(a) and (b) show the multifractal spectra and the second derivative of $h(q)$, calculated by applying MF-DFA with quadratic detrending in a signal length of 1500 samples, using segment sizes between 32 and 128 samples. Note that there is a tendency for more organised signals (C0) to show a spectrum with a zig-zag shape, whereas fragmented signals present a spectrum with a bell-like shape. A zig-zag type spectrum corresponds to higher amplitude in $\Delta^2 h(q)$. Highly fractionated signals present a narrow spectrum which tends toward zero in $\Delta^2 h(q)$.

Figures 3(c) and (d) show the multifractal spectra and $\Delta^2 h(q)$ calculated through WTMM using a wavelet Gaussian order 3, a minimum scale of 2, and maximum scales of 64. A zig-zag shape is observed in the more organised signals from class 1 and class 2. Class 0 signals show the highest amplitude in $\Delta^2 h(q)$. The width of the spectrum is not so marked as in the MF-DFA based spectrum.

Table 2 shows the results for multifractal feature analysis applied to the entire database with signal length $N = 1500$ ms. A comparison with three different types of measures is shown, (i) *hFI* presents a better performance than the other multifractal indexes, such as *AR*, h_{max} and Δh. Note that *hFI* achieves a better sensitivity and specificity when calculated from the MF-DFA spectrum than when calculated from WTMM; (ii) the sensitivity and the specificity of *hFI* is better than fractal measures such as the correlation dimension and the fractal dimension; (iii) the results of *hFI* are better than other well established methods such as DF and CFE_{mean}.

Our results indicate that discrimination between regular activity (Class C0) and fractionated EGM (C1, C2 and C3) can be accurately resolved by *AR*. However, when distinguishing between an intermediate level of fractionation and a high level of fractionation, *hFI* presents a better relation between specificity and sensitivity than *AR*.

Figure 4 shows boxplots for *hFI* and two established methods: CFE_{mean} and DF. Figure 4(a) shows the boxplot for normalised *hFI*, note that high fractionated electrograms are packed in a narrow windows near to zero. For better visualisation of the *hFI* distribution in the boxplot, Figure 4(b) shows the $\log_{10} hFI_n$. Note that the distance between the distributions is maximised and the dispersion is minimised. Figures 4(c) and (d) show the boxplots of well-established

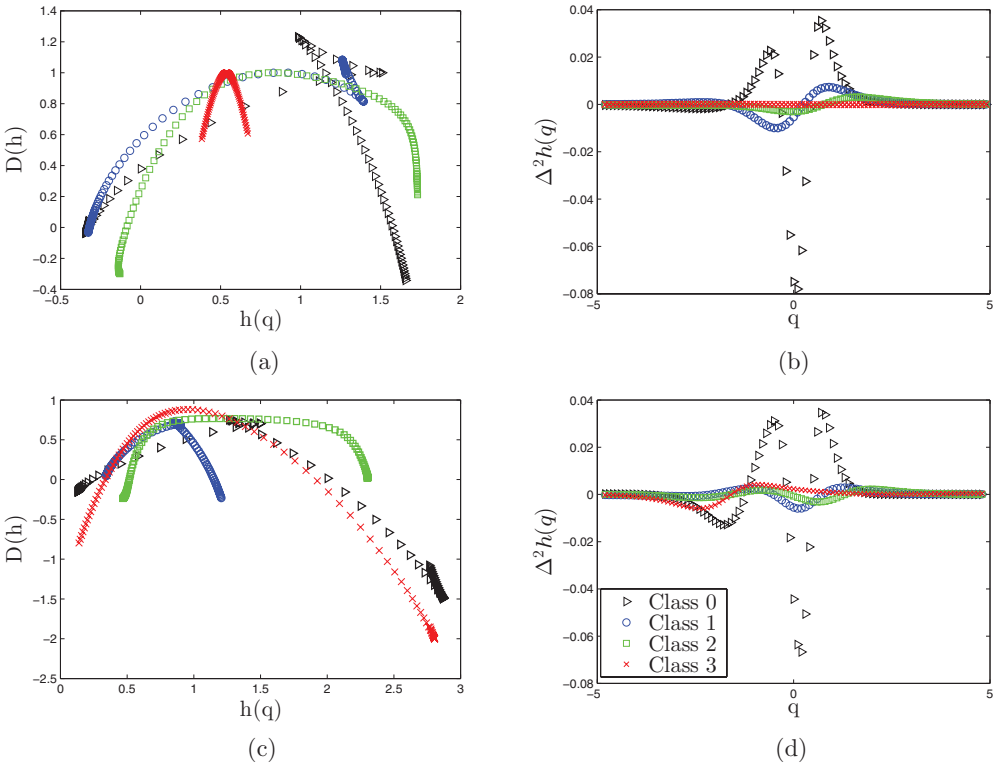


Figure 3. Top: MF-DFA sigularity spectrum of four signals from the EGM database. Non-fractionated signals (class 0) and signals with a mild level of fractionation (class 1) have a zig-zag and asymmetric shape in the spectrum, while the spectrum of the most fractionated signals (class 3) is thin and symmetrical. Bottom: WTMM singularity spectrum of four signals from the EGM database. Note that higher classes of fractionation present a bell-like shape. However, there is no clear trend in the width of the spectrum.

methods. Considering $Q1_i$ and $Q3_i$, where $i = 1, 2, 3$ is the class, the best separation between classes is achieved if $Q1_i$ and $Q3_{i+1}$ do not overlap. Note that only *hFI* fulfills this property.

In order to study the robustness of multifractal analysis under variations of signal length N, the indexes were computed using MF-DFA with the following N values: 500, 750, 1000, 1500, 2000, 3000 and 5000. Figures 5(a) and (b) show the performance of features versus signal length N. The results show that the performance is similar when windows larger than 500 points are used, i.e. $N > 500$, see figures 5(b) and (a). To extract information about the dynamics of the EGM signal, the analysis must include one or more activation wave in the time windows in order to have enough relevant information in the segment.

The computational time required to calculated *hFI* using the MF-DFA approach is 0.5 s for a signal of length 1500 ms and 1.6 s for a signal of length 5000 ms.

4. Discussion

This study reports the use of multifractal analysis for evaluating the degree of fractionation in AF electrograms. The principal findings of the study are: (i) the shape of the multifractal

Table 2. Sensitivity (Se) and specificity (Sp) of different indexes to discriminated different levels of fractionation: not fractionated (C0), intermediate and high level of fractionation (C2 + C3) and high level of fractionation (C3).

		C0		C2 + C3		C3	
Feature		Se	Sp	Se	Sp	Se	Sp
MF-DFA	*hFI*	87.2	86.9	87.6	88.4	83.3	88.4
	AR	92.1	84.0	74.0	78.4	84.8	66.2
	$h_{\max}$	71.7	88.6	74.8	80.8	74.2	82.6
	Δ	79.0	72.0	84.0	74.0	89.4	83.2
WTMM	*hFI*	78.7	83.2	81.2	82.9	86.4	67.9
	AR	77.1	81.1	77.2	72.6	81.8	65.7
	$h_{\max}$	68.7	59.4	71.6	67.1	77.2	76.7
	Δ	73.6	60.6	73.2	47.6	72.7	47.5
Correlation dim.		87.0	80.0	84.0	80.8	84.0	78.3
Fractal dim.		78.2	75.4	81.2	62.3	77.3	76.3
Dominant frequency		52.6	75.4	56.8	67.8	50.0	87.0
CFE mean		93.5	78.3	83.6	76.4	74.2	85.1

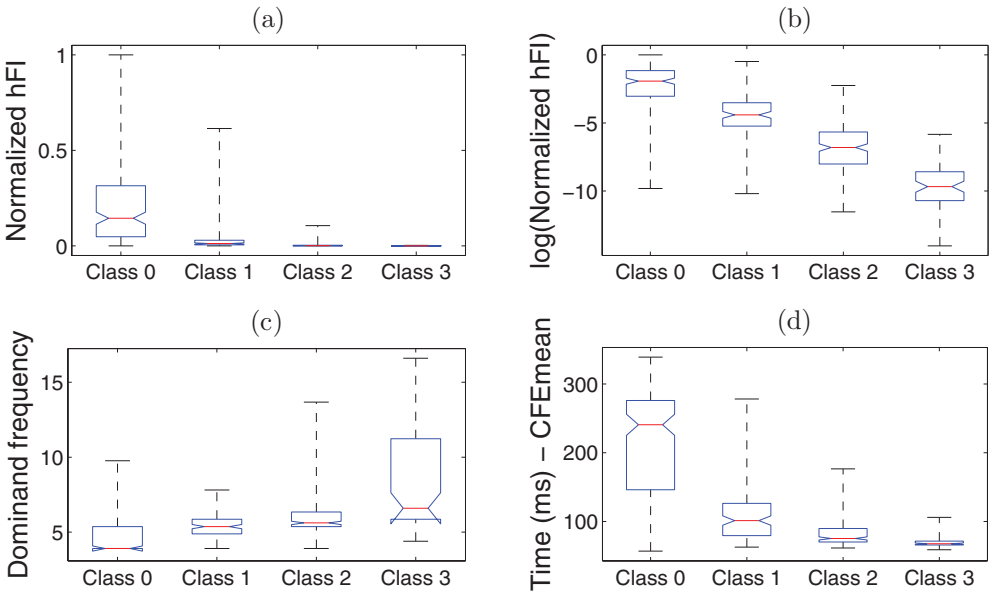


Figure 4. (a) Boxplot of hFI calculated using MF-DFA. (b) Boxplot of $\log_{10}(hFI)$. (c) Boxplot of DF. (d) Boxplot of $\mathrm{CFE}_{\mathrm{mean}}$. Note that better separation between medians and quantiles are achieved using *hFi* or log(*hFI*).

spectrum contains information for grading fractionation in AF electrograms. (ii) A new measure, *hFI*, has been proposed to offer accurate quantification in a single descriptor of the information contained in the spectrum shape.

A multifractal approach has been used to differentiate between healthy and pathological conditions in numerous physiological signals related to physiological control or self-regulated processes, such as: inter-intervals in human walking (Hausdorff 2007), neuronal activity

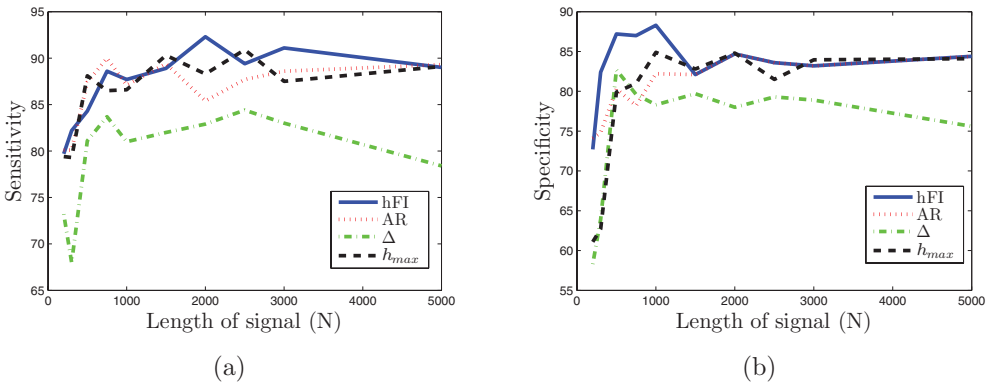


Figure 5. (a) Sensitivity versus length of the signals. (b) Specificity versus length of the signals.

(Zheng *et al* 2005) and heart rhythm (Ivanov *et al* 1999, Goldberger *et al* 2002). MF has also been used, although to a lesser extent, to study complexity from surface electrocardiograms (Wang *et al* 2001, Yang *et al* 2007). However, to the best of our knowledge multifractal analysis has not been reported in the study of fractionation in EGM signals during AF.

We implemented two approaches, MF-DFA and WTMM, to study multifractal properties. The results in table 2 indicate that *hFI* calculated from MF-DFA outperforms the same measure calculated by WTMM. This result is consistent with Kantelhardt *et al* (2002), who proposed and tested MF-DFA and concluded that multifractal scaling behaviour determined by MF-DFA performs better than WTMM for short series and negative moments. Accordingly, our findings suggest that fractionated EGM signals are characterised mainly by weak singularities, which are related to negative values of q. In addition, note that the multifractal analysis was evaluated using short time windows (1500 ms). The computational time is less than the length of the time window, allowing the practical application of this tool. A short time evaluation is important for developing a method capable of tracking unstable fibrillatory conduction.

Several studies have used monofractal measures such as correlation dimension, DFA or fractal dimension to analyse electrograms during AF (Hoekstra *et al* 1995, Novak *et al* 2009). Our findings show that the MF-based feature *hFI* performs better than other monofractal indexes and other fractal measures. In physiological signals, the fractal properties are not homogeneous; single global estimates such as DFA or the correlation dimension of scaling are therefore not enough. Local scaling properties in EGM signals change with time, and these signals are characterised by different local Hurst exponents. Multifractal analysis could therefore be a better descriptor of global singularity distributions. In addition, the multifractal framework offers improved extraction of hidden information about the singularities and fluctuations of EGM signals. Therefore, multifractal analysis can help in understanding fractionation phenomena in AF electrograms. Furthermore, *hFI* can help in differentiating between a bell-like shape and a zig-zag shape. This information can be used for grading fractionation.

hFI can capture information about the spectrum width. Non-fractionated and weakly fractionated signals exhibit a wider spectrum. This may be related to differences in the type of singularities present in the activation waves and in the baseline. Highly fractionated electrograms show a narrower spectrum, indicating that class C3 signals manifest some monofractal properties; the explanation may lie in the dynamics of fibrillatory conduction patterns during AF.

Our results are consistent with some studies which have suggested that fibrillation is a form of spatio-temporal chaos (Garfinkel *et al* 1997). Moreover, Gray *et al* (1998) showed that

transmembrane signals recorded from the surface of rabbit and sheep hearts exhibited attractors in the phase space during fibrillation, and they suggest that identifying phase singularities could help in the analysis of spatio-temporal fibrillation patterns.

The performance of the methods was tested using databases scored by skilled EPs. Furthermore, it was also tested against commonly established methods, such as CFE_{mean} and DF, which were calculated using our datasets (section 2.7). In comparison with well-established methods, *hFI* exhibits sensitivity and specificity values higher than CFE_{mean} and DF. Several authors have discussed the practical usefulness of the analysis of DF because this index is sensitive to the method of signal acquisition, the signal-to-noise ratio (SNR), the far-field ventricular depolarisation and the complexity of the signal (Singh *et al* 2010, Elvan *et al* 2009). Additionally, the DF index was developed to represent the average cycle length. However, when the signal is highly fractionated, the DF peak may not represent the cycle length. On the other hand, CFE_{mean} has been used as a gold standard method based on the basis of a study by Hunter *et al* (2009), who compared several well-established methods for detecting CFEAs and found that the best performance was achieved by CFE_{mean} (Hunter *et al* 2009). CFE_{means} is one of the most used index because it is implemented in the commercial system Ensity NavX (St. Jude Medical). However, the performance of the CFA_{mean} depends on the parameter settings and on the precision with which the deflections are estimated, using time interval criteria. This issue is especially important in highly disorganised signals, where it is difficult to differentiate between activation waves and other peaks. This could be one of the reasons why ablation guided by CFAEs is a disputed method.

Some studies suggest that only a high level of fractionation is related to arrhythmogenic substrates. The performance of *hFI* is notably higher in the discrimination of class C3 signals in comparison with CFE_{means} and *DI*. This is important because Class C3 is the highest fragmented group and could be associated with critical sites in persistent AF. The degree of fractionation of the A-EGM signals is, in reality, assumed to be naturally continuous. Nevertheless a discrete set of levels of fractionation is used in this study due to the impossibility of classifying signals by experts on a smoother scale. Even so, hFI could be used as a variable to measure the fractionation on a continuous scale to discriminate the critical signals without relying on a specific scale to categorise the level of fractionation.

Other studies have reported specificity and sensitivity values higher than 90% in the discrimination of fractionated signals (Barbaro *et al* 2000, Hunter *et al* 2009). However, the comparison of performance using different database could be imprecise. There are different considerations in the databases used. Other authors used only two or three classes, while we used four. We used a multicentre database and we considered signals recorded using different catheters, while other studies used signals recorded under the same conditions. Therefore, we consider that implementation and calculation of all indexes using the same database is the best way to make a comparison. We do not consider the performance reported in studies using classifier based on machine learning (Nollo *et al* 2008) for comparison purposes, because they used several indexes. Therefore, *hFI* could be used as a feature in these kind of systems to improve the results. Moreover, according to the boxplot distribution shown in figure 4(b), we recommend the use of log(*hFI*).

Although there is apparently a random pattern of activation in AF, there should exist a hidden dynamic behind these patterns. Recently, other complexity and entropy measures, such as Shannon entropy, recurrence quantification analysis and approximate entropy (ApEn) have been used by various authors to discriminate CFAEs, and to offer a promise as an approach for locating sites that represent arrhythmogenic substrates (Ng *et al* 2010, Ganesan *et al* 2013, Ugarte *et al* 2014). Our results provide additional evidence that fractionation phenomena could be evaluated using nonlinear dynamics theory to provide a better understanding of

the fractionation mechanism. Our analysis suggests that the use of tools based on nonlinear dynamics is promising for studies of the fractionation phenomena in AF, and could improve the accuracy in grading fractionation, especially in the presence of highly-fractionated signals. Further studies are needed to relate fractionation and arrhythmogenic substrates and their nonlinear and multifractal properties.

5. Conclusion

The multifractal spectrum is a promising tool for studying the dynamics behind the fractionation phenomenon in AF electrograms. In this work we have defined *hFI* as a new measure that improves the extraction of information about the shape of the spectrum. *hFI* presents the relationship between the power of $\Delta^2 h(q)$ and the information existing in the spectrum. Our results indicate that *hFI* is capable of grading fractionation better than other fractal and multifractal indexes.

The index proposed here can be used for detecting different levels of fractionation in EGM signals in order to improve the localisation of possible target sites for ablation in the treatment of atrial fibrillation. Further studies are needed to test the clinical relevance of our method.

Acknowledgments

We would like to acknowledge Prof O Dossel and his laboratory at the Institute of Biomedical Engineering at the Karlsruhe Institute of Technology in Germany, and also Dr A Luik from the Department of Cardiology and the Department of Internal Medicine at the Karslruhe City Hospital in Germany for providing the databases of AF signals. We also appreciate the knowledge and data provided by Dr D Wichterle and his colleagues at the Institute of Clinical and Experimental Medicine in Prague, Czech Republic.

This work has been supported by COLCIENCIAS—Republica de Colombia, project No. 121056933647, and by Czech Science Foundation postdoctoral GACR research project GACR P103/11/P106.

References

Almeida T P, Salinet J L, Chu G S, Ng G A and Schlindwein F S 2013 Different definitions of complex fractionated atrial electrograms do not concur with the clinical perspective *Comput. Cardiol.* **2013** 1055–8

Barbaro V, Bartolini P, Calcagnini G, Censi F, Morelli S and Michelucci A 2001 Mapping the organization of atrial fibrillation with basket catheters. Part I: validation of a real-time algorithm *Pacing Clin. Electrophysiol.* **24** 1082–8 (PMID: 11475823)

Barbaro V, Bartolini P, Calcagnini G, Morelli S, Michelucci A and Gensini G 2000 Automated classification of human atrial fibrillation from intraatrial electrograms *Pacing Clin. Electrophysiol.* **23** 192–202

Berenfeld O and Jalife J 2011 Complex fractionated atrail electrograms: is this the beast to tame in atrial fibrillation? *Circ. Arrhythm. Electrophysiol.* **4** 426–8

Botteron G and Smith J 1995 A technique for measurement of the extent of spatial organization of atrial activation during atrial fibrillation in the intact human heart *IEEE Trans. Biomed. Eng.* **42** 579–86

Chang S L, Chen Y C, Hsu C P, Kao Y H, Lin Y K, Lin Y J, Wu T J, Chen S A and Chen Y J 2013 Electrophysiological characteristics of complex fractionated electrograms and high frequency activity in atrial fibrillation *Int. J. Cardiol.* **168** 2289–99

Eke A, Herman P, Kocsis L and Kozak L R 2002 Fractal characterization of complexity in temporal *Physiol. Meas.* **23** R1–38

Elvan A, Linnenbank A C, van Bemmel M W, Misier A R R, Delnoy P P H M, Beukema W P and de Bakker J M T 2009 Dominant frequency of atrial fibrillation correlates poorly with atrial fibrillation cycle length *Circ. Arrhythmia Electrophysiol.* **2** 634–44

Everett T, Kok L C, Vaughn R, Moorman R and Haines D 2001 Frequency domain algorithm for quantifying atrial fibrillation organization to increase defibrillation efficacy *IEEE Trans. Biomed. Eng.* **48** 969–78

Faes L, Nollo G, Antolini R, Gaita F and Ravelli F 2002 A method for quantifying atrial fibrillation organization based on wave-morphology similarity *IEEE Trans. Biomed. Eng.* **49** 1504–13

Ganesan A N *et al* 2013 Bipolar electrogram Shannon entropy at sites of rotational activation: implications for ablation of atrial fibrillation *Circ. Arrhythmia Electrophysiol.* **6** 48–57

Garfinkel A *et al* 1997 Quasiperiodicity and chaos in cardiac fibrillation *J. Clin. Invest.* **15** 305–14

Goldberger A L, Amaral L A N, Hausdorff J M, Ivanov P C, Peng C K and Stanley H E 2002 Fractal dynamics in physiology: alterations with disease and aging *Proc. Natl Acad. Sci. USA* **99** 2466–72

Gray R, Pertsov A and Jalife J 1998 Spatial and temporal organization during cardiac fibrillation *Nature* **392** 75–8

Hausdorff J M 2007 Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking *Hum. Mov. Sci.* **26** 555–89

Hoekstra B P T, Diks C G H, Allessie M A and DeGoede J 1995 Nonlinear analysis of epicardial atrial electrograms of elelectrical induced atrial fibrillation in man *J. Cardiovasc. Electrophysiol.* **6** 419–40

Hunter R J, Diab I, Thomas G, Duncan E, Abrams D, Dhinoja M, Sporton S, Earley M J and Schilling R J 2009 Validation of a classification system to grade fractionation in atrial fibrillation and correlation with automated detection systems *Europace* **11** 1587–96

Ihlen E A F 2012 Introduction to multifractal detrended fluctuation analysis in Matlab *Front. Physiol.* **3** 141

Ivanov P C, Nunes Amaral L A, Goldberger A L, Havlin S, Rosenblum M, Struzik Z and Stanley E 1999 Multifractality in human heartbeat dynamics *Nature* **399** 461–5

Kalifa J *et al* 2006 Mechanisms of wave fractionation at boundaries of high-frequency excitation in the posterior left atrium of the isolated sheep heart during atrial fibrillation *Circulation* **113** 626–33

Kantelhardt J W, Zschiegner S A, Koscielny-bunde E, Bunde A, Havlin S and Stanley H E 2002 Multifractal detrended fluctuation analysis of nonstationary time series *Physica* A **316** 87–101

Konings K T, Smeets J L, Penn O C, Wellens H J and Allessie M A 1997 Configuration of unipolar atrial electrograms during electrically induced atrial fibrillation in humans *Circulation* **95** 1231–41

Kremen V, Lhotska L, Macas M, Cihak R, Vancura V, Kautzner J and Wichterle D 2008 A new approach to automated assessment of fractionation of endocardial electrograms during atrial fibrillation *Physiol. Meas.* **29** 1371–81

Lin Y J *et al* 2013 Nonlinear analysis of fibrillatory electrogram similarity to optimize the detection of complex fractionated electrograms during persistent atrial fibrillation *J. Cardiovasc. Electrophysiol.* **24** 280–9

Makowiec D, Rynkiewicz A, Gaaaska R, Wdowczyk-Szulc J and Zarczynska-Buchowiecka M 2011 Reading multifractal spectra: aging by multifractal analysis of heart rate *Europhys. Lett.* **94** 68005

Merkwirth C, Parlitz U, Wedekind I, Engster D and Lauterborn W 2009 TSTOOL (www.physik3.gwdg.de/tstool/index.html)

Michelucci A, Bartolini P, Calcagnini G, Censi F, Colella A, Morelli S, Padeletti L, Pieragnoli P and Barbaro V 2001 Mapping the organization of atrial fibrillation with basket catheters. part ii: regional patterns in chronic patients *Pacing Clin. Electrophysiol.* **24** 1089–96

Muzy J F, Bacry E and Arneodo A 1993 Multifractal formalism for fractal signals: the structure-function approach versus the wavelet-transform modulus-maxima method *Phys. Rev.* E **47** 875–83

Nademanee K, McKenzie J, Kosar E, Schwab M, Sunsaneewitayakul, Vasavakul B, Khunnawat C and Ngarmukos T 2004 A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate *J. Am. Coll. Cardiol.* **43** 2044–53

Narayan S M, Krummen D E, Shivkumar K, Clopton P, Rappel W J and Miller J M 2012 Treatment of atrial fibrillation by the ablation of localized sources: confirm *J. Am. Coll. Cardiol.* **60** 628–36

Navoret N, Jacquir S, Laurent G and Binczak S 2013 Detection of complex fractionated atrial electrograms using recurrence quantification analysis *IEEE Trans. Biomed. Eng.* **60** 1975–82

Ng J, Borodyanskiy A I, Chang E T, Villuendas R, Dibs S, Kadish A and Goldberger J 2010 Measuring the complexity of atrial fibrillation electrograms *J. Cardiovasc. Electrophysiol.* **21** 649–55

Nollo G, Marconcini M, Faes L, Bovolo F, Ravelli F and Bruzzone L 2008 An automatic system for the analysis and classification of human atrial fibrillation patterns from intracardiac electrograms *IEEE Trans. Biomed. Eng.* **55** 2275–85

Novák D, Kremen V, Cuesta D, Schmidt K, Chudácek V and Lhotská L 2009 Discrimination of endocardial electrogram disorganization using a signal regularity analysis *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2009** 1812–5

Pavlov A N and Anishchenko V S 2007 Multifractal analysis of complex signals *Phys.—Usp.* **50** 819–34

Qiusheng L, Weixin X and Cheng L 2012 Identification of aircraft targets based on multifractal spectrum features *Proc. of the IEEE Int. Conf. Signal Process.* pp 1821–4

Sanders P *et al* 2005 Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans *Circulation* **112** 789–97

Scherr D, Dalal D, Cheema A, Cheng A, Henrikson C A, Spragg D, Marine J E, Berger R D, Calkins H and Dong J 2007 Automated detection and characterization of complex fractionated atrial electrograms in human left atrium during atrial fibrillation *Heart Rhythm* **4** 1013–20

Schilling C 2012 *Analysis of Atrial Electrograms* (*Karlsruhe Transactions on Biomedical Engineering* vol 17) (Karlsruhe: KIT Scientific Publishing)

Schilling C, Keller M, Scherr D, Oesterlein T, Haissaguerressaguerre M, Schmitt C, Dossel O and Luik A 2015 Fuzzy decision tree to classify complex fractionated atrial electrograms *Biomed. Tech.* **60** 245–55

Sevcik C 1998 A procedure to estimate the fractal dimension of waveforms *Complexity Int.* **5** 1–19

Singh J P, Ptaszek L M and Verma A 2010 Elusive atrial substrate: complex fractionated atrial electrograms and beyond *Heart Rhythm* **7** 1886–90

Takahashi Y *et al* 2008 Characterization of electrograms associated with termination of chronic atrial fibrillation by catheter ablation *J. Am. Coll. Cardiol.* **51** 1003–10

Ugarte J P, Orozco-Duque A, Tobón C, Kremen V, Novak D, Saiz J, Oesterlein T, Schmitt C, Luik A and Bustamante J 2014 Dynamic approximate entropy electroanatomic maps detect rotors in a simulated atrial fibrillation model *Plos One* **9** e114577

Vehel J L and Legrand P 2004 Signal and image processing with FracLab *Proc. of FRACTAL04, Complexity and Fractals in Nature, 8th Int. Multidisciplinary Conf.* 321–2

Wang Y, Zhu Y S, Thakor N V and Xu Y H 2001 A short-time multifractal approach for arrhythmia detection based on fuzzy neural network *IEEE Trans. Biomed. Eng.* **48** 989–95

Wells J L Jr, Karp R B, Kouchoukos N T, MacLean W A, James T N and Waldo A L 1978 Characterisation of atrial fibrillation in man: studies follwing open heart surgery *Pacing. clin. Electrophysiol.* **1** 426–38 (PMID: 95635)

Yang X, Ning X and Wang J 2007 Multifractal analysis of human synchronous 12-lead ECG signals using multiple scale factors *Physica A: Stat. Mech. Appl.* **384** 413–22

Zheng Y, Gao J, Sanchez J C, Principe J C and Okun M S 2005 Multiplicative multifractal modeling and discrimination of human neuronal activity *Phys. Lett.* A **344** 253–64

Zlochiver S, Yamazaki M, Kalifa J and Berenfeld O 2008 Rotor meandering contributes to irregularity in electrograms during atrial fibrillation *Heart Rhythm* **5** 846–54

PLOS | ONE

RESEARCH ARTICLE

Dynamic Approximate Entropy Electroanatomic Maps Detect Rotors in a Simulated Atrial Fibrillation Model

Juan P. Ugarte[1]*, Andrés Orozco-Duque[1], Catalina Tobón[2], Vaclav Kremen[3,4], Daniel Novak[3], Javier Saiz[5], Tobias Oesterlein[6], Clauss Schmitt[7], Armin Luik[7], John Bustamante[1]

1. Centro de Bioingeniería, Universidad Pontificia Bolivariana, Medellín, Colombia, **2.** GI^{2}B, Instituto Tecnológico Metropolitano, Medellín, Colombia, **3.** Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Prague, Czech Republic, **4.** Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic, **5.** I3BH, Universitat Politècnica de València, Valencia, Spain, **6.** Institute of Biomedical Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, **7.** Medizinische Klinik IV, Staedtisches Klinikum Karlsruhe, Karlsruhe, Germany

*juan.ugartem@alfa.upb.edu.co

OPEN ACCESS

Citation: Ugarte JP, Orozco-Duque A, Tobón C, Kremen V, Novak D, et al. (2014) Dynamic Approximate Entropy Electroanatomic Maps Detect Rotors in a Simulated Atrial Fibrillation Model. PLoS ONE 9(12): e114577. doi:10.1371/journal.pone.0114577

Editor: Alexander V. Panfilov, Gent University, Belgium

Received: June 13, 2014

Accepted: November 11, 2014

Published: December 9, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. The data are available at https://github.com/andresfod/Atrial_Electrograms_ApEn.

Funding: JPU, CT, and JB were partially supported by Departamento Administrativo de Ciencia, Tecnología e Innovación de la Republica de Colombia (www.colciencias.gov.co), project #121056933647; AO was supported by the Programa de Formación de Investigadores Francisco Jose de Caldas (www.colciencias.gov.co); VK and DN were partially supported by research project #MSM6840770012 Interdisciplinary Biomedical Engineering Research II from the Ministry of Education (www.msmt.cz), Youth and Sports of the Czech Republic, and VK was partially supported by post-doctoral research project GACR #P103/11/P106 of the Czech Science Foundation (www.gacr.cz). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

There is evidence that rotors could be drivers that maintain atrial fibrillation. Complex fractionated atrial electrograms have been located in rotor tip areas. However, the concept of electrogram fractionation, defined using time intervals, is still controversial as a tool for locating target sites for ablation. We hypothesize that the fractionation phenomenon is better described using non-linear dynamic measures, such as approximate entropy, and that this tool could be used for locating the rotor tip. The aim of this work has been to determine the relationship between approximate entropy and fractionated electrograms, and to develop a new tool for rotor mapping based on fractionation levels. Two episodes of chronic atrial fibrillation were simulated in a 3D human atrial model, in which rotors were observed. Dynamic approximate entropy maps were calculated using unipolar electrogram signals generated over the whole surface of the 3D atrial model. In addition, we optimized the approximate entropy calculation using two real multi-center databases of fractionated electrogram signals, labeled in 4 levels of fractionation. We found that the values of approximate entropy and the levels of fractionation are positively correlated. This allows the dynamic approximate entropy maps to localize the tips from stable and meandering rotors. Furthermore, we assessed the optimized approximate entropy using bipolar electrograms generated over a vicinity enclosing a rotor, achieving rotor detection. Our results suggest that high approximate entropy values are able to detect a high level of fractionation and

to locate rotor tips in simulated atrial fibrillation episodes. We suggest that dynamic approximate entropy maps could become a tool for atrial fibrillation rotor mapping.

Introduction

Catheter ablation based on mapping procedures has revolutionized the treatment of atrial fibrillation (AF). Electroanatomical mapping for guided AF ablation provides a 3D reconstruction of the cardiac chambers together with electrical information obtained from electrograms (EGM). Mappings of activation waves, voltage, dominant frequency and complex fractionated atrial electrograms (CFAE) are used to localize target sites for ablation. However, there are limitations with these techniques, which depend heavily on the expertise of the electrophysiologist [1].

CFAE mapping is still a debated technique [2]. CFAE is a physiopathological concept that was introduced by Nademanee [3]. However, this concept is broadly and unclearly defined, and involves inherent subjectivity [4]. This can lead to incorrect detection of target sites for ablation, mistaking EGM that are fractionated and functional in nature [5]. It also makes studies difficult to compare. While the concept of CFAE has made a relevant contribution to the study of AF, it may fail to describe the wide range of EGM fractionation that occurs in specific cases. In addition, inconsistent results have been found using the CFAE concept. Taking this into account, recent studies have helped to understand the concept of CFAE as a nonlinear phenomenon for quantifying various CFAE patterns, without using cycle length criteria [6–8].

Studies have shown that sites representing AF substrates are characterized by a high degree of disorganization in EGM [9], and, accordingly, methods for EGM signal processing are being designed to quantify the degree of fractionation of EGM [7, 8]. The relationship between CFAE and the rotor tip has been reported in recent studies [7, 10–13], but automatic rotor mapping methods have not been fully established.

The rotor hypothesis described in [14] proposes that in a significant number of patients, a rotor or a small number of rotors are drivers which maintain the arrhythmia. A rotor is a vortex of a spiral wave rotating around an unexcitable core. Narayan et al have provided evidence that human AF can be sustained by localized rotors [15, 16].

Based on the rotor hypothesis, which surmises that sustained AF depends on uninterrupted periodic activity of the discrete reentrant site, and on evidence that a high degree of irregularity is present in EGM signals at the rotor tip, we hypothesized that 1) the level of fractionation on EGM can be measured using a non-linear index such as Approximate Entropy (ApEn), and 2) high levels of ApEn can be used to localize the rotor tip.

Materials and Methods

We present dynamic ApEn maps as a new tool for rotor detection. ApEn values for each EGM recorded from the atrial surface are used to construct a color map. We assess the method in a 3D computational model of human atria. Fig. 1A schematically illustrates the methodology used in this work, as follows. First, two AF episodes were simulated: a chronic AF episode in which two stable rotors were found, and a chronic AF episode in which a meandering rotor was found. Second, ApEn measurements, using standard parameters, are calculated over virtual EGM recorded from the 3D model to construct dynamic ApEn maps. Third, the ApEn parameters were optimized using real measured bipolar EGM from multi-center databases. Fourth, dynamic ApEn maps were generated using optimized parameters. Next, the rotors tips are identified by analyzing local activation time maps and dynamic ApEn maps. The details of each step are presented in the following sections.

Chronic atrial fibrillation model

A realistic 3D model of human atria including the main anatomical structures, fiber orientation, electrophysiological and conduction heterogeneity and anisotropy, was developed in an earlier work [17]. It includes 52906 hexahedral elements and 100554 nodes. The Courtemanche-Ramirez-Nattel-Kneller membrane formalism [18, 19] was implemented to reproduce the human atrial cellular electrical activity. The monodomain model of the electrical propagation of the action potential along the tissue is described by a reaction-diffusion equation, and is solved using a finite element method [17].

To reproduce atrial electrical remodeling, changes in the maximum conductance and kinetics of different ionic channels of human atrial cells observed in experimental studies of chronic AF [20–22] have been incorporated into the atrial cellular model. The following parameters were altered [23]: the maximum conductance of I_{K1} was increased by 100%, while the maximum conductance values of I_{CaL} and I_{to} were decreased by 70% and by 50%, respectively.

Simulation protocol

Two AF episodes were generated by the S1–S2 protocol as follows: a train of stimuli with a basic cycle length of 1000 ms was applied for a period of 5 seconds in the sinus node area to simulate the sinus rhythm (S1). After the last beat of the sinus stimulus, a burst of 6 ectopic beats (S2) to high frequency were delivered into the right superior pulmonary vein, for the first AF episode; and they were delivered into the posterior wall of the left atrium near to the right pulmonary veins, for the second episode.

Virtual electrograms

Unipolar EGM in different points of the atria surface under conditions of uniform intracellular anisotropic resistivity was simulated, as previously described [24].

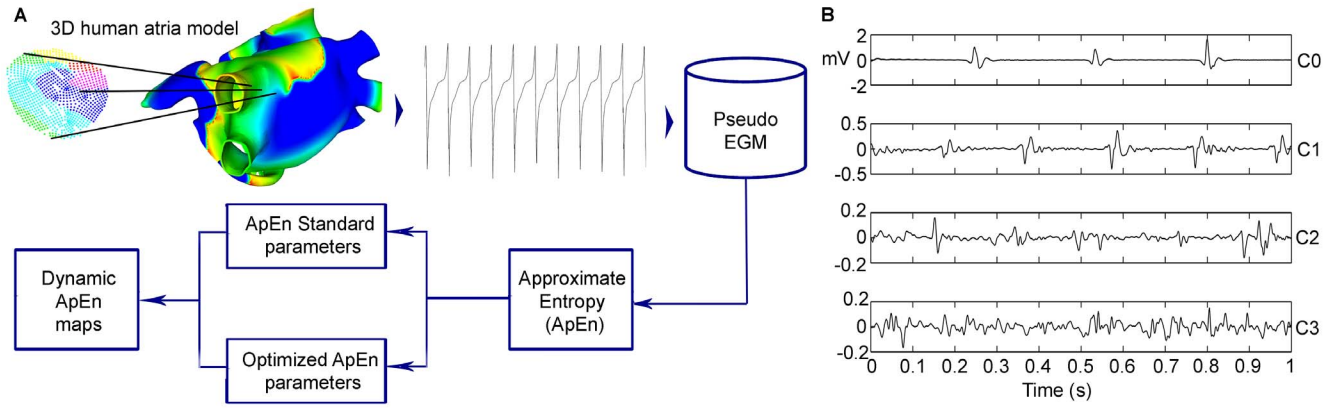


Fig. 1. Experimental setup. A: Simulated episode of chronic AF in a 3D model of human atria. A local activation time map was constructed as an alternative method for detecting rotors. A pseudo-EGM signal was calculated from the 3D model. ApEn was calculated in pseudo-EGM signals recorded over the whole atrial surface. ApEn maps were constructed in order to observe the relation between ApEn values and rotor locations. B: Examples of EGM signals of DB-CZ-DE. Representatives from the four levels of complexity proposed for the purposes of the study are shown from C0 to C3. These signals were used for ApEn parameter optimization.

doi:10.1371/journal.pone.0114577.g001

The extracellular potential (ϕ_e) is given by the following equation:

$$\phi_e(r) = -\frac{1}{4\pi}\frac{\sigma_i}{\sigma_e}\iiint \nabla V_m(r')\cdot\nabla\left[\frac{1}{r'-r}\right]dv \quad (1)$$

where ∇V_m is the spatial gradient of transmembrane potential V_m, σ_i is the intracellular conductivity, σ_e is the extracellular conductivity, r is the distance from the source point (x,y,z) to the measuring point (x',y',z') and dv is the differential volume. EGM signals were recorded at 1 *kHz*. Bipolar EGM were calculated by subtracting two 1-mm-spaced adjacent unipolar EGM.

Numerical and computational methods

A hexahedral mesh was built from the three-dimensional anatomical model using Femap from Siemens PLM software. Equations were numerically solved using EMOS software [25]. EMOS is a parallel code (mpibased) that implements the finite element method and Operator Splitting for solving the monodomain model. The time step was fixed to 0.001 ms. Simulation of 10 seconds of atrial activity took 14 hours on a computing node with two 6-core Intel Xeon X5650 clocked at 2.66 GHz and 48GB DDR3 RAM.

Isochrone activation maps

A method was developed as an arrangement of the isochrone map method [26], where the local activation times (LAT) are represented in a color map. To build an activation map, it is necessary to detect the local activation waves. An algorithm based on the Continuous Wavelet Transform was implemented to detect local activation waves. After a peak has been detected, an isochrone map is constructed with the LAT information. In order to ensure that one complete activation cycle is scanned, it is necessary to visualize the LAT minimum for a period of 100 ms. The

points at which the activation waves are spinning can be observed; these points correspond to the rotor tip. This procedure was used as a gold standard for comparison with our results.

Approximate entropy and parameter optimization

Approximate Entropy (ApEn) is a nonlinear statistic proposed by Pincus [27]. It quantifies the degree of complexity of signals. The calculation of ApEn depends on three parameters: number of data points N, embedding dimension m and threshold r. $ApEn(m,r,N)$ allows to measure regularity by calculating the probability that patterns of length m remain close on next incremental comparisons within a signal of length N, with $m<N$ [28].

ApEn is theoretically defined as the value dependent on m and r, considering $N\rightarrow\infty$. This value cannot be reached but can be approximated. The approximation is well suited when a significant number of patterns, determined by m, are acquired [29]. Pincus stated that small values of m are needed in order to converge to the real value of ApEn [28]. Specifically, he suggested $m=2$, $r\in[0.1,0.25]$ as standard parameters.

We evaluate $ApEn(2,0.1,1000)$ and $ApEn(2,0.1,500)$ from standard parameters. Furthermore, we propose m and r values derived from an optimization procedure using a dataset of EGM. This dataset has already been applied to other studies aimed at developing signal processing tools for CFAE [30, 31].

Dataset

We used two different EGM databases, independently recorded from AF patients, and independently evaluated. The databases were ranked by different electrophysiology teams, with different equipment, from two different countries. The 542 signals in the databases were classified using the same criteria divided into four classes: fractionated signals were categorized by experts into three levels of fractionation (C1, C2 and C3), and nonfractionated EGM signals were considered as level 0 (C0). The four fractionation classes, see Fig. 1B, are:

- C0: Nonfractionated EGM and also high frequency EGM.
- C1: Fractionated EGM with periodic activity.
- C2: A mixture of periodic fractionated and periodic nonfractionated EGM.
- C3: High frequency EGM with continuous activity. No regular activation can be seen.

The entire database constitutes a retrospective-offline analysis. For further information about the acquisition and classification of this database, refer to [32, 33]. The database is available at https://github.com/andresfod/Atrial_Electrograms_ApEn.

Optimization process

ApEn parameters m and r were set using real data from the database, to obtain optimal values. The variation of m was limited to 5, as recommended by Pincus [27]. Parameter r varied from 0.02 to 0.6, increasing in steps of 0.02. The database

was organized into four levels of fractionation (C0, C1, C2 and C3). Each level contains the entire set of signals of the corresponding level of fractionation. The first 500 and 1000 ms of each EGM were considered. The numbers of EGM signals were as follows: 175 signals in C0, 117 signals in C1, 184 signals in C2, and 66 signals in C3. From now on, we refer to this database organization as DB-CZ-GE. For each combination of m and r, *ApEn* was calculated for each EGM from DB-CZ-GE and a boxplot was constructed, assigning a box to each level.

In the interest of minimizing the scatter of each class and maximizing the distances between the ApEn measures of the classes, two criteria were considered in the optimization procedure: interclass percentile distance d_1 and interclass minimum-maximum distance d_2, which were defined as follows:

$$d_1 = \sum_{i=1}^{3} [Q_1(C_{i+1}) - Q_3(C_i)] \tag{2}$$

$$d_2 = \sum_{i=1}^{3} [min(C_{i+1}) - max(C_i)] \tag{3}$$

where Q_1 and Q_3 are the first and third quantile, respectively.

To select the parameters to be used for ApEn, an optimization function $\mathcal{J}_{bp}$ was constructed by equal weighting of criteria d_1 and d_2 from (2) and (3):

$$\mathcal{J}_{bp} = d_1 + d_2 \tag{4}$$

In order to validate the *ApEn* parameters obtained in the optimization procedure and to avoid overtraining, K-cross validation was performed. Randomly generated partitions were selected from full DB-CZ-GE. $K = 10$ subsets were formed.

Dynamic ApEn mapping

In order to generate a dynamic ApEn map over the surface of the atrial model, the standard ApEn parameters $m = 2$ and $r = 0.1$, as suggested by Pincus [27, 28, 34] and the parameters of ApEn obtained in the optimization procedure, were used. ApEn was calculated for virtual unipolar EGM signals, using moving windows of 500 and 1000 points, without overlapping. The range of ApEn, over the entire set of virtual EGM, was applied to a color scale, where red color corresponds to max (*ApEn*) and blue color corresponds to 0. Each virtual EGM is related to an element in the model. A 1000-point window was applied for tracking stable rotors. A 500-point window was applied for tracking meandering rotors. Unipolar EGM over the entire surface of the atria were analyzed, with the exception of the meandering rotor case, in which an observation area was selected in order to evaluate the behavior of the optimized ApEn during the presence and absence of the tip of the rotor.

Bipolar-EGM based dynamic ApEn maps were also constructed for the stable rotor case. An observation area, in which the rotor tip is anchored, was selected. The ApEn values of bipolar EGM were calculated using the optimized parameters and the 1000-point window. Shannon Entropy (ShEn) maps were also generated, through a 4-second window and bins of 0.01 mV, in accordance with the work of Ganesan et al [7]. The ShEn performance in unipolar EGM was also assessed.

Results

CFAE mapping of stable rotors during AF simulation applying ApEn with standard parameters

An AF propagation pattern over the atria was generated in the 3D model of human atria, see the action potential propagation in S1 Video. During AF activity initiated by an ectopic focus applied into the right superior pulmonary vein, two stable rotors were observed in the simulation. One is located in the posterior wall of the left atrium, near the left pulmonary vein, named R1, and the other is located in the superior vena cava, named R2. Fig. 2A shows the action potential wavefronts delimited by contour lines from the interval between 1 s and 2 s of AF simulation. Rotors R1 and R2, and a block line located over the inferior right pulmonary vein, named B1, have been marked. 42835 EGM were calculated in the whole atrial surface of the 3D model, over a four-second window.

For all EGM signals, the ApEn measurements were calculated using the standard parameters ($m=2$, $r=0.1$). The ApEn values were used to generate a color map over the anatomic structure of the atria in a 3D model. In this manner, a dynamic ApEn map was obtained and the frame corresponding to the time interval from 1 s to 2 s is depicted in Fig. 2B. Red color areas, corresponding to high ApEn values (0.3285 ± 0.0202), include B1 and the coronary sinus (CS) and also R1 and R2.

Optimization of ApEn parameters (m and r)

The optimization procedure resulted in $m=3$, $r=0.38$, for $N=1000$ and $m=3$, $r=0.30$, for $N=500$. Fig. 3A shows the boxplots for $ApEn(3,0.38,1000)$ and for $ApEn(3,0.30,500)$, respectively, applied to DB-CZ-GE. In addition, the Spearman correlation coefficient r_S between ApEn and the corresponding level of fractionation was calculated.

CFAE mapping of stable rotors during AF simulation applying ApEn with optimized parameters

A dynamic $ApEn(3,0.38,1000)$ map was generated using unipolar EGM, and the frame corresponding to the time interval 1 s to 2 s is depicted in Fig. 2C, see S1 Video. The dynamic ApEn map reveals two areas of high ApEn values (0.3285 ± 0.0202), corresponding to rotors R1 and R2. In addition, the third area with lower ApEn values (0.2724 ± 0.0238) corresponds to B1, taking into account

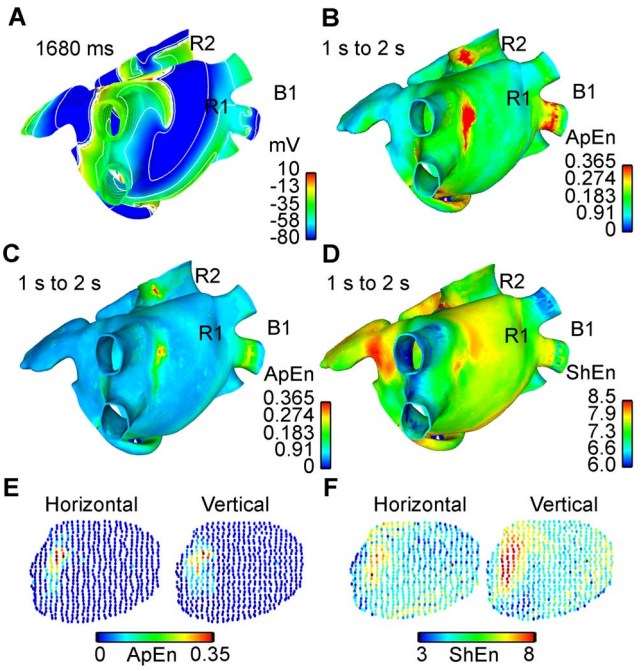


Fig. 2. Comparison between tools for rotor mapping. A. Action potential wavefront delimited by contour lines over the 3D Human Atria Model extracted from the interval between 1 s and 2 s of simulation. The spinning wavefronts around one point define stable rotors R1 and R2. Line block B1 can be seen at the right inferior pulmonary vein. B. Dynamic ApEn map calculated using standard parameters and unipolar EGM. C. Dynamic ApEn map calculated from the optimized parameters obtained in our work, using unipolar EGM. D. Shannon entropy map, using unipolar EGM. Note that map C shows better sensitivity for localizing rotor tips. E. Dynamic ApEn map calculated from optimized parameters using bipolar EGM with horizontal and vertical orientation. The region corresponds to the vicinity of rotor R1. F. ShEn map calculated using the bipolar EGM obtained from the vicinity of rotor R1.

doi:10.1371/journal.pone.0114577.g002

that passive activation regions have ApEn values lower than 0.1. A ShEn map was also generated and is shown in Fig. 2D. Red color areas, corresponding to high ShEn values (8.375 ± 0.125), include small vicinities near to R1 and near to B1, along other zones, e.g. CS, the inferior wall of the left atria, the left and right appendage and the lateral wall near to the left appendage. The LAT method was applied as a reference method, see S2 Video. Fig. 3B shows isochronic maps from R1 and R2. The rotor tip is defined by the converging points of the colored waves. These waves represent the local activation. Green color indicates early activation points. Table 1 shows the spatial location of R1 and R2 found using the LAT method and the dynamic ApEn mapping method. The last column corresponds to the Euclidean distance between the spatial coordinates from the rotor tip, identified using both methods. The maximum distance is 1.66 mm. Unipolar EGM extracted from rotor sites R1 and R2, and near to block area B1, are shown in Fig. 3C. ApEn values are also presented. The EGM corresponding to rotor activity (top) have low voltage and irregular morphology. The EGM corresponding to the block line (below left) presents fractionation, but activation

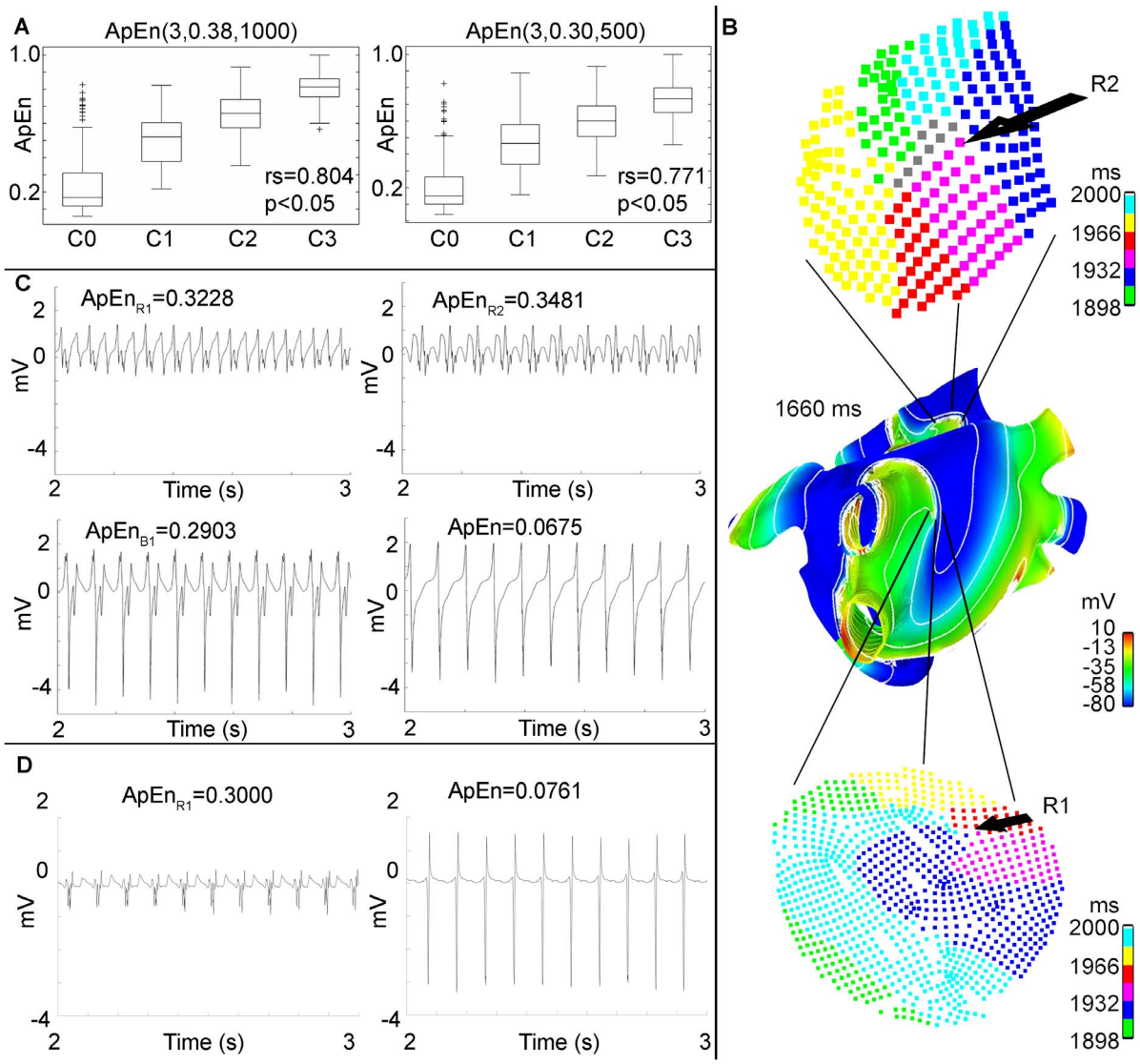


Fig. 3. Localization of stable rotors. A: Results of the optimization procedure. Boxplots of ApEn normalized values using optimized parameters: $ApEn(1000,3,0.38)$ (left) and $ApEn(500,3,0.30)$ (right). The Spearman correlation coefficient calculated over DB-CZ-GE for each boxplot is shown. B: Activation isochronic maps corresponding to R1 (below) and R2 (top). The rotor tip is indicated where the colors converge. C: EGM generated by the model in the areas of stable rotors and the block for the time interval between 2 s and 3 s. EGM corresponding to the R1, R2, B1 and plane activation wavefront areas. D: Bipolar EGM corresponding to R1 and plane activation wavefront area. ApEn values for each EGM are shown.

doi:10.1371/journal.pone.0114577.g003

patterns are visible and their amplitudes are similar to non-fractionated EGM (below right). The highest ApEn values correspond to R1 and R2.

A circular area containing R1 was selected. Bipolar EGM were calculated using horizontal and vertical bipole orientation. 636 EGM were obtained for each orientation. Fig. 2E shows the dynamic *ApEn*(3,0.38,1000) maps for horizontal and vertical bipolar orientation, in the time interval between seconds 1 and 2. High ApEn values correspond with the tip of the rotor in both cases. Fig. 2F shows the ShEn maps. The vertical bipolar orientation map presents a region of

Table 1. Spatial rotor localization.

	Coordinates			Coordinates			Euclidean
	LAT method (mm)			*ApEn* method (mm)			distance (mm)
Rotor tip R1							
Interval (s)	x	y	z	x	y	z	d
0 to 1	4.6	32.2	−167.2	4.4	32.4	−168.1	0.91
1 to 2	4.5	32.2	−167.2	4.6	32.2	−167.2	0.00
2 to 3	3.3	33.1	−167.1	4.5	32.4	−168.1	1.66
3 to 4	3.3	33.2	−167.9	4.5	32.4	−168.1	1.44
Rotor tip R2							
Interval (s)	x	y	z	x	y	z	d
0 to 1	2.6	9.8	−186.8	2.7	10.4	−187.5	0.95
1 to 2	3.5	9.8	−186.6	3.2	10.2	−187.1	0.66
2 to 3	3.2	10.1	−187.1	3.2	10.2	−187.1	0
3 to 4	4.0	9.5	−186.3	3.2	10.2	−187.1	1.32

Comparison between ApEn and the LAT method. Spatial location of $R1$ over the 3D atria anatomical structure. The coordinates were found by two methods: the LAT method, as a reference, and the $ApEn$ map method. Euclidean distance is used to compare the performance of the two methods for four time intervals 1 s in length.

doi:10.1371/journal.pone.0114577.t001

high ShEn values broader than the horizontal orientation map, though both contain the tip of the rotor. Bipolar EGM and its ApEn values are shown in Fig. 3D. The EGM corresponding to R1 (left) have low voltage and an irregular morphology. The EGM corresponding to a region outside the rotor tip (right) presents high amplitude, high frequency activation patterns and a regular morphology.

CFAE mapping of a meandering rotor during AF simulation applying ApEn with optimized parameters

During AF activity initiated by an ectopic focus applied in the posterior wall of the left atrium near to the right pulmonary veins, a meandering rotor was observed in the posterior wall, see the action potential propagation in S3 Video. 933 unipolar EGM signals of 9 seconds length were obtained from the AF simulation. The ApEn measurement was calculated using the parameters found to be optimal ($m=3$, $r=0.30$). The ApEn values were used to generate a color map in a 3D model, see S3 Video. The left and centre maps in Fig. 4 show the evolution of the action potential during a time interval when the rotor tip meanders. Figs. 4A and 4C show the rotor activity, while Fig. 4B shows an action potential map with a plane wavefront over the zone. The dynamic ApEn(3, 0.30, 500) maps shown on the right side of Fig. 4 present high ApEn values only in intervals corresponding to Figs. 4A and 4C. An EGM localized over the area where the rotor meanders is also shown; fractionation of the EGM is observed when the rotor is in the zone.

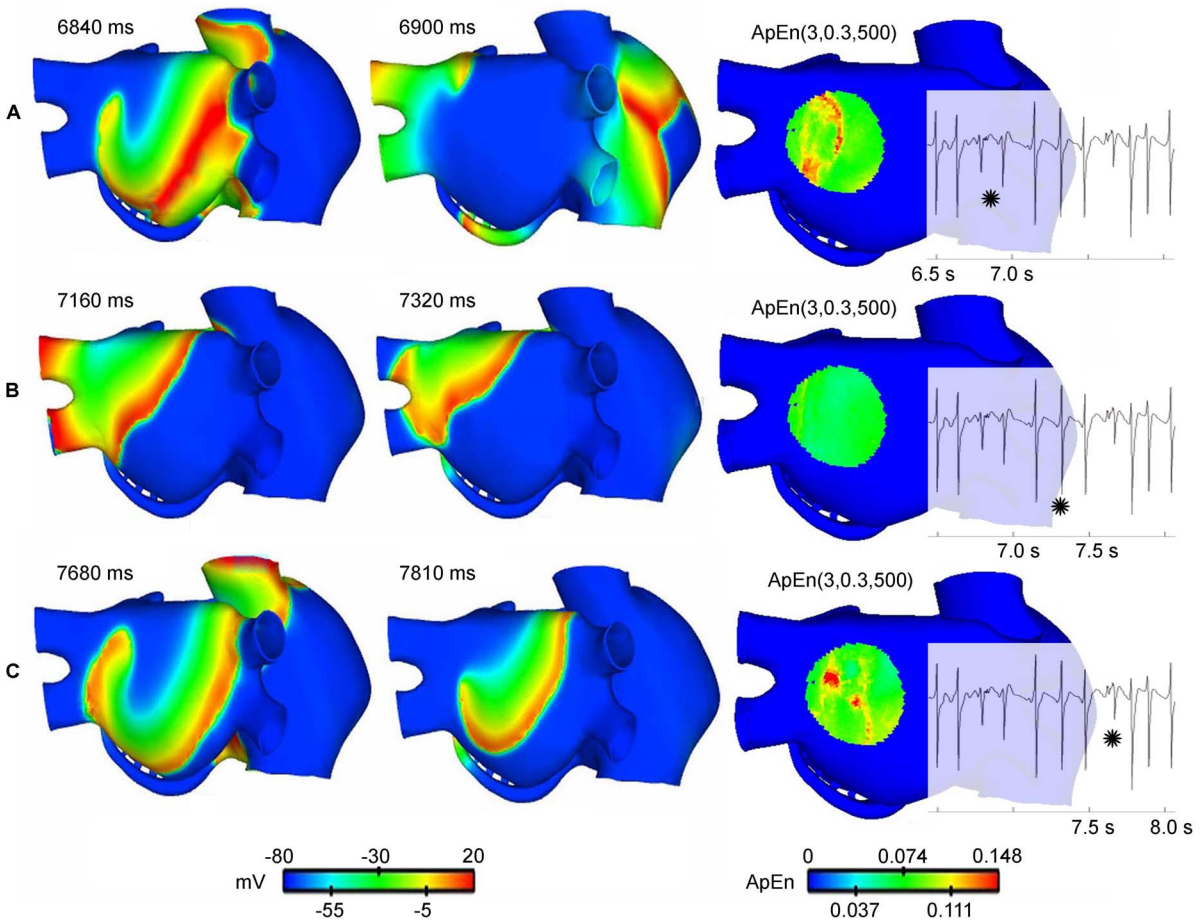


Fig. 4. Meandering rotor tracking. The left and centre snapshots in A, B and C show the evolution of the action potential at three time instants. Meandering rotational activity is present in A and C. The snapshot on the right corresponds to the dynamic ApEn(3,0.30,500) maps. High ApEn values (red color) correspond to the presence of a meandering rotor. EGM is also shown. The star marks the 500-point interval corresponding to the evolution of the action potential. Fragmentation is generated in the presence of a rotor (A and C).

doi:10.1371/journal.pone.0114577.g004

Discussion

The principal findings of our work can be summarized as follows:

- ApEn values with optimal parameters are strongly and positively correlated with the fractionation levels defined in the AF EGM database.
- High ApEn values were found in EGM from stable and meandering rotor tips in a simulated episode of AF.
- We have developed a methodology for localizing the tip of the rotor during AF, using dynamic ApEn maps.

Dynamic ApEn maps for detecting rotor tip sites

A method based on dynamic ApEn maps is presented in this study as a basis for proposing a new tool for automatic detection of ablation target sites. We have

presented evidence that the method is able to detect rotors during two simulated AF episodes. We used the continuous ApEn value for building color maps. Fig. 2C shows the localization of the stable rotors, and Fig. 4 shows the localization of the meandering rotor. The red areas in the ApEn map correspond to the location of rotor tips. This is an indicator that high ApEn values are related with rotor activity and, more precisely, with rotor tips. The mechanism behind the unipolar fractionation in our simulations can be explained as follows: when the rotor turns around the pivot point (tip), the tip is affected by wavefronts from the rotor head. When the wavefront passes near to the tip in each rotation cycle, several electrotonic potentials are observed, consequently, fractionation arises and there is irregularity. The ApEn response to the occurrence of such fractionation is as follows: if the rotor is stable, the unipolar EGM, registered at the tip, will present fractionation as long as the rotor tip remains harbored to that point. This causes ApEn to increase over the entire observing window. If the rotor meanders, fractionated complexes will be observed when the rotor tip passes over that specific point, therefore increments of the ApEn will be expected in that instant of time. Our results are consistent with other works, in which fractionated unipolar EGM were observed at pivot points [35] and unipolar EGM symmetry is affected by the wavefront curvature [36]. Moreover, the action potential map, shown in Fig. 2A, indicates a lobule B1 that does not depolarize when a wavefront arrives. This irregular activity increases the ApEn value but not to rotor levels, as is shown in Fig. 3C. This ApEn level corresponds to the yellow areas. Umapathy et al [11] reported that CFAE were located in sites of wave breaks and in the region of a rotor tip during experiments in a murine HL-1 atrial monolayer model. Their results agree with ours in the 3D model. Furthermore, we show the ability of dynamic ApEn maps to locate and to distinguish between these substrates.

We found that high ApEn values are related with arrhythmogenic substrates, such as rotors and block lines. Discrimination between active and passive CFAE is still an open question [5]. One important question is how to differentiate CFAE according to the substrate that generates it. Using 2D computer models and cell cultures, Navoret et al detected CFAE using amplitude criteria, number of deflections and cycle length. They established a relationship between the detected CFAE and the presence of shock waves and rotors, but they failed to differentiate them [37]. Another study reported an algorithm that classifies CFAE and non-CFAE, but fails to distinguish between active CFAE (whose ablation restored the sinus rhythm) and passive CFAE (whose ablation did not restore the sinus rhythm) [8]. Our study showed that the dynamic ApEn map, calculated using standard ApEn parameters (*ApEn*(2,0.1,1000)), identifies R1, R2, B1 and CS having high ApEn values, however it is not possible to differentiate between them. On the other hand, the ApEn dynamic map, calculated with optimized parameters ApEn(3,0.38,1000), assigns scaled values to zones of interest: rotors R1 and R2 have the highest ApEn value, followed by intermediate ApEn values in block line area B1 and the wave collision in the CS, and lower ApEn values for fibrillatory EGM with an organized activity. These results suggests that the optimization process for ApEn improves the detection and discrimination between rotor tips,

block-lines and wave collisions, and this could help in solving the question discussed above.

Variability and irregularity of fractionated EGM during AF

We hypothesized that the AF EGM has variability at the tip of the rotor, but we also expected irregularity (e.g. fractionation of the EGM). So we used ApEn to quantify the variability and the irregularity at the same time in the following manner: ApEn was calculated from segments of equal length but shorter than the entire signal. Thus, we ensure a quantification of the variability by calculating time consecutive measures of regularity. That is why the time series information aspect is important. However, time series information by itself is not enough to localize rotors. Spatial information (i.e. the spatial coordinates of each EGM) allows a rotor to be situated over the atria.

This study has taken into consideration the morphological features of both unipolar and bipolar EGM under two provisos: 1) Fractionation has been described in unipolar and bipolar EGM [35, 38]. 2) Fractionated EGM have an irregular morphology. We found that ApEn is related with the degree of fractionation: ApEn values tend to be higher if the degree of fragmentation is higher. Therefore, the use of bipolar signals and unipolar signals does not affect the performance of ApEn. This is evidenced in the results for the location of the rotors using unipolar and bipolar EGM in the computational model using ApEn optimized with the bipolar EGM of the multicenter DB-CZ-GE.

ApEn was developed as a measure of regularity to quantify levels of complexity within time series [39, 40]. Since we hypothesized that fractionation of EGM could be graded, with the highest level at the tip of the rotor, we chose ApEn to quantify levels of fractionation, as this is a well-characterized measure in other cardiac applications, e.g. RR intervals, heart rate variability signals and ECG signals [41]. Hoekstra et al [42] showed that the complexity of EGM increases with the type of fibrillation, based on the chaotic spatio-temporal activation patterns of the right atria. In a study on animal models, Ganesan et al [7] found that there are some similarities in the visual appearance of EGM signals with higher ShEn and fractionated signals. Novak et al [43] compared various measures from the theory of nonlinear dynamics that can help to provide an objective description of the level of fractionation and the complexity of CFAE signals. Thus, non-linear tools can provide useful information about the occurrence of arrhythmogenic substrates.

Figs. 3C and 3D show four unipolar and 2 bipolar EGM morphologies observed in the computational model, respectively. The EGM from rotors (top in Fig. 3C and left in Fig. 3D) have low voltage and irregular morphology whereas that, the EGM from sites with a plain wavefront (bottom in Fig. 3C and right in Fig. 3D) are regular. Furthermore, the ApEn values graded the bipolar EGM from DB-CZ-GE according to the level of fractionation (Fig. 3A). Again, the increasing irregularity in the four fractionation classes in DB-CZ-GE can be visually verified in Fig. 1B. Thus, we demonstrate that ApEn is well correlated with the perception

of morphological regularity of EGM. This feature of ApEn was already reported in the work of Anier et al [44], though in electroencephalograms. Additionally, we found that there are different fractionated unipolar EGM morphologies or levels of fractionation, depending on the wave propagation pattern in the 3D model.

Other rotor detection tools and fractionation measures

We developed the dynamic ApEn maps tool considering fractionation of EGM as non-linear dynamics and considering that degrees of complexity can be identified within fractionated EGM. To the best of our knowledge, no other tools have been developed that take these two features into account. However, there are other similar studies regarding rotor detection. Ganesan et al [7] were able to relate high ShEn values with the tip of the rotor in 2D arrays using bipolar EGM and the directional information they contain. We were able to reproduce these results using bipolar signals from the 3D model corresponding to a zone that encloses R1 (Fig. 2F). We have also applied optimized ApEn to the bipolar EGM, successfully detecting the tip of R1 (Fig. 2E). Both approaches define an area enclosing the rotor tip, where the ShEn values decrease more slowly than ApEn in the area surrounding the tip. Thus, ApEn gains in specifying the tip of the rotor in our simulations. Furthermore, ApEn offers reduced dependence on the orientation of the record. We assessed the ShEn map using unipolar EGM (Fig. 2D), in which substrate detection was not accomplished. These results suggest that dynamic ApEn maps could be used in unipolar and bipolar EGM for the task of rotor mapping.

Narayan et al [26] used focal impulse and rotor modulation to guide ablation with better results than when using ablation without this approach. They used the LAT method to localize the rotor tip. We compared the dynamic ApEn maps with the LAT method, and obtained differences less than 2 mm. The use of isochronic activation maps is limited by the high computational cost for constructing the maps, and the fact that they need to observe several frames in order to study transient rotational waves – a procedure that is likely to take considerable time. This implies that meandering rotors are hard to detect using isochronic activation maps. Dynamic ApEn maps can detect a stable rotor, see S1 Video, where the red sites do not change over time, or they can track a meandering rotor tip, see S2 Video, where the red sites change over time. This technique can help to reduce the uncertainty in the location of the rotor from the activation maps.

Zlochiver et al [12] studied the regularity of EGM in the presence of stable rotors. Spectral analysis was applied to the singular value decomposition in order to measure the contribution of the meandering frequency and the rotor tip frequency. They also assessed regularity using the periodicity index. Rotor tips were detected in 2D computational arrays, through low periodicity index values calculated from unipolar EGM. Fractionation of the EGM was observed at the rotor tip. These results support and agree with ours, though the two works apply different approaches: Zlochiver et al exploit the fact that under fractionation conditions, the periodicity of the signal is lost, which can be observed in the

singular components. We consider fractionation as a non-linear phenomenon, and we asses the morphological features of the signal using ApEn. Although the work of Zlochiver et al reports the possibility of detecting stable rotors, it would be interesting to asses whether the tool that they have developed would work with meandering rotors, as ours does. We can see that they need a signal of 2000 seconds, which would initially limit the temporal resolution to tracking a meandering rotor (we used a 500-point window).

Existing measurements of CFAE, implemented in electroanatomical mapping devices, are based on Nademanee's definition of CFAE. There are studies demonstrating that CFAE are not suitable for ablation of AF. Narayan et al [16] graded CFAE applying the interval confidence level (ICL) algorithm [45] which is implemented in Carto XP. They conclude that most CFAE are not able to localize AF sources. Ganesan et al [7] compared the ShEn results with CFAE analysis (CFE-mean) performed in NavX. They observed a consistent but weak inverse correlation between ShEn and CFE-mean. It is worth pointing out that both ICL and CFE-mean algorithms use cycle length and voltage amplitude criteria under the definition of CFAE given by Nademanee [3]. They are sensitive to changes in the morphology of the signal, which can have different appearances depending on the contact between the catheter and the substrate [10], or can be influenced by far-field artifacts or by noise [46]. Our results suggest that rotors can be localized by considering different levels of fractionation and non-linear tools. These observations invite to enhancing the Nademanee's definition of CFAE. His definition is based on cycle length and amplitude criteria, which are not sufficient for describing the levels of fractionation. It may be better to consider defining CFAE as a nonlinear dynamic phenomenon. This last statement is supported by other authors [4, 6, 47, 48]. It could be more effective to consider fractionation as a complex phenomenon, the description of which should be resolved from various fractionation levels, rather than using a CFAE/non-CFAE classification.

Although the remodeling conditions that are used can reproduce the action potential phenotype observed in patients with permanent AF, electrical remodeling is not the only process accompanying chronic AF. Indeed, many of these patients have significant structural remodeling with fibrosis, which contributes to short APD and increases the complexity of the arrhythmia. The 3D anatomical model of human atria does not take into account the real thickness of the atrial walls. Our results were obtained using a specific virtual atria model. Although our model includes a great number of anatomical and morphological details, it corresponds to a particular set of parameters (electrophysiology, anatomy, fiber direction, anisotropy and heterogeneity, among others). In addition, although there are also inter-subject differences in fiber orientation, we have tried to model the most common fiber orientation observed experimentally for the different parts of the atrial model. Regarding EGM analysis, although we assessed the performance of ApEn and ShEn using bipolar EGM, it was executed over a small area due to the high complexity of defining a bipolar record configuration over the whole atria. The behavior of both ApEn and ShEn, calculated using bipolar EGM over broader regions, including other arrhythmo-

genic mechanisms besides rotors, should be studied. In order to fully validate the results of our optimization procedure with DB-CZ-GE, although good performance has been achieved for rotor detection in the 3D model, future studies should involve collecting more AF EGM samples from other databases in order to achieve equally distributed fractionation levels. Additionally, real unipolar EGM must be included.

In conclusion, we have provided evidence for postulating dynamic ApEn maps as a supporting tool in AF ablation procedures. ApEn calculation over atrial EGM signals can be used to build an electroanatomical map with continuous ApEn values represented in colors that can identify rotor tip locations without pre-specifying thresholds. This property makes the method adaptable to specific electrophysiological cases. A combination of all this could help to improve ablation procedures. We suggest that dynamic ApEn color maps could become a tool for AF rotor mapping. The methodology proposed here needs to be validated by experimental models and by clinical studies to evaluate ablation procedures guided by dynamic ApEn maps.

Supporting Information

S1 Video. Stable rotors. Left: Simulation of an AF episode induced by six transitory ectopic beats applied in the ostium of the right pulmonary vein, sustained by multiple reentrant waves. Two rotors can be observed: in the posterior wall of the left atria posterior wall, and in the superior cava vein. Right: A dynamic ApEn map using optimized parameters. Red areas correspond to the tips of rotors R1 and R2.
doi:10.1371/journal.pone.0114577.s001 (MP4)

S2 Video. Isochrone map. Isochrone maps for rotors R1 and R2 for the time interval between 1500 ms and 2000 ms.
doi:10.1371/journal.pone.0114577.s002 (MP4)

S3 Video. Meandering rotor. Left: Simulation of an AF episode induced by six transitory ectopic beats applied in the posterior wall of the left atrium near to the right pulmonary veins, sustained by multiple reentrant waves. One meandering rotor can be observed in the posterior wall of the left atria. Right: A dynamic ApEn map using optimized parameters. Note that, when a plane activation wavefront is present within the observation area, the dynamic ApEn map shows low values (green). Red zones appear when the rotor tip is within the observation area (e.g. the time interval between 2500 ms and 3000 ms).
doi:10.1371/journal.pone.0114577.s003 (MP4)

Author Contributions

Conceived and designed the experiments: JPU AO CT JS JB. Performed the experiments: JPU AO CT VK DN JB. Analyzed the data: JPU AO CT VK DN AL JB. Contributed reagents/materials/analysis tools: VK DN JS TO CS AL. Wrote the

paper: JPU AO CT. Drafted the article or revised it critically for important intellectual content: JPU AO CT VK DN JS TO CS AL JB.

References

1. **Gupta A, Maheshwari A, Thakur R, Lokhandwala Y** (2002) Cardiac mapping: utility or futility? Indian Pacing and Electrophysiology Journal 2: 20–32.
2. **De Bakker J, Wittkampf F** (2010) The pathophysiologic basis of fractionated and complex electrograms and the impact of recording techniques on their detection and interpretation. Circ Arrhythm Electrophysiol 3: 204–213.
3. **Nademanee K, McKenzie J, Kosar E, Schwab M, Sunsaneewitayakul B, et al.** (2004) A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate. Journal of the American College of Cardiology 43: 2044–53.
4. **Reddy V** (2009) Atrial fibrillation: unanswered questions and future directions. Cardiology Clinics 27: 201–216.
5. **Jadidi A, Duncan E, Miyazaki S, Lellouche N, Shah A, et al.** (2012) Functional nature of electrogram fractionation demonstrated by left atrial high-density mapping. Circulation Arrhythmia and electrophysiology 5: 32–42.
6. **Ciaccio E, Biviano A, Whang W, Garan H** (2012) Identification of recurring patterns in fractionated atrial electrograms using new transform coefficients. BioMedical Engineering OnLine 11: 4.
7. **Ganesan A, Kuklik P, Lau D, Brooks A, Baumert M, et al.** (2013) Bipolar electrogram shannon entropy at sites of rotational activation: implications for ablation of atrial fibrillation. Circ Arrhythm Electrophysiol 6: 48–57.
8. **Navoret N, Jacquir S, Laurent G, Binczak S** (2013) Detection of complex fractionated atrial electrograms using recurrence quantification analysis. IEEE transactions on bio-medical engineering 60: 1975–82.
9. **Skanes A, Mandapati R, Berenfeld O, Davidenko J, Jalife J** (1998) Spatiotemporal periodicity during atrial fibrillation in the isolated sheep heart. Circulation 98: 1236–1248.
10. **Navoret N, Jacquir S, Laurent G, Binczak S** (2012) Relationship between complex fractionated atrial electrogram patterns and different heart substrate configurations materials numerical cell models method the modelled catheter is a thermocool irrigated tip. Computing in Cardiology 39: 893–396.
11. **Umapathy K, Masse S, Kolodziejska K, Veenhuyzen G, Chauhan V, et al.** (2008) Electrogram fractionation in murine HL-1 atrial monolayer model. Heart rhythm: the official journal of the Heart Rhythm Society 5: 1029–35.
12. **Zlochiver S, Yamazaki M, Kalifa J, Berenfeld O** (2008) Rotor meandering contributes to irregularity in electrograms during atrial fibrillation. Heart rhythm 5: 846–854.
13. **Orozco-Duque A, Ugarte JP, Tobón C, Saiz J, Bustamante J** (2013) Approximate entropy can localize rotors, but not ectopic foci during chronic atrial fibrilation: A simulation study. Computing in Cardiology 40: 903–906.
14. **Jalife J, Berenfeld O, Skanes A, Mandapati R** (1998) Mechanisms of atrial fibrillation: mother rotors or multiple daughter wavelets, or both. Cardiovasc Electrophysiol 9: S2–12.
15. **Narayan S, Patel J, Mulpuru S, Krummen D** (2012) Focal impulse and rotor modulation ablation of sustaining rotors abruptly terminates persistent atrial fibrillation to sinus rhythm with elimination on follow-up A video case study. Heart rhythm 9: 1436–1439.
16. **Narayan S, Shivkumar K, Krummen D, Miller J, Rappel WJ** (2013) Panoramic electrophysiological mapping but not electrogram morphology identifies stable sources for human atrial fibrillation. Circ Arrhythm Electrotrophysiol 6(1): 58–67.
17. **Tobón C, Ruiz-Villa C, Heidenreich E, Romero L, Hornero F, et al.** (2013) A three-dimensional human atrial model with fiber orientation. electrograms and arrhythmic activation patterns relationship. PLoS ONE 8: e50883.

18. **Courtemanche M, Ramirez R, Nattel S** (1998) Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. The American Journal of Physiology 275: H301–21.

19. **Kneller J, Zou R, Vigmond E, Wang Z, Leon L, et al.** (2002) Cholinergic atrial fibrillation in a computer model of a two-dimensional sheet of canine atrial cells with realistic ionic properties. Circulation Research 90: E73–87.

20. **Bosch R, Zeng X, Grammer J, Popovic K, Mewis C, et al.** (1999) Ionic mechanisms of electrical remodeling in human atrial fibrillation. Cardiovascular research 44: 121–31.

21. **Workman A, Kane K, Rankin A** (2001) The contribution of ionic currents to changes in refractoriness of human atrial myocytes associated with chronic atrial fibrillation. Cardiovascular research 52: 226–35.

22. **Van Wagoner D** (2003) Electrophysiological remodeling in human atrial fibrillation. Pacing and clinical electrophysiology 26: 1572–1575.

23. **Zhang H, Garratt CJ, Zhu J, Holden A** (2005) Role of up-regulation of ik1 in action potential shortening associated with atrial fibrillation in humans. Cardiovasc Res 66: 493–502.

24. **Roberge F, Vinet A, Victorri B** (1986) Reconstruction of propagated electrical activity with a two-dimensional model of anisotropic heart muscle. Circ Res 58: 461–475.

25. **Heidenreich Ea, Ferrero JM, Doblaré M, Rodríguez JF** (2010) Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology. Annals of biomedical engineering 38: 2331–45.

26. **Narayan S, Krummen D, Rappel WJ** (2012) Clinical mapping approach to diagnose electrical rotors and focal impulse sources for human atrial fibrillation. Journal of Cardiovascular Electrophysiology 23: 447–454.

27. **Pincus S** (1991) Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America 88: 2297–301.

28. **Pincus S, Goldberger A** (1994) Physiological time-series analysis: what does regularity quantify? American Journal of Physiology Heart and Circulatory Physiology 266: H1643–H1656.

29. **Alcaraz R, Rieta J** (2010) A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. Biomedical Signal Processing and Control 5: 1–14.

30. **Kremen V, Lhotská L, Macas M, Cihák R, Vancura V, et al.** (2008) A new approach to automated assessment of fractionation of endocardial electrograms during atrial fibrillation. Physiological measurement 29: 1371–81.

31. **Kremen V, Kordik P, Lhotska L** (2009) Comparison of several classifiers to evaluate endocardial electrograms fractionation in human. In: Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE. pp. 2502–2505.

32. **Kremen V** (2008) Automated assessment of endocardial electrograms fractionation in human. Ph.D. thesis, Czech Technical University in Prague Faculty of Electrical Engineering Department of Cybernetics.

33. **Schilling C** (2012) Analysis of Atrial Electrograms, volume 17. Karlrsruhe: KIT Scientific Publishing.

34. **Pincus S** (1992) Approximating Markov chains. Proceedings of the National Academy of Sciences of the United States of America 89: 4432–6.

35. **Konings K, Smeets J, Penn O, Wellens H, Allessie M** (1997) Configuration of unipolar atrial electrograms during electrically induced atrial fibrillation in humans. Circulation 95: 1231–1241.

36. **Jacquemet V, Virag N, Ihara Z, Dang L, Blanc O, et al.** (2003) Study of Unipolar Electrogram Morphology in a Computer Model of Atrial Fibrillation. Journal of Cardiovascular Electrophysiology 14: S172–S179.

37. **Navoret N, Xu B, Jacquir S, Binczak S** (2010) Comparison of complex fractionated atrial electrograms at cellular scale using numerical and experimental models. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference 2010: 3249–52.

38. **Nademanee K, Lockwood E, Oketani N, Gidney B** (2010) Catheter ablation of atrial fibrillation guided by complex fractionated atrial electrogram mapping of atrial fibrillation substrate. Journal of Cardiology 55: 1–12.

39. **Pincus S, Singer B** (1996) Randomness and degrees of irregularity. Proceedings of the National Academy of Sciences of the United States of America 93: 2083–8.

40. **Yentes J, Hunt N, Schmid K, Kaipust J, McGrath D, et al.** (2012) The appropriate use of approximate entropy and sample entropy with short data sets. Annals of Biomedical Engineering 41: 349–365.

41. **Richman J, Moorman J** (2000) Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 278: H2039–2049.

42. **Hoekstra B, Diks C, Allessie M, DeGoede J** (1995) Nonlinear analysis of epicardial atrial electrograms of electrically induced atrial fibrillation in man. J Cardiovasc Electrophysiol 6: 419–440.

43. **Novák D, Kremen V, Cuesta D, Schmidt K, Chudácek V, et al.** (2009) Discrimination of endocardial electrogram disorganization using a signal regularity analysis. Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference 2009: 1812–5.

44. **Anier A, Lipping T, Ferenets R, Puumala P, Sonkajärvi E, et al.** (2012) Relationship between approximate entropy and visual inspection of irregularity in the EEG signal, a comparison with spectral entropy. British Journal of Anaesthesia 109: 928–34.

45. **Scherr D, Dalal D, Cheema A, Cheng A, Henrikson C, et al.** (2007) Automated detection and characterization of complex fractionated atrial electrograms in human left atrium during atrial fibrillation. Heart Rhythm The Official Journal of the Heart Rhythm Society 4: 1013–1020.

46. **Grzęda R, Noujaim FS, Berenfeld O, Jalife J** (2009) Complex fractionated atrial electrograms: properties of time- domain vs. frequency-domain methods. NIH Public Access 6: 1475–1482.

47. **Hunter RJ, Diab I, Tayebjee M, Richmond L, Sporton S, et al.** (2011) Characterization of fractionated atrial electrograms critical for maintenance of atrial fibrillation: A randomized, controlled trial of ablation strategies (the CFAE AF trial). Circulation Arrhythmia and electrophysiology 4: 622–9.

48. **Berenfeld O, Jalife J** (2011) Complex fractionated atrail electrograms: is this the beast to tame in AF. NIH Public Access 4: 426–428.

31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Discrimination of endocardial electrogram disorganization using a signal regularity analysis

D. Novák, V. Křemen, D. Cuesta, K. Schmidt, V. Chudáček, L. Lhotská

***Abstract*— Measures from the theory of nonlinear dynamics were applied on complex fractionated atrial electrograms (CFAEs) in order to characterize their physiological dynamic behavior. The results were obtained considering 113 short term atrial electrograms (A-EGMs) which were annotated by three experts into four classes of fractionation according to A-EGMs signal regularity. The following measures were applied on A-EGM signals: General Correlation Dimension, Approximate Entropy, Detrended Fluctuation Analysis, Lempel-Ziv Complexity, and Katz-Sevcik, Variance and Box Counting Fractal Dimension. Assessment of disorganization was evaluated by a Kruskal Wallis statistical test. Except Detrended Fluctuation Analysis and Variance Fractal Dimension, the CFAE disorganization was found statistically significant even for low significant level $\alpha = 0.001$. Moreover, the increasing complexity of A-EGM signals was reflected by higher values of General Correlation Dimension of order 1 and Approximate Entropy.**

I. INTRODUCTION

Endocardial sites generating complex fractionated atrial electrograms (CFAEs) have been reported as ablative targets for the treatment of atrial fibrillation (AF) [1]. In order to identify those sites, a great effort has been made to describe patterns of activation in AF [2] and to quantify general characteristics of CFAEs either in time- or frequency-domain [3]. However, the process of a CFAE identification is highly dependent on the operator's judgment. Moreover, it is not clear if CFAE are a random process of a local atrial electrogram disorganization or a reproducible physiological effect [4]. This study is aimed at applying a signal regularity analysis for the description of spatio-temporal changes of the A-EGMs fractionation levels. This analysis enables to investigate A-EGMs nonlinear dynamics and confirm the hypothesis that high degrees of A-EGMs fractionation are also reflected by higher values of the regularity measures.

II. METHODOLOGY

A. Experimental Data Set

Atrial bipolar electrograms were collected during left-atrial endocardial mapping using 4-mm irrigated-tip ablation catheter (NaviStar, Biosense-Webster) in 12 patients (9 males, aged 56 $\pm$ 8 years) with persistent AF. Sampling frequency was 977Hz. 113 of the A-EGMs were manually selected and cropped by an expert. For the purposes of the study (see Fig. 1), four Classes of Fractionation (CF) were set. The dataset ranking, which was done by three independent experts, resulted in a total of 339 rankings (113*3 = 339). The three independent experts never disagreed in their ranking by more than one neighboring CF. Therefore, the most prevailing pattern was chosen as a final score. The four CFs enabled to get a uniform dataset of A-EGMs with a significant number of samples in each class ($class_1 : C_1 = 22$, $class_2 : C_2 = 42$, $class_3 : C_3 = 36$, $class_4 : C_4 = 13$), so that such dataset could be used in a regularity analysis [5],[6].

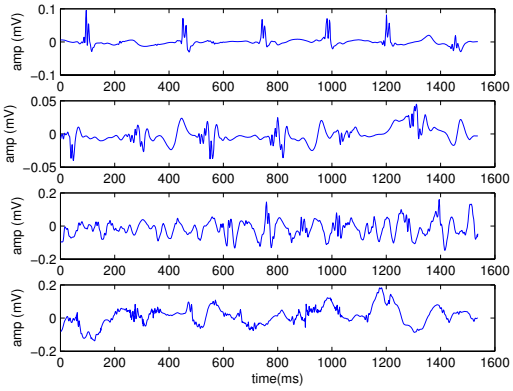


Fig. 1. Epochs of four complex fractionated atrial electrograms. These are representatives of each CF used in the study. From top to bottom: i) $class_1$: Organized activity, ii) $class_2$: Mild degree of fractionation. iii) $class_3$: Intermediate degree of fractionation. iv) $class_4$: High degree of fractionation.

B. Regularity Signal Measures

Before attempting to calculate the fractal dimension for endocardial electrograms, it is important to establish evidence that these waveforms may be characterized as fractals. Fractal dimension values are related to the regularity of a pattern, or the quantity of information embedded in a pattern in terms of morphology, entropy, spectra or variance [7]. Considering morphological properties of A-EGMs, it appears that these signals possess valid fractal dimension values, mainly because of two reasons. First, the signals do not self-cross. By looking at any one of A-EGMs waveforms in Fig.1, it is apparent that in order to scale it, a different scaling factor is required for each axis. This indicates that the A-EGM waveform is self-affine. Second, the signals exhibit clear quasi-periodicity because they emerge from natural repetitive processes (heart beats specifically). In the next subsections, we will describe the signal regularity measures used in this study.

D. Novák, V. Křemen, V. Chudáček, L. Lhotská are with Department of Cybernetics, Czech Technical University in Prague, Czech Republic, xnovakd1@labe.felk.cvut.cz. D. Cuesta is with Department of Computer Science,Polytechnic University of Valencia, Spain. K. Schmidt is with Faculty of Biomedical Engineering, Department of Biomedical Technology, Czech Technical University in Prague.

C. General Correlation Dimension

1) Parameter Selection and Signal Embedding: Only a time series of one variable was used as an input, but this time-series data was used to reconstruct a multidimensional embedding space [7]. It is necessary to determine the following three parameters i) *time delay* between sampled values, ii) *time interval* between successive vectors and iii) *embedding dimension*.

The *time delay* between sampled values was estimated using the auto mutual information. The first minimum of the auto mutual information was preferred value for attractor reconstruction from time series [7].

The *time interval between successive vectors* was set to be equal to the sampling time. In this case the number of vectors was equal to the number of samples in the time series minus the embedding dimension [7].

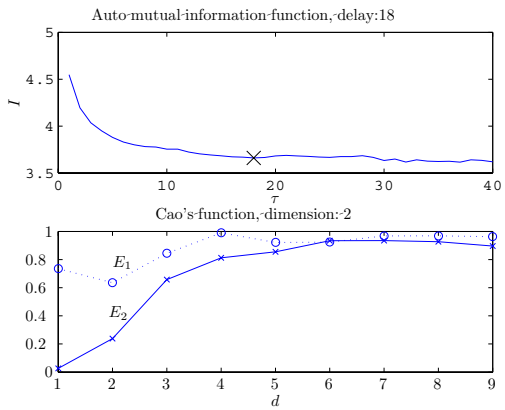


Fig. 2. Auto-mutual information and Cao's method functions $E_1(d)$ and $E_2(d)$ for a signal in $class_1$. The first minimum of the auto mutual function appears at time delay $\tau = 15$. regarding embedding dimension, the optimal value is found in two dimensional space.

To determine the optimal *embedding dimension* from the time series the reliable and commonly used Cao's method was used [8]. We used two measures: $E_1(d)$ which is the average measure of a ratio of the Euclidian distance between the reconstructed vector and its nearest neighbor and $E_2(d)$ which is useful to distinguish deterministic signals from stochastic signals. We were looking for the first minimum either in $E_1(d)$ or $E_2(d)$ as shown in Fig.2. In practical computations, it was difficult to resolve whether the $E_1(d)$ was slowly increasing or had stopped changing if d was sufficiently large. Therefore, the $E_2(d)$ quantity was considered [8].

2) Correlation Dimension: The correlation dimension is computed most efficiently by the correlation sum:

$$C(d,r) = \frac{1}{N_{pairs}} \sum_{i=d}^{N} \sum_{j<i-w} \Theta(r - (y_i - y_j)) \tag{1}$$

where y_i are m-dimensional delay vectors. $N_{pairs} = (N - d - w)(N - d - w + 1)/2$ is the number of pairs of points covered by the sums. Θ is the Heaviside step function: $\Theta(x)$ is zero for $x < 0$ and one for $x \geq 0$. The summation counts the number of pairs (y_i, y_j) for which the distance $|y_i - y_j|$ is less than r. All pairs of points in (3) whose time indices differ by less than w were ignored in order to exclude temporally correlated points. The fractal dimension of order 2,D_2, is than defined as:

$$D_2 = \lim_{r \to 0} \frac{\log C(r)}{\log r} \tag{2}$$

3) General Correlation Dimension: GCD are a class of metrics to characterize the fractality [7]. It is based on counting the number of points in a box. Let B_i denote the ith box, and let $P_i = \mu(\mathcal{B}_i)/\mu(\mathcal{A})$ be the normalized measure of this box, where $\mathcal{A}$ is the fractal whose dimension has to be computed and μ is the population mean of a set. In other words, P_i is the probability for a randomly chosen point on the attractor to be in $\mathcal{B}_i$, and it is usually estimated by counting the number of points that are in the ith box and dividing by the total number of points. If the attractor had been embedded in dimension d, hypercubes of dimension d would be used. The generalized dimension is defined by:

$$GCD_q = \frac{1}{q-1} \lim_{r \to 0} \frac{\log \sum_i P_i^q}{\log r} \tag{3}$$

D. Approximate Entropy

ApEn is a statistic measure that quantifies the unpredictability of fluctuations in a time series such as an instantaneous heart rate time series [9]. Informally, given N points, the family of statistics $ApEn(m, r, N)$ is approximately equal to the negative average natural logarithm of the conditional probability that two sequences that are similar for m points remain similar, that is, within a tolerance r, at the next point. Thus, a low value of ApEn reflects a high degree of regularity.

E. Detrended Fluctuation Analysis

DFA quantifies the presence or absence of long-range correlations. This technique is a modification of root-mean-square analysis of random walks applied to nonstationary data. Briefly, the time series to be analyzed (with N samples) is first integrated [9]. Next, the integrated time series is divided into boxes of equal length, n. In each box of length n, a least squares line is fit to the data (representing the trend in that box). The y coordinate of the straight line segments is denoted by $y_n(k)$. Next, the integrated time series, $y(k)$ is detrended by subtracting the local trend, $y_n(k)$, in each box. The root-mean-square fluctuation of this integrated and detrended time series is calculated.

F. Lempel-Ziv Complexity

LZC and its variants have been used widely to identify nonrandom patterns in biomedical signals obtained across distinct physiological states [10]. In general, LZ complexity measures the rate of generation of new patterns along a sequence and in the case of ergodic processes is closely related to the entropy rate of the source.

G. Katz-Sevcik Fractal Dimension

The Katz-Sevcik algorithm for obtaining a fractal dimension is largely based on morphology, and is calculated as [11] $KFD = \log_{10} n / \log_{10} n + log_{10} \frac{g}{L}$ where n is the number of increments between samples of the signal over which KFD is calculated; L is the sum of all the distances between successive increments; and g is the value of the maximum distance measured from the beginning of the first increment. It has to be pointed out that before calculating KFD, normalization along the y- and x-axes of the signal is performed.

H. Variance Fractal Dimension

The VFD is determined by the Hurst exponent, H, whose calculation was derived from the properties of fractional Brownian motion [12]. This calculation is based on the power law relationship between the variance of the amplitude increments of a signal, $\mathcal{C}(t)$, which was produced by a dynamical process over a time increment $\triangle t = \mid t2 - t1 \mid$, with $\mathcal{C}(t2) - \mathcal{C}(t1)$ denoted as $\triangle\mathcal{C}$. The power law is as follows: $Var[\triangle\mathcal{C}] \sim \triangle t^{2H}$ where the Hurst exponent is:

$$H = \lim_{\triangle t \to 0} \left(\frac{\log_2 Var[\triangle\mathcal{C}]}{\log_2 \triangle t} \right) \tag{4}$$

The VFD for a process with embedding Euclidean dimension, E (equal to 1 for A-EGMs signal), is determined by: $VDF = E + 1 - H$.

I. Box Counting method

An established approach to compute the fractal dimension of a set is the box-counting method (BFD) [13]. In detail, for a set of N points in $\mathcal{R}^d$, and a partition of the space in grid cells of length l, the fractal dimension DB can be derived from:

$$BFD = -\lim_{l \to 0} \frac{\log_{10} N(l)}{\log_{10}(l)} \tag{5}$$

where $N(l)$ represents the number of cells occupied by at least one point.

J. Statistical Evaluation

Non-parametric Kruskal Wallis test was applied for regularity measures comparison. The test was used in order to cope with a smaller number of A-EMG signals, specially in class 4, where 13 signals are only available.

III. Results and Discussion

The reliability of the A-EGM fractionation assessment using regularity measures is summarized in Table I. Mean and standard deviation values are reported for each class.

The optimum parameter setting for ApEn, DFA and BFD was found by a trial and error method for 10 runs of the experiment. In case of ApEn, tolerance parameter r was increased by 0.05 while keeping pattern length m fixed. Considering DFA, parameters for defining the slope of two DFA curves were increased by 4 in each run. Finally, BFD dimension dim_{BFD} was incremented by 1 in each run. The final parameters that resulted in minimum p-values across 10 runs were following: GCD: embedding dimension-variable for each signal, Cao's method was applied, time delay: variable for each signal, auto-mutual information was applied; ApEn: tolerance $r = 0.1$, pattern length $m = 2$; DFA: $fast = 2, mid = 32, slow = 64$; BFD: dimension $dim_{BFD} = 5$; LZC: binary coding was used.

TABLE I
Comparison of A-EGM fractionation

Method	C1	C2	C3	C4	Kruskal
GCD_{-6}	0.034 0.029	0.061 0.049	0.078 0.079	0.13 0.078	3.2e-7
GCD_{-5}	0.057 0.021	0.071 0.065	0.11 0.099	0.17 0.076	4.9e-6
GCD_1	**1.22** **0.27**	**1.50** **0.41**	**1.80** **0.30**	**2.16** **0.13**	**8.4e-11**
ApEn	**0.28** **0.088**	**0.46** **0.096**	**0.57** **0.070**	**0.71** **0.050**	**6.9e-13**
DFA	1.24 0.26	1.31 0.22	1.41 0.16	1.41 0.15	1.8e-3
LZC	0.37 0.040	0.47 0.061	0.51 0.071	0.60 0.054	9.1e-13
KFD	1.36 0.020	1.38 0.021	1.38 0.028	1.41 0.020	2.4e-8
VFD	1.78 0.17	1.67 0.17	1.64 0.14	1.65 0.13	0.1e-3
BFD	1.43 0.040	1.47 0.036	1.50 0.042	1.54 0.038	1.5e-9

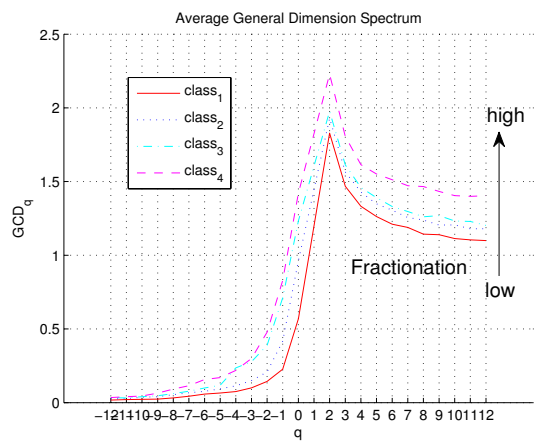


Fig. 3. The average generalized dimension spectrum for all classes of A-EGM fractionation.

Fig.3 shows the average generalized dimension spectrum. The spectrum should be convex, monotonically increasing [7]. The differences among classes are more apparent from general dimension $q \geq -2$. Especially, GCD values of $class_3$ are bigger than values of $class_4$ for generalized dimension $q = -4$. GCD measure reaches peak at $q = 2$ indicating that higher embedding dimensions $q > 2$ do not provide a reliable measure of discrimination due to the appearance of numerical errors in higher dimensional space. The GCD values for $q = -6, -5, 2$ were reported in Table I in order to point out the importance of using higher negative correlation dimensions.

Our assumption about fractal nature of A-EGM signals is confirmed by very low p-values in Table I. Even if significance level is set to $\alpha = 1e-3$, most fractal dimensions are statistically significant in differentiating A-EGM classes disorganization. The exception are DFA and VFD measures. Furthermore, LZC,VFD and BFD measures have lower discrimination capability compared to GCD_1 and ApEn measures as can be seen in Table I. The main reason is data insufficiency; A-EGMs are very short segments of 1.5s duration sampled with 977Hz which accounts for 1537 values. First, ApEn is able to work properly if input dataset contains at least 1000 samples [9]. Second, the good discrimination was mainly achieved by careful selection of algorithm parameters, especially in case of the GCD measure where auto-mutual infirmation and Cao's approach were applied.

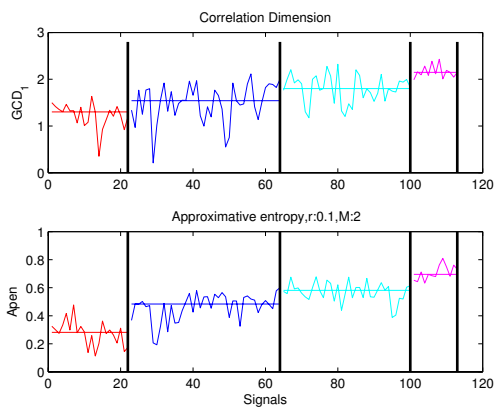


Fig. 4. Assessment of A-EGM fractionation using Correlation Dimension and Approximate Entropy. Horizontal line represents mean value for each class.

We compared two best performing dimensions according to the results of the statistical tests: General Correlation Dimension of order 1 and Approximate Entropy - see bold values in Table I. Fractal dimensions were calculated for all A-EGM signals in the dataset. General correlation dimension of order 1 is shown in upper part of Fig. 4. The first and the last classes are well separated by GCD_1 values. Note that intermediate classes 2 and 3 partially overlap. In case of ApEn, the separation of the 2nd and 3th class is slightly better compared to GCD_1 dimension as can be seen in lower part of Fig. 4.

IV. CONCLUSIONS

In the era of catheter ablation of AF, the initial attempts to describe A-EGMs during AF were predominantly based on frequency-domain analysis of atrial signals [3]. Not only dominant frequency (DF) but also the level of A-EGMs fractionation may be a clinically important descriptor of local atrial signal [1]. Sites with highly fractionated A-EGMs almost fully encompass the sites with high DF whereas the opposite is not always true. Therefore it is very important to develop other measures which could assess A-EGMs regularity, especially differentiating low A-EGMs fractionation (LF) from high fractionation (HF).

This study revealed the presence of nonlinear dynamics in A-EGM across all selected fractionation levels. Fig. 4 documents the fact that with increasing complexity of A-EGM's signals the value of regularity measures increases too. Across all levels of A-EGM fractionation, the discrimination using regularity measures was found statistical significant at significance level $\alpha = 0.001$. The proposed complex measures were successful in separating class C_1(LF) from class C_4 (HF) of the A-EGM signals. However, separation between intermediate classes 2 and 3 was not so clear. One of the reasons is that the classes were defined by averaging the classification by 3 experts to obtain semi-continuous scale of fractionation. In some cases it was very difficult to assign A-EGM signals either to class 2 or class 3.

Regarding future studies, the selected complex measures (GCD_1 and ApEn) can be used as additional features for automated and operator independent system that facilitates AF substrate ablation [5].

ACKNOWLEDGMENT

The project was supported by Ministry of Education, Youth and Sport of the Czech Republic with the grant No. MSM6840770012 entitled "Transdisciplinary Research in Biomedical Engineering II" and by Spanish Ministry of Science and Innovation, grant TEC2008-05871/TEC.

REFERENCES

[1] K. Nademanee and et al, "A new approach for catheter ablation of atrial fibrillation: mapping of the electrophysiologic substrate," *J Am Coll Cardiol.*, vol. 43, pp. 2044–2053, 2004.

[2] K. Kumagai, "Patterns of activation in human atrial fibrillation," *Heart Rhythm*, vol. 4, no. 3, pp. S7–S12, 2008.

[3] J. Ng, J. Gold, and J. Goldberger, "Understanding and interpreting dominant frequency analysis of af electrograms," *Journal of Cardiovascular Electrophysiology*, vol. 18, no. 6, pp. 680–685, 2007.

[4] B. H. et al, "Nonlinear analysis of epicardial atrial electrograms of electrically induced atrial fibrillation in man," *J Cardiovasc Electrophysiol.*, vol. 6, no. 6, pp. 419–440, 1995.

[5] V. Křemen, L. Lhotská, R.Čihák, V.Vančura, J. Kautzner, and D. Wichterle., "A new approach to automated assessment of endocardial electrograms fractionation in human left atrium during atrial fibrillation," *Physiol. Measurement*, vol. 29, pp. 1371–1381, 2008.

[6] V. Křemen, "Automated assessment of endocardial electrograms fractionation in human," Ph.D. dissertation, Czech Technical University in Prague, 2008.

[7] R. Hilborn, *Chaos and Nonlinear Dynamics*. Oxford University Press, 2000.

[8] L.Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, no. 1-2, pp. 43–50, 1997.

[9] K.Ho and et al., "Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics," *Circulation*, vol. 96, no. 3, pp. 842–848, 1997.

[10] J.Hu, J.Gao, and J.Principe, "Analysis of biomedical signals by lemepl-ziv complexity: the effect of finite data size," *IEEE Trans. Biomed. Engg.*, vol. 53, no. 12, pp. 2606–2609, 2006.

[11] C. Sevcik, "On fractal dimension of waveforms," *Chaos, Solitons & Fractals*, vol. 28, no. 2, pp. 579–580, 2006.

[12] J. Gnitecki and Z.Moussavi, "Variance fractal dimension trajectory as a tool for heart sound localization in lung sounds recording," in *Proc. IEEE Eng. Med. Biol Soci. (EMBS)*, 2003, pp. 2420–2423.

[13] S. Liebovitch and T. Toth, "A fast algorithm to determine fractal dimensions by box-counting," *PHYS-LETT-A*, vol. 141, no. 8-9, pp. 386–390, 1989.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2008 (17.01.2008)

PCT

(10) International Publication Number
WO 2008/007236 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/IB2007/052095

(22) International Filing Date: 5 June 2007 (05.06.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
06115092.6 7 June 2006 (07.06.2006) EP

(71) Applicant *(for all designated States except DE, US)*: **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(71) Applicant *(for DE only)*: **PHILIPS INTELLECTUAL PROPERTY & STANDARDS GMBH** [DE/DE]; Lübeckertordamm 5, 20099 Hamburg (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCHMIDT, Ralf** [DE/DE]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **HARRIS, Matthew** [AU/DE]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **NOVAK, Daniel** [CZ/CZ]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **PERKUHN, Michael** [DE/DE]; c/o Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agents: **SCHOUTEN, Marcus, M.** et al.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

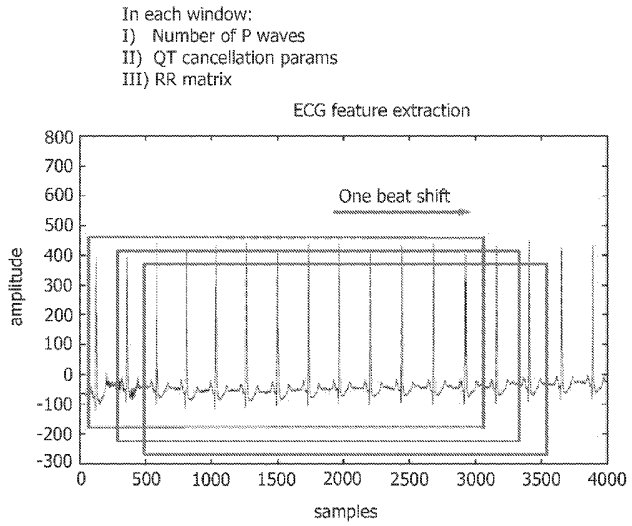
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: ATRIAL FIBRILLATION DETECTION



(57) Abstract: A method of detecting atrial fibrillation from an ECG is presented, comprising a combination of at least two of the steps of analyzing R-R interval sequences to produce a measure of the irregularity of the RR interval sequence; canceling the QRST portion of the ECG, and analyzing the resulting signal; and analyzing the ECG signal preceding the QRS complex, to determine the presence or absence of P-waves. This is followed by a step of using a classifier to classify the ECG into one of two classes, namely "AF present" and "AF absent", based on a range or ranges of results of the measurement steps which are determined in advance. The invention improves patient monitoring.

WO 2008/007236 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Atrial fibrillation detection

This invention relates to the detection of atrial fibrillation, on the basis of ECG readings from electrodes. It may be used for example with readings derived from electrodes incorporated in "wearable" electrode systems.

Atrial fibrillation ("AF") is a common heart arrhythmia with a prevalence of approximately 0.4 to 1% of the general population, and its prevalence increases with age. It is responsible for the highest number of hospital admissions due to arrhythmias, and consequently, it is desirable to be able to monitor the condition of patients, using portable devices which are capable of producing reliable indications of arrhythmia, without producing false positives.

Electrocardiograph (ECG) signals show a characteristic pattern of electrical impulses that are generated by the heart. Different waves are identifiable in the ECG signal – the P wave is from the atrial excitation and the QRS complex and T-wave are from the ventricular excitation and relaxation, respectively, as illustrated in Figure 1.

The ST segment usually appears as a straight, level line between the QRS complex and the T wave. Elevated or depressed ST segments may mean the heart muscle is damaged or not receiving enough blood, a sign that a myocardial infarct may have occurred.

The T wave corresponds to the period when the lower heart chambers are relaxing electrically and preparing for their next muscle contraction.

AF is a heart rhythm which is usually characterized by QRS complexes with normal morphology and with irregular arrival times. This can be caused by a diseased atrium which disrupts the normal passage of electrical stimulus from the sinus node through the atrium to the ventricles. One example of AF is depicted in Figure 2.

AF can be either chronic or intermittent. Intermittent AF is referred to as paroxysmal AF. AF is difficult to detect, particularly if it is paroxysmal, since a sample ECG recording from paroxysmal AF subject may not contain any actual episodes of AF. It is therefore preferable to monitor paroxysmal patients on a regular basis without causing them any discomfort. For this purpose a wearable measurement system that is incorporated in a textile has been developed [1]. If an indication of AF is detected by a suitable method, and

preferably confirmed by a cardiologist, a drug administration or other suitable therapeutic intervention can be provided to manage AF treatment.

AF detection is most often based upon R-R analysis. Attempts have been made to detect AF based on R-R interval sequences using a variety of statistical methods [2]. Another indicator for AF is the absence of clear P-waves before the QRS complex. In such cases, it may be possible to diagnose AF on the basis of a lack of regularly occurring P-waves [3], [4]. Another possible approach is to apply QRST cancellation [5], so as to remove the ventricular activity from the signal, and then calculate the power spectrum of the remainder ECG.

Because of the difficulty of obtaining consistent results from the various different methods of detection and subsequent analysis of the detected ECGs, it would be preferable to find a more universally-applicable method of detecting and monitoring the condition. Preferably this should combine the best features of existing methods, while appropriately weighting the significance of the detected signals from each one.

Accordingly, the present invention provides a method of detecting atrial fibrillation from an ECG comprising a combination of at least two of the steps of:

(a) analyzing R-R interval sequences to produce a measure of the irregularity of the RR interval sequence;

(b) canceling the QRST portion of the ECG, and analyzing the resulting signal; and

(c) analyzing the ECG signal preceding the QRS complex, to determine the presence or absence of P-waves;

followed by a step of using a classifier to classify the ECG into one of two classes, namely “AF present” and “AF absent”, based on a range or ranges of results of the selected measurement steps which are determined in advance.

Accordingly, in one embodiment of the invention the selected measurements comprise the steps of (a) analyzing R-R interval sequences; and (c) determining the presence or absence of P-waves, as set out above. In another embodiment the selected measurements are (a) analyzing R-R intervals and (b) analyzing the signal remaining after QRST cancellation, while in a third embodiment the selected measurements are (b) canceling the QRST portion of the ECG and analyzing the resulting signal; and (c) determining the presence or absence of P-waves. Of course it will also be appreciated that all three steps (a), (b) and (c) may also be used in combination.

AF detection can only be carried out on clean (i.e. relatively noise free) signals. Detecting noisy ECG segments (to then be discarded from further treatment) can be done by a combination of threshold detection and the identification of high frequencies which are usually characteristic of noise.

The initial collection of the data may, for example, be carried out using a wearable belt including three integrated dry electrodes based on carbon loaded rubber. This allows the device to be easily worn by a patient, and can be arranged to transmit signals wirelessly, for example by means of bluetooth, to an external PC or other portable computing device.

Accordingly, the invention also extends to apparatus for use in the detection and/or monitoring of atrial fibrillation, comprising a wearable device incorporating electrodes adapted for contacting the skin, and means for transmitting detected signals to a computer system which is arranged to detect an AF condition by the method of the invention as outlined above.

The device preferably incorporates a wireless transmission system such as "Bluetooth".

Preferably, the wearable device is integrated into an item of clothing such as a belt or shirt, so that it can be held in suitably good contact with the patient's skin. The computer system may be a PC or a hand-held device such as a notebook computer, PDA or "smartphone".

Some embodiments of the invention will now be described by way of example with reference to the accompanying drawings in which:

Fig. 1 is a simplified diagram of a typical ECG signal;

Fig. 2 is an example of an AF episode detected by the methods of the present invention;

Figs. 3a and 3b illustrate the process of noise detection;

Fig. 4 shows a decision tree algorithm;

Fig. 5 illustrates a "sliding window" technique in feature generation; and

Fig. 6 illustrates a wearable measuring device.

The data may, for example, be collected using a wearable measuring device comprising a belt 2 with three integrated dry electrodes 4, and incorporating a miniaturized ECG amplifier indicated at 6, as illustrated in Figure 6. The electrodes, which are based on carbon-loaded rubber, are fixed into the belt using a thermal moulding process. The position of the belt is preferably around the chest to obtain an optimal ECG signal. The battery capacity preferably allows measuring for at least 7 days continuously in a typical operation mode. Data is transmitted to a PC via the bluetooth interface. Figure 2 shows typical example of AF acquired by the wearable system.

In case of paroxysmal patients, a P wave template can be selected from the normal sinus segment for each patient. Consequently, this P wave template is compared to the P wave candidate before QRS complex. In case of an AF segment, the P wave may disappear, resulting in possible indication of AF. The P wave general template is generated from healthy patients.

The patient is rested during measurement and if any significant noise is present the noisy segments are rejected, using known methods of noise detection. For example this can be done by identifying high frequency regions of the signal (normally indicative of noise) and applying threshold detection.

This process is illustrated by Figures 3a and 3b. It can be seen that the parts with saturation noise and high frequency noise have been successfully detected.

Feature Extraction

Before the feature extraction itself, as a preliminary step of the ECG signal, the fiducial points must be detected, for example by using the modified Pan-Tompkins algorithm [8]. There are three important feature groups used in detection of atrial fibrillation, features using RR interval information, features using P-wave morphology, and features using QRST cancellation. In the preferred embodiment of the invention a combination of features from at least two of these groups is used.

1. The first feature group relates to the RR intervals. A measure of the irregularity of the RR intervals can be obtained from the RR interval transition matrix used in [2]. This matrix represents the relative frequency of transitions between intervals whose lengths are either short, regular or long.

2. The second feature group is a test for the presence of a P wave. In normal sinus rhythm, the P wave can be observed before QRS complex while in a case of AF, there is no P wave present. The P wave detection can be done, for example, using template matching in which a correlation coefficient is used as a dissimilarity measure between the P

wave candidate and a template. A threshold must be chosen to allow acceptance of very similar beats. In this way, each QRS complex can be labeled as having/not having a preceding P wave.

3. Finally, the last feature group consists of the frequency domain properties of ECG remainder obtained after QRST cancellation. The remainder electrocardiogram after the ventricular component has been removed represents the atrial activity component of the signal. Fiducial points of ventricular complexes are marked, preferably using a method based on the algorithm presented by Pan and Tompkins [6]. Basically, the average beat is aligned with the fiducial points of all dominant beat windows and subtracted. The absolute powers in the frequency bands of PSD spectrum extending from 10, 20, 30, 40, 60, 80Hz to 125Hz are estimated (e.g. P20 is the summation of the power found in frequency bands between 20 and 125 Hz). Ratios of high frequency (from F to 125Hz) to low frequency (extending from 0Hz to F Hz) are also calculated. As a percentage they are expressed as:

$$P(F) = \frac{\sum_{f=F}^{125HzP}}{\sum_{f=0}^{FHzP}} \quad \text{x } 100$$

Since the AF wave has a random character, the entropy of the remainder may be calculated as well.

Feature Selection

Using the above described methods may result in a large number of features. In order to reduce the dimension of the feature space it is possible to use a decision tree algorithm. Preferably the two most significant features are retained by looking at the first levels of the resulting decision tree. One simplified example of the decision tree process is shown in Figure 4 where two features of the R-R interval analysis and the P wave template matching are selected.

The features are extracted in a sliding window consisting of 30 beats, as shown in Figure 5. This approach results in a one-to-one correspondence between features and beats in the stream. In this way, each beat is labeled individually, rather than in groups.

Classification

It is possible to use various different classifiers to analyze the resulting waveforms.

For example there are quadratic classifiers, normal densities based linear classifiers or normal densities based quadratic classifiers. There are Bayes normal classifiers, where in the first case one assumes equal covariance matrices resulting in a linear discriminant function (LDC). In the second case the covariances matrices are different for each category resulting in a quadratic discriminant function (QDC). Other alternative classifiers are a k-nearest neighbor classifier (e.g. 3-KNN) or a neural network, such as a back propagation neural network. As an example this may comprise one hidden layer of 10 neuron units and one output neuron unit (10-ANN). Other possibilities include Support Vector Machines (SVM) or a C4.5 decision tree.

Analysis of Results

Feature extraction is preferably performed automatically using a decision tree structure. It can also be performed manually by looking at different scattered plots and statistical parameters such as the correlation matrix. In one preferred embodiment of the invention, two features as an input for classifier are selected, using automatic analysis, which comprise one feature from the R-R interval analysis and one feature from the group of P template matching (number of found P waves in the window of 30 beats long).

The QRST cancellation implemented in a preferred embodiment of the invention subtracts the mean beat computed for the whole record [5]. When several QT beat morphologies are presented in the signal, the cancellation technique may be inadequate. Due to big differences in even interpersonal ECG morphology two or three beat templates are preferably computed using an unsupervised approach such as hierarchical clustering.

In this way atrial fibrillation detection can be reliably achieved using simple features combined with a suitable classifier. Most algorithms requires only time or morphology information for AF classification. The approach of the present invention combines both methods.

REFERENCES:

1. J Muhlsteff, O Such, and R Schmidt, "Wearable approach for continuous ecg – and activity patient-monitoring," in *26th Annual International Conference of the IEEE EMBS*, vol. 1, 2004, pp. 2184-2187.

2. G Moody and R Mark, "A new method for detecting atrial fibrillation using r-r intervals," in *Computers in Cardiology*, vol. 10, 1983 pp. 227-230.

3. P Stafford, P Denbigh, and R Vincent, "Frequency analysis of the p wave:comparative techniques," in *Pace*, vol. 18, 1995, pp. 261-270.

4. P Stafford and J Kolvekar, S Cooper, "Signal averaged p wave compared with standard electrocardiography or echocardiography for prediction of atrial fibrillation after coronary bypass grafting," in *Heart*, vol. 77, 1997, pp. 417-422.

5. J Slocum, A Sahakian, and S Swiryn, "Diagnosis of atrial fibrillation from surface electrocardiograms based on computer-detected atrial activity." *Journal of Electrocardiology*, vol. 25, no. 1, pp. 1-8, 1992.

6. J Pan and W Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 32. 230-236, 1985.

CLAIMS:

1. A method of detecting atrial fibrillation from an ECG comprising a combination of at least two of the steps of:

(a) analyzing R-R interval sequences to produce a measure of the irregularity of the RR interval sequence;

(b) canceling the QRST portion of the ECG, and analyzing the resulting signal; and

(c) analyzing the ECG signal preceding the QRS complex, to determine the presence or absence of P-waves;

followed by a step of using a classifier to classify the ECG into one of two classes, namely "AF present" and "AF absent", based on a range or ranges of results of the measurement steps which are determined in advance.

2. A method of detecting atrial fibrillation according to claim 1, in which two measurement steps are selected, comprising:

(a) analyzing R-R interval sequences; and

(c) determining the presence or absence of P-waves.

3. A method of detecting atrial fibrillation according to claim 1, in which two measurement steps are selected, comprising:

(a) analyzing R-R interval sequences; and

(b) canceling the QRST portion of the ECG and analyzing the resulting signal.

4. A method of detecting atrial fibrillation according to claim 1 in which two measurement steps are selected, comprising:

(b) canceling the QRST portion of the ECG and analyzing the resulting signal; and

(c) determining the presence or absence of P-waves.

5. A method according to any one of claims 1, 2 or 4 in which the power spectrum of the resulting signal is analyzed after cancellation of the QRST portion.

6. A method according to any one of claims 1 to 5 in which the measurement features for classification are selected in advance using a decision tree algorithm.

7. A method according to any one of the preceding claims in which the classifier comprises a quadratic classifier, a normal densities based quadratic classifier, a normal densities based linear classifier, or a k-nearest neighbor classifier.

8. A method according to one of claims 1 to 6 in which the classifier comprises a neural network, a C4.5 decision tree, or a Support Vector machine (SVM).

9. A method according to any one of claims 1 to 3 in which the R-R feature used as input to the classifier comprises a value derived from an RR interval transition matrix representing the relative frequency of transitions between intervals whose lengths are either short, regular or long.

10. A method according to any one of claims 1, 2 or 4 in which the P-wave feature used as an input to the classifier comprises a value derived using template matching which gives a measurement of dissimilarity measure between the actual P-wave and a template.

11. Apparatus for detecting and/or monitoring atrial fibrillation, comprising a wearable device incorporating electrodes for collecting signals via the skin of a patient, and means for transmitting detected signals to a computer system, the computer system being programmed to analyze the detected signals by a method in accordance with any preceding claim.

12. Apparatus according to claim 11 in which the signal transmitting means comprises a wireless system.

13. Apparatus according to claim 11 or claim 12 in which the wearable device comprises a belt, shirt or harness.

14. Apparatus according to any one of claims 11 to 13 in which the computer system comprises a mobile device which can also be worn, or carried, by the patient.

15. Apparatus according to claim 14 in which the mobile device is adapted to send signals to an external monitoring means.

16. A method of treating a subject known to be prone to atrial fibrillation, comprising: regularly monitoring the incidence of AF using the detection method of any one of claims 1 to 10 and a wearable device according to any one of claims 11 to 15, and controlling therapeutic interventions to the subject in accordance with the results of the monitoring procedure.

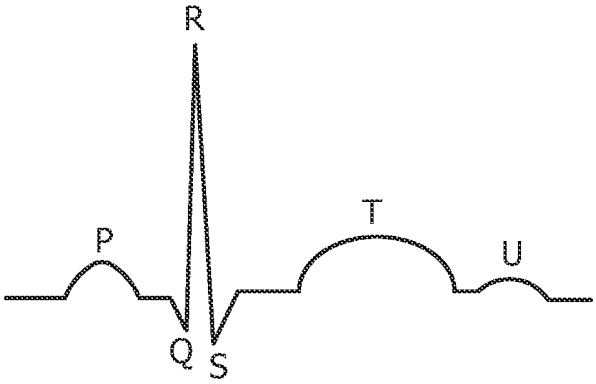


FIG. 1

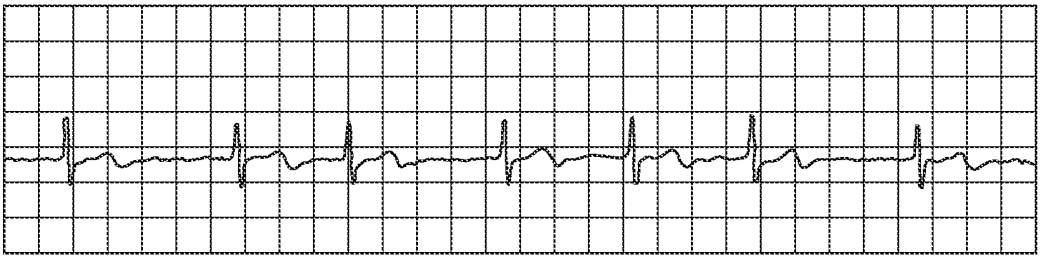


FIG. 2

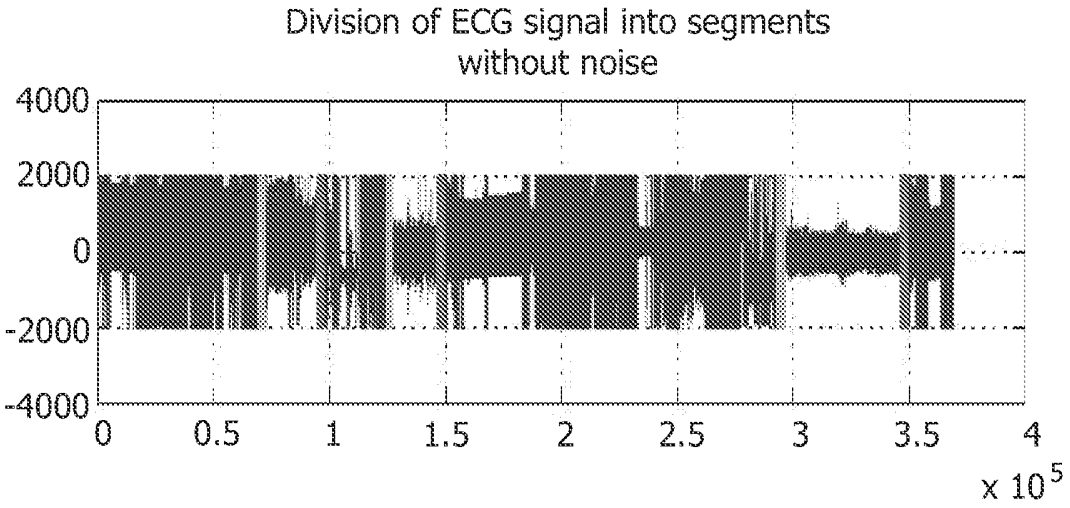

Division of ECG signal into segments
without noise
4000
2000
0
-2000
-4000
0
0.5
1
1.5
2
2.5
3
3.5
4
x 10^5

FIG. 3a

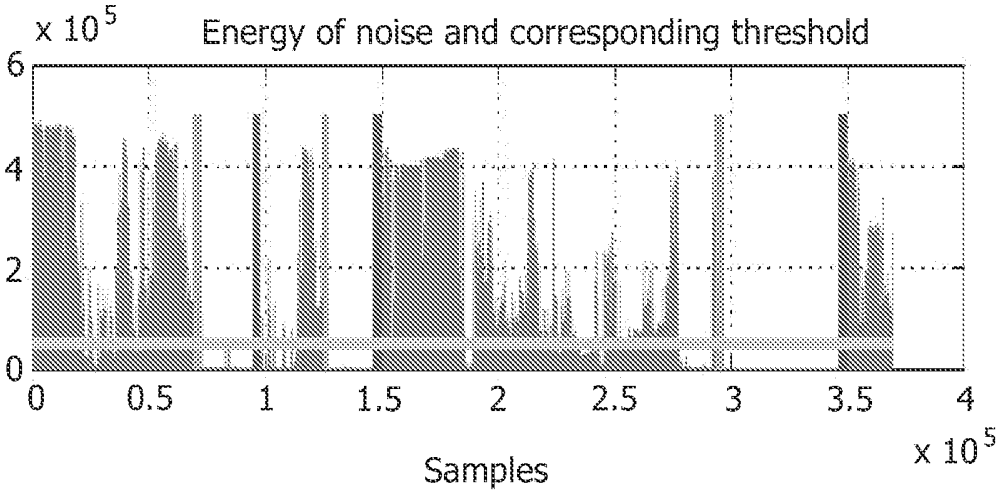

x 10^5
Energy of noise and corresponding threshold
6
4
2
0
0
0.5
1
1.5
2
2.5
3
3.5
4
x 10^5
Samples

FIG. 3b

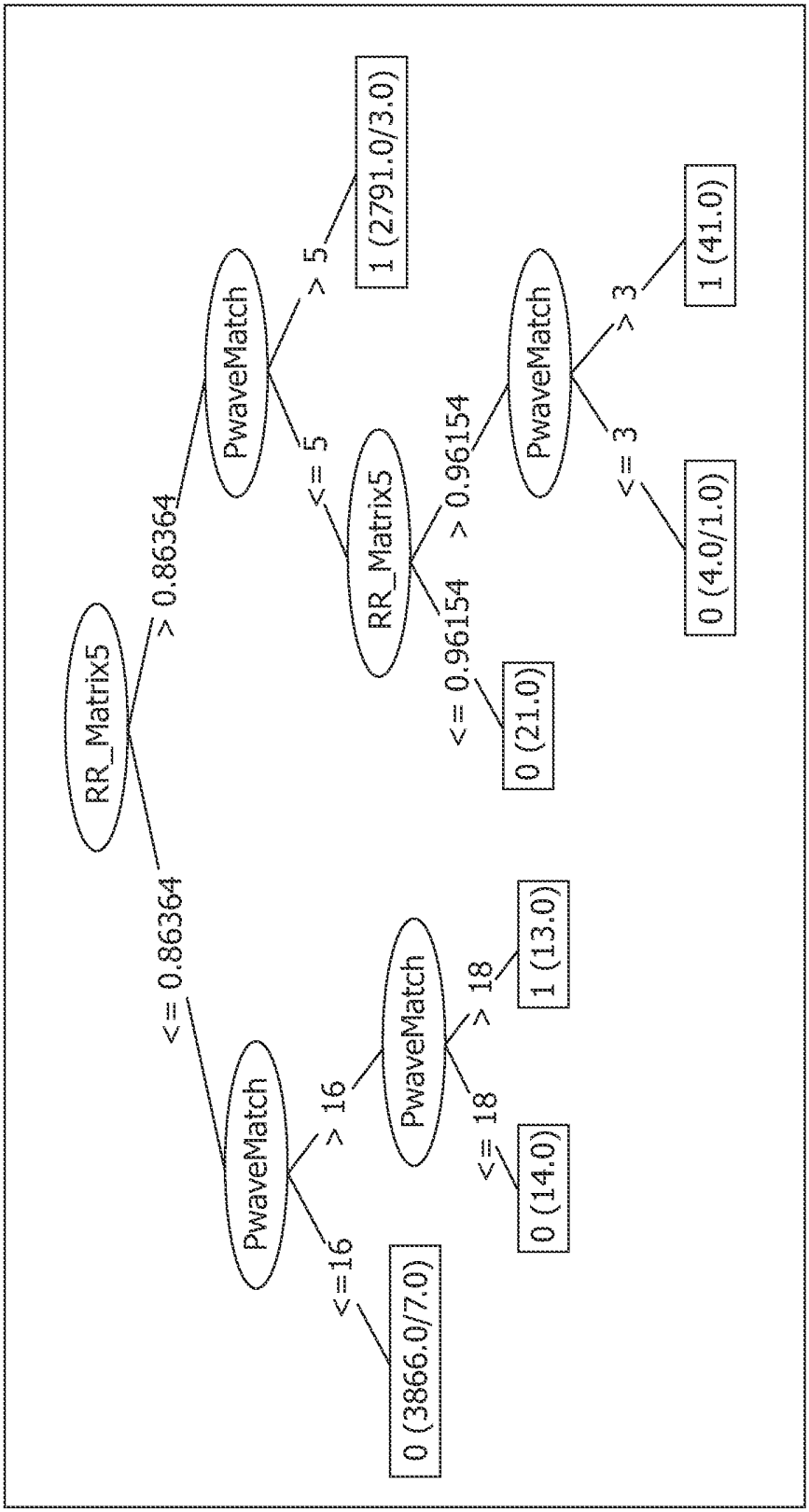

RR_Matrix5
> 0.86364
<= 0.86364
PwaveMatch
> 5
<= 5
1 (2791.0/3.0)
RR_Matrix5
> 0.96154
<= 0.96154
PwaveMatch
> 3
<= 3
1 (41.0)
0 (4.0/1.0)
0 (21.0)
PwaveMatch
> 16
<=16
PwaveMatch
> 18
<= 18
1 (13.0)
0 (14.0)
0 (3866.0/7.0)

FIG. 4

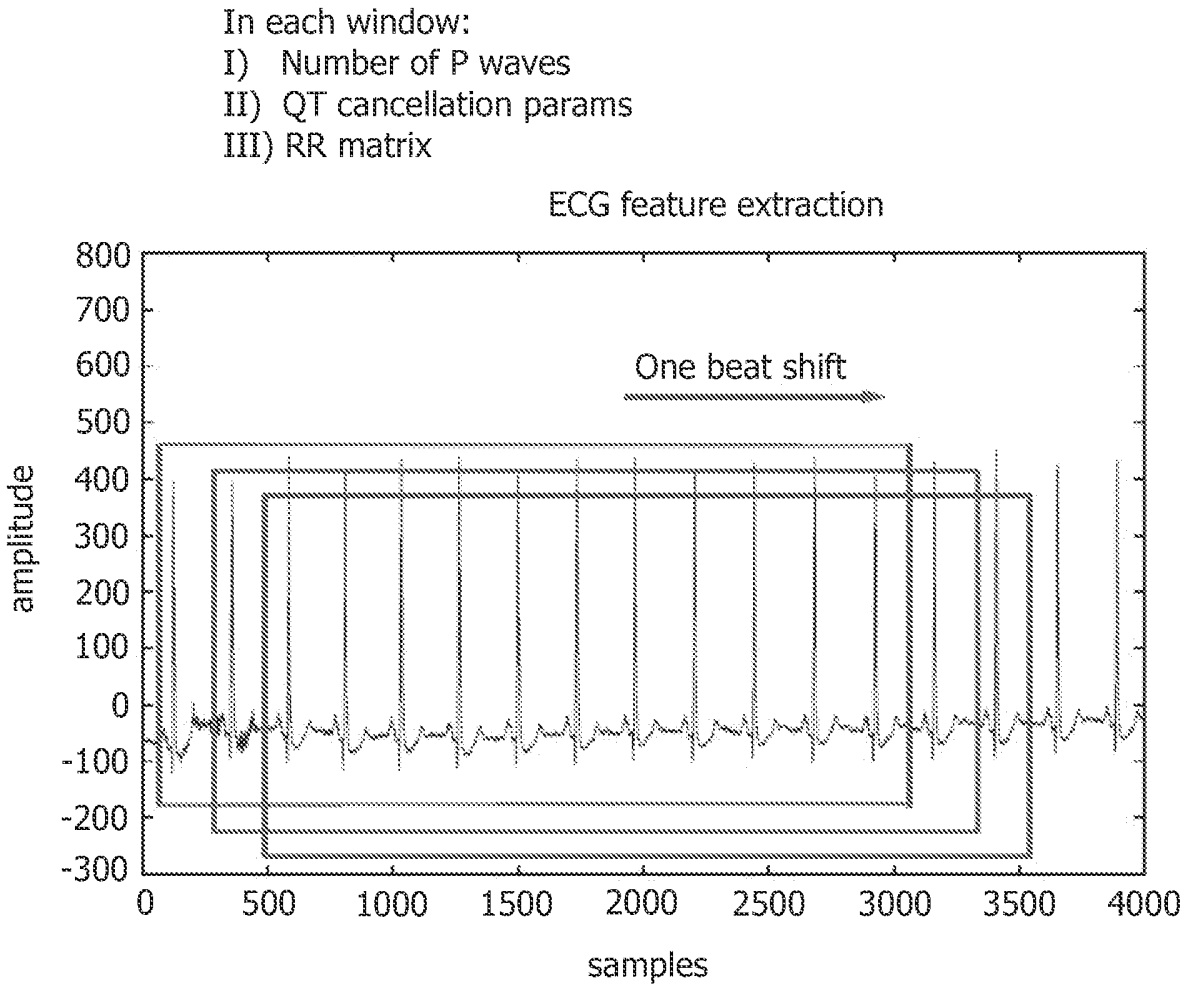

In each window:
I) Number of P waves
II) QT cancellation params
III) RR matrix
ECG feature extraction
One beat shift
amplitude
samples
800
700
600
500
400
300
200
100
0
-100
-200
-300
0
500
1000
1500
2000
2500
3000
3500
4000

FIG. 5

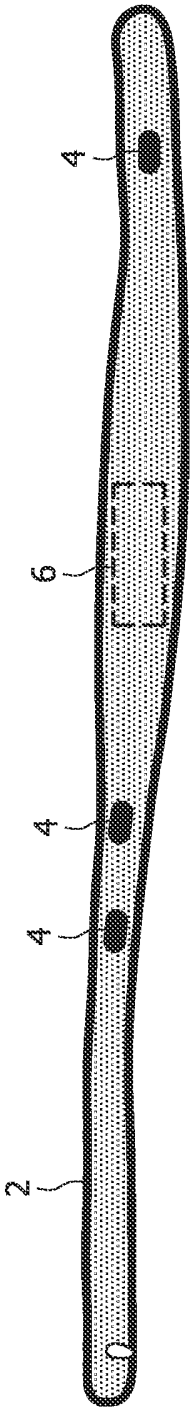


FIG. 6

Chapter 3

Analysis of single-neuron recordings in Parkinson desease

Distinct populations of neurons respond to emotional valence and arousal in the human subthalamic nucleus

Tomáš Sieger[a,b,1], Tereza Serranová[a,1], Filip Růžička[a], Pavel Vostatek[b], Jiří Wild[b], Daniela Šťastná[c], Cecilia Bonnet[a], Daniel Novák[b], Evžen Růžička[a], Dušan Urgošík[c], and Robert Jech[a,2]

[a]Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine and General University Hospital, Charles University in Prague, 128 21, Prague, Czech Republic; [b]Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, 166 27, Prague, Czech Republic; and [c]Department of Stereotactic and Radiation Neurosurgery, Na Homolce Hospital, 150 30, Prague, Czech Republic

Edited by Joseph E. LeDoux, New York University, New York, NY, and approved January 27, 2015 (received for review June 10, 2014)

Both animal studies and studies using deep brain stimulation in humans have demonstrated the involvement of the subthalamic nucleus (STN) in motivational and emotional processes; however, participation of this nucleus in processing human emotion has not been investigated directly at the single-neuron level. We analyzed the relationship between the neuronal firing from intraoperative microrecordings from the STN during affective picture presentation in patients with Parkinson's disease (PD) and the affective ratings of emotional valence and arousal performed subsequently. We observed that 17% of neurons responded to emotional valence and arousal of visual stimuli according to individual ratings. The activity of some neurons was related to emotional valence, whereas different neurons responded to arousal. In addition, 14% of neurons responded to visual stimuli. Our results suggest the existence of neurons involved in processing or transmission of visual and emotional information in the human STN, and provide evidence of separate processing of the affective dimensions of valence and arousal at the level of single neurons as well.

emotion | basal ganglia | subthalamic nucleus | single neuron | arousal

At one time, the subthalamic nucleus (STN), which is an important target for deep brain stimulation (DBS) in the treatment of motor symptoms in Parkinson's disease (PD), was considered an important regulator of motor function (1, 2); however, the occurrence of postoperative neuropsychiatric complications has expanded interest in the nonmotor function of the STN (3, 4). Animal and human studies have already demonstrated the additional functional role of the STN in emotional and motivational processes (5–12). In addition, recent functional MRI studies found STN activation in response to emotional stimuli in healthy subjects (13, 14). Therefore, we hypothesized that emotional activity-related neurons should exist in the STN. Participation of the STN in processing emotion has not yet been investigated directly at the single-neuron level in humans. Nonetheless, single-neuron activity related to a priori defined emotional categories (e.g., positive vs. negative) has been detected in a few human brain regions, including the hippocampus, amygdala, and prefrontal and subcallosal cortex (15–18).

It has been proposed that emotional behavior is organized along two psychophysiological dimensions: emotional valence, varying from unpleasant to pleasant, and arousal, varying from low to high (19). The individual assessment of these dimensions is closely correlated with somatic and autonomic measures of emotions (20, 21). Contrary to a priori categories, they can better reflect emotional characteristics of the stimulus in an individual context, and they take into account interindividual differences based on specific behavioral determinants, such as affective disposition and personality traits (22).

In this study, we aimed to detect single-neuron firing pattern changes in the STN that are related to emotional arousal and valence from the individual ratings of emotionally charged and neutral pictures presented to PD patients undergoing DBS electrode implantation. It has been shown that different features of tasks are linked to neuronal activity in different frequency bands. Whereas beta band oscillations (13–30 Hz) restricted to the dorsolateral (sensorimotor) part of the STN are linked mainly to motor functions and their alteration in PD (23–25), the gamma band oscillations (30–100 Hz) perhaps have a more general meaning. Along with motor functions, they are modulated by picture perception and early emotional arousal (26, 27). Because we were interested in the affective content of visual processing, we focused on the alpha oscillations (8–12 Hz), because they repeatedly showed emotion-related behavior in local field potential (LFP) recordings (7, 28, 29). We used the power spectra bands, which are well known in descriptions of continuous LFP and EEG signals that we adopted for analysis of the discrete single-neuron signal from the STN during the task with affective picture presentation.

In this study, we compared the individual alpha firing activity of single neurons with specific affective experiences expressed in subjective ratings of the emotional valence and arousal of each presented picture, and then mapped these neurons into the STN model (30). A neuron was classified as affective if its history-adjusted (and category-adjusted) activity in the alpha band correlated with these ratings.

Studies of spatiotemporal dynamics of emotions (affective picture or facial emotion processing) have observed early and late changes that have been attributed to different stages of emotional processing

Significance

The involvement of the subthalamic nucleus (STN) in affective processing has been suggested with the appearance of neuropsychiatric side effects of deep brain stimulation in Parkinson's disease (PD), but direct evidence has been lacking. In our study, we recorded single-neuron activity from the STN during affective picture presentation to PD patients intraoperatively. We discovered two spatially distinct populations of "affective" neurons responding to the emotional dimensions of the stimuli: valence (pleasantness-unpleasantness) and arousal (intensity). As previously believed, neural circuits underlying these two affective dimensions are functionally segregated. Here we observed separated emotional processing even at the single neuron level. These results extend our knowledge regarding the emotional role of the STN and the neural basis of emotions.

Author contributions: T. Sieger, T. Serranová, and R.J. designed research; T. Serranová, F.R., E.R., D.U., and R.J. performed research; T. Sieger and P.V. contributed new reagents/analytic tools; T. Sieger, T. Serranová, F.R., P.V., J.W., D.Š., C.B., and D.N. analyzed data; and T. Sieger, T. Serranová, and R.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]T. Sieger and T. Serranová contributed equally to this work.

[2]To whom correspondence should be addressed. Email: jech@cesnet.cz.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1410709112/-/DCSupplemental.

Table 1. Patients' and normative ratings of emotional stimuli used

	Patients' rating		Normative rating	
Category	Valence, mean (SD)	Intensity, mean (SD)	Valence, mean (SD)	Intensity, mean (SD)
Negative	3.1 (1.6)	5.1 (2.6)	3.4 (0.7)	5.2 (1.1)
Neutral	5.2 (1.0)	2.6 (1.7)	5.0 (0.2)	2.8 (0.3)
Positive	6.0 (1.3)	4.0 (2.1)	6.6 (0.8)	5.2 (1.1)

Patients' ratings represent subjective ratings assessed at 1 mo after bilateral insertion of the permanent electrode into the STN after overnight withdrawal of levodopa in the DBS-off condition. Normative ratings are those available from the IAPS (80).

(31–33). Therefore, for our analysis, we arbitrarily split the picture observation period of 2 s into two time windows. Within the early window (0–500 ms), which may contain early emotional images confounded more by perceptual and attentional processes (31, 33–37), we searched for activity related to the affective picture presentation in contrast to the black screen periods preceding each picture. For emotional activity, we searched in the late window, starting at 500 ms after the visual stimulus onset, because this can be better related to emotional processing after the conceptual knowledge of the presented emotion (represented here in individual ratings of the emotional valence and arousal) is built (32).

Results

We recorded single-neuron activity in the STN from 13 PD patients intraoperatively performing an affective task consisting of a presentation with pleasant, unpleasant, and neutral pictures displayed for 2,000 ms, preceded by a black screen with a white fixating cross presented for 3,500–5,500 ms. We acquired 97 microelectrode recordings obtained from 47 sites in the STN, in which a total of 125 neurons were detected. The activity of 35 neurons was related to eye movements, and these neurons were excluded from further analysis. The remaining 90 neurons (71 in the left hemisphere) were searched for early perceptual and emotional characteristics. Normative and postoperatively recorded individual valence and arousal ratings for each picture category are presented in Table 1.

The alpha band activity of 15 of 90 neurons (17%) during the late period of picture presentation epochs (500–2,000 ms after stimulus onset) was related to the emotional content of the presented pictures expressed in individual valence or arousal ratings ($P < 0.05$, uncorrected). The activity of six neurons (7%) correlated with the valence ratings (four neurons negatively, two neurons positively; Fig. 1). The activity of other nine neurons (10%) correlated with the arousal ratings (seven neurons positively, two neurons negatively; Fig. 2). Ten of these neurons were located in the left STN, and the other five were located in the right STN, with no interhemispheric differences ($\chi^2 = 0.854$; df = 1; $P = 0.36$). These 15 emotion-related neurons were more than was expected by chance (test in binomial distribution with a false-positive rate of 0.1; $P < 0.05$). Fig. S1 shows how alpha band activity was derived in one selected neuron associated with the arousal rating.

In addition, 13 neurons (14%) demonstrated significantly altered alpha band activity between the black screen (2,000-ms duration) and the early picture presentation (window 0–500 ms after stimulus onset) ($P < 0.05$). Only one neuron exhibited altered alpha band activity in both the early and late time windows.

The locations of the neurons sensitive to emotional content are depicted in Fig. 3. The valence-related neurons in the STN were located more posteriorly compared with the arousal-related neurons ($P < 0.05$, permutation test). The anterior–posterior difference in the mean position of the neuronal populations was 1.9 mm.

In post hoc analyses, we searched for emotion-related neuronal activity in other frequency bands. Four neurons were related to arousal and no neurons were related to valence in the beta band, but their numbers were insignificant ($P = 0.98$, binomial test). In the gamma band, seven neurons were related to arousal and no neurons were related to valence, again not significant ($P = 0.81$, binomial test). No overlaps of beta and gamma emotion-related neurons with alpha emotion-related neurons were observed.

To support the specificity of emotion-related neurons located in the STN, we analyzed the activity of 32 other eye movement-unrelated neurons in other basal ganglia, including 18 neurons from the substantia nigra pars reticulata (recorded from patients 1, 3, 5, 6, 7, 11, 12, and 13; Table 2) and 14 neurons from the globus pallidus (Table S1). None of these neurons was related to individual valence or arousal ratings of the presented pictures.

Discussion

Using perioperative microrecordings from the STN of patients with PD, we analyzed changes in the firing patterns of single neurons in relationship to visually presented emotional material, and found a relatively large proportion of neurons with activity related to emotional and early perceptual processing. In addition, we showed how easy it is to transform the single-neuron action potentials to a pseudocontinuous signal to perform spectral analysis typical for conventional electroencephalography. Using this approach, we documented the impact of a visual emotional task on single-neuron activity in the alpha band similar to those previously shown with LFPs (7, 28).

Affective Neurons in the STN. In this study, 17% of the STN neurons analyzed for activity in the alpha band responded to emotional stimuli. Different neurons responded to changes in emotional valence or in arousal ratings. Figs. S2 and S3 support this finding visually by showing that the valence-related neurons were mostly uncorrelated with arousal and vice versa. As for the character of changes in neuronal activity, both increases and decreases were observed in both populations of neurons, suggesting a further level of specialization within each emotional dimension. There is a large body of evidence suggesting that behavioral responses to emotional valence and arousal are mediated by different brain circuits. The independence of valence

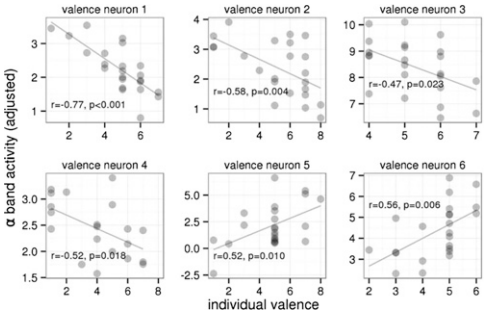


Fig. 1. Relationship of the single-neuron alpha band activity during emotional picture presentation (in the interval of 500–2,000 ms after picture onset) to individual valence ratings of the presented pictures in six neurons of the STN in patients with PD, in which the relationship was significant (as identified by linear models; *Experimental Procedures*). In none of these neurons was the activity related to individual arousal ratings (Fig. S2). The horizontal axis shows the individual ratings of the pictures' valence, varying from 1 (unpleasant) to 9 (pleasant). The vertical axis shows the alpha band neuronal activity adjusted for the past activity (Experimental Procedures). For visualization purposes, correlation coefficients and their significance are included.

NEUROSCIENCE

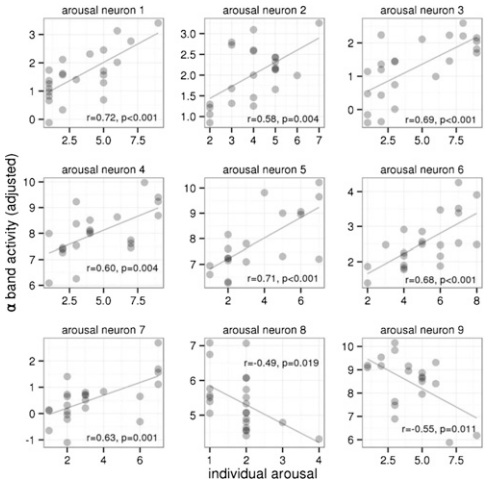


Fig. 2. Relationship of the single-neuron alpha band activity during emotional picture presentation (in the interval of 500–2,000 ms after picture onset) to individual arousal ratings of the presented pictures in nine neurons of the STN in patients with PD, in which the dependency was significant (as identified by linear models; *Experimental Procedures*). In none of these neurons was the activity related to individual valence ratings (Fig. S3). The horizontal axis shows the individual ratings of the pictures' arousal, varying from 1 (low) to 9 (high). The vertical axis shows the alpha band neuronal activity adjusted for the past activity and picture categories (*Experimental Procedures*). For visualization purposes, correlation coefficients and their significances are included.

and arousal has already been demonstrated for a variety of physiological reactions (21, 38, 39) and in affect-related cognitive processing (40). Functional imaging and animal studies also have demonstrated their functional segregation, with several brain regions associated with affective valence (eg, orbitofrontal cortex, mesolimbic dopamine system) and others associated with affective arousal (eg, amygdala, mesencephalic reticular activating system) (41–45); however, there is also evidence indicating that the two emotional dimensions are not fully independent (46), and that some subcortical regions may code the overall emotional value of a stimulus (47).

The neuronal activity in the STN that we observed in the late window (500–2,000 ms) may reflect the formation of conceptual knowledge related to emotional valence and arousal, because this is in line with the late neuronal response (625–1,500 ms) related to different valences of stimuli already described in the amygdala (18). We may speculate that the information represented in ratings of emotional valence and arousal in the late time window depends on processes involving the orbitofrontal and ventromedial prefrontal cortex, which provide significant input to the STN and play a major role in stimulus subjective valuation, representation of hedonic pleasure and value-based decision making (48–51). In addition, the previous passage of emotional information from the ventral basal ganglia involving input from the amygdala to the dorsal basal ganglia, including the dorsal portion of the STN, also may be a reason for late emotional activity (52, 53).

As expected, we did not find a statistically sufficient number of neurons responding to the emotional valence and arousal in the beta and gamma frequency bands. This corresponds with the negative results of previous LFP studies (7, 29) and further corroborates the functional specialization of different frequencies within the STN. In addition, we found no neurons responding to emotional content in the globus pallidus or substantia nigra pars reticulata, suggesting that the finding of affective neurons is specific to the STN, explained by its central position in the corticobasal ganglia circuit (54–57) and its connections to both the cortical and the subcortical components of the reward and limbic circuits (53, 58–63).

Despite the previously suggested emotional network asymmetry of both hemispheres (64, 65), the roles of the right and left STN in affective processes are unclear (66, 67). We did not detect any interhemispheric difference in proportion of the emotion-related neurons between the left and right STN; however, this result could have been affected by the limited number of recordings obtained from the right STN.

Localization of Affective Neurons Within the STN. A once widely held assumption is that the STN is divided into three functional zones within the STN: limbic, associative, and sensorimotor regions, residing in the anteromedial, middle, and dorsolateral portions, respectively, of the STN (58, 68, 69). This concept has been challenged by several recent electrophysiologic, neuroanatomic, and neuroimaging studies showing incomplete separation of the subthalamic territories (56, 63, 70, 71). Therefore, it is not surprising that the affective neurons were found in the sensorimotor regions, suggesting that motor and nonmotor regions overlap in the STN. In addition, the affective neurons were distributed differently, with the valence-related neurons located more posteriorly and the arousal-related neurons located more anteriorly, supporting the idea that emotional processing is spatially segregated within the STN (Fig. 3). In this study, however, because we were limited by the parkinsonism-improving implantation procedure, the microrecording was restricted predominantly to the lateral portion of the STN; therefore, we cannot rule out the presence or different spatial distribution of affective neurons in the medial part of the STN (Fig. 3).

Perceptual Neurons in the STN. Fourteen percent of the subthalamic neurons responded in the alpha band firing activity during the early time window (0–500 ms), suggesting their connection with perceptual processing. Neuronal short-latency activity changes related to visual perception have already been found in animal STNs (72–74), and have been confirmed in humans by distortion of visual evoked potentials due to STN DBS (75). The difference in neural activity between the fixation and picture viewing periods is not necessarily evidence of visual processing, however; it also may reflect other processes, such as an engagement of selective attention, a shift from gaze fixation to scanning eye movements, or other cognitive functions intervening between vision and action, including memory involvement, target selection,

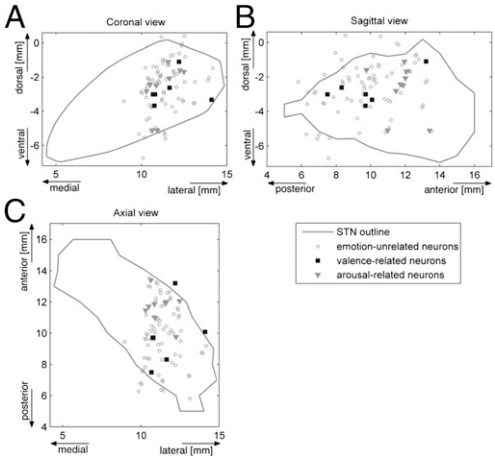


Fig. 3. Locations of STN neurons related to emotional content of the presented pictures in coronal (*A*), sagittal (*B*), and axial (*C*) views of the STN model (30) (Fig. S4). The valence-related neurons were located more posteriorly compared with the arousal-related neurons.

Table 2. Descriptive data of PD patients with STN neuronal activity explored

Patient	Age, y	DD, y	Preoperative levodopa, mg	UPDRS-III score	Neurons, n	Emotion-related neurons, n
1	64	14	1,375	31	8	1
2	61	14	1,200	37	5	0
3	46	15	1,000	40	10	1
4	63	30	1,250	50	2	1
5	53	12	700	37	7	1
6	69	9	750	47	2	0
7	49	12	1,550	65	4	1
8	53	11	1,663	45	5	1
9	64	17	1,500	31	11	2
10	42	9	740	33	16	3
11	55	19	1,980	35	8	2
12	60	14	1,060	18	7	0
13	43	9	1,100	34	5	2

Age refers to age on the day of surgery; DD, PD duration; preoperative levodopa, dose/day in mg, including levodopa-equivalent dosage of dopamine agonist (patient 4 was also treated with mianserin, and patients 6, 7, 12, and 13 were also treated with citalopram); UPDRS-III, motor score of the Unified Parkinson´s Disease Rating Scale in the off-medication condition; neurons, number of subthalamic neurons unrelated to eye movements; emotion-related neurons, number of neurons responding to emotional stimuli.

saccade choice, and content valuation (76). On the other hand, the neuronal activity in the early time window also could be affected by early emotional and motivational activity.

The STN is anatomically connected to subcortical centers that contain visually responsive neurons (i.e., superior colliculus, pulvinar, amygdala, substantia innominate, and nucleus accumbens) involved in the visual encoding of emotional stimuli (73, 77). Given that the visual, attentional, and emotional systems are intensively interconnected, some proportion of the affective neurons also might be expect to respond in the early time window; however, here only one of the neurons was activated during both the early and late time windows. Therefore, we can speculate that distinct populations of neurons are involved at different stages of processing of the visual emotional material within the dorsolateral part of the STN.

Limitations. Several factors could affect our results and reduce the inferences that can be drawn with regards to the physiology of emotional processing and the role of the STN in the limbic circuits. One such factor is that the study was conducted with PD patients, who are known to have a widespread central nervous system pathology (78) and to experience problems in emotional processing (79), and thus the number of neurons responding to emotional stimuli in the STN might be different from that in healthy subjects. Their number is rather low, but nonetheless is comparable to that reported in previous relevant single-neuron studies on emotion in humans (15, 18). Another factor that might have contributed to the relatively low number of neurons is that the study was limited to the routine trajectory of intraoperative microrecording exploration targeting the lateral sensorimotor part of the STN, which has shown less reactivity to emotional stimuli than the ventromedial part (67). Moreover, emotional pictures were selected according to normative ratings that were acquired in a healthy, younger population with a different cultural background and were partially heterogeneous with respect to visual and emotional content. Finally, some our PD patients rated the stimuli less variable along the dimensions of emotional valence and arousal, making the mathematical model less sensitive (80).

Conclusion

Early-perceptual and late-emotional single-neuron activity in the human STN corroborates the STN's participation in nonmotor circuits. The STN was previously shown to participate in different components of emotional processing, including emotion recognition and subjective feelings (4, 54). We confirm the importance of the STN as a hub within the limbic circuitry involved in both emotional valence and arousal processing as in two functionally and spatially segregated systems. This, together with our finding of several neurons involved separately in perceptual and emotional processing, support the complex role of the STN. Our results thus extend our knowledge of the STN's role in limbic circuits and contribute to a better understanding of the affective disturbances seen in PD patients treated with subthalamic stimulation.

Experimental Procedures

Subjects. Thirteen PD patients (11 men, 2 women; mean age, 55.5 ± 8.7 y; age range, 42–69 y) undergoing bilateral electrode implantation for the STN DBS due to motor fluctuations and/or disabling dyskinesias were enrolled. The subjects had mean duration of PD of 14.2 ± 5.6 y (range, 9–30 y) and a mean motor score on the Unified Parkinson's Disease Rating Scale (UPDRS-III) in the off-medication condition of 38.7 ± 11.4 (range, 18–65) (Table 2). We also included another four patients undergoing bilateral electrode implantation for the globus pallidus interna DBS due to PD (Table S1), to study neuronal activity outside of the STN. All patients met the United Kingdom Brain Bank Criteria for diagnosis of PD. The study was approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic, and all patients provided written informed consent for participation in the study.

Four days before surgery, dopamine agonists were substituted with equivalent doses of levodopa. All other anti-PD medications (e.g., amantadine, anticholinergics) had been suspended earlier during the preparation for surgery. Levodopa was withdrawn at least 12 h before the surgery. Five of the patients were receiving antidepressant therapy (1 on mianserin, 4 on citalopram), which was not discontinued. Patients with dementia and/or depression had been excluded by a routine psychiatric examination and neuropsychological testing (e.g., mini-mental state examination, Mattis dementia rating scale, Beck depression inventory).

Affective Task. Emotionally charged pictures of three categories were selected from the International Affective Picture System (IAPS) (80). The "pleasant" category included pictures with erotic themes (e.g., people, romantic couples) and adventure (e.g., exotic landscapes, animals, sports); the "unpleasant" category included pictures of victims (e.g., mutilations) and threats (e.g., human or animal attacks, aimed guns); and the "neutral" category comprised pictures of household objects, buildings, plants, and neutral faces, and scenes. Out of 144 unique pictures, six different variants of the task, each containing 24 pictures, were compiled involving 8 pictures from each category. Pleasant and unpleasant pictures were selected in a way such that they represented emotional stimuli scaled from weak to strong according to normative emotional valence and arousal. In addition, the pictures were organized pseudorandomly, so that no more than two pictures from one category followed in sequence. Each picture was presented for 2 s, preceded by a black screen with a white cross in the center for various durations (3,500–5,500 ms). Patients were instructed to fix their eyes on the cross on the black screen and to simply watch the pictures presented and remain motionless until the end of the task.

Surgery and Intraoperative Microrecording. DBS electrodes (model 3389; Medtronic) were implanted bilaterally under local anesthesia, guided by stereotactic magnetic resonance, microelectrode recordings (MERs), and macroelectrode intraoperative stimulation as described elsewhere (81, 82). Presurgical planning was done with the SurgiPlan software system (Elekta) and was based on coregistration of preoperative frameless 3.0-T MRI using T1-weighted images (MP-RAGE sequence; 160 sagittal slices, 1.0 mm thick, with x-y resolution 1 × 1 mm; TR, 2,300 ms; TE, 4.4 ms; FA, 10 degrees) and T2-weighted images (28 axial slices and 21 coronal slices, 2 mm thick, with x-y resolution 0.9 × 0.9 mm; TR, 2,430 ms; TE, 80 ms) with preoperative frame-based 1.5-T T1-weighted images (MP-RAGE sequence; 160 sagittal slices, 1.25 mm thick, with x-y resolution 1 × 1 mm; TR, 2,500 ms; TE, 3.1 ms; FA, 45 degrees) obtained immediately before the surgical procedure with the stereotactic Leksell frame attached.

The central trajectory of the exploratory microelectrode was aimed at the STN center near the anterior part of the red nucleus. Extracellular single-neuron activity was mapped by the MER using parallel insertion of five tungsten microelectrodes spaced 2 mm apart in a "Ben-gun" configuration to select sites for macroelectrode intraoperative stimulation. The five parallel microelectrodes were advanced simultaneously with a motor microdrive in 0.5-mm steps, beginning at 10 mm above the target. Depending on the length of STN-positive signals, the MERs were extended ∼2–3 mm beyond the STN. Four out of five channels of the Leadpoint recording system (Medtronic) were used for the MERs, filtered with a 500-Hz high-pass filter and a 5-kHz low-pass filter, sampled at 24 kHz, and stored for off-line processing. Single-channel electrooculography was performed for analysis of eye movement-related neuronal activity (83).

In up to six regions with an easily classifiable neuronal pattern specific for STN, the neuronal activity was recorded during the affective task presentation with a unique variant of affective pictures in each position. The number of positions depended on the time course of the surgery and on the patient's preference, clinical condition, and compliance. Patients were observed during the affective task, and if there appeared to be any distracting discomfort or sleepiness during surgery, then the experiment was abbreviated or not performed. The affective task was presented on a 17″ computer screen placed approximately 55 cm in front of the patient's eyes, with the patient laying motionless in the supine position, as is customary for this surgical procedure. The MER signals were acquired in 2-s epoch intervals recorded during both the picture presentation (PIC epoch) and the black screen presentation (FIX epoch), producing a sequence of 48 MER epochs (FIX_1, PIC_1, ..., FIX_{24}, PIC_{24}) for a total duration of 96 s.

Rating of Emotional Valence and Arousal. A subjective rating of the emotional content of the pictures was not elicited during the surgery, to avoid possible contamination of neuronal activity by voluntary movements during the rating process. Emotional valence and arousal ratings for each picture in the task were assessed before the initiation of chronic DBS at 4–5 wk after implantation, with a sufficient delay after surgery to allow cessation of any transitory microlesion effect related to penetration of the DBS electrode (84). No patient had any change in medication regimen after the surgery. Patients were assessed under similar conditions as during surgery in the off-stimulation and off-medication states (after withdrawal of dopaminergic treatment for at least 12 h).

Each picture was presented on a touch-sensitive screen for a 6-s period. The patients were instructed to look at each picture and to rate it along the dimensions of emotional valence and arousal by self-paced touching of appropriate symbols on two independent visual scales presented on-screen after picture offset. The scales were designed according to the original IAPS scales. Valence was rated on a scale of 1–9, with 9 representing the most pleasant stimulus, and arousal was also rated on a scale of 1–9, with 9 the most arousing stimulus.

Data Analysis. WaveClus (85), an unsupervised spike detection and sorting tool that has performed reasonably well on single-channel MERs (86), was used to extract the series of action potentials of single neurons from MER signals. Neurons related to eye movements were excluded from further analysis (83). For other neurons, the alpha band activity expressing the magnitude of 8–12 Hz periodic increases and decreases in the intensity of neuronal firing was computed as described below. The number of action potentials in 5-ms segments was calculated and concatenated to form a discrete signal representing the instantaneous intensity of firing. The signal was standardized to zero mean, and fast Fourier transform was carried out, applying a Hann window of length 100 with 75% overlap. The alpha band (8–12 Hz) spectral component of the signal was then extracted (Fig. S1), and alpha band activity was defined as the mean power of the alpha band spectral component, subjected to the square root transform to stabilize variance.

To detect neurons with emotion-related activity, we built linear models of the alpha band activity obtained during PIC epochs in the 500–2,000 ms interval after picture onset. To find valence-related neurons, we built a model of the alpha band activity during PIC epochs for each neuron using three covariates: (*i*) valence ratings, (*ii*) alpha band activity in the previous FIX epoch, and (*iii*) alpha band activity in the previous PIC epoch. The latter two covariates were included to adjust for past activity, given that a strong serial correlation was present in alpha band activity (*SI Experimental Procedures* and Fig. S5).

To find arousal-related neurons, we created another model of alpha band activity during PIC epochs for each neuron using five covariates: (*i*) arousal rating, (*ii*) alpha band activity in the previous FIX epoch, (*iii*) alpha band activity in the previous PIC epoch, and (*iv* and *v*) two categorical covariates adjusting for the picture category (positive vs. neutral, negative vs. neutral). We hypothesized that pictures of the same arousal but of distinct categories could provoke activities of different intensity. A neuron was considered to be related to valence (arousal) if the valence (arousal) covariate in the respective model was significant. Interhemispheric differences in the number of neurons were tested using the χ^2 test of proportions.

To detect neurons sensitive to visual stimuli, we analyzed differences in the alpha band activity between the FIX epoch and the 0–500 ms interval of the following PIC epoch using the paired *t* test.

Each neuron was finally mapped into reference STN space by assessing the position of the neuron within the patient's STN and then aligning each STN with the model (30). To assess the position of a neuron within the STN, the preoperative STN-delineating frame-based MRI used for presurgical planning was coregistered with a frameless postoperative MRI displaying the position of the permanent electrode being in a known position relative to the microelectrode used for the MERs. To align each STN with the model, 12 points anatomically delineating the STN and the anterior and posterior commissures were identified in both the model and the preoperative T2-weighted MRI, and then fitted to one another by a linear transform (*SI Experimental Procedures* and Tables S2 and S3). A permutation test was used to assess the difference in the relative location of valence-related and arousal-related neurons.

Data processing and analyses were performed with MATLAB R2007b (MathWorks) and R statistical software (87).

ACKNOWLEDGMENTS. We thank Markéta Fialová, Anna Rezková, and Martin Voleman for administrative and technical support. This work was supported by the Czech Ministry of Education [Grant LH13256 (VES13-KontaktII)], the Czech Science Foundation (Grant Project 309/09/1145), the Czech Ministry of Health (Grant IGA MZ ČR NT12282-5/2011; Research Project MŠM 0021620849), the Charles University in Prague (Research Project PRVOUK-P26/LF1/4), and the European Social Fund (within the framework of realizing the project "Support of Inter-Sectoral Mobility and Quality Enhancement of Research Teams at Czech Technical University in Prague"; CZ.1.07/2.3.00/30.0034).

1. Okun MS (2012) Deep-brain stimulation for Parkinson's disease. *N Engl J Med* 367(16):1529–1538.
2. Alexander GE, Crutcher MD, DeLong MR (1990) Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog Brain Res* 85:119–146.
3. Voon V, Kubu C, Krack P, Houeto JL, Tröster AI (2006) Deep brain stimulation: Neuropsychological and neuropsychiatric issues. *Mov Disord* 21(Suppl 14):S305–S327.
4. Castrioto A, Lhommée E, Moro E, Krack P (2014) Mood and behavioural effects of subthalamic stimulation in Parkinson's disease. *Lancet Neurol* 13(3):287–305.
5. Dujardin K, et al. (2004) Subthalamic nucleus stimulation induces deficits in decoding emotional facial expressions in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 75(2):202–208.
6. Schneider F, et al. (2003) Deep brain stimulation of the subthalamic nucleus enhances emotional processing in Parkinson disease. *Arch Gen Psychiatry* 60(3):296–302.
7. Kühn AA, et al. (2005) Activation of the subthalamic region during emotional processing in Parkinson disease. *Neurology* 65(5):707–713.
8. Schroeder U, et al. (2004) Facial expression recognition and subthalamic nucleus stimulation. *J Neurol Neurosurg Psychiatry* 75(4):648–650.
9. Baunez C, Amalric M, Robbins TW (2002) Enhanced food-related motivation after bilateral lesions of the subthalamic nucleus. *J Neurosci* 22(2):562–568.
10. Rouaud T, et al. (2010) Reducing the desire for cocaine with subthalamic nucleus deep brain stimulation. *Proc Natl Acad Sci USA* 107(3):1196–1200.
11. Serranová T, et al. (2011) Subthalamic nucleus stimulation affects incentive salience attribution in Parkinson's disease. *Mov Disord* 26(12):2260–2266.
12. Serranová T, et al. (2013) Sex, food and threat: Startling changes after subthalamic stimulation in Parkinson's disease. *Brain Stimul* 6(5):740–745.
13. Karama S, Armony J, Beauregard M (2011) Film excerpts shown to specifically elicit various affects lead to overlapping activation foci in a large set of symmetrical brain regions in males. *PLoS ONE* 6(7):e22343.
14. Frühholz S, Grandjean D (2012) Towards a fronto-temporal neural network for the decoding of angry vocal expressions. *Neuroimage* 62(3):1658–1666.
15. Kawasaki H, et al. (2005) Analysis of single-unit responses to emotional scenes in human ventromedial prefrontal cortex. *J Cogn Neurosci* 17(10):1509–1518.
16. Laxton AW, et al. (2013) Neuronal coding of implicit emotion categories in the subcallosal cortex in patients with depression. *Biol Psychiatry* 74(10):714–719.
17. Ojemann JG, Ojemann GA, Lettich E (1992) Neuronal activity related to faces and matching in human right nondominant temporal cortex. *Brain* 115(Pt 1):1–13.
18. Wang S, et al. (2014) Neurons in the human amygdala selective for perceived emotion. *Proc Natl Acad Sci USA* 111(30):E3110–E3119.
19. Russell JA (2003) Core affect and the psychological construction of emotion. *Psychol Rev* 110(1):145–172.
20. Lang PJ, Bradley MM, Cuthbert BN (1998) Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biol Psychiatry* 44(12):1248–1263.

21. Lang PJ, Greenwald MK, Bradley MM, Hamm AO (1993) Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30(3):261–273.
22. Hamann S, Canli T (2004) Individual differences in emotion processing. *Curr Opin Neurobiol* 14(2):233–238.
23. Degos B, Deniau JM, Chavez M, Maurice N (2009) Chronic but not acute dopaminergic transmission interruption promotes a progressive increase in cortical beta frequency synchronization: Relationships to vigilance state and akinesia. *Cereb Cortex* 19(7): 1616–1630.
24. Bergman H, Wichmann T, Karmon B, DeLong MR (1994) The primate subthalamic nucleus, II: Neuronal activity in the MPTP model of parkinsonism. *J Neurophysiol* 72(2):507–520.
25. Jenkinson N, Brown P (2011) New insights into the relationship between dopamine, beta oscillations and motor function. *Trends Neurosci* 34(12):611–618.
26. Huebl J, et al. (2014) Oscillatory subthalamic nucleus activity is modulated by dopamine during emotional processing in Parkinson's disease. *Cortex* 60:69–81.
27. Brown P, Williams D (2005) Basal ganglia local field potential activity: Character and functional significance in the human. *Clin Neurophysiol* 116(11):2510–2519.
28. Brücke C, et al. (2007) The subthalamic region is activated during valence-related emotional processing in patients with Parkinson's disease. *Eur J Neurosci* 26(3):767–774.
29. Huebl J, et al. (2011) Modulation of subthalamic alpha activity to emotional stimuli correlates with depressive symptoms in Parkinson's disease. *Mov Disord* 26(3):477–483.
30. Morel A (2007) *Stereotactic Atlas of the Human Thalamus and Basal Ganglia* (Informa Healthcare, London), p 160.
31. Bradley MM, Codispoti M, Lang PJ (2006) A multi-process account of startle modulation during affective perception. *Psychophysiology* 43(5):486–497.
32. Adolphs R (2002) Neural systems for recognizing emotion. *Curr Opin Neurobiol* 12(2): 169–177.
33. Guillory SA, Bujarski KA (2014) Exploring emotions using invasive methods: Review of 60 years of human intracranial electrophysiology. *Soc Cogn Affect Neurosci* 9(12): 1880–1889.
34. Schupp HT, Junghöfer M, Weike AI, Hamm AO (2004) The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology* 41(3):441–449.
35. Junghöfer M, Bradley MM, Elbert TR, Lang PJ (2001) Fleeting images: A new look at early emotion discrimination. *Psychophysiology* 38(2):175–178.
36. D'Hondt F, et al. (2010) Early brain-body impact of emotional arousal. *Front Hum Neurosci* 4:33.
37. Burgmer M, et al. (2013) Early affective processing in patients with acute posttraumatic stress disorder: Magnetoencephalographic correlates. *PLoS ONE* 8(8):e71289.
38. Cacioppo JT, Petty RE, Losch ME, Kim HS (1986) Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *J Pers Soc Psychol* 50(2):260–268.
39. Vrana SR, Spence EL, Lang PJ (1988) The startle probe response: A new measure of emotion? *J Abnorm Psychol* 97(4):487–491.
40. Kuhbandner C, Zehetleitner M (2011) Dissociable effects of valence and arousal in adaptive executive control. *PLoS ONE* 6(12):e29287.
41. Faure A, Reynolds SM, Richard JM, Berridge KC (2008) Mesolimbic dopamine in desire and dread: Enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. *J Neurosci* 28(28):7184–7192.
42. Anders S, Lotze M, Erb M, Grodd W, Birbaumer N (2004) Brain activity underlying emotional valence and arousal: A response-related fMRI study. *Hum Brain Mapp* 23(4):200–209.
43. Colibazzi T, et al. (2010) Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion* 10(3):377–389.
44. Small DM, et al. (2003) Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39(4):701–711.
45. Anderson AK, et al. (2003) Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci* 6(2):196–202.
46. Bradley MM, Codispoti M, Cuthbert BN, Lang PJ (2001) Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion* 1(3):276–298.
47. Winston JS, Gottfried JA, Kilner JM, Dolan RJ (2005) Integrated neural representations of odor intensity and affective valence in human amygdala. *J Neurosci* 25(39): 8903–8907.
48. Roy M, Shohamy D, Wager TD (2012) Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn Sci* 16(3):147–156.
49. Berridge KC, Kringelbach ML (2013) Neuroscience of affect: Brain mechanisms of pleasure and displeasure. *Curr Opin Neurobiol* 23(3):294–303.
50. Grabenhorst F, Rolls ET (2011) Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cogn Sci* 15(2):56–67.
51. Smith KS, Tindell AJ, Aldridge JW, Berridge KC (2009) Ventral pallidum roles in reward and motivation. *Behav Brain Res* 196(2):155–167.
52. Humphries MD, Prescott TJ (2010) The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Prog Neurobiol* 90(4):385–417.
53. Haber SN, Knutson B (2010) The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4–26.
54. Péron J, Frühholz S, Vérin M, Grandjean D (2013) Subthalamic nucleus: A key structure for emotional component synchronization in humans. *Neurosci Biobehav Rev* 37(3): 358–373.
55. Baunez C, Yelnik J, Mallet L (2011) Six questions on the subthalamic nucleus: Lessons from animal models and from stimulated patients. *Neuroscience* 198:193–204.
56. Haynes WI, Haber SN (2013) The organization of prefrontal-subthalamic inputs in primates provides an anatomical substrate for both functional specificity and integration: Implications for basal ganglia models and deep brain stimulation. *J Neurosci* 33(11):4804–4814.
57. Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal "hyperdirect" pathway. *Neurosci Res* 43(2):111–117.
58. Parent A, Hazrati LN (1995) Functional anatomy of the basal ganglia, II: The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Res Brain Res Rev* 20(1):128–154.
59. Groenewegen HJ, Berendse HW (1990) Connections of the subthalamic nucleus with ventral striatopallidal parts of the basal ganglia in the rat. *J Comp Neurol* 294(4): 607–622.
60. Winter C, et al. (2008) High-frequency stimulation of the subthalamic nucleus modulates neurotransmission in limbic brain regions of the rat. *Exp Brain Res* 185(3): 497–507.
61. Turner MS, Lavin A, Grace AA, Napier TC (2001) Regulation of limbic information outflow by the subthalamic nucleus: Excitatory amino acid projections to the ventral pallidum. *J Neurosci* 21(8):2820–2832.
62. Ghashghaei HT, Hilgetag CC, Barbas H (2007) Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* 34(3):905–923.
63. Lambert C, et al. (2012) Confirmation of functional zones within the human subthalamic nucleus: Patterns of connectivity and sub-parcellation using diffusion-weighted imaging. *Neuroimage* 60(1):83–94.
64. Wager TD, Phan KL, Liberzon I, Taylor SF (2003) Valence, gender, and lateralization of functional brain anatomy in emotion: A meta-analysis of findings from neuroimaging. *Neuroimage* 19(3):513–531.
65. Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF (2012) The brain basis of emotion: A meta-analytic review. *Behav Brain Sci* 35(3):121–143.
66. Frühholz S, Grandjean D (2013) Multiple subregions in superior temporal cortex are differentially sensitive to vocal expressions: A quantitative meta-analysis. *Neurosci Biobehav Rev* 37(1):24–35.
67. Eitan R, et al. (2013) Asymmetric right/left encoding of emotions in the human subthalamic nucleus. *Front Syst Neurosci* 7:69.
68. Karachi C, et al. (2005) The pallidosubthalamic projection: An anatomical substrate for nonmotor functions of the subthalamic nucleus in primates. *Mov Disord* 20(2): 172–180.
69. Joel D, Weiner I (1997) The connections of the primate subthalamic nucleus: Indirect pathways and the open-interconnected scheme of basal ganglia-thalamocortical circuitry. *Brain Res Brain Res Rev* 23(1-2):62–78.
70. Brunenberg EJ, et al. (2012) Structural and resting state functional connectivity of the subthalamic nucleus: Identification of motor STN parts and the hyperdirect pathway. *PLoS ONE* 7(6):e39061.
71. Espinosa-Parrilla JF, Baunez C, Apicella P (2013) Linking reward processing to behavioral output: Motor and motivational integration in the primate subthalamic nucleus. *Front Comput Neurosci* 7:175.
72. Matsumura M, Kojima J, Gardiner TW, Hikosaka O (1992) Visual and oculomotor functions of monkey subthalamic nucleus. *J Neurophysiol* 67(6):1615–1632.
73. Coizet V, et al. (2009) Short-latency visual input to the subthalamic nucleus is provided by the midbrain superior colliculus. *J Neurosci* 29(17):5701–5709.
74. Rolland M, Carcenac C, Overton PG, Savasta M, Coizet V (2013) Enhanced visual responses in the superior colliculus and subthalamic nucleus in an animal model of Parkinson's disease. *Neuroscience* 252:277–288.
75. Jech R, et al. (2006) Deep brain stimulation of the subthalamic nucleus affects resting EEG and visual evoked potentials in Parkinson's disease. *Clin Neurophysiol* 117(5): 1017–1028.
76. Shires J, Joshi S, Basso MA (2010) Shedding new light on the role of the basal ganglia-superior colliculus pathway in eye movements. *Curr Opin Neurobiol* 20(6):717–725.
77. Tamietto M, de Gelder B (2010) Neural bases of the non-conscious perception of emotional signals. *Nat Rev Neurosci* 11(10):697–709.
78. Braak H, et al. (2002) Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages). *J Neurol* 249(Suppl 3):):III/1–III/5.
79. Péron J, Dondaine T, Le Jeune F, Grandjean D, Vérin M (2012) Emotional processing in Parkinson's disease: A systematic review. *Mov Disord* 27(2):186–199.
80. Lang PJ, Bradley MM, Cuthbert BN (2008) *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual* (University of Florida, Gainesville, FL), Technical Report A-8.
81. Gross RE, Krack P, Rodriguez-Oroz MC, Rezai AR, Benabid AL (2006) Electrophysiological mapping for the implantation of deep brain stimulators for Parkinson's disease and tremor. *Mov Disord* 21(Suppl 14):S259–S283.
82. Pollak P, et al. (2002) Intraoperative micro- and macrostimulation of the subthalamic nucleus in Parkinson's disease. *Mov Disord* 17(Suppl 3):S155–S161.
83. Sieger T, et al. (2013) Basal ganglia neuronal activity during scanning eye movements in Parkinson's disease. *PLoS ONE* 8(11):e78581.
84. Jech R, et al. (2012) The subthalamic microlesion story in Parkinson's disease: Electrode insertion-related motor improvement with relative cortico-subcortical hypoactivation in fMRI. *PLoS ONE* 7(11):e49056.
85. Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* 16(8):1661–1687.
86. Wild J, Prekopcsak Z, Sieger T, Novak D, Jech R (2012) Performance comparison of extracellular spike sorting algorithms for single-channel recordings. *J Neurosci Methods* 203(2): 369–376.
87. R Core Team (2013) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).

Supporting Information

Sieger et al. 10.1073/pnas.1410709112

SI Experimental Procedures

Mapping of STN Neurons. To map subthalamic neurons in the reference STN space, the position of each neuron in patients' individual STN was assessed, followed by fitting each STN onto the STN model (1). To determine the position of neurons in the patient's STN, first both the frameless preoperative 3.0-T T2-weighted MRI visualizing the STN (*Experimental Procedures*) and the frameless postoperative 1.5-T T1-weighted MRI displaying the susceptibility artifact of the permanent DBS electrode were automatically coregistered with the frame-based preoperative 1.5-T T1-weighted MRI used for neurosurgical planning (*Experimental Procedures*) in each patient using SurgiPlan software (Elekta), thereby placing all preoperative and postoperative images into one stereotactic space. The accuracy of the merge was always confirmed using the SurgiPlan "movable lens," enabling the user to visually examine two overlapping images by sweeping across the screen to convert from one to the other. The frameless postoperative 1.5-T T1-weighted MRI was acquired at ~2–5 d after DBS implantation (MP-RAGE sequence, 160 axial slices, 1.65-mm thick, with x-y resolution 0.9 × 0.9 mm; TR, 2,140 ms; TE, 3.93 ms; and FA, 15 degrees).

Second, the position of the five parallel trajectories of microelectrodes used for perioperative microrecording was reconstructed based on the location of the T1-weighted MRI susceptibility artifact of the permanent DBS electrode and the known relative position of the trajectories with respect to the stimulation electrode from perioperative documentation. Finally, the position of each recorded neuron along each reconstructed trajectory was manually determined in the SurgiPlan software based on the known depth at which the neuron was identified.

To align each STN with the model, 12 points anatomically delineating the STN and 2 additional points, the anterior commissure (AC) and posterior commissure (PC), were identified in both the reference model (1) and the preoperative 3-T T2-weighted MRI, and then coregistered with the 14 corresponding points of the model using a linear mapping approach (see below). The reference model consisted of coronal slices of the STN manually digitized from the atlas (1) (Fig. S4*A*). Each individual STN was delineated manually in preoperative axial and coronal projections; the delineating points are defined in Table S2 and depicted in Fig. S4. The anteroposterior direction in both the model and the MRI was defined by the line connecting the AC and PC. All planes of view were used for accurate delineation of STN borders in Surgiplan.

The linear mapping involved aligning the centroid of the 12 corresponding points of each individual STN (together with AC and PC) with the analogous centroid in the STN model, followed by scaling in the mediolateral (S_{ML}), anteroposterior (S_{AP}), and ventrodorsal (S_{VD}) directions, and rotating around those three directions (α_{ML}, α_{AP}, and α_{VD}). The scaling and rotation parameters were found by numerical optimization, minimizing the sum of square distances between the pairs of the delineating points in the patient's STN and the model starting from $S_{ML} = S_{AP} = S_{VD} = 100\%$ and $\alpha_{ML} = \alpha_{AP} = \alpha_{VD} = 0°$. A summary of the optimized parameters, given in Table S3, shows a reasonable match. On average, a patient's STN had to be scaled by <10% and rotated by <6° around each axis.

The accuracy of the mapping procedure was validated by comparing the borders of each STN, as identified in microelectrode recordings, with MRI-based STN reconstruction using the fitted model. The mean distance between points at which microelectrodes entered (and exited) the STN and the corresponding points interpolated on the STN model was 0.69 mm (IQR, 0.70 mm) and ranged from 0.01 to 2.54 mm, showing good accuracy of the mapping that is comparable to the results of other authors (2–4).

Several reasons might account for the observed mismatches between the microrecording-defined and MRI-defined STN borders in our study. First, the position of the microelectrode was inferred indirectly from the position of the DBS macroelectrode in postoperative MRI. Thus, it is not possible to rule out, among other factors, minor brain shifts in some patients during removal of the microelectrode and DBS macroelectrode insertion. Second, the delineation of the STN, which is hypointense in the T2-weighted MRI owing to the presence of iron, is not always clear on the scans, because there are other iron-containing structures running close to the STN (4). Moreover, the distribution of iron deposition is uneven in the nucleus, and thus T2 hypointensity does not utterly correspond to the nuclei (5), and the T2-weighted MRI can be locally distorted by the presence of iron (6).

Serial Correlation in Alpha Band Activity. For each neuron, linear models were built explaining the present activity in the interval of 500–2,000 ms after the emotional picture onset (=PIC epoch) by particular type of past activity, and the significance of the explanatory past activity and the R^2 coefficient of determination expressing the percentage of the present activity explained by the past activity were computed for each model. The observed R^2 coefficients were compared with those expected assuming the present activity to be independent of past activity.

Serial correlation in alpha band activity was found, as illustrated by the increased R^2 coefficients of determination in models predicting the present activity in terms of past activity (Fig. S5). Among 90 STN neurons, the past activity in the interval of 0–2,000 ms during the blank screen presentation with a fixation cross (=FIX_i epoch), PIC_{i-1}, FIX_{i-1}, and PIC_{i-2} epochs predicted the activity in PIC_i epochs in a significant number of 12, 10, 12, and 10 neurons, respectively ($P < 0.05$, binomial test). Serial correlation was detected over periods of up to 13 s.

1. Morel A (2007) *Atlas of the Human Thalamus and Basal Ganglia* (Informa Healthcare, London), p 160.
2. Bardinet E, et al. (2009) A three-dimensional histological atlas of the human basal ganglia, II: Atlas deformation strategy and evaluation in deep brain stimulation for Parkinson disease. *J Neurosurg* 110(2):208–219.
3. Hamani C, et al. (2005) Correspondence of microelectrode mapping with magnetic resonance imaging for subthalamic nucleus procedures. *Surg Neurol* 63(3):249–253, discussion 253.
4. Schlaier JR, et al. (2013) The influence of intraoperative microelectrode recordings and clinical testing on the location of final stimulation sites in deep brain stimulation for Parkinson's disease. *Acta Neurochir (Wien)* 155(2):357–366.
5. Dormont D, et al. (2004) Is the subthalamic nucleus hypointense on T2-weighted images? A correlation study using MR imaging and stereotactic atlas data. *AJNR Am J Neuroradiol* 25(9):1516–1523.
6. Menuel C, et al. (2005) Characterization and correction of distortions in stereotactic magnetic resonance imaging for bilateral subthalamic stimulation in Parkinson disease. *J Neurosurg* 103(2):256–266.

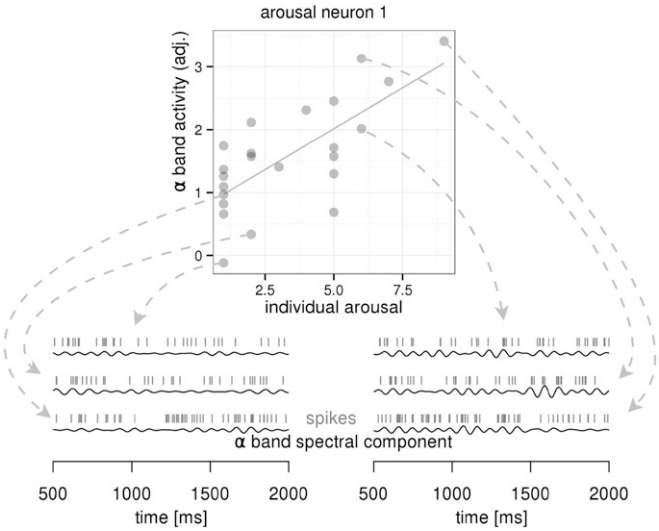


Fig. S1. Examples of neuronal firing and the alpha band spectral component of the instantaneous intensity of firing in six selected epochs recorded in the arousal-related neuron 1 (Fig. 2). The spikes are shown as vertical bars, the derived alpha band component of the firing intensity is plotted as a continuous curve beneath the spikes. Note that the amplitude and power of the alpha band spectral components (and thus the alpha band neuronal activity derived from it (*Experimental Procedures*) are much lower for the traces on the left side than for the traces on the right side.

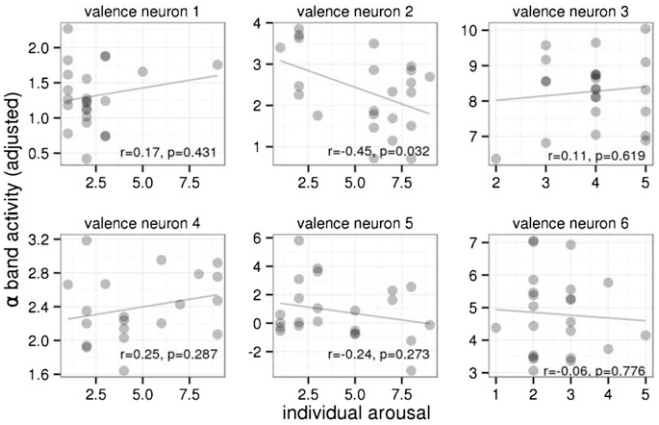


Fig. S2. Relationship of the single-neuron alpha band activity during emotional picture presentation (in the interval of 500–2,000 ms after picture onset) to the individual arousal ratings of the presented pictures in the six valence-related neurons of the STN (Fig. 1). The horizontal axis shows the individual ratings of the pictures' arousal, varying from 1 (low) to 9 (high). The vertical axis shows the alpha band neuronal activity adjusted for past activity and picture category (*Experimental Procedures*). For visualization purposes, correlation coefficients and their significance are included. Despite the fact that the correlation was significant in one neuron, none of these neurons was determined to be related to the individual arousal ratings using linear models.

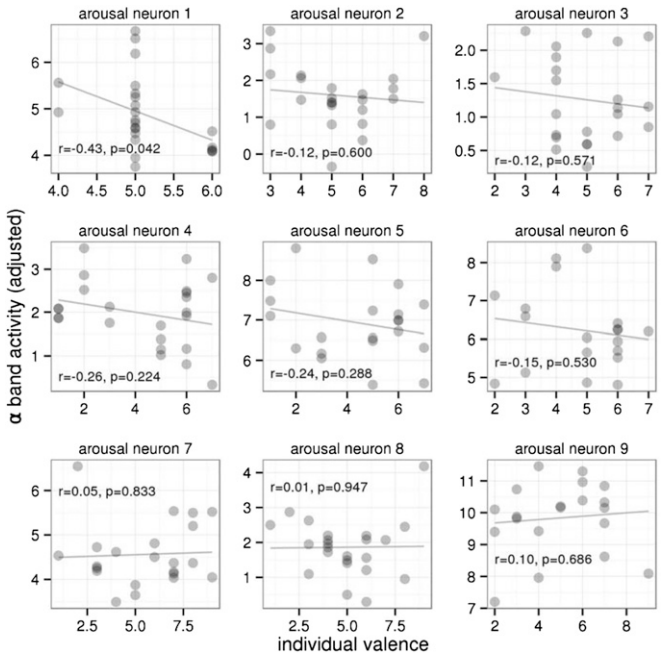


Fig. S3. Relationship of the single-neuron alpha band activity during emotional picture presentation (in the interval of 500–2,000 ms after picture onset) to the individual valence ratings of the presented pictures in the nine arousal-related neurons of the STN (Fig. 2). The horizontal axis shows the individual ratings of the pictures' valence, varying from 1 (unpleasant) to 9 (pleasant). The vertical axis shows the alpha band neuronal activity adjusted for past activity (*Experimental Procedures*). For visualization purposes, correlation coefficients and their significance are included. Despite the fact that the correlation was significant in one neuron, none of these neurons was determined to be related to the individual valence ratings using linear models.

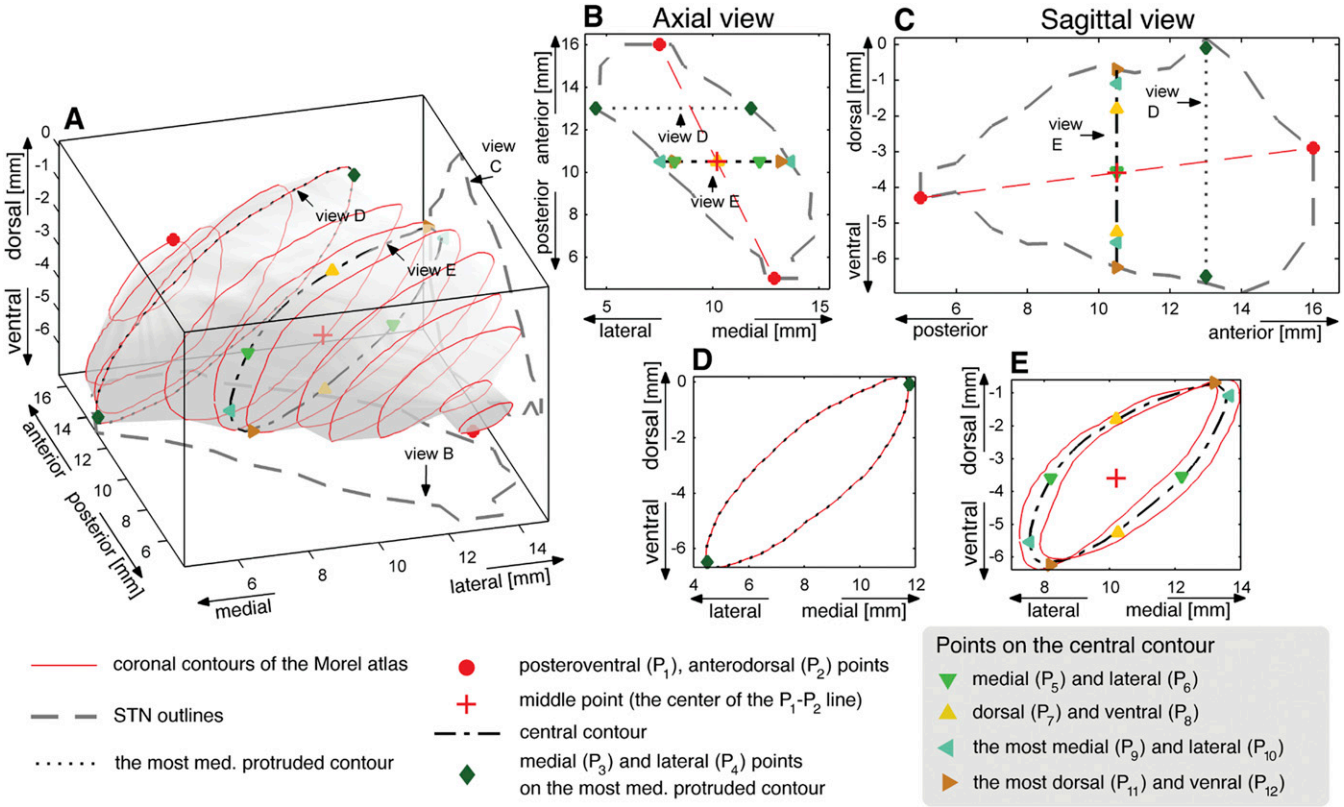


Fig. S4. Delineating points of the STN used to map each patient's STN onto the reference STN model (1). (*A*) Three-dimensional overview of the delineating points overlaid on the coronal contours of the model. (*B*) Axial view of the middle portion of the STN. (*C*) Sagittal view of the middle portion of the STN. (*D*) The most medially protruded slice in the anterior part of the STN. (*E*) The middle coronal slice of the STN model, defined by interpolating the two nearest model contours. Table S2 presents the exact definitions of the delineating points.

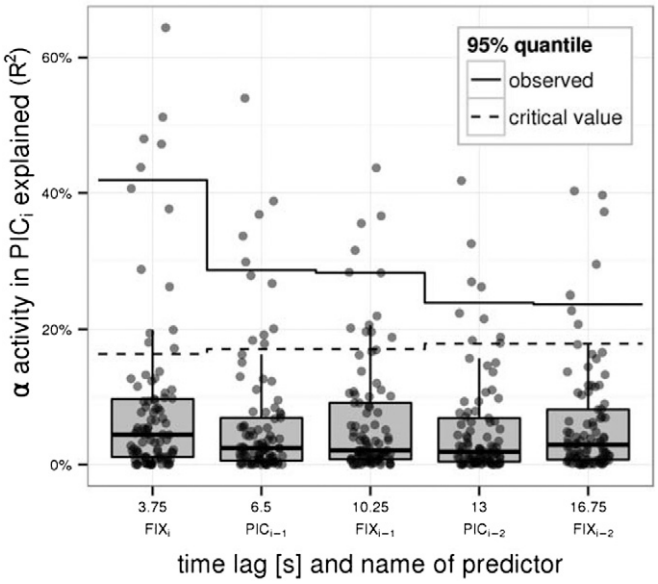


Fig. S5. Serial correlation in alpha band activity as expressed by the ability of past activity to predict present activity. The percentage of activity during the interval of 500–2,000 ms after picture onset (=PIC$_i$ epochs) on the vertical axis explained by particular types of past activity on the horizontal axis is shown for each of 90 STN neurons. The observed percentages are summarized by boxplots and 95% quantiles (solid line) and contrasted with their critical value (dashed line), computed assuming the present activity to be independent of the past activity. Covariates predicting the activity in PIC$_i$ epochs included, from left to right, the 0–2,000 ms interval of blank screen presentation with the fixating cross (=FIX$_i$ epochs) immediately preceding the PIC$_i$ epochs, the previous PIC$_{i-1}$ epochs, the FIX$_{i-1}$ epochs preceding the previous PIC$_{i-1}$ epochs, and so on. Serial correlation was present over periods of up to 13 s.

Table S1. Description of patients with PD in whom globus pallidus neuronal activity was explored

Patient	Age, y	DD, y	Levodopa, mg	UPDRS-III	H-Y	GP neurons, n
1	63	12	2,100	29	2	3
2	53	9	2,180	30	3	7
3	28	3	920	13	2	2
4	44	10	1,325	23	2	2

Age refers to age on the day of surgery; DD, PD duration; levodopa, dose/day in mg, including the levodopa-equivalent dosage of dopamine agonist (patient 1 was also treated with citalopram, and patient 4 was also treated with bupropion); UPDRS-III, motor score of the Unified Parkinson's Disease Rating Scale in the off-medication condition; H-Y, Hoehn and Yahr stage in the off-medication condition; GP neurons, neurons in the globus pallidus unrelated to eye movements.

Table S2. Definition of the points anatomically delineating the STN

STN delineating point	Position
P_1	Most ventral point in the posterior boundary
P_2	Most dorsal point in the anterior boundary
P_3	Most medial point in the in the most medially protruded slice in the anterior part of the STN
P_4	Most lateral point in the most medially protruded slice in the anterior part of the STN
P_5	Medial intersection of the axial plane crossing the middle point* and the middle contour†
P_6	Lateral intersection of the of the axial plane crossing the middle point* and the middle contour†
P_7	Dorsal intersection of the sagittal plane crossing the middle point* and the middle contour†
P_8	Ventral intersection of the sagittal plane crossing the middle point* and the middle contour†
P_9	Most medial point on the middle contour†
P_{10}	Most lateral point on the middle contour†
P_{11}	Most dorsal point on the middle contour†
P_{12}	Most ventral point on the middle contour†

These points were used to map the patients' STN onto the reference Morel STN atlas.

*The middle point is the center of the line connecting points P_1 and P_2.

†The middle contour is the coronal contour intersecting the middle point. In the Morel atlas, in which there was no contour intersecting the middle point, the middle contour was based on the interpolation of the two nearest contours.

Table S3. Summary of the parameters of the linear mapping between each individual STN and the reference STN model (1) in PD

Linear mapping parameter	Mean (SD)	Range (min; max)
S_{ML}, %	91 (8)	75; 109
S_{AP}, %	98 (5)	88; 106
S_{AP}, %	98 (8)	85; 117
α_{ML}, °	−1.3 (3.0)	−7.5; 5.4
α_{AP}, °	5.2 (5.1)	−2.3; 18.3
α_{VD}, °	1.4 (5.3)	−7.7; 8.7

Mean (SD) and minimum and maximum values for scaling in the mediolateral (s_{ML}), anteroposterior (s_{AP}), and ventrodorsal (s_{VD}) directions and rotation along these three directions (α_{ML}, α_{AP}, and α_{VD}).

Basal Ganglia Neuronal Activity during Scanning Eye Movements in Parkinson's Disease

Tomáš Sieger[1,2], Cecilia Bonnet[1], Tereza Serranová[1], Jiří Wild[2], Daniel Novák[2], Filip Růžička[1], Dušan Urgošík[1,3], Evžen Růžička[1], Bertrand Gaymard[4,5], Robert Jech[1]*

1 Department of Neurology and Center of Clinical Neuroscience, Charles University in Prague, 1st Faculty of Medicine and General University Hospital, Prague, Czech Republic, **2** Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic, **3** Department of Stereotactic and Radiation Neurosurgery, Na Homolce Hospital, Prague, Czech Republic, **4** CRICM UPMC/INSERM UMR_S975, ˆCNRSˆUMR7225, ICM, Pitie-Salpetriere Hospital, Paris, France, **5** Pierre et Marie Curie Paris-6 University, Paris, France

Abstract

The oculomotor role of the basal ganglia has been supported by extensive evidence, although their role in scanning eye movements is poorly understood. Nineteen Parkinsons disease patients, which underwent implantation of deep brain stimulation electrodes, were investigated with simultaneous intraoperative microelectrode recordings and single channel electrooculography in a scanning eye movement task by viewing a series of colored pictures selected from the International Affective Picture System. Four patients additionally underwent a visually guided saccade task. Microelectrode recordings were analyzed selectively from the subthalamic nucleus, substantia nigra pars reticulata and from the globus pallidus by the WaveClus program which allowed for detection and sorting of individual neurons. The relationship between neuronal firing rate and eye movements was studied by crosscorrelation analysis. Out of 183 neurons that were detected, 130 were found in the subthalamic nucleus, 30 in the substantia nigra and 23 in the globus pallidus. Twenty percent of the neurons in each of these structures showed eye movement-related activity. Neurons related to scanning eye movements were mostly unrelated to the visually guided saccades. We conclude that a relatively large number of basal ganglia neurons are involved in eye motion control. Surprisingly, neurons related to scanning eye movements differed from neurons activated during saccades suggesting functional specialization and segregation of both systems for eye movement control.

Citation: Sieger T, Bonnet C, Serranová T, Wild J, Novák D, et al. (2013) Basal Ganglia Neuronal Activity during Scanning Eye Movements in Parkinson's Disease. PLoS ONE 8(11): e78581. doi:10.1371/journal.pone.0078581

Editor: Bing Hou, Beijing Institute of Radiation Medicine, China

Received April 3, 2013; **Accepted** September 16, 2013; **Published** November 6, 2013

Funding: This work was supported by the Czech Science Foundation: grant project 309/09/1145, by the Czech Ministry of Health: IGA MZ ÄŒR NT12282-5/2011, IGA MZ ÄŒR NT12288-5/2011, by the Czech Ministry of Education: research project MÅ M 0021620849, MÅ M 6840770012 Transdisciplinary Research in the Area of Biomedical Engineering II, by the Charles University in Prague: research project PRVOUK-P26/LF1/4. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jech@cesnet.cz

Introduction

In everyday life we scan the environment with a series of eye movements, pointing the fovea towards objects of interest and the most salient areas of the scene. The pattern of such eye movements (EM) carried out while exploring an image, also called scanning EM, is composed of a succession of small saccades and fixations, corresponding to successive re-allocation of attention from one detail to another [1,2]. Therefore, scanning EM can be considered as internally triggered EM, as the subject moves the gaze around a complex visual image actively searching for information relevant to current motivations and goals. The visual scanpath is generated by complex parallel strategies [3] and depends on planning, visuospatial attention, spatial working memory and emotional state [4,5]. Scanning EM have mostly been the domain of psychiatric research which has focused on the behavioral aspects of the eye scanning path rather than to pathophysiological origin and scanning EM control [6].

The structures and mechanisms involved in scanning EM are still poorly understood. At the subcortical level, an involvement of the basal ganglia during scanning EM was suggested by early research using regional cerebral blood flow in healthy controls and schizophrenic patients [7]. The importance of the basal ganglia in EM control was further confirmed by animal studies [8,9,10,11,12], which discovered neurons co-activated during EM by single cell recordings in several regions of the basal ganglia and brainstem [9,11,13]. However, subcortical neuronal activity during scanning EM is still unknown and has never been studied in animals or in humans before. Several human studies supported the participation of the basal ganglia in EM control but just with results based on reflexive and voluntary saccades analyzed from oculographic recordings [14,15,16,17,18,19] or local field potentials [20]. The only evidence of human EM-related neurons was obtained from the subthalamic nucleus during saccade tasks and smooth pursuit movements in patients with Parkinson's disease [21].

In our study, we systematically searched for basal ganglia neurons participating in scanning EM. We took advantage of intraoperative microelectrode recordings of single neuronal activity routinely used to identify the basal ganglia based on specific electrophysiological pattern [22]. We have focused on the subthalamic nucleus (STN), substantia nigra pars reticulata (SNr) and globus pallidus (GP) – i.e. nuclei in which EM-related activity was previously reported [11,13] and which are easily accessible

during the implantation procedure for deep brain stimulation in Parkinsons disease (PD).

Besides EM-related neurons firing selectively when a specific position, velocity or acceleration of the eyeballs is reached, we expected to find less specialized neurons with activity depending on two or more kinematic features simultaneously. This comes from the hypothesis of functional overlap based on neuronal convergence along the striato-pallido-thalamic projection and assuming compression of information when travelling from larger to smaller nuclei [23]. Findings of STN neurons showing co-activation during various eye movement tasks are in agreement with this theory [9,21]. On the other hand, there is a segregation hypothesis which expects different neuronal populations to selectively respond to specific kinematic parameters or to fire only during a specific kind of the EM. Indeed, functional and anatomical segregation between various EM tasks has been previously observed at different levels involving the cortex, basal ganglia or cerebellum [4,8,24]. Therefore, in a subgroup of patients, we additionally studied the basal ganglia neurons during externally triggered EM using a visually guided saccade task. To further elucidate the function of neurons related to EM, we explored temporal relations of EM kinematic parameters with respect to their preceding and following activity, which may suggest their involvement in execution or control processes.

Methods

Ethics statement

The study was approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic and was conducted according to the Declaration of Helsinki.

Patients

Nineteen PD patients were enrolled consecutively from 2008 to 2011 (15 men, 4 women; mean age: 54.5, SD 9.8, range 28–69 years; mean PD duration: 13.8, SD 6.1, range 3–30 years; Hoehn-Yahr stage 2-4; mean motor score of the Unified Parkinsons Disease Rating scale – UPDRS III in OFF condition: mean 36.5, SD 13.6, range 10–65). All of them were suffering from motor fluctuations and/or disabling dyskinesias (demographic details in Table 1) and were indicated for treatment with deep brain stimulation due to motor fluctuations and dyskinesias. All of them met the UK Brain Bank Criteria for diagnosis of PD [25] and all gave their written informed consent for participation. Patients with dementia and/or depression had been excluded by a routine psychiatric examination and neuropsychological testing (Mini-mental state examination, Mattis dementia rating scale, Beck depression inventory). As a normal cognitive state was requested to fulfill the general indication criteria for implantation surgery, all patients understood the nature of the experiment. They had been informed that procedures related exclusively for study purposes could be skipped if desired. It had been emphasized that they were allowed to forego the experiment at any time before or during the surgery. Four days before surgery, dopamine agonists were substituted by equivalent doses of levodopa. Other anti-PD medication (amantadine, anticholinergics) was suspended earlier for the surgery preparation. Levodopa was withdrawn at least 12 hours before the surgery.

Surgery and intraoperative microrecording

Implantation of the deep brain stimulation system was performed separately in two steps: (i) stereotactic insertion of the permanent quadripolar electrode into the STN bilaterally and (ii) implantation of connection leads and the neurostimulator to the subclavial region. The Leksell frame and SurgiPlan software system (Elekta, Stockholm, Sweden) were employed in the stereotactic procedure. Pre-surgical planning was based on 1.5 T MRI with direct visualization of the target. The central trajectory was intentionally focused on the STN center near the anterior part of the red nucleus (15 patients) or to the posteroventrolateral portion of the GP interna (4 patients). The first surgery was performed while awake under local anesthesia. The extracellular neuronal activity was mapped by conventional microelectrode recordings (MER) using parallel insertion of five tungsten microelectrodes spaced 2 mm apart in a "Ben-gun" configuration to select sites for the macroelectrode intraoperative stimulation [26,27]. Four out of five channels of the Leadpoint recording system (Medtronic, MN) were used for the MER, filtered with 500 Hz high pass filter and 5 kHz low pass filter, sampled at 24 kHz and stored for off-line processing. As the firing pattern of the external globus pallidus could not always be distinguished from the internal globus pallidus, we classified both areas as one structure – GP. Up to six recording positions in the STN, SNr or GP were used for the EM tasks in each patient. The number of positions depended on the time course of the surgery, patients' clinical conditions and compliance. Tasks were not performed if patients demonstrated discomfort from being in the supine position or exhibited painful symptoms relating to the off-medication state as well as increased fatigue or sleepiness during surgery. Immediately after the procedure, the position of each permanent electrode was verified by two orthogonal X-ray images co-registered with a presurgical MRI plan. No dislocation larger than 1 mm was found in any patient.

Eye movement recording

Eye movements during scanning and visually guided EM tasks were recorded using electrooculography (EOG), a technique measuring the position of the eye in terms of the electric potential induced by the eye dipole. Technical constraints during surgery (limited space around the stereotactic frame and a limited number of recording channels) did not allow for more elaborate recordings than the use of one single-channel EOG. The signal was band-pass filtered in the range of 0.1–20 Hz and recorded using the Leadpoint recording system simultaneously with MER acquisition through a pair of surface electrodes attached near the outer canthus and the lower lid of the left eye. This setup enabled the orthogonal projection of the eye position on the axis connecting the two EOG electrodes. All eye movements except those which were orthogonal to the axis could be recorded with this technique.

Tasks

The EM tasks were presented on a 17"-computer screen placed approximately 55 cm in front of the eyes of patients lying in supine position.

The scanning EM task. The goal of this task was to induce self-initiated free-direction scanning EM. The task consisted of a presentation of a series of photographs selected from the International Affective Picture System (IAPS, Figure 1A) [28], depicting objects, persons, animals and landscapes. To avoid showing the same picture more than once, six unique variants of the test, each containing 24 pictures, were prepared. Each picture was presented for a period of 2 s and was preceded by a black screen for various durations (3500–5500 ms) with a white cross in the center. Patients were asked to fix their eyes on the cross on the black screen and then to simply watch the pictures presented. The MER and EOG signals were acquired in 2 s epoch intervals recorded both during the picture presentation and the black screen. The task lasted approximately for 2.5 minutes.

Table 1. Description of patients with Parkinson's disease.

patient	Age [years]	DD [years]	levodopa [mg]	UPDRS III	H-Y	DBS target	task	neurons
1	64	14	1375	31	2.0	STN	SEM	12
2	61	14	1200	37	2.5	STN	SEM	7
3	46	15	1000	40	3.0	STN	SEM	15
4	63	30	1250	50	3.0	STN	SEM	3
5	53	12	700	37	2.5	STN	SEM	14
6	69	9	750	47	3.0	STN	SEM	5
7	49	12	1550	65	4.0	STN	SEM	7
8	59	12	600	30	2.5	STN	SEM	8
9	63	14	1350	21	2.0	GPi	SEM	4
10	53	10	750	42	4.0	GPi	SEM	11
11	53	11	1663	45	2.5	STN	SEM	5
12	57	26	2000	59	4.0	STN	SEM	12
13	28	3	720	10	2.0	GPi	SEM	6
14	64	17	1500	31	2.5	STN	SEM	12
15	53	12	1000	40	4.0	STN	SEM	9
16	44	10	1130	23	2.0	GPi	SEM, VGS	2
17	42	9	740	33	3.0	STN	SEM, VGS	20
18	55	19	1980	35	2.0	STN	SEM, VGS	16
19	60	14	1060	18	2.0	STN	SEM, VGS	15

Age – age on the day of surgery; STN – subthalamic nucleus; GPi – globus pallidus interna; DD – Parkinsons disease duration; Levodopa – dose/day in mg including levodopa equivalent dosage of dopamine agonist; patient 4 was also treated with mianserin; patients 6, 7, 8, 9, 10 with citalopram and 16 with bupropion; UPDRS III – motor score of the Unified Parkinsons Disease Rating Scale in OFF medication condition; H-Y – Hoehn and Yahr stage in OFF medication condition; DBS target – nucleus chosen for bilateral deep brain stimulation; SEM – scanning eye movement task; VGS – visually guided saccade task; neurons – number of neurons identified in the basal ganglia.

doi:10.1371/journal.pone.0078581.t001

The visually guided saccade task. The goal of the task was to induce externally generated horizontal saccades (Figure 1B). Initially, a black screen with a central white cross was shown for a pseudorandom period of 2, 2.25, or 2.5 seconds. Subsequently, a peripheral target, a small white square, was presented for 1 s, 17 degrees laterally from the central fixation cross, pseudorandomly to the left (5 trials) or right (5 trials). Patients were instructed to initially fixate on the central cross and then to track the lateral target as fast as possible. The MER and EOG signals of 2 s durations were recorded during all 10 trials. The task lasted for 32.5 seconds.

Data analysis

Microelectrode recordings. WaveClus [29], an unsupervised spike detection and sorting tool, which performed reasonably well on the single channel MER [30], was used to extract the series of action potentials of individual neurons from MER signals (Figure 2). Instantaneous firing rate (IFR) of each neuron was estimated by convolving the series of action potentials with the causal kernel function $\alpha^2 * t * exp(-\alpha * t)$ defined for positive time t, where $1/\alpha$ was empirically set to 20 ms.

Each neuron was then mapped relative to the border of the STN, GPi and SNr identified by intraoperative MER. One-dimensional positions along the dorso-ventral microelectrode trajectory were determined using this technique (Figure S1).

EM recordings. EOG signals were rated manually and those contaminated with technical or major blinking artifacts, usually represented by large amplitude changes oversaturating the recording channel, were excluded from further analyses. As we presumed that neuronal activity could be related not only to the position of the eye, but also to its motion and the dynamics of the motion [21,31] we characterized EM by: i) the eye position (POS), defined by the EOG signal itself, ii) the eye velocity (VELOC), defined as the derivative of POS, and iii) the acceleration of the eye (ACCEL), defined as the derivative of VELOC. The derivative of the signal was defined in terms of the differences between successive samples in a low-pass filtered signal computed using a sliding rectangular window with the cutoff frequency of 12.5 Hz. The maximum and typical amplitude of the EM was extracted in each recording position in each task for each patient. While the maximum amplitude was defined as the extreme value in VELOC, the typical amplitude was defined as the median peak exceeding ± 1 SD of the VELOC.

To identify neurons whose activity was associated with EM, the relationships between IFR and POS, IFR and VELOC, and IFR and ACCEL were assessed. A neuron was considered connected to EM if its IFR was related to at least one of POS, VELOC, and ACCEL at the Bonferroni-corrected significance level of $p<0.05$. The relationships between IFR and the EM characteristics were analyzed using cross-correlation, which could reveal not only the link between concurrent IFR and EM, but also the link of IFR to preceding and following EM (Figure 3A-C). The maximal cross-correlation lag considered was ± 500 ms with steps of 2.5 ms. Biased estimates of correlation coefficients were computed to diminish uncertainty in estimates of correlation coefficients over longer lags. The cross-correlation coefficient between two signals was defined as the extreme correlation coefficient between the signals over all the lags considered. The lag in which the extreme cross-correlation was reached was called the *optimal EM-to-IFR cross-correlation lag*. The statistical significance of the cross-correla-

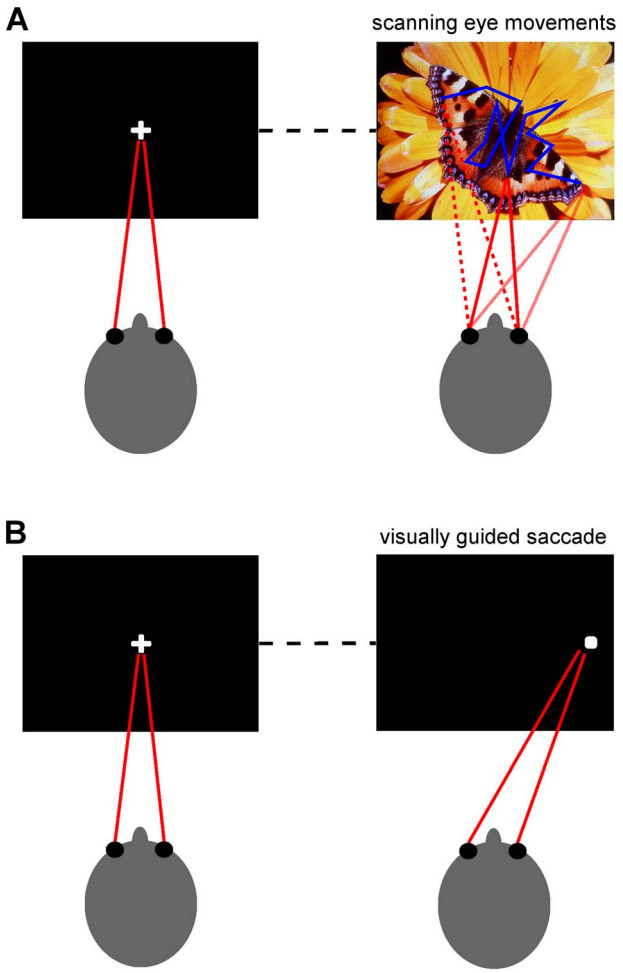


Figure 1. Eye movement (EM) tasks employed in the study. A - The scanning EM task. After the presentation of the black screen with a central cross, a photograph chosen from the International Affective Picture System was presented for 2 s. Patients were asked to initially fix their eyes on the cross (left picture) and then simply watch the photograph (right picture). In total, 24 pictures were consecutively used during the task. The blue line highlights a possible eye scanpath. **B - The visually guided saccade task** consisted of a presentation of 10 pairs of indifferent central (left picture) and lateral GO (right picture) targets positioned pseudorandomly on the left/right side of the screen. Patients were instructed to initially fixate the central cross and then track to the lateral targets as fast as possible.
doi:10.1371/journal.pone.0078581.g001

tion coefficient between two signals was assessed with Monte-Carlo simulations [32,33] using original and surrogate signals generated by randomly changing the phases of the spectral representation of the original signal.

The binomial test, Pearson's correlation coefficient test, Fisher exact test, two-sample proportion test, likelihood ratio test comparing Poisson regression models of dependence and independence in a 2-by-2-by-2 contingency table and paired t-test were used for statistical analysis. Data processing and analyses were performed in MATLAB (R2007b, The MathWorks, Natick, MA) and "R" software [34].

Results

We acquired 137 pairs of MER and EOG signals from 91 recording positions: 97 MERs were assigned to the STN, 21 to the GP and 19 to the SNr according to their firing pattern. In total, 183 neurons were detected using the spike sorting procedure, out of which 130 were located in the STN, 23 in the GP and 30 in the SNr (Table 2).

Neuronal activity related to scanning eye movements

Thirty seven (20%) out of 183 neurons identified in the basal ganglia during the scanning EM task were related to at least one of the EM kinematic parameters (POS, VELOC, ACCEL) (Table 3). Their proportion was higher than the expected false positive rate in each of the analyzed nuclei (binomial test, $p<0.001$): 26/130 neurons (20%) in the STN, 5/23 neurons (22%) in the GP and 6/30 neurons (20%) in the SNr. Locations of the EM-related neurons are depicted in the Figure S1. In the STN, the ratio of the EM-related neurons was higher in the ventral part (0 to 1 mm from the ventral STN border) compared to the rest of the nucleus (proportion test, $\chi^2 = 2.722$, df = 1, $P<0.05$).

The firing rate of the neurons relating to eye position (POS) significantly correlated with fluctuations of the EOG (Pearson's r = 0.89 (STN), 0.91 (GP), 0.86 (SNr); df = 18, $p<0.001$) (Figure 4). A relatively large number of neurons were related to more than one kinematic parameter (likelihood ratio test, D = 42.2 (STN), 19.8 (GP), 28.0 (SNr); df = 3, $p<0.001$).

As follows from cross-correlation analysis, the firing rate of the neurons was related either to concurrent, previous, or future EM (Figure 3). However, none of the nuclei predominantly contained any kind of the time-related neurons.

Neuronal activity related to visually guided saccades

There were 10/46 neurons (22%) whose activity was related to visually guided saccades in the STN, 1/2 of the neurons were in the GP and 2/5 were in the SNr. A description of neurons related to all EM kinematic parameters (POS, VELOC, ACCEL) is shown in Table 4.

Eye movements in the scanning and saccadic tasks

As both the scanning EM and visually guided saccades tasks were executed by only four patients, 19 relevant recording positions were analyzed. Neurons related to scanning EM were usually not activated in the visually guided saccades task and vice versa. Out of 46 STN neurons found in these patients, ten neurons related to scanning EM, ten neurons related to visually guided saccades and only two were activated during both tasks. These neuronal populations seemed to be independent in each of the two tasks as no evidence against the null hypothesis of independence was found (Fisher exact test, $p = 1.0$) although the test had enough power to reject the null hypothesis had the number of co-activated neurons been higher. In the GP and SNr, an insufficient number of neurons were detected for proper assessment of independence in neuronal activity between the two tasks. However, no GP or SNr neurons were co-activated during both tasks.

Descriptive analyses of the EM amplitude revealed that the maximal amplitude of the scanning EM and visually guided saccades were nearly identical. As requested by the visually guided task, patients executed large saccades, while small EM predominated in the scanning task where large EM occurred only rarely. The amplitude of the typical EM made during the visually guided saccades task was greater than during the scanning task (t = 5.7, df = 18, $p<0.001$). On average, the median saccade amplitude was 2.6 times larger in the visually guided task than in the scanning EM task.

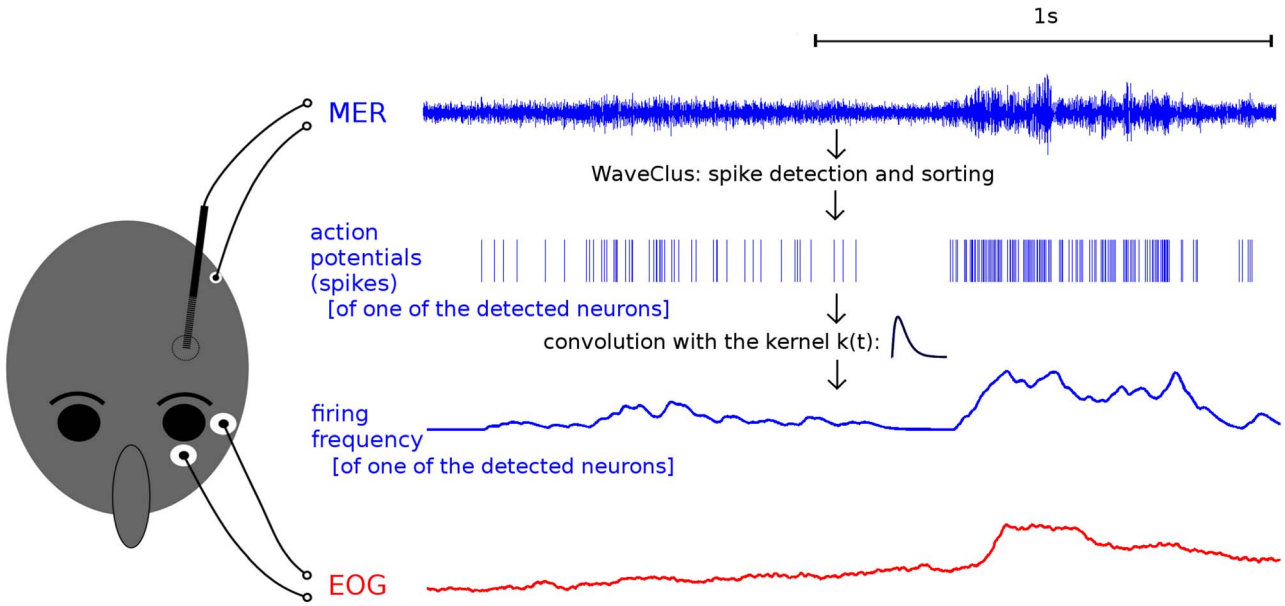


Figure 2. Microelectrode recording (MER) and electrooculography (EOG) signal acquisition and processing. Action potentials of individual neurons were identified using the WaveClus algorithm in the MER signal. The instantaneous firing rate (IFR) was then estimated by convolving a series of extracted action potentials generated by a single neuron with a causal kernel function. Finally, the IFR was correlated with the eye movement kinematic parameters derived from the EOG.
doi:10.1371/journal.pone.0078581.g002

Discussion

We showed that nuclei of the basal ganglia (namely, STN, GP and SNr) contain neurons whose firing rates correlated with eye movements during the scanning EM task. The proportion of EM-related neurons was relatively high reaching 20-22% in each of those nuclei (Table 3). Despite technical limitations due to the single-channel EOG recording we found relationships between different kinematic parameters of the EM and the firing rate in many neurons (Table 3, Figure 4). These findings point to the role of the basal ganglia in the static and dynamic representation of the EM, a role of importance for the maintenance of accuracy in goal-directed movements.

Eye movement activity in basal ganglia

Our single unit records from the STN showed that the proportion of EM-related neurons was higher in its ventral part (Figure S1). A 20% share of oculomotor neurons in the ventral part of the STN has already been noted in monkeys [10] and in humans [21]. However, those were solely neurons involved in saccadic EM. As suggested by our results, the SNr and GP are probably as equally important for control of voluntary scanning EM as the STN. We consider this as one of the major outcomes of our study because in both of these nuclei, the oculomotor activity had previously been noted during EM only in animals [11,13,35].

The role of the STN in EM has been largely explored in deep brain stimulation treated patients with Parkinson's disease. A high intensity STN neurostimulation resulted in contraversive eyeball deviations [36,37], similar to STN inactivation after locally injected GABA in animals [38] or after unilateral traumatic striato-subthalamic lesion [39]. An electrode penetration to the STN has an impact on the EM parameters as well. It causes a transitory microlesion [40] prolonging the latency of reflexive saccades [41] which are already prolonged due to Parkinson's disease [42]. Unlike microlesion, deep brain stimulation has an opposite effect on the STN as the latency of visually initiated reflexive saccades become shorter and normalized [41,43,44] while their gain is growing [45]. In addition, the STN deep brain stimulation improves some of the parameters of voluntary saccades [18,46] and suppresses interruptive saccades during fixation [47].

The significance of the STN in EM control is also well documented by other studies. The STN participates in the initiation of voluntary EM and in the inhibition of automatic EM [9,46], probably reflecting the influence of a hyperdirect pathway connecting the SMA and the motor cortex [48,49,50] including the supplementary eye field [51] with the STN, bypassing slower projections through the basal ganglia [52,53]. By the hyperdirect pathway, the motor plan can be rapidly implemented at the STN level and interfere with automatic EM [9,11]. The STN influence is then propagated by the following two main outputs [54,55]. The first is an excitatory glutamatergic projection to the SNr [56], whose activity is reduced or increased during saccades or smooth pursuit movements [12,13,35]. The SNr subsequently sends ipsi- as well as contralateral projections to the superior colliculus [57] which is an important nucleus involved in the control of automatic reflexive saccades [8]. The second glutamatergic output from the STN projects to the internal part of the GP [56] through which the oculomotor pattern can be further modified. The GP is more than just a skeletomotor structure as confirmed by several findings of EM-related neurons in its external and internal role during visually guided saccades [11] and anti-saccades in animals [58]. Moreover, bilateral pallidotomy affects the fixation [17] and reduces the velocity of self-initiated saccades [15]. On the other hand, deep brain stimulation of the GP interna modifies other parameters of automatic as well as voluntary saccades (Fawcett et al., 2005). Hence the fact that during the scanning EM task we found EM-related neurons in the STN, SNr and GP was not surprising.

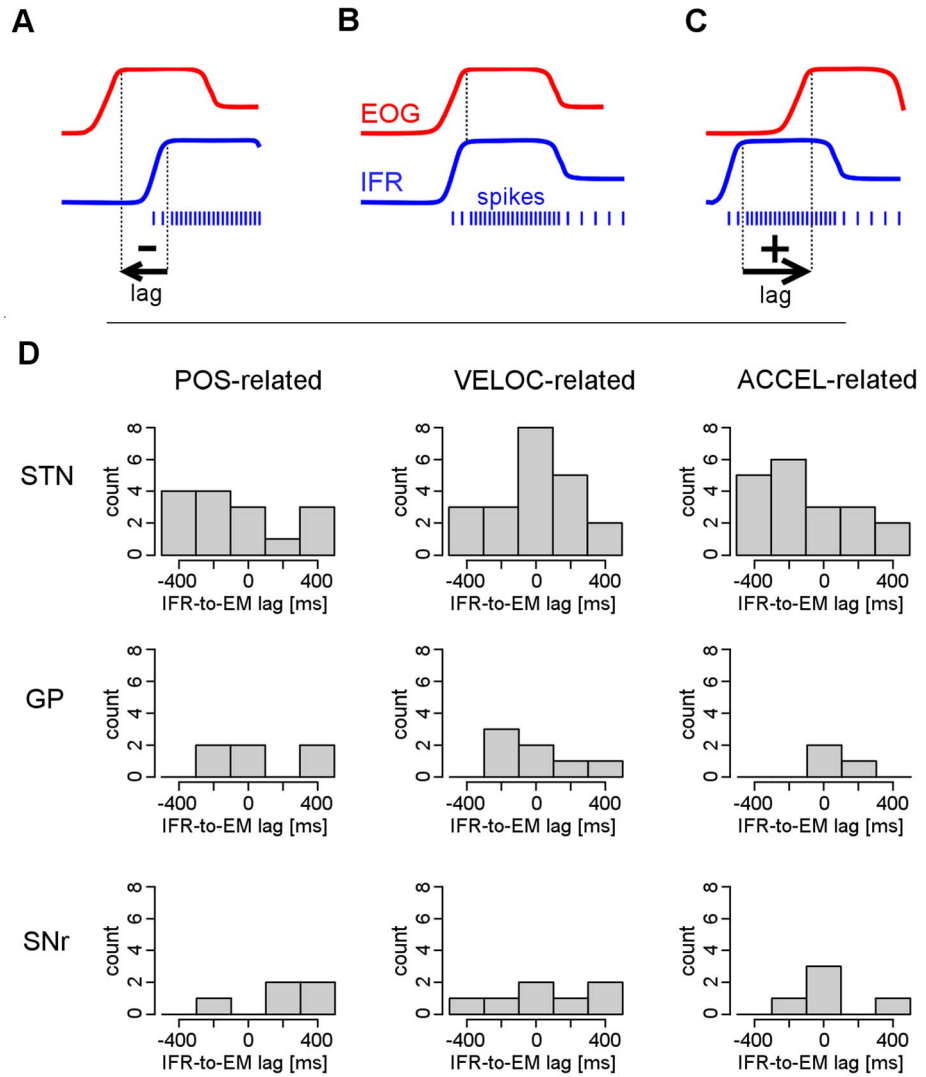


Figure 3. Time lag of neuronal activity with respect to electrooculography (EOG). A, B, C - Explanation of the cross-correlation procedure in three examples. Action potentials of three hypothetical neurons along with corresponding instantaneous firing rate (IFR) were correlated with the theoretical EOG signal. Figure A – the IFR correlates with the past EOG signal suggesting a sensory function of the neuron. Figure B – the IFR correlates with the concurrent EOG signal suggesting an executive function of the neuron. Figure C – the IFR correlates with the future EOG signal suggesting a preparatory function of the neuron. The time lag of the IFR in which the maximal (and significant) correlation with EM is reached is called the *optimal IFR-to-EM cross-correlation lag*. This lag is negative in A, zero in B and positive in C. Figure D - Frequency histograms of the optimal instantaneous firing rate (IFR) to eye movement cross-correlation lags in all eye movement-related neurons during the scanning eye movement task across the subthalamic nucleus (STN), globus pallidus (GP), and substantia nigra pars reticulata (SNr) considering kinematic parameters of the electrooculography (POS, VELOC, ACCEL, in columns). No significant differences in the locations of these distributions were found.
doi:10.1371/journal.pone.0078581.g003

Table 2. Numbers of microelectrode recordings and neurons detected.

	STN	GP	SNr	Total
MER count	97	21	19	137
neuron count (SEM task)	130	23	30	183
neuron count (SEM & VGS task)	46	2	5	53

MER count – number of microelectrode recordings obtained in each nucleus; SEM – scanning eye movement task; VGS – visually guided saccade task; neuron count – number of neurons identified in each nucleus during the SEM task (patients 1-19) and during both the SEM and VGS tasks (patients 16-19); STN – subthalamic nucleus; GP – globus pallidus; SNr – subtantia nigra pars reticulata.
doi:10.1371/journal.pone.0078581.t002

Segregation and convergence in eye movement control

Scanning EM are an important tool in the exploration of complex visual stimuli [6,59]. Their trajectory is made up of a sequence of variably large saccades and fixations with the visual field maintained for tens to hundreds of milliseconds. As a result, a certain detail is steadily projected on the fovea. This is followed by a saccade, a rapid voluntary movement, by means of which the fovea moves on to a new point of interest while information from the other parts of the retina is being concurrently assessed in search of another point of fixation. This distributed parallel processing has been recently confirmed by the sequential scanning task [60]. As expected, in four patients where both tasks were used, the median amplitude of scanning EM was smaller than that of the saccades in the visually guided task. At the same time, the

Table 3. Number of neurons related to eye movements in the scanning eye movement task.

	STN (130 neurons)	GP (23 neurons)	SNr (30 neurons)	Total (183 neurons)
EM-related neurons†	**26 (20%)*** **	**5 (22%)*** **	**6 (20%)*** **	**37 (20%)*** **
POS-related	15 (12%)**	6 (26%)***	5 (17%)*	26 (14%)***
VELOC-related	21 (16%)***	7 (30%)***	7 (23%)***	35 (19%)***
ACCEL-related	19 (15%)***	3 (13%)	5 (17%)*	27 (15%)***
POS & VELOC-related	10 (8%)	4 (17%)*	5 (17%)*	19 (10%)**
POS & ACCEL-related	7 (5%)	3 (13%)	3 (10%)	13 (7%)
VELOC & ACCEL-related	10 (8%)	3 (13%)	4 (13%)	17 (9%)*
POS & VELOC & ACCEL-related	7 (5%)	3 (13%)	3 (10%)	13 (7%)

EM-related neurons – the number of eye movement-related neurons associated with at least one kinematic parameter (†Bonferroni-corrected number of neurons for three kinematic parameters). Neurons functionally associated with one or more kinematic parameters (POS – eye position; VELOC – eye velocity; ACCEL – eye acceleration) are reported for each nucleus separately (STN – subthalamic nucleus; GP – globus pallidus; SNr – substantia nigra pars reticulata). Number of neurons significantly greater than expected 5% false positivity rate is denoted: *($p<0.05$), **($p<0.01$) ***($p<0.001$).
doi:10.1371/journal.pone.0078581.t003

amplitudes of largest EM executed in both tasks were similar. This is in agreement with previous studies, indicating that the amplitudes of scanning EM follow a heavily skewed distribution towards low values, with relatively rare movements of larger amplitude [61].

From what structures and in which way the scanning movements are controlled is still poorly understood. Since they are under voluntary control, they can be seen as a model with internally generated movements – unlike reflexive saccades which are initiated by external stimuli. Internally and externally triggered movements are generally subject to different control and executive mechanisms [62,63]. Hence, we assumed that both oculomotor systems are functionally segregated even at basal ganglia level. This hypothesis proved to be correct because in a subgroup of patients engaged in tasks which involved scanning as well as visually guided saccades, we observed that different EM-related neurons were involved in each of the tasks (Table 4). The principle of functional segregation in the control of voluntary and automatic EM had already been previously implied in connection with the interpretation of deep brain stimulation effects [16] or cerebellar lesions [24]. Animal studies have identified spatially segregated functional territories for the control of saccadic EM in the basal ganglia [64,65]. In primates, the majority of visuo-oculomotor neurons were found in the ventral part of the STN, one third of them being active during reflexive externally triggered saccades and another third being active predominantly during internally triggered (memory guided) saccades [10]. Our results go even further in terms of this specialization hierarchy. Apart from the segregation of populations of EM neurons for scanning movements and visually guided saccades, we identified a higher degree of segregation in all three nuclei neurons. In fact, some neurons responded exclusively to a specific kinematic parameter of the EM associated with an increasing or decreasing firing rate depending on whether or not the eye had reached a particular position, velocity or acceleration of movement (Figure 3).

Some of our results conform to the opposite principle arising from the convergence of cortico-striato-pallido-thalamic projection, i.e. from input nuclei which are larger, to output nuclei which are smaller [23,66] implying that initially complex information undergoes compression and simplification on its way to the output [67,68]. Indeed, a small percentage of the STN neurons showed the same neuronal activity in both types of tasks (Table 4). The convergence theory is supported by our observation of 5–8% of STN neurons, whose activity correlated with several kinematic parameters simultaneously (Table 3, 4) suggesting the presence of universal oculomotor neurons. This is in agreement with previous findings of STN neurons which become activated by switching from automatic to voluntary controlled EM [9], with the STN neurons activated from saccades and also during passive movements of the limb [21], with the SNr neurons activated during both pursuit and saccadic EM [13], or with anatomical connections documenting overlap between saccadic and pursuit oculomotor system at the brainstem level [69]. The functional convergence is further supported by the STN deep brain stimulation joint effect on the oculomotor and motor system of the neck and trunk in Parkinson's disease, marked by simultaneously improved orienting eye-head movements [45] or by improved oculomotor performance associated with body turning [70].

Time relation between EOG and neuronal activity

In our study, the eye-movement neurons in the STN, SNr or GP were not firing solely in a particular phase of the scanning EM task. In all three nuclei, these neurons became active 200–400 ms before EM, in its course and also 200–400 ms after its onset (Figure 3D). While STN neuronal activity expressed in saccade-related potentials already began 0.8–1.8 s before the saccade, suggesting the involvement of nonspecific readiness non-motor mechanisms [20], single unit neuronal STN and SNr activity culminated within 250 ms after the saccade onset [21] suggesting monitoring or sensory function. Our results are more in agreement with observations of the STN showing modified neuronal activity before, during and after the saccade [10]. This means that scanning EM-related neurons of the STN could be involved in all the preparatory, executive and monitoring phases of EM. This cannot be concluded for GP and SNr due to a relatively low amount of data.

Limitations

As there were several limitations we should interpret our results with caution. The main problem arised from the impossibility of using infra-red oculography or two-channel EOG during surgery. While their use would definitely have improved the accuracy of the kinematic parameters during EM, they would also have interfered with the established implantation procedure. The use of single-channel EOG, which failed to capture the full extent of free-

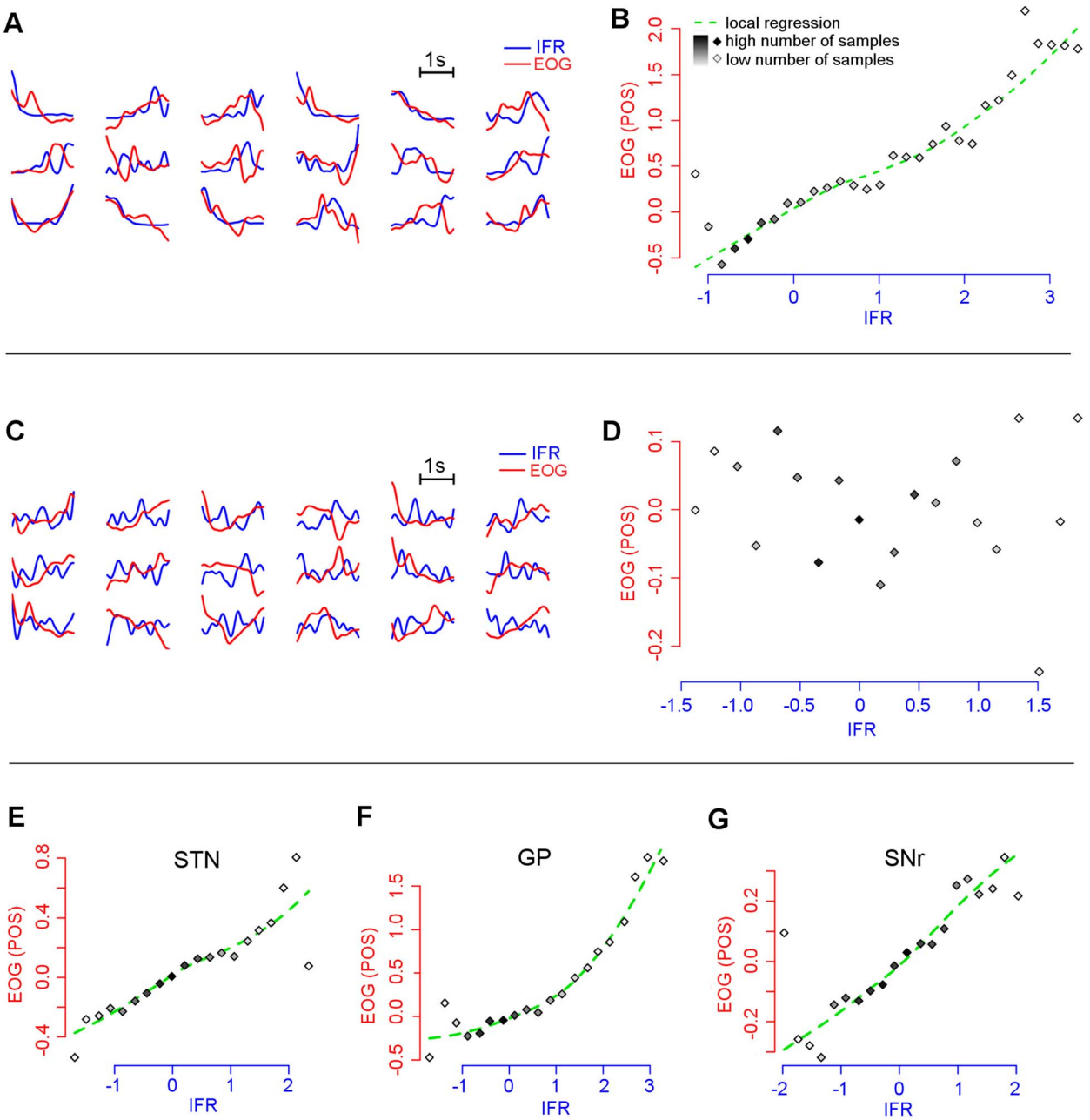


Figure 4. Neuronal activity during the scanning movement task. Example of neuron related (A, B) and unrelated (C, D) to eye movements based on correlation analysis of the instantaneous firing rate (IFR) and eye position (POS) derived from the electrooculography (EOG). All eye movement-related neuronal populations in the STN, GP and SNr are plotted in figures E, F, and G. Figures A, C show the IFR (blue) and EOG (red) pairs recorded during epochs of the task involving both the black screen and pictures presentations. Figures B, D, E, F, G show the dependency of the normalized eye position (POS) derived from the electrooculography (EOG) on the normalized, sorted and binned amplitude of the instantaneous firing rate (IFR). While the IFR from a single neuron was used on figures B and D; the IFR from all eye sensitive neurons were used on figures E, F, and G for each nucleus separately. The amplitudes of the POS signals which correlated negatively with the IFR signal were reversed. The number of signal samples in each bin is expressed by different shades of grey in the diamond glyphs.
doi:10.1371/journal.pone.0078581.g004

direction EM and yielded no more than EM projection into a one-dimensional space, is clearly a limitation which to some extent compromised the sensitivity of our study. Another limitation is connected with the assessment of neuronal activity during the oculomotor tasks based on just correlation analysis. Neuronal firing does not have to relate to EM activity alone but it may also reflect visual perception, planning, visuo-spatial attention or other cognitive processing which coincide with oculomotor activity. In addition, our results could be affected by the fact that our data was obtained from patients with Parkinson's disease in whom abnormal saccadic EM were repeatedly reported [42,44,46,71,72,73,74]. Whether any abnormalities exist in Parkinson's disease during scanning EM also is not clearly known since, with the exception of one study which showed a deficit in trans-saccadic working memory [75], no-one has systematically

Table 4. Eye movement-related neurons detected in the scanning eye movement task and/or visual guided saccade tasks.

	STN (46 neurons)			GP (2 neurons)			SNr (5 neurons)		
	SEM	VGS	Both	SEM	VGS	Both	SEM	VGS	Both
EM-related neurons†	**10**	**10**	**2**	**0**	**1**	**0**	**1**	**2**	**0**
POS-related	4	9	0	0	0	0	0	2	0
VELOC-related	9	4	1	0	0	0	1	0	0
ACCEL-related	8	11	3	0	1	0	2	0	0
POS & VELOC-related	3	4	0	0	0	0	0	0	0
POS & ACCEL-related	2	4	0	0	0	0	0	0	0
VELOC & ACCEL-related	4	2	0	0	0	0	1	0	0
POS & VELOC & ACCEL-related	2	2	0	0	0	0	0	0	0

EM-related neurons – the number of eye movement-related neurons associated with at least one kinematic parameter (†Bonferroni-corrected number of neurons for three kinematic parameters) identified from patients 16-19 which performed both the scanning eye movement task (SEM) and visual guided saccade task (VGS) in the subthalamic nucleus (STN), globus pallidus (GP) and substantia nigra pars reticulata (SNr). Neurons functionally associated with one or more kinematic parameters (POS – eye position; VELOC – eye velocity; ACCEL – eye acceleration) are reported for each nucleus separately.
doi:10.1371/journal.pone.0078581.t004

focused on scanpath or other parameters of complex exploratory EM in these patients.

Conclusions

As our results showed, the STN, SNr and GP contain neuronal populations related to scanning EM. Their representation reached about 20% in each of the three nuclei. Basal ganglia are thus not limited to previously described saccade control and perhaps play a more general role in EM circuitry. Oculomotor systems responsible for the execution and monitoring of scanning EM and visually guided saccades are mostly segregated as suggested by neurons involved exclusively in one of two EM tasks or by neurons selectively co-activated in association with a specific kinematic parameter. However, some functional overlap of the two oculomotor systems does exist, albeit confined to small groups of neurons conforming to the complementary convergence principle. Further studies combining clinical and electrophysiological approaches are needed to clarify the role of the basal ganglia in automatic and voluntary oculomotor behavior. We should emphasize, that the large representation of basal ganglia neurons showing activity during all phases of the EM is also an argument for taking them into account when designing new tasks using single unit microrecording. Many visual, ocular or motor experiments are potentially oculomotor in their nature which may compromise results if the EM-related neuronal activity was not considered.

Supporting Information

Figure S1 Positions of the eye movement-related neurons along dorso-ventral microelectrode trajectory within the basal ganglia. A – length of the subthalamic nucleus (STN), B – length of the globus pallidus (GP) and C – length of the substantia nigra pars reticulata (SNr) explored intraoperatively by the five microelectrodes in both the left and right hemispheres and projected to one-dimensional space aligned to the ventral border of the STN and GPi and to the dorsal border of the SNr. Position of each neuron along the dorso-ventral axis is shown in each subject. The proportion of eye movement-related neurons (EM) was significantly higher in the ventral part of the STN.
(TIFF)

Acknowledgments

We are grateful to all the patients who participated in the study. We thank Markéta Fialová for administrative support, Martin Voleman for technical support, Henri Bonnet and Valerie Reeves for reviewing the manuscript.

Author Contributions

Conceived and designed the experiments: RJ T. Serranová T. Sieger. Performed the experiments: RJ DU T. Sieger FR. Analyzed the data: T. Sieger RJ T. Serranová JW DN CB. Contributed reagents/materials/analysis tools: T. Sieger RJ DU ER T. Serranová JW DN CB BG. Wrote the paper: CB T. Sieger RJ. Critical revision of the manuscript: T. Sieger CB T. Serranová BG ER RJ DN.

References

1. Araujo C, Kowler E, Pavel M (2001) Eye movements during visual search: the costs of choosing the optimal path. Vision Res 41: 3613–3625.
2. Burman DD, Segraves MA (1994) Primate frontal eye field activity during natural scanning eye movements. J Neurophysiol 71: 1266–1271.
3. Wolfe JM, Vo ML, Evans KK, Greene MR (2011) Visual search in scenes involves selective and nonselective pathways. Trends Cogn Sci 15: 77–84.
4. Mort DJ, Perry RJ, Mannan SK, Hodgson TL, Anderson E, et al. (2003) Differential cortical activation during voluntary and reflexive saccades in man. Neuroimage 18: 231–246.
5. Zihl J, Hebel N (1997) Patterns of oculomotor scanning in patients with unilateral posterior parietal or frontal lobe damage. Neuropsychologia 35: 893–906.
6. Toh WL, Rossell SL, Castle DJ (2011) Current visual scanpath research: a review of investigations into the psychotic, anxiety, and mood disorders. Compr Psychiatry 52: 567–579.
7. Tsunoda M, Kurachi M, Yuasa S, Kadono Y, Matsui M, et al. (1992) Scanning eye movements in schizophrenic patients. Relationship to clinical symptoms and regional cerebral blood flow using 123I-IMP SPECT. Schizophr Res 7: 159–168.
8. Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol Rev 80: 953–978.
9. Isoda M, Hikosaka O (2008) Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. J Neurosci 28: 7209–7218.
10. Matsumura M, Kojima J, Gardiner TW, Hikosaka O (1992) Visual and oculomotor functions of monkey subthalamic nucleus. J Neurophysiol 67: 1615–1632.
11. Shin S, Sommer MA (2010) Activity of neurons in monkey globus pallidus during oculomotor behavior compared with that in substantia nigra pars reticulata. J Neurophysiol 103: 1874–1887.
12. Sato M, Hikosaka O (2002) Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. J Neurosci 22: 2363–2373.

13. Basso MA, Pokorny JJ, Liu P (2005) Activity of substantia nigra pars reticulata neurons during smooth pursuit eye movements in monkeys. Eur J Neurosci 22: 448–464.
14. Averbuch-Heller L, Stahl JS, Hlavin ML, Leigh RJ (1999) Square-wave jerks induced by pallidotomy in parkinsonian patients. Neurology 52: 185–188.
15. Blekher T, Siemers E, Abel LA, Yee RD (2000) Eye movements in Parkinson's disease: before and after pallidotomy. Invest Ophthalmol Vis Sci 41: 2177–2183.
16. Fawcett AP, Gonzalez EG, Moro E, Steinbach MJ, Lozano AM, et al. (2010) Subthalamic Nucleus Deep Brain Stimulation Improves Saccades in Parkinson's Disease. Neuromodulation 13: 17–25.
17. O'Sullivan JD, Maruff P, Tyler P, Peppard RF, McNeill P, et al. (2003) Unilateral pallidotomy for Parkinson's disease disrupts ocular fixation. J Clin Neurosci 10: 181–185.
18. Rivaud-Pechoux S, Vermersch AI, Gaymard B, Ploner CJ, Bejjani BP, et al. (2000) Improvement of memory guided saccades in parkinsonian patients by high frequency subthalamic nucleus stimulation. J Neurol Neurosurg Psychiatry 68: 381–384.
19. Temel Y, Visser-Vandewalle V, Carpenter RH (2008) Saccadic latency during electrical stimulation of the human subthalamic nucleus. Curr Biol 18: R412–414.
20. Fawcett AP, Cunic D, Hamani C, Hodaie M, Lozano AM, et al. (2007) Saccade-related potentials recorded from human subthalamic nucleus. Clin Neurophysiol 118: 155–163.
21. Fawcett AP, Dostrovsky JO, Lozano AM, Hutchison WD (2005) Eye movement-related responses of neurons in human subthalamic nucleus. Exp Brain Res 162: 357–365.
22. Hutchison WD, Allan RJ, Opitz H, Levy R, Dostrovsky JO, et al. (1998) Neurophysiological identification of the subthalamic nucleus in surgery for Parkinson's disease. Ann Neurol 44: 622–628.
23. Selemon LD, Goldman-Rakic PS (1990) Topographic intermingling of striatonigral and striatopallidal neurons in the rhesus monkey. J Comp Neurol 297: 359–376.
24. Alahyane N, Fonteille V, Urquizar C, Salemme R, Nighoghossian N, et al. (2008) Separate neural substrates in the human cerebellum for sensory-motor adaptation of reactive and of scanning voluntary saccades. Cerebellum 7: 595–601.
25. Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 55: 181–184.
26. Pollak P, Krack P, Fraix V, Mendes A, Moro E, et al. (2002) Intraoperative micro- and macrostimulation of the subthalamic nucleus in Parkinson's disease. Mov Disord 17 Suppl 3: S155–161.
27. Gross RE, Krack P, Rodriguez-Oroz MC, Rezai AR, Benabid AL (2006) Electrophysiological mapping for the implantation of deep brain stimulators for Parkinson's disease and tremor. Mov Disord 21 Suppl 14: S259–283.
28. Lang PJ, Bradley MM, Cuthbert BN (2005) International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Gainesville, FL: Florida TRA-Uo.
29. Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 16: 1661–1687.
30. Wild J, Prekopcsak Z, Sieger T, Novak D, Jech R (2012) Performance comparison of extracellular spike sorting algorithms for single-channel recordings. J Neurosci Methods 203: 369–376.
31. Kitama T, Omata T, Mizukoshi A, Ueno T, Sato Y (1999) Motor dynamics encoding in cat cerebellar flocculus middle zone during optokinetic eye movements. J Neurophysiol 82: 2235–2248.
32. Simpson DM, Infantosi AF, Rosas DA (2001) Estimation and significance testing of cross-correlation between cerebral blood flow velocity and background electro-encephalograph activity in signals with missing samples. Med Biol Eng Comput 39: 428–433.
33. Manly BEJ (1997) Randomization, bootstrap and Monte Carlo methods in biology. London: Chapman & Hall.
34. R-Core-Team (2012) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
35. Hikosaka O, Wurtz RH (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. I. Relation of visual and auditory responses to saccades. J Neurophysiol 49: 1230–1253.
36. Sauleau P, Pollak P, Krack P, Pelisson D, Vighetto A, et al. (2007) Contraversive eye deviation during stimulation of the subthalamic region. Mov Disord 22: 1810–1813.
37. Shields DC, Gorgulho A, Behnke E, Malkasian D, DeSalles AA (2007) Contralateral conjugate eye deviation during deep brain stimulation of the subthalamic nucleus. J Neurosurg 107: 37–42.
38. Baron MS, Wichmann T, Ma D, DeLong MR (2002) Effects of transient focal inactivation of the basal ganglia in parkinsonian primates. J Neurosci 22: 592–599.
39. Bartanusz V, Daniel RT, Villemure JG (2005) Conjugate eye deviation due to traumatic striatal-subthalamic lesion. J Clin Neurosci 12: 92–94.
40. Jech R, Mueller K, Urgosik D, Sieger T, Holiga S, et al. (2012) The subthalamic microlesion story in Parkinson's disease: electrode insertion-related motor improvement with relative cortico-subcortical hypoactivation in fMRI. PLoS One 7: e49056.
41. Antoniades CA, Buttery P, FitzGerald JJ, Barker RA, Carpenter RH, et al. (2012) Deep brain stimulation: eye movements reveal anomalous effects of electrode placement and stimulation. PLoS One 7: e32830.
42. Pinkhardt EH, Jurgens R, Lule D, Heimrath J, Ludolph AC, et al. (2012) Eye movement impairments in Parkinson's disease: possible role of extradopaminergic mechanisms. BMC Neurol 12: 5.
43. Temel Y, Visser-Vandewalle V, Carpenter RH (2009) Saccadometry: a novel clinical tool for quantification of the motor effects of subthalamic nucleus stimulation in Parkinson's disease. Exp Neurol 216: 481–489.
44. Antoniades CA, Carpenter RH, Temel Y (2012) Deep brain stimulation of the subthalamic nucleus in Parkinson's disease: similar improvements in saccadic and manual responses. Neuroreport 23: 179–183.
45. Sauleau P, Pollak P, Krack P, Courjon JH, Vighetto A, et al. (2008) Subthalamic stimulation improves orienting gaze movements in Parkinson's disease. Clin Neurophysiol 119: 1857–1863.
46. Yugeta A, Terao Y, Fukuda H, Hikosaka O, Yokochi F, et al. (2010) Effects of STN stimulation on the initiation and inhibition of saccade in Parkinson disease. Neurology 74: 743–748.
47. Wark HA, Garell PC, Walker AL, Basso MA (2008) A case report on fixation instability in Parkinson's disease with bilateral deep brain stimulation implants. J Neurol Neurosurg Psychiatry 79: 443–447.
48. Nambu A, Takada M, Inase M, Tokuno H (1996) Dual somatotopical representations in the primate subthalamic nucleus: evidence for ordered but reversed body-map transformations from the primary motor cortex and the supplementary motor area. J Neurosci 16: 2671–2683.
49. Nambu A, Tokuno H, Takada M (2002) Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. Neurosci Res 43: 111–117.
50. Baudrexel S, Witte T, Seifried C, von Wegner F, Beissner F, et al. (2011) Resting state fMRI reveals increased subthalamic nucleus-motor cortex connectivity in Parkinson's disease. Neuroimage 55: 1728–1738.
51. Huerta MF, Kaas JH (1990) Supplementary eye field as defined by intracortical microstimulation: connections in macaques. J Comp Neurol 293: 299–330.
52. Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9: 357–381.
53. Bauswein E, Fromm C, Preuss A (1989) Corticostriatal cells in comparison with pyramidal tract neurons: contrasting properties in the behaving monkey. Brain Res 493: 198–203.
54. Deniau JM, Hammond C, Chevalier G, Feger J (1978) Evidence for branched subthalamic nucleus projections to substantia nigra, entopeduncular nucleus and globus pallidus. Neurosci Lett 9: 117–121.
55. Kita H, Kitai ST (1987) Efferent projections of the subthalamic nucleus in the rat: light and electron microscopic analysis with the PHA-L method. J Comp Neurol 260: 435–452.
56. Windels F, Bruet N, Poupard A, Urbain N, Chouvet G, et al. (2000) Effects of high frequency stimulation of subthalamic nucleus on extracellular glutamate and GABA in substantia nigra and globus pallidus in the normal rat. Eur J Neurosci 12: 4141–4146.
57. Beckstead RM, Edwards SB, Frankfurter A (1981) A comparison of the intranigral distribution of nigrotectal neurons labeled with horseradish peroxidase in the monkey, cat, and rat. J Neurosci 1: 121–125.
58. Yoshida A, Tanaka M (2009) Enhanced modulation of neuronal activity during antisaccades in the primate globus pallidus. Cereb Cortex 19: 206–217.
59. Noton D, Stark L (1971) Eye movements and visual perception. Sci Am 224: 35–43.
60. Trukenbrod HA, Engbert R (2012) Eye movements in a sequential scanning task: evidence for distributed processing. J Vis 12.
61. von Wartburg R, Wurtz P, Pflugshaupt T, Nyffeler T, Luthi M, et al. (2007) Size matters: saccades during scene perception. Perception 36: 355–365.
62. Jahanshahi M, Jenkins IH, Brown RG, Marsden CD, Passingham RE, et al. (1995) Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. Brain 118 (Pt 4): 913–933.
63. Wiese H, Stude P, Nebel K, de Greiff A, Forsting M, et al. (2004) Movement preparation in self-initiated versus externally triggered movements: an event-related fMRI-study. Neurosci Lett 371: 220–225.
64. Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, et al. (2010) Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. Nat Rev Neurosci 11: 760–772.
65. Cui DM, Yan YJ, Lynch JC (2003) Pursuit subregion of the frontal eye field projects to the caudate nucleus in monkeys. J Neurophysiol 89: 2678–2684.
66. Haber SN, Calzavara R (2009) The cortico-basal ganglia integrative network: the role of the thalamus. Brain Res Bull 78: 69–74.
67. Filion M, Tremblay L, Matsumura M, Richard H (1994) [Dynamic focusing of informational convergence in basal ganglia]. Rev Neurol (Paris) 150: 627–633.
68. Ramanathan S, Hanley JJ, Deniau JM, Bolam JP (2002) Synaptic convergence of motor and somatosensory cortical afferents onto GABAergic interneurons in the rat striatum. J Neurosci 22: 8158–8169.
69. Yan YJ, Cui DM, Lynch JC (2001) Overlap of saccadic and pursuit eye movement systems in the brain stem reticular formation. J Neurophysiol 86: 3056–3060.

70. Lohnes CA, Earhart GM (2012) Effect of subthalamic deep brain stimulation on turning kinematics and related saccadic eye movements in Parkinson disease. Exp Neurol 236: 389–394.
71. Corin MS, Elizan TS, Bender MB (1972) Oculomotor function in patients with Parkinson's disease. J Neurol Sci 15: 251–265.
72. DeJong JD, Jones GM (1971) Akinesia, hypokinesia, and bradykinesia in the oculomotor system of patients with Parkinson's disease. Exp Neurol 32: 58–68.
73. Shibasaki H, Tsuji S, Kuroiwa Y (1979) Oculomotor abnormalities in Parkinson's disease. Arch Neurol 36: 360–364.
74. Ventre J, Zee DS, Papageorgiou H, Reich S (1992) Abnormalities of predictive saccades in hemi-Parkinson's disease. Brain 115 (Pt 4): 1147–1165.
75. Hodgson TL, Tiesman B, Owen AM, Kennard C (2002) Abnormal gaze strategies during problem solving in Parkinson's disease. Neuropsychologia 40: 411–422.

Journal of Neuroscience Methods 203 (2012) 369–376

Contents lists available at SciVerse ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Performance comparison of extracellular spike sorting algorithms for single-channel recordings

Jiri Wild[a,*], Zoltan Prekopcsak[c], Tomas Sieger[a,b], Daniel Novak[a,**], Robert Jech[b]

[a] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Karlovo nam. 13, 121 35 Praha 2, Czech Republic
[b] Department of Neurology, 1st Faculty of Medicine and General Teaching Hospital, Charles University in Prague, Katerinska 30, 128 21 Praha 2, Czech Republic
[c] Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Magyar Tudosok Korutja 2, H-1117 Budapest, Hungary

ARTICLE INFO

Article history:
Received 24 March 2011
Received in revised form 14 October 2011
Accepted 17 October 2011

Keywords:
Spike sorting
Action potential
WaveClus
OSort
MClust
KlustaKwik
Mutual information
Micro-recording
Neuronal activity

ABSTRACT

Proper classification of action potentials from extracellular recordings is essential for making an accurate study of neuronal behavior. Many spike sorting algorithms have been presented in the technical literature. However, no comparative analysis has hitherto been performed. In our study, three widely-used publicly-available spike sorting algorithms (WaveClus, KlustaKwik, OSort) were compared with regard to their parameter settings. The algorithms were evaluated using 112 artificial signals (publicly available online) with 2–9 different neurons and varying noise levels between 0.00 and 0.60. An optimization technique based on Adjusted Mutual Information was employed to find near-optimal parameter settings for a given artificial signal and algorithm. All three algorithms performed significantly better ($p < 0.01$) with optimized parameters than with the default ones. WaveClus was the most accurate spike sorting algorithm, receiving the best evaluation score for 60% of all signals. OSort operated at almost five times the speed of the other algorithms. In terms of accuracy, OSort performed significantly less well ($p < 0.01$) than WaveClus for signals with a noise level in the range 0.15–0.30. KlustaKwik achieved similar scores to WaveClus for signals with low noise level 0.00–0.15 and was worse otherwise. In conclusion, none of the three compared algorithms was optimal in general. The accuracy of the algorithms depended on proper choice of the algorithm parameters and also on specific properties of the examined signal.

1. Introduction

Classifying neuronal action potentials is a technical challenge that is a prerequisite for studying many types of brain function. Accurate detection of the activity of individual neurons can be difficult to achieve due to the large amount of background noise and the complexity in distinguishing the action potentials of one neuron from others. Even if the activity of several neurons is recorded with only a single electrode, spike sorting allows the researcher to measure the activity of the individual neurons separately. Although there are many spike sorting software packages (including commercial packages), we are not aware of any objective comparison of them that discusses adjustments to their parameters and their impact on spike sorting accuracy.

1.1. Spike sorting algorithms

Most unsupervised spike sorting algorithms employ three principal steps (Fig. 1). In the first step, spikes are detected with an automatic spike detection method. In the second step, a set of features is extracted from each spike – principal component analysis (PCA) in Adamos et al. (2008) or the wavelet transform (Quiroga et al., 2004) are usually used in this step. Finally, the spikes represented by their features are assigned to different neurons by an unsupervised learning algorithm (e.g., a clustering algorithm). We should mention that these steps are sometimes combined (Franke et al., 2009; Herbst et al., 2008), but most spike sorting algorithms handle the three steps independently.

We focus on stages 2 and 3, as there are already a number of comparative studies in the field of spike detection (Lewicki, 1998; Adamos et al., 2008; Gibson et al., 2008), and because the studied spike sorting algorithms are modular, thus allowing the researcher to choose freely which spike detection algorithm to use. The spike detection part was omitted by providing the algorithms with reference spike times.

The idea of recording multiple neurons and then grouping the action potentials by the source neuron is not new. It was first

* Corresponding author. Tel.: +420 224357666.
** Principal corresponding author.
E-mail addresses: wildjiri@fel.cvut.cz, jirka.wild@gmail.com (J. Wild), prekopcsak@tmit.bme.hu (Z. Prekopcsak), siegetom@fel.cvut.cz (T. Sieger), xnovakd1@labe.felk.cvut.cz (D. Novak), jech@cesnet.cz (R. Jech).

0165-0270/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2011.10.013

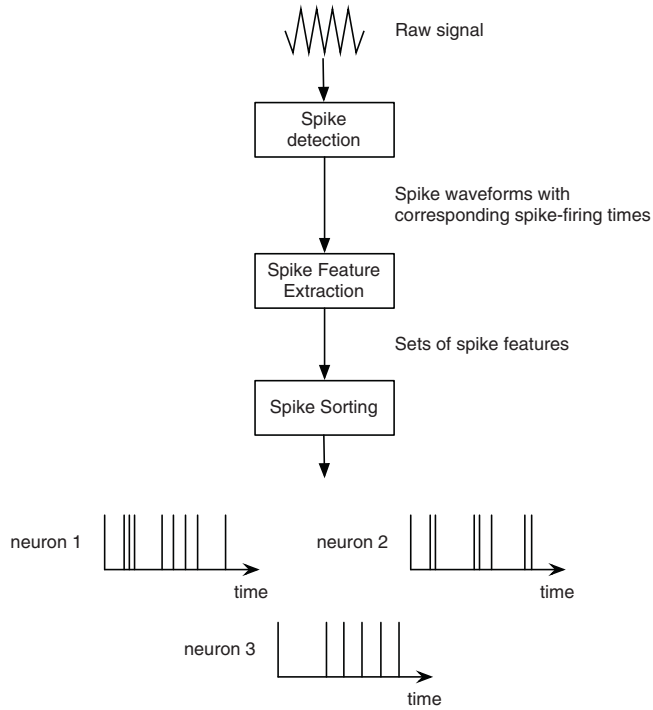


Fig. 1. Three principal stages of unsupervised spike sorting algorithms.

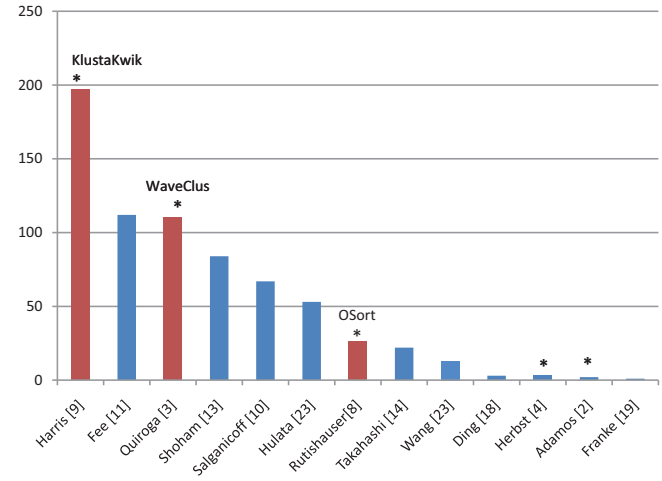


Fig. 2. Citation histogram of spike sorting algorithms as of January 2011. The approaches are tagged by an asterisk if the source code is available.

proposed in the 1960s (Gerstein and Clark, 1964), and since then numerous approaches to the problem have been developed.

Given a lower-dimensional representation of the spikes and disregarding the times at which the spikes occurred, the spike sorting problem reduces to a clustering problem. Therefore, most of the better known clustering algorithms have been applied to spike sorting: k-means clustering (Salganicoff et al., 1998), hierarchical clustering (Fee et al., 1996), superparamagnetic clustering (Quiroga et al., 2004), as well as mixtures of Gaussians (Sahani, 1999) and mixtures of t-distributions (Shoham et al., 2003). The method used in Fee et al. (1996) grouped multiple classes according to whether the interspike interval histogram of the group showed a significant number of spikes in the refractory period.

Takahashi et al. (2003a,b) combined independent component analysis (ICA) and the efficiency of an ordinary spike sorting technique (k-means clustering) to solve spike overlapping and non-stationarity problems of tetrode recordings with no limitation on the number of single neurons to be separated. Adamos et al. (2010) attempted to resolve overlapping spikes by introducing a hybrid scheme that combines the robust representation of spike waveforms to facilitate the reliable identification of contributing neurons with efficient data learning to enable the precise decomposition of coactivations.

Fee et al. (1997) described a procedure for efficiently sorting spikes in the presence of noise that is anisotropic, i.e., dominated by particular frequencies, and whose amplitude distribution may be non-Gaussian, such as occurs when spike waveforms are a function of the interspike interval. Support vector machines were used in Ding and Yuan (2008) to solve the superposition spike problem.

Herbst et al. (2008) combined the spike detection and classification steps into a single computational procedure using a Hidden Markov Model framework. Detection and classification was also merged in Franke et al. (2009), where a method of linear filters was inspected to find a new representation of the data and to optimally enhance the signal-to-noise ratio. By incorporating direct feedback, the algorithm adapted to nonstationary data. Delescluse and Pouzat (2006) used Markov chain Monte Carlo in order to estimate and make use of the firing statistics as well as the spike amplitude dynamics of the Purkinje cells. Online spike-sorting approaches suitable for HW implementation were addressed in Gibson et al. (2010) and Rutishauser (2006). Adamos et al. (2008) performed a comparative study focused on PCA using synthetic data on which correlated and white Gaussian noise processes are superimposed, and the KlustaKwik (Harris, 2000) clustering approach was used. Wang et al. (2006), proposed a robust approach employing an automatic overlap decomposition technique based on the relaxation algorithm that required simple fast Fourier transforms. Hulata et al. (2002) used a simple k-means technique for spike sorting while applying the wavelet packets decomposition framework in an extraction step.

The following approaches dealt with the quality of the spike sorting process. Schmitzer-Torbert et al. (2005) introduced two measures: L-ratio and Isolation Distance. The two measures quantified how well separated the spikes of one cluster were from other spikes. Joshua et al. (2007) described the isolation score, which measured the overlap between the noise (non-spike) and the spike clusters. The measure of Tankus et al. (2009) was based on visual features of the spike waveform and an automatic adaptive algorithm that learned the classification by a given human and could apply similar visual characteristics for classifying new data.

1.2. Comparative scheme

This paper describes a comparative analysis od the three most cited spike-sorting approaches with a publicly available source-code: WaveClus (Quiroga et al., 2004), OSort (Rutishauser, 2006) and KlustaKwik (Harris, 2000). The citation index was used as a measure for selecting the algorithm – see Fig. 2. Emphasis was put on involving one algorithm (Rutishauser, 2006) that can be used for real-time analysis.

The papers on WaveClus and KlustaKwik did not make direct comparisons with any other spike sorting method. They merely made comparisons between different versions of the same algorithm. OSort was compared with both methods, but from the perspective of online spike sorting (Rutishauser, 2006). We are convinced there is a need to evaluate them within a common framework, in order to determine which one to use for a specific task.

Lewicki (1998) presented an extensive review on spike sorting in 1998, but did not include any quantitative experiments, and dozens of new algorithms have been proposed since that review appeared. Gibson et al. (2008) compared several spike detection and feature extraction methods, but they did not include a

Table 1
Summary of the properties of each spike sorting algorithm.

	WaveClus	KlustaKwik	OSort
Features	Wavelet transform	PCA	Raw data points
Clustering method	Superparamagnetic clustering	AutoClass	Template matching
User-tunable parameters	20	10	2
Real-time use	No	No	Yes
Open source	Yes	Yes	Yes
GUI available	Yes	Yes	Yes (Mclust)
Version tested	2.0	1.6	2.1

comparison of the clustering algorithm, because the goal of the paper was only to reduce the data for hardware implementation.

In summary, very few quantitative comparisons of spike sorting methods have been made, and there are no standard criteria for evaluating them. We propose in this paper an evaluation framework aimed at providing a fair comparison of spike sorting methods in more optimal terms.

2. Materials and methods

The objective of the study is to compare the three most widely-used publicly-available spike sorting algorithms (WaveClus, KlustaKwik, OSort) with regard to their parameter settings. We observed that even a small change in the parameters of a spike sorting algorithm may have a dramatic impact on their accuracy. Therefore a comparison between spike sorting algorithms and non-optimal parameters could be biased. To overcome this weakness, we employ an optimization technique on artificial signals to find near-optimal parameter settings. Using these settings, we compared the algorithms on various types of artificial signals, focusing on single-channel recordings (similar to extracellular signals recorded using a single micro electrode).

2.1. Spike sorting algorithms

The most important properties of all three spike sorting algorithms selected in the previous section are summarized in Table 1. There follows. A more detailed description of the algorithms that have been used follows.

2.1.1. WaveClus

WaveClus is an unsupervised spike detection and sorting algorithm that combines the wavelet transform (localizing distinctive spike features) with superparamagnetic clustering (SPC), which is a method used in statistical mechanics (Quiroga et al., 2004). It enables clustering of the data without assumptions such as low variance or Gaussian distributions. In the first step, spikes are detected with an automatic amplitude threshold on the high-pass filtered data. In the second step, a small set of wavelet coefficients from each spike is chosen as the input for the clustering algorithm. Finally, SPC classifies the spikes according to the selected set of wavelet coefficients (Quiroga et al., 2004). WaveClus is one of the most widely-used spike sorting algorithms, and it has a large number of parameters for fine-tuning the method (see Table 2 for details). WaveClus version 2.0 was used for the comparison.

2.1.2. OSort

OSort is an implementation of a template-based, unsupervised online spike sorting algorithm. The estimation of the number of neurons present, as well as the assignment of each spike to a neuron, is based on a distance metric between two spikes (Rutishauser, 2006). Based on this distance, a threshold is used: (i) to decide how many neurons are present and (ii) to assign each spike uniquely to a neuron cluster, or to a noise cluster if unsortable. The threshold is calculated from the noise properties of the signal and is equal to the squared average standard deviation of the signal, calculated with a sliding window. The main advantage of OSort over its competitors is that it can be used online, thus enabling realtime spike sorting during an experiment (Rutishauser, 2006). OSort version 2.1 was used for the comparison.

2.1.3. KlustaKwik

KlustaKwik is a software for unsupervised classification of multidimensional data. It is employed in the MClust toolbox which enables both manual and automatic spike sorting on single-electrode, stereotrode and tetrode recordings. KlustaKwik fits a mixture of Gaussians with unconstrained covariance matrices and automatically chooses the number of mixture components. PCA is used to extract spike features for the clustering and a penalty term for selecting the number of clusters is implemented. The penalty is based on the ability to specify Bayesian information con-

Table 2
List of parameters impacting the spike sorting accuracy for each algorithm. The parameter names were taken directly from the original source code of each algorithm author.

WaveClus	
force_auto	Automatically force membership of spikes assigned to noise cluster using template matching
inputs	Number of wavelet coefficients to use as features for clustering
KNearNeighb	Number of data points used for the nearest neighbors interactions in the SPC
min_clus_stop	Minimum size of a cluster (cluster will be deleted if the number of spikes it contains is lower than this value)
mintemp	SPC minimum temperature – a lower temperature value groups all data into a single cluster, while higher values allow the data to split into more clusters
scales	Number of wavelet decomposition levels used
SWCycles	Number of Monte Carlo iterations used by SPC
template_type	Type of template matching method used – template matching is used for spike sorting speed up in the case of large number of spikes or for assigning spikes in the noise cluster to the existing clusters (if force_auto is set)
KlustaKwik	
noDim	Number of PCA dimensions used for clustering
MinClusters	The random initial assignment will have no less than *MinClusters* clusters. The final number may be different, since clusters can be split or deleted during the course of the algorithm
PenaltyMix	Amount of Bayesian information content (BIC) or Akaike information content (AIC) to use as a penalty for more clusters. Default of 0 sets to use all AIC. Use 1.0 to use all BIC (this generally produces fewer clusters)
OSort	
minNrSpikes	Minimum size of a cluster (cluster will be deleted if the number of spikes it contains is lower than this value)
correctionFactorThreshold	Value correcting a signal noise estimate used as a clustering threshold

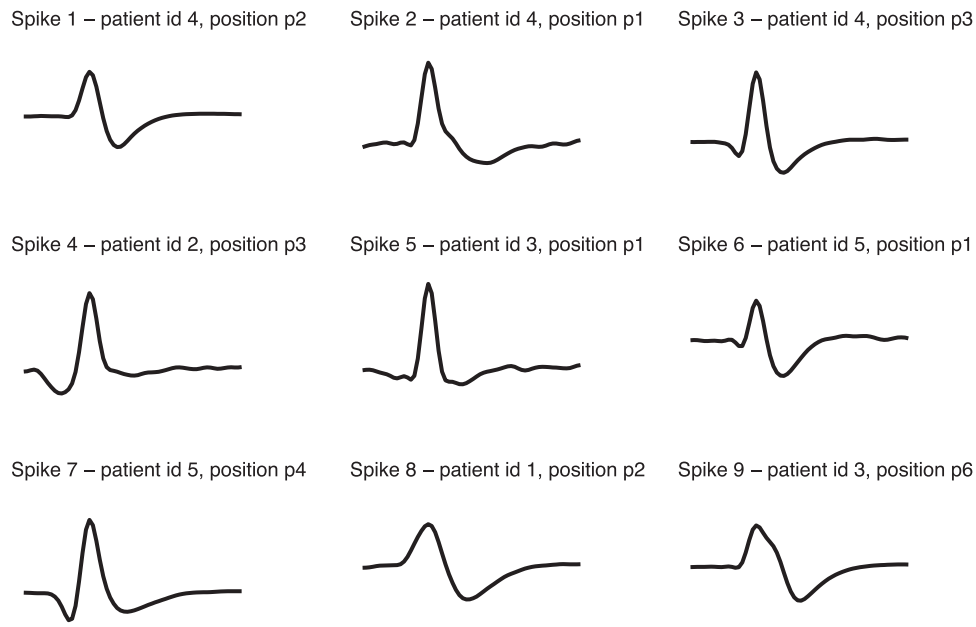


Fig. 3. Waveforms of 9 real spikes, used for artificial signal generation. Each spike represents a different neuron.

tent (Cheeseman and Stutz, 1996). KlustaKwik allows a variable number of clusters to be fitted. The program periodically checks if splitting any cluster would improve the overall score. KlustaKwik also checks to see if deleting any cluster and reallocating its points would improve the overall score. The splitting and deletion features often allow the program to escape from local minima, reducing sensitivity to the initial number of clusters, and reducing the total number of starts needed for a data set (Harris, 2000). KlustaKwik version 1.6 was used for the comparison.

2.2. Test data

For the purposes of comparison we used two sets of artificial data: previously published data (Quiroga et al. (2004), referred to as QQ after Quian Quiroga) and data generated by our own method (referred to as JW, publicly available online – http://nit.felk.cvut.cz/~wildj1/ssc). Both of these data sets were obtained simulating extracellular signals recorded using a single micro electrode.

Our artificial data was generated by superimposing real spikes at random times onto a noise background. Since several aspects of signals affect spike sorting, we used a wide range of signals of different characteristics (signal noise level, number of neurons) to maximize the objectivity and discriminability of our results.

A total of 9 real spikes (64 samples) shown in Fig. 3 were picked manually from extracellular tungsten micro-electrode recordings during a Deep Brain Stimulation operation from the sub-thalamus nuclei (STN) of 5 patients. Each spike was deduced from a different position in the STN, thus eliminating the possibility of extracting two separate spikes of the same neuron.

To generate a signal with n neurons, spikes $1 \ldots n$ were used as a template for each neuron. Each template was first scaled to 75–125% (uniform distribution) of its maximal amplitude to mimic the different spatial distance from each neuron to the electrode and was placed at random positions in the signal, while maintaining a neuronal refractory period of 3 ms. The contribution of different neurons was independent, such that spikes of different neurons might have coincided with each other in the signal, simulating the situation of several neurons firing at the same time.

The noise background for longer signals (60, 960 s) was generated in the same way as for the QQ data (Quiroga et al., 2004) using over 2000 different spikes (some of which might be from the same neuron), thus simulating the activity of many distant neurons in the brain. For shorter signals (20 s), a spike-less part of a raw signal recorded from STN was used as a noise background to approximate real signals more closely. The noise was then scaled, so that its standard deviation σ lies within the range, and was then superimposed on the previously generated signal to get the final artificial record.

Twenty-two QQ signals (60 s) and another 90 JW signals with 2–9 neurons generated using the described procedure were used to evaluate the spike sorting algorithms on a large variety of signals with different properties. The JW signals were split into three groups according to their length – 40 short JW signals (20 s), 40 long JW signals (60 s) and 10 very long JW signals (960 s). The signals with the same number of neurons differed in the standard deviation of the noise that was superimposed on the signal element. However, as it was very difficult to estimate (and compare) the standard deviation of the noise component in the case of real signals, all the JW and QQ data was labeled using a straightforward noise estimation method (see Section 2.3).

2.3. Noise level estimation

The noise level n_l was defined as the reciprocal value to the signal-to-noise ratio SNR (Smith, 1999)

$$n_l = \frac{1}{SNR} = \left(\frac{A_{\text{noise}}}{A_{\text{signal}}}\right)^2 \tag{1}$$

where A_{signal} represents the root mean square (RMS) amplitude calculated from all the spikes extracted using spike detection, and A_{noise} accounts for RMS computed from the rest of the signal. As the estimated noise level was normalized, it was easier for comparison across signals with different amplitude ranges, as opposed to standard deviation. An inevitable drawback of the method described here was that the estimation was slightly biased as, in theory, A_{signal} had to be calculated only from the useful signal (spikes), whereas in the real case the spikes themselves were corrupted by noise.

For illustration purposes, Fig. 4 depicts the same 250 ms long signal with four different noise levels (0.05, 0.15, 0.25, 0.35). On

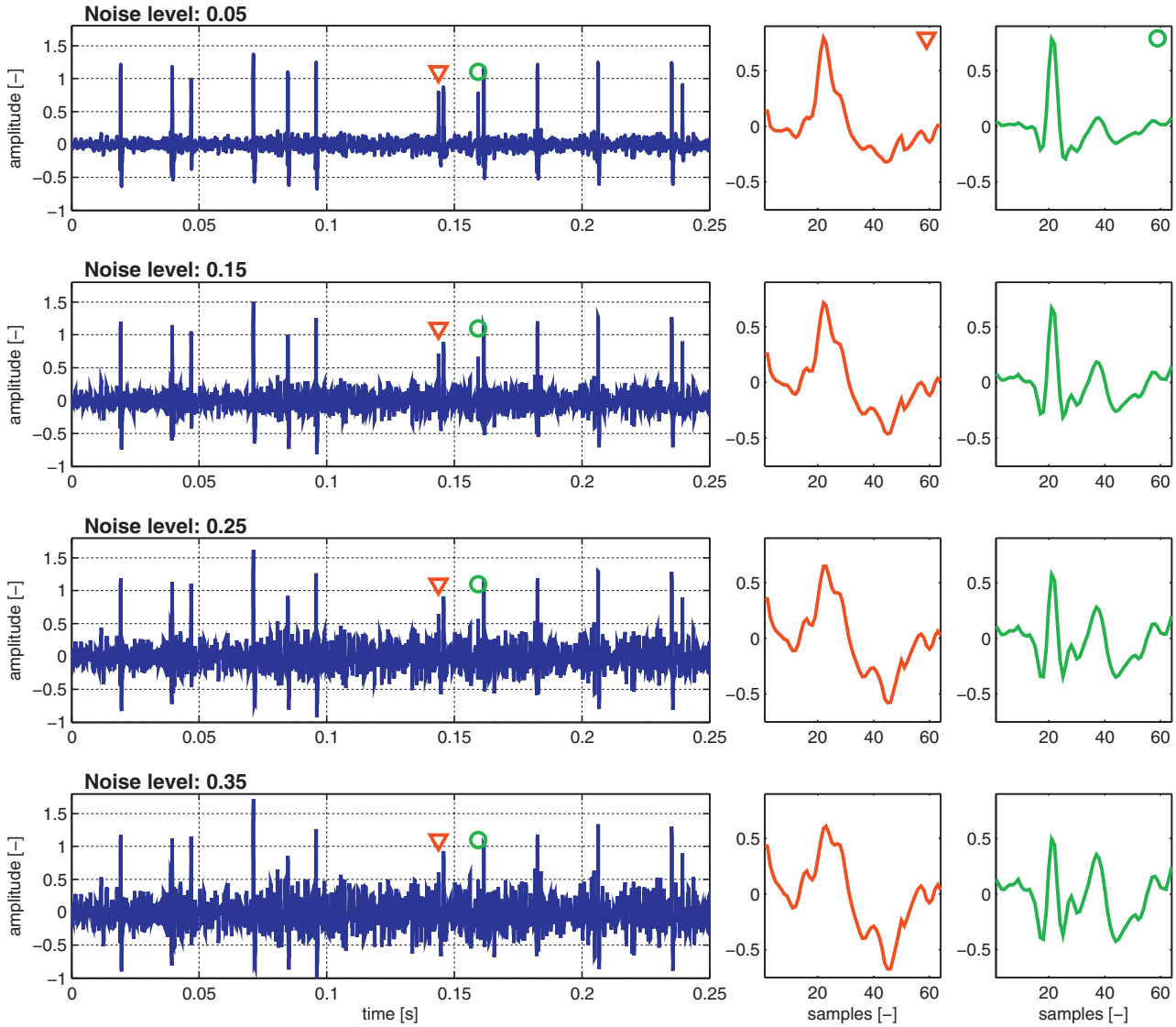


Fig. 4. Example of the same 250 ms-long signal with different noise levels ranging from 0.05 to 0.35. The spikes marked in the signal by a triangle and a circle each belonged to a different neuron and are shown in greater detail on the right side – in the case of a higher noise level at 0.25 and 0.35, a new noisy spike could be misleadingly detected.

the right side of each signal there is a detail of two spikes (marked in the signal by a triangle and a circle), each belonging to a different neuron. This is an example to illustrate of how much the noise affected the shape of the spikes.

2.4. Performance rating function

In order to asses the accuracy of different spike sorting algorithms and to provide an objective function for optimization, a performance measure was needed. As the experiments were performed using artificial data, the true clustering of the spikes was available. In machine learning research, many measures have been proposed for this type of clustering evaluation task (Warrens, 2008; Vinh et al., 2009, 2010), and some of them have already been used for spike sorting evaluation (Kretzberg et al., 2009; Gasthaus, 2008). Recently, Vinh et al. (2010) showed that Adjusted Mutual Information (AMI) had the best properties among all these clustering evaluation measures, so this measure was used for the evaluation.

AMI is an information theoretic measure which usually provides a value between 0 and 1. The value is 0 if the clustering provides information about the true clustering just by chance, and it is 1 if all information is revealed, meaning that the two clusterings are the same. Hence, AMI can be considered as the ratio of true information in a spike sorting result. Several AMI values and their corresponding clustering are shown in Fig. 5.

2.5. Spike sorting parameters

All of the spike sorting algorithms discussed in this paper have a number of parameters (OSort – 2 parameters; KlustaKwik – 9 parameters; WaveClus – 13 parameters) that can be adjusted in order to improve the spike sorting accuracy. However, it was very difficult to set these parameters correctly using manual methods.

Although all the parameters were documented, it was an almost impossible task to find out how to operate them so that the algorithm would perform better on a given signal. The parameter search was thus formulated as an automatic optimization problem: given a set of algorithm parameters $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, find a solution for $arg\,max_{\mathbf{x}} f(\mathbf{x})$, where the $f(\mathbf{x})$ objective function is the value of the performance rating function (the AMI score) for the spike sorting results obtained with parameter vector $\mathbf{x}$. As artificial signals were used in this study, the AMI could be calculated for the parameter space and the optimal solution could be identified by an exhaustive search. Gradient descent and genetic algorithms were

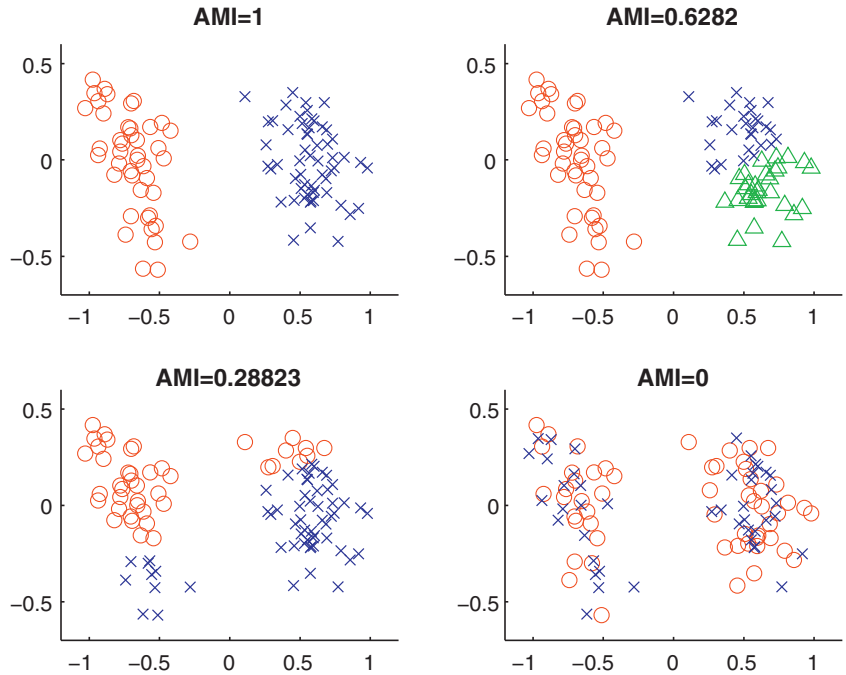


Fig. 5. Several clustering results with their corresponding AMI values. The correct clustering is presented at the top on the left with different shapes for each cluster. At the top on the right, one cluster is further split, so AMI is reduced. At the bottom on the left, the number of clusters is correct, but there is a wrong split. At the bottom on the right, there is random clustering, so the AMI value is zero.

also considered, but the objective function changed significantly with only a small change in the parameters, so only an exhaustive search guaranteed finding the global optima.

While employing the exhaustive search, only some of the algorithm parameters proved to have an impact on the spike sorting accuracy. The Table 2 summarizes the names of these parameters for all three algorithms. A complete annotated list of all parameters is available online at http://nit.felk.cvut.cz/~wildj1/ssc or at each algorithm author's website.

In order to make a fair comparison between algorithms with different numbers of parameters, all signals were split into two parts. The first part was used for optimization to find the ideal parameters, and the second part was utilized to evaluate the spike sorting accuracy with these parameters.

2.6. Technical equipment used

All calculations and statistical analyses were performed using MatLab (Mathworks, Natick, MA). The spike sorting results for the different algorithms were calculated using a Dell Precision workstation running 32-bit Linux Mint with a 2.13 GHz Intel Core 2 Duo E6400 2.13 GHz and 2 GB of DDR2 RAM.

2.7. Statistical methods

For each artificial signal the AMI scores were calculated for each spike sorting algorithm, using either optimized or default parameters. For the spike sorting evaluation, the signals and their corresponding AMI scores were grouped according to the algorithm used and the signal noise level. Each group was visualized as a simplified boxplot showing the median and the lower and upper quartiles. The range between these quartiles is referred to as the spread. Differences between group medians were assessed using the two-sided Wilcoxon signed-rank test. Bonferroni corrections for multiple comparisons were applied whenever appropriate.

For the comparison between the optimized parameters, and the default parameters the AMI scores were grouped according to the algorithm and parameters that were used (either optimized or default). For visualization, the simplified boxplots were used as described above. Significant differences between the medians of the groups were assessed in the same way as for the spike sorting evaluation, using the two-sided Wilcoxon signed rank test.

3. Results and discussion

The algorithms were compared in two main aspects. First, the spike sorting accuracy was measured with AMI (one AMI score for each signal and algorithm). The results correspond to the evaluation part of the signals, unless otherwise stated. Second, the speeds of these algorithms were compared to give some impression of the number of spikes that can be processed within a reasonable time.

3.1. Optimized parameters

As was already discussed in Section 2.5, the parameters were optimized on one part of the signal and evaluated on the other half. It was important to see whether this optimization really yielded better results than the default parameters of the algorithm. Fig. 6 shows the spike sorting accuracy results using near-optimal parameters in comparison with the results employing the default parameters. JW short, long and QQ signals with noise levels ranging from 0.0 to 0.6 were used for this comparison. Although the spread of the AMI values was quite high, mainly due to the noise level diversity of signals used, it could be clearly seen that the optimization improved all three algorithms ($p < 0.01$).

3.2. Spike sorting accuracy

Our main assumption was that increasing noise levels have a negative effect on spike sorting accuracy. We therefore present our results depending on noise levels. Fig. 7 shows the spike sorting accuracy of WaveClus, KlustaKwik and OSort on short (10s) JW signals with noise levels ranging from 0.0 to 0.6. For signals with noise level between 0.00 and 0.15, WaveClus was the most accurate

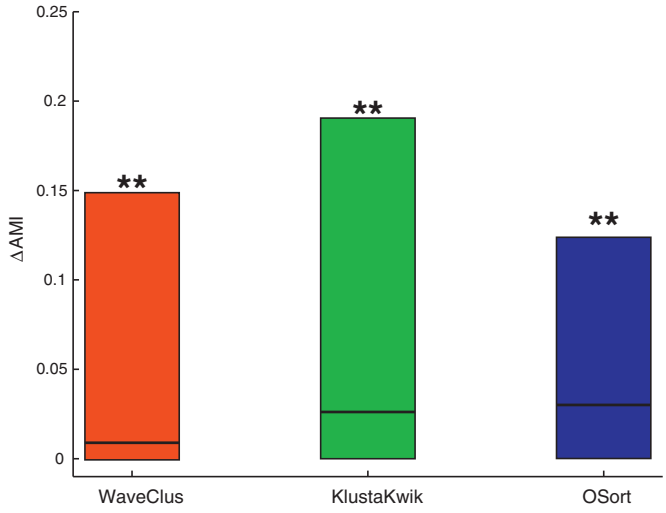


Fig. 6. Comparison of the accuracy of the algorithms when used with default and optimized parameters. The *y*-axis represents the difference between the achieved AMI score while using optimized parameters and while using default parameters). Symbol ** indicates that the medians of the marked boxplots are significantly different from zero ($p < 0.01$).

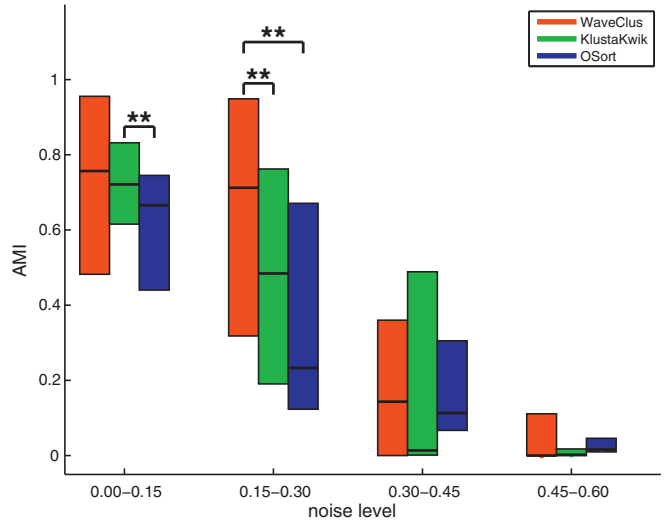


Fig. 8. Performance of spike sorting algorithms using long (30 s) artificial JW and QQ signals with noise levels binned and optimized parameters. The *y*-axis represents the AMI score of each algorithm along with its spread. Symbol ** indicates that the medians of the marked boxplots are significantly different ($p < 0.01$ corrected for 3 comparisons).

algorithm, with a median AMI of 0.7. However, because of its large spread the difference between WaveClus and KlustaKwik or OSort was not significant.

With added noise, the median AMI of all respective algorithms decreased, with both WaveClus and KlustaKwik proving to be significantly better than OSort ($p < 0.05$ and $p < 0.01$ at noise level 0.15–0.30), both having a better AMI score than OSort for 80% of 10 s signals within the respective noise level range. For signals with noise levels above 0.30 all three algorithms had very poor accuracy, indicating that signals with such a high noise level were beyond their capabilities.

Fig. 8, which depicts the same experiment as Fig. 7, only with longer JW and QQ signals (30 s), gave us somewhat similar results for WaveClus and OSort. WaveClus performed best in all cases, though it was significantly better ($p < 0.01$) than both of its competitors only for noise level 0.15–0.30 (it had a better AMI score for

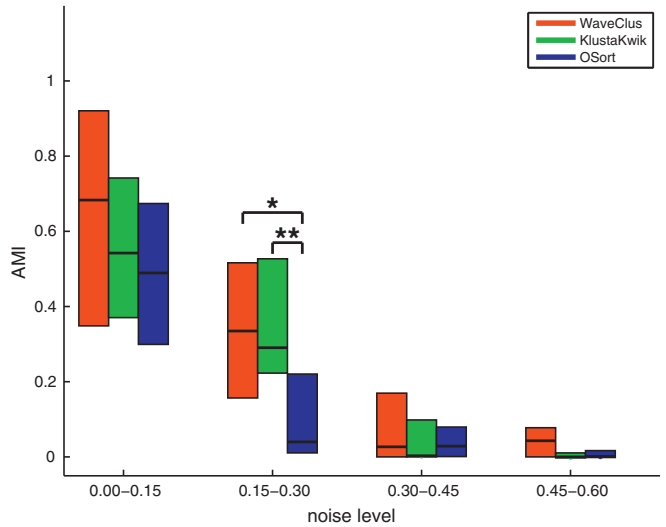


Fig. 7. Performance of spike sorting algorithms using short (10 s) artificial JW signals with noise levels binned and optimized parameters. The *y*-axis represents the AMI score of each algorithm along with its spread. Symbols * and ** indicate that the medians of the marked boxplots are significantly different ($p < 0.05$ and $p < 0.01$, corrected for 3 comparisons).

89% of the respective signals). KlustaKwik was significantly better than OSort for noise level 0.00–0.15, though with higher noise levels KlustaKwik had a larger spread than OSort. Again, noise level above 0.30 was too high for the algorithms to give reasonable results.

Some spikes were visually investigated in order to explain the effect of the noise levels. Judging from Fig. 4, the spike shape (in this case) remained almost unchanged for noise levels 0.05 and 0.15 but at 0.25 and 0.35 the spike shape did not seem like the shape at 0.05. This had a direct negative effect on the spike sorting accuracy of OSort, as shown in Figs. 7 and 8, in comparison with WaveClus because OSort used raw spike shapes (without any filtering) and a simple distance measure for sorting.

3.3. Spike sorting time complexity

In a real world scenario, the speed of an algorithm may be of considerable importance. For example, if a certain algorithm can be run online, it will help researchers to gather sorted spiking data from micro electrodes in real time. Of these three algorithms, only OSort is online, which means that it processes spikes one-by-one as they come. For the other two algorithms, the whole spike sorting process needs to be re-run with all previous data to cluster the new spikes so they are more targeted for offline analysis when new spikes are not coming in. Even for large-scale offline analysis, it would be good to know the computational demand of the algorithms.

Ten very long signals (960 s) with noise level 0.15 were used for evaluating the time complexity of each algorithm. The 960 s signals were cut into shorter signals, with the number of spikes varying from 100 to 19,460. The parameters for each spike sorting algorithm were optimized on the first part (1400 spikes) of each 960 s signal, and remained unchanged for all the other parts originating from this signal. As only the speed of the algorithms was measured and not their actual accuracy, parameter optimization of each individual signal part was unnecessary.

The results of the speed test are shown in Fig. 9. OSort was the fastest algorithm, with an average speed of 1100 spikes/s, whereas the average speed for KlustaKwik and WaveClus was 200 and 100 spikes/s respectively.

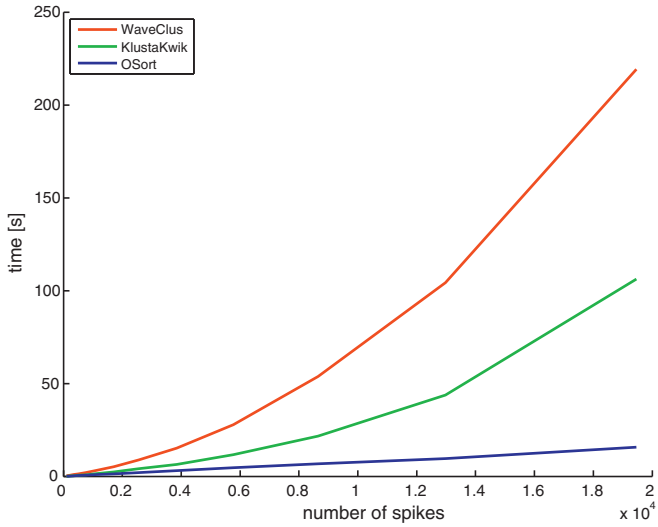


Fig. 9. Relation between number of spikes and time needed for the spike sorting algorithms to run. Long JW signals (960 s, noise level 0.15) cut into several parts were used. The parameters were optimized using the first signal segment with 1400 spikes and remained the same for all other segments.

4. Conclusion

Three widely-used publicly-available spike sorting algorithms were compared (WaveClus, KlustaKwik, OSort) with regard to their parameter settings, using single-channel artificial data with different noise levels and different number of neurons. To avoid biased results, an optimization technique was employed based on Adjusted Mutual Information to find near-optimal parameter settings for our artificial signals. When using the near-optimal parameters, each algorithm improved its spike sorting accuracy as opposed to when only the default parameters were used ($p < 0.01$). Using these settings, an objective comparison of the three algorithms was made.

WaveClus was the best performing spike sorting algorithm. The accuracy of KlustaKwik was comparable to that of WaveClus at a lower noise level (0.00–0.15), and worse otherwise. Although OSort performed less well than both WaveClus and KlustaKwik, it sorted spikes at more than five times faster, and can thus be recommended for real-time signal processing with a low amount of noise present (below noise level 0.15). Where there is high noise (noise level greater than 0.3), none of the three algorithms provided reasonable results.

As our artificial data is publicly available online, we believe that our framework can be extended to further spike sorting algorithms, thus providing an objective comparison platform for neuroscience researchers.

Acknowledgements

This work has been supported by Czech Science Foundation research program 309/09/1145 and by Grant Agency of the Czech Technical University in Prague Grant no. SGS10/279/OHK3/3T/13. The second author has been supported by Visegrad Fund scholarship no. 51000739.

References

Adamos D, Kosmidis E, Theophilidis K. Performance evaluation of PCA-based spike sorting algorithms. Comput Methods Prog Biomed 2008;91:232–44.
Adamos D, Laskaris N, Kosmidise E, Theophilidis G. Nass: an empirical approach to spike sorting with overlap resolution based on a hybrid noise-assisted methodology. J Neurosci Methods 2010;190(1):129–42.
Cheeseman P, Stutz J. Bayesian classification (autoclass): theory and results. AAAI Press, MIT Press; 1996.
Delescluse M, Pouzat C. Efficient spike-sorting of multi-state neurons using inter-spike intervals information. J Neurosci Methods 2006;150(1):16–29.
Ding W, Yuan J. Spike sorting based on multi-class support vector machine with superposition resolution. Med Biol Eng Comput 2008;46:139–45.
Fee MS, Mitra PP, Kleinfeld D. Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-gaussian variability. J Neurosci Methods 1996;69(2):175–88.
Fee MS, Mitra PP, Kleinfeld D. Erratum: automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-gaussian variability. J Neurosci Methods 1997;71(2):233.
Franke F, Natora M, Boucsein C, Munk MHJ, Obermayer K. An online spike detection and spike classification algorithm capable of instantaneous resolution of overlapping spikes. J Comput Neurosci 2009;29:127–48.
Gasthaus J. Spike sorting using time-varying DIRICHLET process mixture models. Master's thesis. University College London; 2008.
Gerstein GL, Clark W. Simultaneous studies of firing patterns in several neurons. Science 1964;143(20):1325–7.
Gibson S, Judy J, Markovic D. Comparison of spike-sorting algorithms for future hardware implementation. In: Engineering in medicine and biology society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE; 2008. p. 5015–20.
Gibson S, Judy J, Markovic D. Technology-aware algorithm design for neural spike detection, feature extraction, and dimensionality reduction. IEEE Trans Neural Syst Rehabil Eng 2010;18(5):469–78.
Harris KD. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. J Neurophysiol 2000;84(1):401–14.
Herbst J, Gammeter S, Ferrero D, Hahnloser R. Spike sorting with hidden Markov models. J Neurosci Methods 2008;174(1):126–34.
Hulata E, Segev R, Ben-Jacob E. A method for spike sorting and detection based on wavelet packets and shannons mutual information. J Neurosci Methods 2002;117:1–12.
Joshua M, Elias S, Levine O, Bergmana H. Quantifying the isolation quality of extracellularly recorded action potentials. J Neurosci Methods 2007;163:267–82.
Kretzberg J, Coors T, Furche J. Comparison of valley seeking and T-distributed EM algorithm for spike sorting. BMC Neurosci 2009;10(1):47.
Lewicki M. A review of methods for spike sorting: the detection and classification of neural action potentials. Network-Comp Neural 1998;9(4):53–78.
Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 2004;16(8):1661–87.
Rutishauser U. Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. J Neurosci Methods 2006;154:204–24.
Sahani M. Latent variable models for neural data analysis. Ph.D. thesis. Pasadena, California: California Institute of Technology; 1999.
Salganicoff M, Sarna M, Sax L, Gerstein GL. Unsupervised waveform classification for multineuron recordings: a real-time, software-based system. J Neurosci Methods 1998;25(3):181–7.
Schmitzer-Torbert N, Jackson J, Henze DHK, Redish A. Quantitative measures of cluster quality for use in extracellular recordings. Neuroscience 2005;131:1–11.
Shoham S, Fellows MR, Normann RA. Robust, automatic spike sorting using mixtures of multivariate T-distributions. J Neurosci Methods 2003;127(2):111–22.
Smith SW. The scientist and engineer's guide to digital signal processing. San Diego, California: California Technical Publishing; 1999.
Takahashi S, Anzai Y, Sakurai Y. A new approach to spike sorting for multi-neuronal activities recorded with a tetrode – how ICA can be practical. Neurosci Res 2003a;46(3):265–72.
Takahashi S, Anzai Y, Sakurai Y. Automatic sorting for multineuronal activity recorded with tetrodes in the presence of overlapping spikes. J Neurophysiol 2003b;89(4):2245–58.
Tankus A, Yeshurun Y, Fried I. An automatic measure for classifying clusters of suspected spikes into single cells versus multiunits. J Neural Eng 2009;6(5):056001.
Vinh N, Epps J, Bailey J. Information theoretic measures for clusterings comparison: is a correction for chance necessary? In: Proceedings of the 26th annual international conference on machine learning. ACM; 2009. p. 1073–80.
Vinh N, Epps J, Bailey J. Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. JMLR 2010;11:2837–54.
Wang G, Zhou Y, Chen A, Zhang P, Liang P. A robust method for spike sorting with automatic overlap decomposition. IEEE Trans Bio-Med Eng 2006;53(6):1195–8.
Warrens M. On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index. J Classif 2008;25(2):177–83.

Proceedings of the 4th International
IEEE EMBS Conference on Neural Engineering
Antalya, Turkey, April 29 - May 2, 2009

Identifying Number of Neurons in Extracellular Recording

D. Novák[1], J. Wild[1], T. Sieger[1], R. Jech[2]
[1]Department of Cybernetics, Czech Technical University in Prague, Czech Republic
[2]Department of Neurology, 1st Medical Faculty, Charles University, Czech Republic

***Abstract*—One of the most difficult aspects of spike sorting is choosing the number of neurons in extracellular recording. The paper proposes a methodology for estimating the number of neurons based on the Gaussian mixture model. The following criteria have been examined: Bayesian selection method, Akaikes information criteria, minimum description length, minimum message length, fuzzy hyper volume, evidence density and partition coefficient. In order to validate the procedure, an experimental comparative study was carried out, comparing the proposed methodology with three spike sorting algorithms. The proposed methodology has an advantage of setting the minimum number of parameters and is very robust to background noise. We conclude that only fuzzy hyper volume and evidence density criteria are able to identify the correct number of neurons across different noise levels.**

***Index Terms*—Spike sorting, Gaussian mixture models, Model selection, Cluster evaluation**

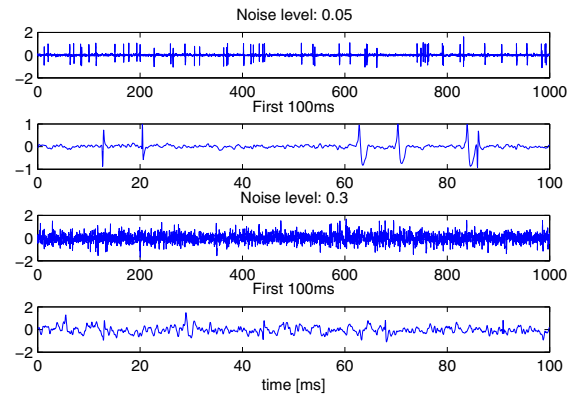


Fig. 1. Simulated recording with low=0.05 and high=0.3 noise level.

I. Introduction

Extracellular recording of multiple neurons in densely packed neural structure as hippocampus (human epilepsy patient) or basal ganglia (deep brains stimulation) is one of the most used technique to reveals the brain's coding representation. Spike sorting algorithms are a must tool for further processing of such acquired data. Throughout the literature several methods of automatically identifying and separating the neurons using unsupervised learning were proposed [1],[2],[3]. An important part of the sorting is determining the number of constituent clusters which best describe the data. A natural choice is to consider that each neuron represented by set of spikes shapes is generated by simple probability distribution and that the whole data set can be described as a weighted sum of these simpler distributions.

II. Methodology

A. Simulation of Data

The same data were applied as in [1] for comparison purposes. First, background noise was generated by superimposing spikes at random times at positions. Next, a train of three distinct spike shapes was superimposed on the noise signal at random times. Amplitude of the three spike classes was normalized to have a peak value of 1. The noise level was determined from its standard deviation, which was equal to 0.05, 0.1, 0.15,0.2, 0.25, 0.3 and 0.35 relative to the amplitude of the spike classes. Two spike trains examples at coarser and finer scale are shown in Fig. 1.

B. Model of Data

We assume that our data are generated by finite mixture models. Indeed in case, such as with stationary spike shapes and uncorrelated noise, the clusters will be nearly spherical in which the model can be very accurate. In less ideal situations, such as correlated noise or non-stationary spike shapes, the partition can be modelled by general Gaussian mixture models [4]. Let $\boldsymbol{X} = [X_1, \ldots, X_d]$ be a d-dimensional random variable, with $\boldsymbol{x} = [x_1, \ldots, x_d]$ representing one particular outcome of $\boldsymbol{X}$. It is said that $\boldsymbol{X}$ follows a K-component finite mixture distribution if its probability density function can be written as

$$p(\boldsymbol{x}|\boldsymbol{\Theta}) = \sum_{k=1}^{K} p(k) p(\boldsymbol{x}|\boldsymbol{\Theta}_k) \tag{1}$$

where $p(1), \ldots, p(K)$ are the mixing probabilities, each $\boldsymbol{\Theta}_k$ is the set of parameters defining the kth component and $\boldsymbol{\Theta} = \{\boldsymbol{\Theta}_1, \ldots, \boldsymbol{\Theta}_k, p(1), \ldots, p(K)\}$ is the complete set of parameters needed to specify the mixture. In this paper, we assume that all the components have d-variate Gaussian distributions (2), with each one characterized by $\boldsymbol{\Theta}_k = \{\mu_k, \sigma_k\}$. Throughout the paper we assume that all data/the feature vectors $\boldsymbol{x}^n$ are independent of each other. Therefore the covariance matrix degenerates into a variance vector σ_k.

$$p(\boldsymbol{x}^n|\boldsymbol{\Theta}_k) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}\sigma_{k,i}} \exp - \frac{(x_i^n - \mu_{k,i})^2}{2\sigma_{k,i}^2} \tag{2}$$

Given a set of N independent and identically distributed samples $\mathcal{X} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^N\}$, the log-likelihood corresponding to a K-component mixture is

$$L(\mathcal{X}|\mathbf{\Theta}_K) = \log \prod_{i=1}^{N} p(\boldsymbol{x}^i|\mathbf{\Theta}) = \sum_{k=1}^{K} p(k)p(\boldsymbol{x}^i|\mathbf{\Theta}_k) \quad (3)$$

The standard method used to fit finite mixture models to observed data is the expectation-maximization (EM) algorithm, which converges to a maximum likelihood (ML) estimate of the mixture parameters [5].

Several selection methods have been proposed to estimate the number of components of a mixture. The methods start by obtaining a set of candidate models (usually by EM) for a range of values of k (from k_{min} to k_{max}) which is assumed to contain the true/optimal k [6]. The number of components is then selected according to

$$\hat{k} = \underset{k}{\operatorname{argmin}} \left\{ \mathcal{C}\left(\hat{\mathbf{\Theta}}(k), k\right),\ k = k_{min}, \ldots, k_{max} \right\} \quad (4)$$

where $\mathcal{C}$ is some model selection criterion, and $\hat{\mathbf{\Theta}}(k)$ is an estimate of the mixture parameters assuming that it has k components. Usually, these criteria have the form $\mathcal{C}$

$$\mathcal{C}\left(\hat{\mathbf{\Theta}}(k), k\right) = -L\left(\mathcal{X}|\hat{\mathbf{\Theta}}(k)\right) + \mathcal{P}(k) \quad (5)$$

where $\mathcal{P}(k)$ is an increasing function penalizing higher values of k. Whilst the first measure decreases with the number of parameters, the second (often referred to as the 'Occam's razor' after the 13th century philosopher) increases as more parameters are estimated using a finite data set. In the following paragraphs we will discuss several model selection criteria that follow the scheme described by (5) and that try to cope with the problems mentioned above.

1) AIC: A number of interpretations of the AIC criterion have been applied to unsupervised learning. We use the following AIC criterion offered in [7]. The AIC is defined as

$$AIC(K) = -\frac{2}{N} L(\mathcal{X}|\mathbf{\Theta}_K)(N - 1 - d - \frac{K_{max}}{2}) + 3N_p \quad (6)$$

where K_{max} is the largest number of components and N_p is the number of parameters in the model [1].

2) BSM: BSM is based again on Bayesian methodology derived in [8]. It has been tested on synthetic and real data sets and according to Roberts [8] it often outperforms the other more heuristic methods. It is defined as:

$$\begin{aligned} BSM(K) = {} & L(\mathcal{X}|\mathbf{\Theta}_K) - Kd \log(2\sigma_{pop}^2) + \log(K-1)! \\ & + \frac{N_p}{2} + \frac{1}{2}\left(\sum_{k=1}^{K} \log \sum_{n=1}^{N} \frac{p(k|\boldsymbol{x}^n)}{p(k)} - \frac{p(K|\boldsymbol{x}^n)}{p(K)} \right) \\ & + 2d \sum_{k=1}^{K} \log(\sqrt{2}Np(k)) - 2\sum_{k=1}^{K}\sum_{i=1}^{d} \log \sigma_{k,l}^2 \quad (7) \end{aligned}$$

[1]We do not take into the account small parameters (less than 10^{-2}).

where $p(k|\boldsymbol{x}^n)$ is the probability that the sample $\boldsymbol{x}^n$ was generated by kth mixture, σ_{pop}^2 is the diagonal element of covariance matrix of $\mathcal{X}$.

3) MDL: The MDL approach originates from an optimal coding viewpoint. The second term of (8) is the average code length for transmitting the model parameters $\mathbf{\Theta}_k$, while the first term is the average code length for transmitting the discrepancy between the model and actual values $\mathcal{X}$ [9].

$$MDL(K) = 2L(\mathcal{X}|\mathbf{\Theta}_K) + N_p \log(N) \quad (8)$$

4) MML: Again the MML is based on an information-theoretic perspective. It was developed extensively by Oliver [7]. The MML expression is defined via

$$\begin{aligned} MML(K) = {} & -L(\mathcal{X}|\mathbf{\Theta}_K) + Kd \log(2\sigma_{pop}^2) - \log(K-1)! \\ & + \frac{N_p}{2} + \frac{N_p}{2} \log \kappa(N_p) - \log K! \\ & + \sum_{k=1}^{K}\sum_{i=1}^{d} \frac{\sqrt{(}2)N_k}{\sigma_{k,i}^2} - \frac{1}{2}\sum_{k=1}^{K} \log p(k) \quad (9) \end{aligned}$$

where $\kappa(N_p)$ is the optimal lattice quantising constant in an N_p-dimensional space. Since optimal lattice constants are not known in some dimensions we used the same linear interpolation as given in [8].

5) FHV: The FHV looks at models with the lowest total volume. It was defined in [10] where it was used as a cluster validity measure. The hypervolume criterion is related to the within-cluster deviation, but due to its original fuzzy characteristics, unlike the square error criterion, it is not a monotone function of k. FHV is defined by

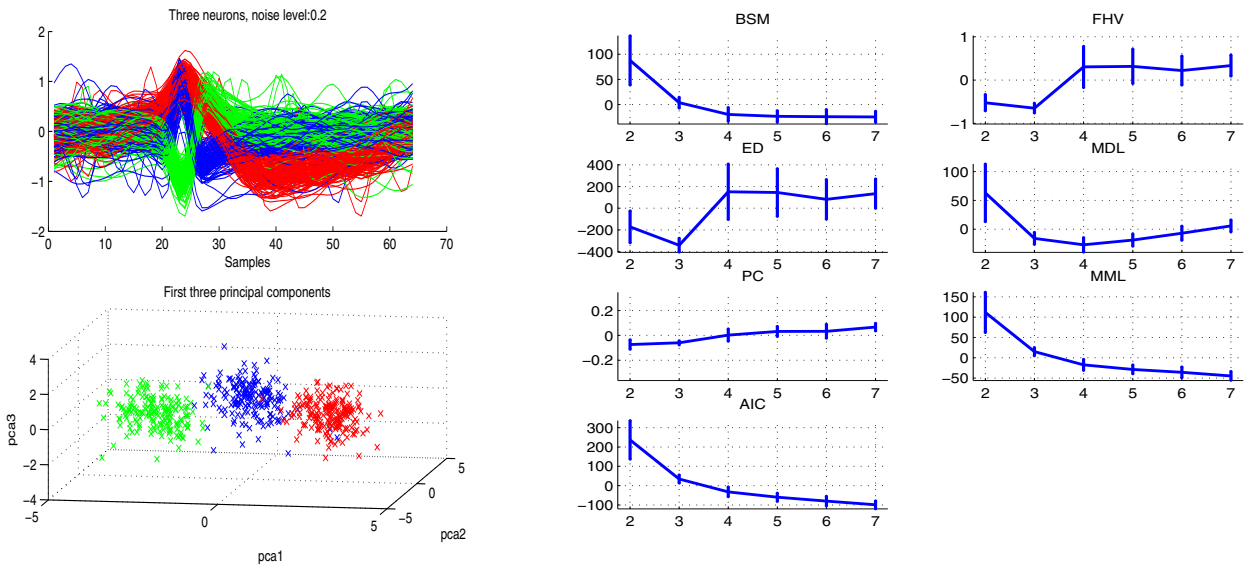
$$FHV(K) = \sum_{k=1}^{K} \prod_{i=1}^{N} \sigma_{k,i}^2 \quad (10)$$

6) ED: This criterion is argued for in [8] and allows (10) to act as a penalty term, in such a way that data models with large values of $FHV(K)$ have correspondingly low prior probabilities $p(k)$.

$$ED(K) = \frac{L(\mathcal{X}|\mathbf{\Theta}_K)}{FHV(K)} \quad (11)$$

7) PC: To calculate PC, we sum the squares of the probability that object $\boldsymbol{x}^n$ belongs to component k as it was defined in [7]

$$PC(K) = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} p(k|x^n) \quad (12)$$



(a) Spike Trains of three neurons

(b) Model selection criterion results

Fig. 2. Sub-figure (a): Three neurons with noise level at 0.2. Number of outliers were introduced by superimposition at background noise, Sub-figure (b): In this example only FHV and ED measures support three-clusters underlying model. Note that the number of clusters is clearly seen in PCA projection on the three biggest principal components of FHV and ED measures.

III. DISCUSSION

We estimated the number of neurons for different noise levels. One example of spikes is shown in Fig. 2(a) along with its PCA representation. The methodology performance for noise level 0.2 is shown in Fig .2(b). Results are presented over one hundred runs of the EM algorithm. Both figures show the mean and the standard deviation (SD) for $K = 2..7$. In all cases the true number is taken as the minimum. The input to GMM models was an amplitude of a spike shape. Only FHV and ED measures detected the correct number of neurons across all noise levels while MDL criteria overestimated the number as can be seen in Fig. 2(b). Note that standard deviation of FHV, ED and PC measures is the smallest for three clusters. The rest of criteria do not penalize more complex models resulting in their monotonic behavior. Therefore, we focused only on FHV and ED measures in the following comparison experiment.

Three commonly used spike sorting algorithms were used for result comparison. The first algorithm we compared against is the well known KlustaKwik clustering algorithm [3]. We used the first 10 principal components, computed using PCA, as features. The minimum number of clusters was set to 3 and the maximum number clusters to 30. The second algorithm we compared against is the WaveClust algorithm developed by Quiroga [1]. The extraction threshold for spike detection was set to 4 times the standard deviation of the background noise. Finally, the last algorithm we compared against is online clustering approach Qsort [2]. The extraction threshold was 2 times the standard deviation of signal energy. The minimum number of spikes to form a cluster was 130 in WaveClust and Qsort. All numerous remaining parameters of the presented algorithms were set to the default values.

Results for all noise level are presented in Table I. The subscript index in FHV and ED indicates that the input to GMM model was the first 3 PCA components of a spike's amplitude. Regarding EM algorithm parameter adjustment, the number of iterations was set up to 100 and the algorithm stopped if the change of the loglikelihood was smaller that 0.001. The results of KlustaKwik are averaged for 10 runs of the algorithm. All measures correctly identified three neurons for noise level till 0.15. WaveClust underestimates the number from noise level 0.2 when spikes start being overlapped. As background noise increases, the spike shapes are more polluted resulting in a failure of all approaches for noise level 0.35. FHV_3 and ED_3 criteria underestimated the correct number from level 0.25 since in the elevated presence of noise 3 principal components are not sufficient for the partition description.

We have validated the number of clusters/neurons detected using the following criteria i) refractory period: less than a small (e.g. 5%) have an InterSpike Interval (ISI) of less than 3 ms, ii) projection test [2]: the minimum distance between projection of every spike and the center of the cluster, iii) visual inspection of cross- and auto-correlation [3]. We present the evaluation results for noise level 0.25 in detail. First, look at the results of projection test distance D. The clusters 1 and 3 are well separated ($D_{13} = 15$), while clusters 1,2 ($D_{12} = 7.7$) and 2,3 ($D_{23} = 8.8$) overlap. As we can observed from Table I, values below $D < 10$ indicate overlapping spikes. Secondly, ISI values reflect influence of

TABLE I
COMPARISON OF FHV AND ED MEASURES WITH SORTING ALGORITHMS

	0.05	0.1	0.15	0.2	0.25	0.3	0.35
FHV	3	3	3	3	3	3	2
FHV_3	3	3	3	3	2	2	2
ED	3	2	3	3	3	3	2
ED_3	3	3	3	3	2	2	2
WaveClust	3	3	3	2	2	2	1
Qsort	3	3	3	3	3	3	2
KlustaKwik	3.6	3.0	2.9	2.6	2.9	2.5	2
D_{13}	39	32	28	10	15	14	-
D_{12}	45	20	28	16	**7.7**	**5.9**	**9.9**
D_{23}	24	37	14	21	**8.8**	**8.8**	-
ISI_1	2.5	2.7	**5.3**	1.5	2.9	**8.4**	-
ISI_2	2.4	1.8	2.5	3.9	3.9	**5.5**	**13.9**
ISI_3	2.2	2.4	1.2	2.8	**7.5**	**8.6**	**13.6**

noise to cluster overlapping. The threshold ISI value of 5% was determined as an indication of overlaps. The threshold is exceeded in case of the first neuron as Fig. 3 depicts as well as in all cases for noise levels 0.3 and 0.35 - see Table I.

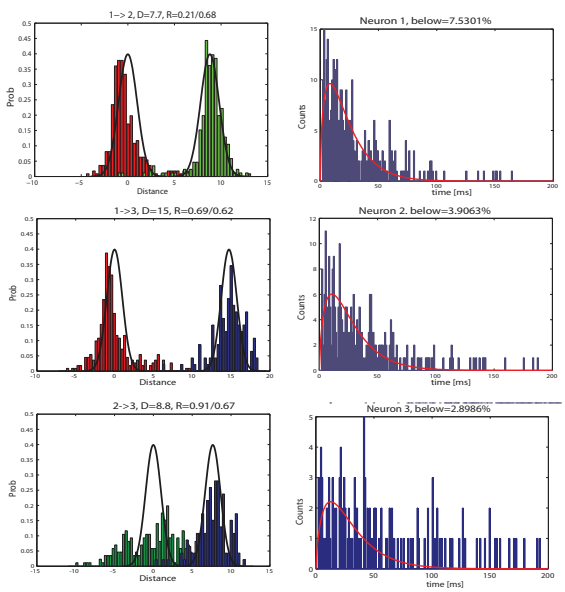


Fig. 3. Projection test for noise level 0.25 and ISI histograms.

It has been proposed that the neurons spike refractory period can be used as a reliable indicator of whether the recording was made from a single cell or from multiple neurons. Conversely, cross-correlation of two clusters with a common refractory period is taken as an indication that the clusters actually represent the same cell [3]. Fig. 4 depicts the auto (ACG)- and cross-correlogram (CCG) for noise level 0.25. In case of two completely unrelated spike trains, the CCG is expected to be flat. However, the high value in CCG of neurons 2 and 3 indicates spikes overlap.

IV. CONCLUSION

We have presented a method for automatic unsupervised determination of the number of neurons in extra-cellular recording. The proposed methodology has an advantage of setting the minimum number of parameters and is very robust to background noise. We shown that two particular measures are helpful in determining the correct number of neurons: Fuzzy hyper volume and Evidence density. Furthermore, the problem of cluster evaluation was addressed. We suggested to use distance of projection test as an indication of spikes overlap.

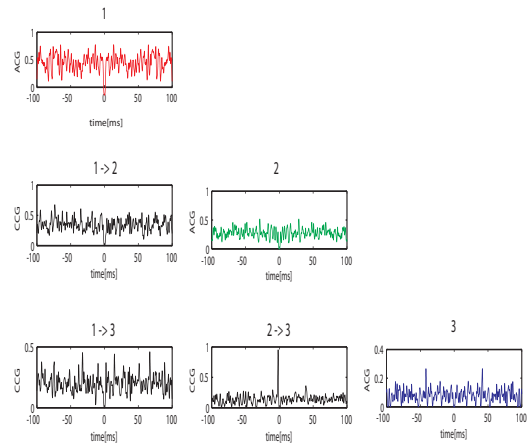


Fig. 4. Auto and cross-correlograms of spike trains for noise level 0.25

ACKNOWLEDGMENT

This work was supported by the research program Czech Science Foundation 309/09/1145. This work was partially supported by Spanish Ministry of Science and Innovation, grant TEC2008-05871/TEC.

REFERENCES

[1] R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Comput.*, vol. 16, no. 8, pp. 1661–1687, 2004.
[2] R. U., S. E., and M. A., "Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo." *J Neuroscience Methods*, 2006.
[3] K. D. Harris, D. A. Henze, J. Csicsvari, H. Hirase, K. D, D. A. Henze, and J. Csicsvari, "Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements," *J Neurophysiol*, vol. 84, pp. 401–414, 2000.
[4] L. M., "A review of methods for spike sorting: the detection and classification of neural action potentials," *Comput. Neural Syst*, 1998.
[5] G. McLachlan and D. Peel, *Finite Mixture Models*. John Wiley & Sons, 2000.
[6] A.T.Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 3, pp. 381–396, March 2002.
[7] J. Oliver, R.A.Baxter, and C. Wallace, "Unsupervised learning using mml," in *Machine Learning:Proceedings of the Thirteen International Conference (ICML 96)*. San Francisco: Morgan Kaufmann Publishers, 1996, pp. 364–372.
[8] S.J.Roberts, D. Husmeier, I. Rezek, and W. Penny, "Bayesian approaches to mixture modelling," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 11, pp. 1133–1142, 1998.
[9] T. Hastie, R. Tibshirani, and J. Friedman, *Data mining, Inference and Prediction*. Springer-Verlag, 2001.
[10] I.Gath and B.Geva, "Unsupervised optimal fuzzy clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 773–781, 1989.

Chapter 4

Metabolomic and statistical data analysis in diabetes mellitus

Perioperative Tight Glucose Control Reduces Postoperative Adverse Events in Nondiabetic Cardiac Surgery Patients

Jan Bláha, Miloš Mráz, Petr Kopecký, Martin Stříteský, Michal Lipš, Michal Matias, Jan Kunstýř, Michal Pořízka, Tomáš Kotulák, Ivana Kolníková, Barbara Šimanovská, Mykhaylo Zakharchenko, Jan Rulíšek, Robert Šachl, Jiří Anýž, Daniel Novák, Jaroslav Lindner, Roman Hovorka, Štěpán Svačina, and Martin Haluzík

Department of Anesthesia, Resuscitation and Intensive Medicine (J.B., P.K., M.S., M.L., M.Ma., J.K., M.P., T.K., I.K., B.S., M.Z., J.R., R.S.), Third Department of Medicine–Department of Endocrinology and Metabolism (M.Mr., S.S., M.H.), and Second Department of Surgery– Department of Cardiovascular Surgery (J.L.), General University Hospital, Charles University in Prague First Faculty of Medicine, 128 08 Prague 1, Czech Republic; Department of Cybernetics (J.A., D.N.), Faculty of Electrical Engineering, Czech Technical University, 121 35 Prague 2, Czech Republic; and Institute of Metabolic Science (R.H.), University of Cambridge, Cambridge CB2 3AD, United Kingdom

Context: Tight glucose control (TGC) reduces morbidity and mortality in patients undergoing elective cardiac surgery, but only limited data about its optimal timing are available to date.

Objective: The purpose of this article was to compare the effects of perioperative vs postoperative initiation of TGC on postoperative adverse events in cardiac surgery patients.

Design: This was a single center, single-blind, parallel-group, randomized controlled trial.

Settings: The setting was an academic tertiary hospital.

Participants: Participants were 2383 hemodynamically stable patients undergoing major cardiac surgery with expected postoperative intensive care unit treatment for at least 2 consecutive days.

Intervention: Intensive insulin therapy was initiated perioperatively or postoperatively with a target glucose range of 4.4 to 6.1 mmol/L.

Main Outcome Measures: Adverse events from any cause during postoperative hospital stay were compared.

Results: In the whole cohort, perioperatively initiated TGC markedly reduced the number of postoperative complications (23.2% vs 34.1%, 95% confidence interval [CI], 0.60–0.78) despite only minimal improvement in glucose control (blood glucose, 6.6 ± 0.7 vs 6.7 ± 0.8 mmol/L, $P < .001$; time in target range, 39.3% ± 13.7% vs 37.3% ± 13.8%, $P < .001$). The positive effects of TGC on postoperative complications were driven by nondiabetic subjects (21.3% vs 33.7%, 95% CI, 0.54–0.74; blood glucose 6.5 ± 0.6 vs 6.6 ± 0.8 mmol/L, not significant; time in target range, 40.8% ± 13.6% vs 39.7% ± 13.8%, not significant), whereas no significant effect was seen in diabetic patients (29.4% vs 35.1%, 95% CI, 0.66–1.06) despite significantly better glucose control in the perioperative group (blood glucose, 6.9 ± 1.0 vs 7.1 ± 0.8 mmol/L, $P < .001$; time in target range, 34.3% ± 12.7% vs 30.8% ± 11.5%, $P < .001$).

Conclusions: Perioperative initiation of intensive insulin therapy during cardiac surgery reduces postoperative morbidity in nondiabetic patients while having a minimal effect in diabetic subjects. (***J Clin Endocrinol Metab* 100: 3081–3089, 2015)**

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received April 13, 2015. Accepted June 10, 2015.
First Published Online June 16, 2015

Abbreviations: BMI, body mass index; CABG, coronary artery by-pass grafting; CI, confidence interval; CKD, chronic kidney disease; EuroSCORE, European System for Cardiac Operative Risk Evaluation; GI, gastrointestinal; ICU, intensive care unit; IIT, intensive insulin therapy; LOS, length of stay; PERI, perioperative; POST, postoperative; RCT, randomized controlled trial; RR, relative risk; TGC, tight glucose control.

Elevated blood glucose is strongly associated with increased morbidity and mortality of patients with a critical illness or undergoing a major surgical procedure (1–4). In 2001, the landmark Leuven trial performed in a surgical intensive care unit (ICU) demonstrated that tight glucose control (TGC) using intravenous intensive insulin therapy (IIT) aimed to maintain euglycemia (4.4–6.1 mmol/L) substantially reduced in-hospital mortality and the number of postoperative complications (5). Several other studies confirmed the positive effects of TGC on selected postoperative outcomes (6–8), whereas other trials on more heterogeneous ICU populations did not show significant benefits (9, 10). The largest multicenter NICE-SUGAR trial even demonstrated increased mortality in patients with TGC, most likely attributable to increased incidence of hypoglycemia (11). A recent meta-analysis including all major randomized trials in ICUs showed a significant benefit of TGC in surgical but not nonsurgical ICU patients (12). Despite these rather inconsistent findings, the need to control elevated glucose levels in critically ill patients is generally accepted, although target ranges are mostly set higher than those in the original Leuven trial.

Although the concept of TGC has been studied intensively, the optimal timing of TGC initiation in surgical patients remains elusive despite the fact that excessive hyperglycemia during surgery was shown to be an independent predictor of perioperative morbidity and mortality (13, 14). Only a few small studies comparing intraoperative vs postoperative TGC initiation were published with rather inconsistent results (15, 16). A recent meta-analysis of 5 randomized controlled trials comparing intensive and conventional insulin therapy during cardiac surgery did not show any benefit of the former except of reduced infection rates (17). It is thus currently unclear whether perioperative initiation of TGC affects patients' outcomes.

To this end, we performed a randomized controlled trial (RCT) comparing the effects of perioperative vs postoperative initiation of TGC on postoperative adverse events in cardiac surgery patients.

Subjects and Methods

Trial design and population

We conducted a single-center, single-blind, parallel-group, RCT involving adult cardiac surgery patients (aged 18–90 years) in an academic tertiary hospital in Prague, Czech Republic, between January 2007 and June 2012. The study was registered at Clinicaltrials.gov (registration number NCT01548963). Eligible participants were all hemodynamically stable patients undergoing major cardiac surgery with expected postoperative ICU treatment for at least 2 consecutive days. Exclusion criteria included allergy to insulin, mental incapacity, language barrier, and refusal to participate in the study. Severe hemodynamic instability during the surgery, patient's rejection of further participation, or loss to follow-up were set as study discontinuation criteria. Informed consent was obtained from all participants before being enrolled into the trial. The study was approved by the Human Ethics Review Board of General University Hospital in Prague, Czech Republic, and was performed in accordance with the guidelines proposed in the Declaration of Helsinki.

Study interventions

Study participants were randomly assigned to 2 treatment groups: the perioperative (PERI) group with intraoperative and postoperative TGC and the postoperative (POST) group with only postoperative TGC. In the PERI group, IIT was initiated at any time from the beginning of cardiac surgery if blood glucose levels exceeded 6.1 mmol/L, whereas in the POST group, IIT was started after the admission to the postoperative ICU, and every time blood glucose during the operation exceeded 10 mmol/L an iv bolus of 1 to 2 IU of rapid-acting insulin was administered to keep glucose values under this threshold. In both groups, TGC lasted until the end of the ICU stay or until oral intake was restored. The target blood glucose range was set at 4.4 to 6.1 mmol/L. For post hoc result analysis and to better reflect the current, more moderate, glucose control approach a second range of less stringent glucose control was defined as 4.4 to 8.3 mmol/L. For allocation of participants to one of the study groups, a simple randomization procedure according to a computer-generated list of random numbers was used. Only the study coordinator and the operation staff were aware of the treatment assignment; the patients themselves and the postoperative ICU staff and outcome assessors and data analysts were kept blinded to the treatment allocation.

IIT protocols

Two protocols for TGC were used during the trial: the primarily used Matias protocol, which is essentially a modified Leuven protocol with the addition of insulin boluses, was later replaced by the computer-based eMPC (enhanced model predictive control) protocol with variable sampling rates, which in previous trials proved to be superior in terms of efficacy as well as safety (18, 19). Both protocols were described in detail elsewhere (18). Human rapid-acting insulin (Actrapid HM; Novo Nordisk) was given via a central venous line as a continuous infusion alone or in combination with insulin boluses (when the Matias protocol was applied). A standard concentration of 50 IU of insulin in 50 mL of 0.9% NaCl was used.

Blood glucose measurement and glucose infusion

Blood glucose samples were obtained from an arterial line whenever possible; otherwise, a central venous line was used. Capillary samples were not used during the ICU stay but became acceptable after the patient was discharged to a standard ward. Blood glucose levels were assessed by a blood gas analyzer (86.9% of samples, ABL 700; Radiometer Medical–) or a standard point-of-care glucometer (13.1% of samples, ACCU-CHEK Inform system; F. Hoffmann La-Roche AG–). The blood sampling rate was guided by the protocol applied.

During operations, blood glucose was measured every 1 hour with the frequency increasing to every 30 minutes in the on-pump period. In all patients, infusion of a 10% glucose solution with a glucose dose of 6.7 g/h was initiated upon the admission to the postoperative ICU and was continued for approximately 18 hours, when oral food intake was reestablished. In patients receiving mechanical ventilation, the glucose infusion lasted for 48 hours and was then replaced by standard enteral nutrition.

Data collection

Patient history and clinical parameters including age, sex, race, height, weight, body mass index (BMI), EuroSCORE (the European System for Cardiac Operative Risk Evaluation), history of diabetes mellitus, and type of surgery were collected prospectively. Blood glucose levels were recorded from the beginning of the operation until the end of the postoperative hospital stay. Perioperative and postoperative adverse events, medication, and nutrition were continuously monitored and documented.

Outcome measures

The primary study endpoint was defined as the number of adverse events from any cause during the postoperative hospital stay and included following newly developed organ dysfunctions: cardiovascular (low cardiac output syndrome, postoperatively initiated inotropic support or intraaortic balloon counterpulsation, acute myocardial ischemia, moderate to severe arrhythmias, and cardiopulmonary resuscitation); respiratory (acute pneumonia, fluidothorax of >300 ml, reintubation, and acute respiratory distress syndrome/acute lung injury), neurological (stroke and transient ischemic attack), gastrointestinal (GI) (ileus, gastric ulcer, GI bleeding, hepatopathy, acute pancreatitis, and need for parenteral nutrition), renal (acute kidney injury defined by RIFLE criteria–stage injury and above), and infections defined by the clinical picture and the need for systemic antibiotic therapy (detailed criteria for all selected adverse events are listed in the Supplemental Material). All events were evaluated according to the prespecified criteria by attending ICU physicians who were blinded to the treatment assignment. Parameters of glucose control (average blood glucose, time in, above, and below the target range, time in hyperglycemia of >8.3 mmol/L, and number of hypoglycemic episodes) and postoperative hospital stay length were set up as secondary endpoints. Severe hypoglycemia was defined as blood glucose of <2.2 mmol/L.

Statistical analysis

To detect an overall difference of 10% in postoperative complications between the treatment groups with a two-sided 0.1% significance level and a power of 99%, a sample size of 2400 patients in the whole cohort was necessary, assuming a baseline postoperative morbidity of 30%. To include this number of patients, a 3-year inclusion period with 800 patients per year was anticipated. One interim analysis was performed after 1400 patients had been enrolled with the P value maintained at 0.1%, confirming the formerly calculated sample size. Numerical data from both groups were compared using the Student t test or the Mann-Whitney rank sum test as appropriate. Categorical data were analyzed with a 2-sample proportion test using standard approximation. The difference between the primary endpoints was expressed as relative risk reduction with a 95% confidence interval (CI). The significance level was set at $P = .05$. To correct for baseline bias, an adjustment analysis using logistic regression or negative binomial regression and ANOVA or a likelihood ratio test, as appropriate, was performed.

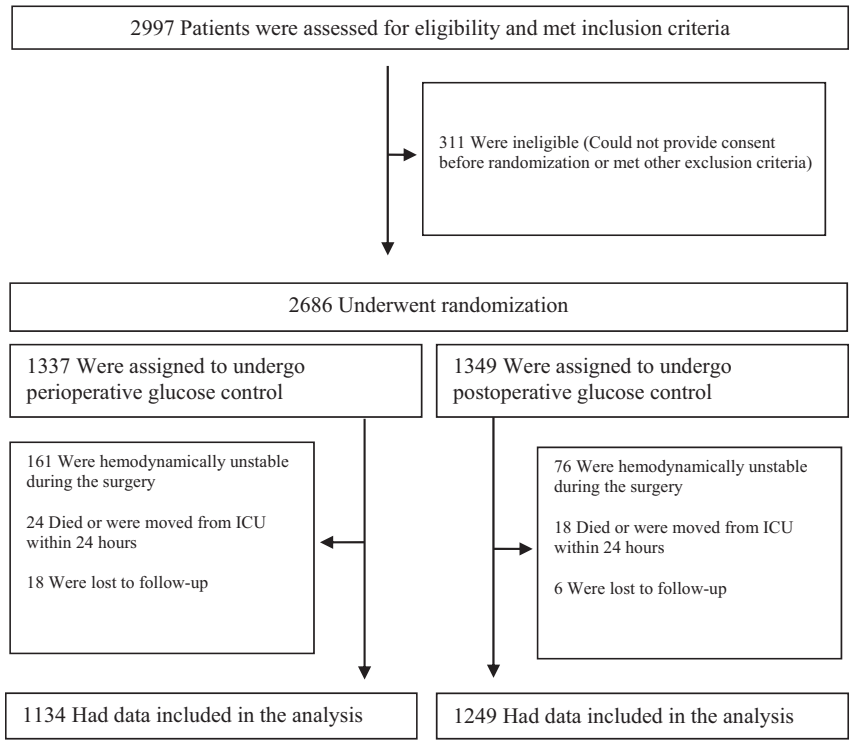


Figure 1. Assessment, randomization and follow-up of study patients.

Results

Baseline characteristics of study subjects

A total of 2383 subjects were randomized into the trial between January 2007 and December 2010, 1134 in the PERI group and 1249 in the POST group. The detailed enrollment process is depicted in the consort diagram (Figure 1). Patients in the POST group were slightly older with a higher prevalence of diabetes mellitus and chronic kidney disease (CKD) together with a higher EuroSCORE. Other baseline parameters including BMI, left ventricular ejection fraction and baseline creatinine were comparable between the two groups (Table 1).

When divided according to the presence of diabetes mellitus, the baseline profiles of nondiabetic sub-

Table 1. Baseline Characteristics of Study Subjects

	Whole Cohort		Nondiabetic Subjects		Diabetic Subjects	
	PERI Group	POST Group	PERI Group	POST Group	PERI Group	POST Group
No. of patients	1134	1249	869	910	265	339
Females, n (%)	323 (28.6)	372 (29.8)	243 (28.0)	263 (28.9)	80 (30.7)	109 (32.2)
Age, y	64.7 ± 11.1	66.6 ± 9.7[a]	64.4 ± 11.5	65.8 ± 10.0	65.8 ± 9.3	68.8 ± 8.3[a]
BMI, kg/m^2	28.4 ± 5.6	28.2 ± 4.3	27.8 ± 5.6	27.9 ± 4.3	30.3 ± 5.4	29.0 ± 4.3[b]
Diabetes mellitus, n (%)	265 (23.4)	339 (27.1)[b]	0	0	265	339
Neurological disease, n (%)	108 (9.5)	106 (8.5)	79 (9.1)	64 (7.0)	29 (10.9)	42 (12.4)
COPD, n (%)	157 (13.8)	193 (15.5)	109 (12.5)	142 (15.6)	48 (18.1)	51 (15.0)
CKD, n (%)	43 (3.8)	78 (6.2)[c]	24 (2.8)	46 (5.1)[b]	19 (7.2)	32 (9.4)
Renal replacement therapy, n (%)	9 (0.8)	14 (1.1)	6 (0.7)	9 (1.0)	3 (1.1)	5 (1.5)
Smoker, n (%)	250 (22.0)	270 (21.6)	200 (23.0)	214 (23.5)	50 (18.9)	56 (16.5)
Baseline creatinine, μmol/L	99.4 ± 61.9	99.3 ± 53.5	95.9 ± 52.6	97.1 ± 48.3	110.9 ± 84.8	105.2 ± 65.1
LV EF (%)	55.8 ± 13.3	55.2 ± 13.8	55.8 ± 13.2	55.8 ± 13.6	55.8 ± 13.6	53.6 ± 14.4[b]
Additive EuroSCORE	3.8 ± 2.2	4.2 ± 2.3[a]	3.8 ± 2.1	4.0 ± 2.3	3.8 ± 2.2	4.7 ± 2.4[a]
Logistic EuroSCORE	7.2 ± 9.6	8.5 ± 12.2[b]	7.3 ± 10.0	8.0 ± 11.8	6.8 ± 8.1	9.8 ± 13.1[a]
Elective surgery, n (%)	1005 (88.6)[c]	1048 (83.9)	758 (87.2)[b]	761 (83.6)	247 (93.2)[c]	287 (84.6)
CABG, n (%)	790 (69.7)	966 (77.3)[a]	564 (64.9)	672 (73.8)[a]	226 (85.3)	294 (86.7)
Aortic valve replacement, n (%)	236 (20.8)	237 (19.0)	211 (24.3)	200 (22.0)	25 (9.4)	37 (10.9)
Mitral valve replacement, n (%)	143 (12.6)	139 (11.1)	121 (13.9)	107 (11.8)	22 (8.3)	32 (9.4)
Other surgery types, n (%)	106 (9.3)[c]	75 (6.0)	100 (11.5)[c]	63 (6.9)	6 (2.3)	12 (3.5)
Off-pump surgery, n (%)	412 (36.3)	506 (40.5)[b]	342 (39.4)	361 (39.7)	70 (26.4)	145 (42.8)[a]
Extracorporeal circulation, n (%)	722 (63.7)[b]	743 (59.5)	527 (60.6)	549 (60.3)	195 (73.6)[c]	194 (57.2)
Extracorporeal circulation duration, min	127.8 ± 82.4	135,1 ± 77.2[b]	136.2 ± 75.7	137.9 ± 82.2	105.0 ± 95.5	127.7 ± 60.1[a]

Abbreviations: COPD, chronic obstructive pulmonary disease; LV EF, left ventricular effusion fraction. Data are expressed as means ± SD or absolute number with relative percentage.

[a] $P < .001$.

[b] $P < .05$.

[c] $P < .01$.

jects in both the PERI and POST groups reflected the situation in the whole cohort with no difference in most of the baseline parameters except of CKD prevalence. In contrast, subjects with diabetes showed increased age and EuroSCORE, decreased BMI and slightly reduced left ventricular ejection fraction, but there was no difference in the number of patients with CKD in the POST group than in PERI group (Table 1).

Types of surgery

Elective operations dominated the spectrum of surgical procedures, with acute operations being performed in approximately 10% to 15% of all subjects. Coronary artery by-pass grafting (CABG) was the most prevalent type of surgery. Other types included aortic, mitral, and tricuspidal valve repair or replacement, thoracic aortic surgery, and pulmonary endarterectomy (Table 1).

Glucose control

During the ICU stay, only minimal differences in the main parameters of glucose control were observed between the two study cohorts, favoring almost exclusively the PERI group, including average blood glucose and time in hyperglycemia as well as the number of hypoglycemic episodes (Table 2). These differences were even less pronounced in the nondiabetic subgroup with comparable average ICU glycemia and time in the target range and reduced number of hypoglycemic episodes. In contrast, subjects with diabetes mellitus in the PERI group showed slightly tighter glucose control as demonstrated by decreased ICU and intraoperative glycemia and longer time spent in the target range, with no significant difference in the occurrence of hypoglycemia, even though the time under the target range was increased. Episodes of severe hypoglycemia (<2.2 mmol/L) were comparably low in all study subgroups (Table 2). During the operation period intravenous insulin was administered to 95.1% of all subjects in the PERI group (94.1% in the nondiabetic and 98.1% in the diabetic subgroup) and to 22.7% of subjects in the POST group (11.9% in the nondiabetic and 51.6% in the diabetic subgroup), respectively.

Perioperative morbidity and mortality

In the whole cohort, the number of patients with postoperatively developed organ complications was significantly reduced in the PERI group compared with that in the POST group (23.2% vs 34.1%; relative risk [RR],

Table 2. ICU Glucose Control (From the Beginning of Operation to the End of ICU Stay)

	Whole Cohort		Nondiabetic Subjects		Diabetic Subjects	
	PERI Group	POST Group	PERI Group	POST Group	PERI Group	POST Group
No. of patients	1134	1249	869	910	265	339
Average blood glucose, mmol/L						
Whole period	6.6 ± 0.7	6.7 ± 0.8[a]	6.5 ± 0.6	6.6 ± 0.8	6.9 ± 1.0	7.1 ± 0.8[a]
Intraoperative period	7.0 ± 1.4	7.4 ± 1.5[a]	6.8 ± 1.1	7.0 ± 1.2[a]	7.7 ± 1.9	8.3 ± 1.8[a]
Time in TGC target range (4.4–6.1 mmol/L), %	39.3 ± 13.7[a]	37.3 ± 13.8	40.8 ± 13.6	39.7 ± 13.8	34.3 ± 12.7[a]	30.8 ± 11.5
Time in GC range (4.4–8.3 mmol/L), %	79.3 ± 13.3[a]	75.8 ± 14.4	82.5 ± 11.1[a]	79.7 ± 12.5	68.8 ± 14.6[a]	65.2 ± 13.9
Time above target range (>8.3 mmol/L), %	14.5 ± 12.2	17.2 ± 13.5[a]	12.5 ± 10.2	13.9 ± 11.8[b]	21.1 ± 15.6	26.1 ± 13.8[a]
Time below target range (<4.4 mmol/L), %	6.2 ± 5.7	7.0 ± 5.8[a]	5.0 ± 5.2	6.4 ± 5.6[a]	10.1 ± 5.7[c]	8.7 ± 5.9
Moderate hypoglycemia (2.2–3.2 mmol/L), no. of measurements/all measurements), %	508/56 319 (0.9)	703/62 855 (1.1)[a]	267/40 766 (0.7)	419/45 100 (0.9)[a]	241/15 553 (1.5)	266/17 755 (1.5)
Severe hypoglycemia <2.2 mmol/L (no. of measurements/all measurements), %	44/56 319 (0.1)	61/62 855 (0.1)	20/40 766 (0.1)	33/45 100 (0.1)	24/15 553 (0.2)	28/17 755 (0.2)

Abbreviation: GC, glucose control. Data are expressed as means ± SD or absolute number with relative percentage.

[a] $P < .001$.

[b] $P < .05$.

[c] $P < .01$.

0.68; 95% CI [CI], 0.60–0.78). This decrease was driven by all types of dysfunctions except respiratory, with neurological and infectious complications showing the maximum reduction (Table 3). The favorable effects of intraoperatively initiated TGC were even more pronounced in nondiabetic subjects, achieving a reduction in risk of developing any kind of postoperative complication of 37% (21.3% of the PERI group vs 33.7% of the POST group; RR, 0.63; 95% CI, 0.54–0.74). Analogous to the whole cohort, only new-onset dysfunctions of the respiratory tract did not differ between the PERI and POST subgroups, whereas cardiovascular, renal, GI, neurological, and infectious complications were decreased to a similar or even greater extent than in the whole group (Table 4). Among subjects with diabetes mellitus, however, no difference could be seen between the PERI and POST groups in the incidence of postoperative complications of all types except cardiovascular (Table 4). With adjustment for baseline differences in age, prevalence of diabetes mellitus and CKD, logistic EUROSCORE, percentage of elective procedures, CABG, off-pump surgery, and extracorporeal circulation between the PERI and POST groups, all the results still retained their significance with the exception of renal complications in the whole cohort and GI adverse events in the nondiabetic group, which both slightly failed to cross the $P < .05$ threshold (RR, 0.55–1.01, $P = .055$ for renal complications and RR, 0.34–1.00, $P = .05$ for GI complications).

Intraoperative initiation of TGC showed no effect on the whole postoperative length of stay (LOS) or the duration of the ICU treatment in the whole cohort (Table 3) or in the nondiabetic subgroup (Table 4), whereas it reduced both ICU and total hospital stay in the diabetic group (Table

Table 3. Perioperative Morbidity and Mortality: Whole Cohort

	PERI Group	POST Group	AD or RR (95% CI)
No. of patients	1134	1249	115
Hospital stay length, days	11.7 ± 8.1	12.2 ± 9.4	0.5 (−0.2 to 1.2)
ICU stay length, hours	117.5 ± 132.1	115.5 ± 117.7	2.0 (−12.2 to 8.1)
Perioperative mortality, no of patients (%)	37 (3.3)	48 (3.8)	0.85 (0.56 to 1.29)
Perioperative morbidity, no. of patients (%)	263 (23.2)	426 (34.1)[a]	0.68 (0.60 to 0.78)
Complications, no. of events (%)			
Cardiovascular	135 (11.9)	257 (20.6)[a]	0.58 (0.48 to 0.70)
Respiratory	72 (6.3)	94 (7.5)	0.84 (0.63 to 1.13)
Renal	88 (7.8)	131 (10.5)[b]	0.74 (0.57 to 0.96)
Gastrointestinal	33 (2.9)	66 (5.3)[c]	0.55 (0.37 to 0.83)
Neurological	30 (2.6)	82 (6.6)[a]	0.40 (0.27 to 0.61)
Infectious	36 (3.2)	89 (7.1)[a]	0.45 (0.31 to 0.65)

Data are expressed as means ± SD or absolute number with relative percentage. The difference between the groups was expressed as absolute difference (AD) for numerical data or relative risk (RR) for categorical data, both with 95% CIs. The AD and RR values are unadjusted.

[a] $P < .001$.

[b] $P < .05$.

[c] $P < .01$.

Table 4. Perioperative Morbidity and Mortality: Nondiabetic and Diabetic Subjects

	Nondiabetic Subjects			Diabetic Subjects		
	PERI Group	POST Group	AD or RR (95% CI)	PERI Group	POST Group	AD or RR (95% CI)
No. of patients	869	910	41	265	339	74
Hospital stay length, days	11.6 ± 7.9	11.6 ± 8.4	0.01 (−0.7 to 0.8)	12.0 ± 8.7	13.6 ± 11.4[a]	1.7 (0.1–3.3)
ICU stay length, hours	120.3 ± 133.7	115.8 ± 118.9	4.5 (−16.4 to 7.3)	108.4 ± 126.5	114.7 ± 114.6[b]	6.3 (2.0–19.0)
Perioperative mortality, no of patients (%)	19 (2.2)	33 (3.6)	0.60 (0.35 to 1.05)	18 (6.8)	15 (4.4)	1.54 (0.79–2.99)
Perioperative morbidity, no. of patients (%)	185 (21.3)	307 (33.7)[c]	0.63 (0.54 to 0.74)	78 (29.4)	119 (35.1)	0.84 (0.66–1.06)
Complications, no. of events (%)						
Cardiovascular	109 (12.5)	193 (21.2)[c]	0.59 (0.48 to 0.73)	26 (9.8)	64 (18.9)[b]	0.52 (0.34–0.80)
Respiratory	56 (6.4)	69 (7.6)	0.85 (0.60 to 1.19)	16 (6.0)	25 (7.4)	0.82 (0.45–1.50)
Renal	54 (6.2)	92 (10.1)[b]	0.61 (0.45 to 0.85)	34 (12.8)	39 (11.5)	1.12 (0.72–1.72)
Gastrointestinal	22 (2.5)	46 (5.1)[b]	0.50 (0.30 to 0.83)	11 (4.2)	20 (5.9)	0.70 (0.34–1.44)
Neurological	8 (0.9)	60 (6.6)[c]	0.14 (0.07 to 0.29)	22 (8.3)	22 (6.5)	1.28 (0.72–2.26)
Infectious	24 (2.7)	60 (6.6)[c]	0.42 (0.26 to 0.67)	12 (4.5)	29 (8.6)	0.53 (0.28–1.02)

Data are expressed as means ± SD or absolute number with relative percentage. The difference between the groups was expressed as absolute difference (AD) for numerical data or relative risk (RR) for categorical data, both with 95% CI. The AD and RR values are unadjusted.

[a] $P < .05$.

[b] $P < .01$.

[c] $P < .001$.

4). Perioperative mortality did not differ significantly between any of the studied groups (Tables 3 and 4).

Discussion

In the present trial, we show that perioperative initiation of TGC reduces postoperative complications and improves outcomes predominantly in nondiabetic patients undergoing cardiac surgery. Although excessive hyperglycemia during surgery is a well-established and independent predictor of perioperative morbidity and mortality (13, 14, 20), only limited data assessing the effects of its lowering are available to date. Lazar et al (15) demonstrated that the administration of glucose-insulin-potassium infusion aimed to maintain blood glucose levels between 6.7 and 10 mmol/L decreases episodes of recurrent ischemia and wound infections and improves 2-year survival compared with a sliding scale insulin protocol with a target range of <13.9 mmol/L in subjects undergoing CABG. Similarly, glucose-insulin-potassium infusion with target glucose levels of 6.0 to 10.0 mmol/L improved myocardial contractile function and decreased inotropic support in a study by Koskenkari et al (21). However, the first RCT to comprehensively assess the value of intraoperative TGC, which included 400 patients receiving either TGC aiming at blood glucose between 4.4 and 5.5 mmol/L or conventional treatment with a glycemic target of <11.1 mmol/L during CABG implantation, failed to show any significant difference in the composite outcome (death, deep sternal wound infection, prolonged infection, cardiac arrhythmias, stroke, and renal failure) between both groups. In fact, intraoperative TGC significantly increased the number of strokes and tended to increase overall mortality, thus raising concerns about the efficacy and safety of TGC during surgical procedures (16).

In contrast to these data, perioperatively initiated TGC in our study markedly decreased postoperative complications with an overall risk reduction of 32%. In addition to confirming the previously established association between perioperative TGC and the reduction in cardiovascular and infectious complications, our data also show a strong beneficial effect of early TGC initiation on other adverse event types including neurological, renal, and GI. Strikingly, this substantial risk reduction was associated with very little overall glucose control improvement in the PERI group throughout the ICU stay. Obviously, the differences in glucose levels were slightly more pronounced in the intraoperative period, but whether this was the primary mechanism by which postoperative complications in the

PERI group were reduced remains questionable (especially considering the fact that the nondiabetic subgroup, which profited the most from perioperative initiation of TGC, showed the least improvement in glucose control and vice versa). The marginally increased incidence of hypoglycemia, a factor associated with higher morbidity and mortality in subjects with TGC, in the POST group might also not fully explain the reduction in postoperative outcomes. It might be speculated that early administration of insulin already during the surgery could have to some extent moderated the developing operation-induced stress response by effects other than glucose lowering, including anti-inflammatory, antioxidant, antithrombotic, and vasodilatory effects (22–24). On the other hand, this would seem to be in contrast with a subanalysis of the Leuven study, which showed that blood glucose, rather than insulin, is responsible for the positive effects of TGC (25). The proposed hypothesis notwithstanding, the exact mechanisms by which intraoperatively initiated IIT reduced the number of postoperative complications remain yet to be fully elucidated. Nevertheless, our data indicate that the intraoperative phase of cardiac surgery might have more significance in postoperative outcomes than previously thought and that the exact timing of insulin infusion might be one of the key elements contributing to the efficacy of the TGC regimen.

Another striking finding of our study was the fact that the reduction in postoperative morbidity connected with intraoperative initiation of TGC was driven predominantly by nondiabetic subjects. These findings are in line with some of the previously published data showing a stronger association of hyperglycemia with increased mortality risk in nondiabetic individuals than in those with diabetes (26, 27) and less benefit of TGC in subjects with diabetes mellitus (6, 28). The reason for this difference might be an adaptive response to hyperglycemia in diabetic patients due to their chronic exposure to higher glucose levels, whereas in nondiabetic individuals such mechanisms are missing (29). Another factor that might have contributed to the different outcomes in individuals with and without diabetes mellitus was the presence of hypoglycemia, as the number of moderately hypoglycemic subjects (2.2–3.2 mmol/L) was significantly reduced in the nondiabetic PERI subgroup compared with that in the POST group, while being conversely increased in diabetic subjects. However, the incidence of severe hypoglycemia was comparable throughout all subgroups. Finally, the fact that >50% of subjects in the diabetic POST subgroup received intravenous insulin during the operation (albeit only in the amount of a single bolus of 1–2 IU and with the target glucose being <10 mmol/L) might have to some extent diluted any positive effects of intraoperative insulin administration in the diabetic PERI group.

Although our study was not powered to assess mortality, this parameter certainly comprises the most important safety signal for any intervention. An increased number of deaths, although statistically nonsignificant (4 vs 0), raised concerns about the safety of intraoperative TGC in the study by Gandhi et al (16). Here we did not find any significant difference in perioperative mortality between any of the PERI and POST subgroups. The length of the postoperative stay as one of the secondary endpoints did not differ between the subgroups either in the whole cohort or in nondiabetic individuals, whereas it was slightly prolonged in the diabetic POST group. These results largely confirm the findings of Gandhi et al (16), who also did not observe any shortening of the LOS in association with intraoperative TGC.

Because TGC per se increases the potential risk of hypoglycemia, safety is one of the primary concerns connected to TGC. Compared with the intensive arms of NICE-SUGAR and the original van den Berghe trial with glucose targets similar to those in our study, both the PERI and POST groups showed lower incidence of severe hypoglycemia (<2.2 mmol/L: 3.2% and 4.2% of subjects in the PERI and POST groups vs 5.1% in van den Berghe et al [5] and 6.8% in NICE-SUGAR [11]) (5, 11). Somehow surprisingly, intraoperative initiation of TGC in our study slightly decreased the number of moderate hypoglycemic episodes (2.2–3.3 mmol/L) and reduced the time spent under the target range compared with results for postoperative initiation in the nondiabetic subgroup, whereas the rates of severe hypoglycemia were comparable across all groups. Our current data do not enable us to unravel the exact mechanisms responsible for this positive effect, although one possible explanation is that in subjects without a previous history of diabetes mellitus, perioperative insulin administration prevented early glycemic rises and stabilized glucose profiles, which resulted in decreased occurrence of hypoglycemia. Nevertheless, the results obtained indicate that intraoperative initiation of TGC is a safe procedure with minimal additional hypoglycemic risk for the patient compared with that for postoperative initiation.

Several limitations might have partially affected the results of the present study and their further applicability. Despite the inclusion efforts, several baseline characteristics differed slightly between the groups including older age, increased prevalence of diabetes mellitus and CKD, lower percentage of elective procedures, higher proportion of CABG and off-pump surgery, and worse prognosis as assessed by the additive and logistic EuroSCORE in the POST group. Nevertheless, the differences between the groups remained valid after statistical adjustment for these baseline

inconsistencies. The higher dropout rate due to hemodynamic instability during surgery in the PERI group as well as administration of low corrective bolus insulin doses during the intraoperative period in part of the POST group might constitute another source of potential bias. It also should be mentioned that the subgroup analysis between diabetic vs nondiabetic patients had not been planned in the original protocol and therefore has to be considered exploratory with all potential limitations. The selected target range reflects the original Leuven trial more than the current, less stringent, recommendations, partially owing to the pre-NICE-SUGAR design of the trial and partially owing to the unclear situation regarding optimal target ranges for different ICU subgroups with some meta-analytic data indicating benefits of tighter glucose targets in cardiac surgery subjects (12, 30). As this trial was not designed to compare different target ranges, we are not able to draw any relevant conclusions about this question. Nevertheless, the lower mortality rates in both the PERI and POST groups than in either the Leuven or the NICE-SUGAR trials suggest that our approach was safe. The absence of postdischarge follow-up due to complicated logistics comprises another limitation of the study. In contrast, the use of 2 different protocols for TGC should not have affected the outcomes, as only 1 protocol was used at a given time and the number of patients treated by a particular protocol was comparable in each group.

In summary, we have demonstrated that intraoperative initiation of TGC using IIT substantially reduces the incidence of postoperative complications without affecting mortality or postoperative LOS in nondiabetic patients undergoing cardiac surgery and has little effect in subjects with diabetes mellitus.

Acknowledgments

We thank Tomas Sieger, PhD, from the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic, for his invaluable help with statistical analysis and adjustment models.

Address all correspondence and requests for reprints to: Martin Haluzík MD, DSc, Third Department of Medicine–Department of Endocrinology and Metabolism, General University Hospital, Unemocnice 1, 128 08 Prague 2, Czech Republic. E-mail: mhalu@lf1.cuni.cz.

This work was supported by the Grant Agency of Charles University in Prague (GAUK 44407), the Ministry of Health of the Czech Republic (Research Grant RVO-VFN64165), and the Charles University in Prague (Research Grant SVV260145).

This trial was registered at Clinicaltrials.gov under registration number NCT01548963.

Disclosure Summary: The authors have nothing to disclose.

References

1. **Furnary AP, Gao G, Grunkemeier GL, et al.** Continuous insulin infusion reduces mortality in patients with diabetes undergoing coronary artery bypass grafting. *J Thorac Cardiovasc Surg*. 2003;125:1007–1021.
2. **Krinsley JS.** Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients. *Mayo Clin Proc*. 2003;78:1471–1478.
3. **Gale SC, Sicoutris C, Reilly PM, Schwab CW, Gracias VH.** Poor glycemic control is associated with increased mortality in critically ill trauma patients. *Am Surg*. 2007;73:454–460.
4. **Ascione R, Rogers CA, Rajakaruna C, Angelini GD.** Inadequate blood glucose control is associated with in-hospital mortality and morbidity in diabetic and nondiabetic patients undergoing cardiac surgery. *Circulation*. 2008;118:113–123.
5. **van den Berghe G, Wouters P, Weekers F, et al.** Intensive insulin therapy in critically ill patients. *N Engl J Med*. 2001;345:1359–1367.
6. **Van den Berghe G, Wilmer A, Milants I, et al.** Intensive insulin therapy in mixed medical/surgical intensive care units: benefit versus harm. *Diabetes*. 2006;55:3151–3159.
7. **Bilotta F, Caramia R, Paoloni FP, Delfini R, Rosa G.** Safety and efficacy of intensive insulin therapy in critical neurosurgical patients. *Anesthesiology*. 2009;110:611–619.
8. **Okabayashi T, Shima Y, Sumiyoshi T, et al.** Intensive versus intermediate glucose control in surgical intensive care unit patients. *Diabetes Care*. 2014;37:1516–1524.
9. **Brunkhorst FM, Engel C, Bloos F, et al.** Intensive insulin therapy and pentastarch resuscitation in severe sepsis. *N Engl J Med*. 2008;358:125–139.
10. **Preiser JC, Devos P, Ruiz-Santana S, et al.** A prospective randomised multi-centre controlled trial on tight glucose control by intensive insulin therapy in adult intensive care units: the Glucontrol study. *Intensive Care Med*. 2009;35:1738–1748.
11. **NICE-SUGAR Study Investigators, Finfer S, Chittock DR, et al.** Intensive versus conventional glucose control in critically ill patients. *N Engl J Med*. 2009;360:1283–1297.
12. **Griesdale DE, de Souza RJ, van Dam RM, et al.** Intensive insulin therapy and mortality among critically ill patients: a meta-analysis including NICE-SUGAR study data. *CMAJ*. 2009;180:821–827.
13. **Doenst T, Wijeysundera D, Karkouti K, et al.** Hyperglycemia during cardiopulmonary bypass is an independent risk factor for mortality in patients undergoing cardiac surgery. *J Thorac Cardiovasc Surg*. 2005;130:1144.e1–1144.e8.
14. **Gandhi GY, Nuttall GA, Abel MD, et al.** Intraoperative hyperglycemia and perioperative outcomes in cardiac surgery patients. *Mayo Clin Proc*. 2005;80:862–866.
15. **Lazar HL, Chipkin SR, Fitzgerald CA, Bao Y, Cabral H, Apstein CS.** Tight glycemic control in diabetic coronary artery bypass graft patients improves perioperative outcomes and decreases recurrent ischemic events. *Circulation*. 2004;109:1497–1502.
16. **Gandhi GY, Nuttall GA, Abel MD, et al.** Intensive intraoperative insulin therapy versus conventional glucose management during cardiac surgery: a randomized trial. *Ann Intern Med*. 2007;146:233–243.
17. **Hua J, Chen G, Li H, et al.** Intensive intraoperative insulin therapy versus conventional insulin therapy during cardiac surgery: a meta-analysis. *J Cardiothorac Vasc Anesth*. 2012;26:829–834.
18. **Blaha J, Kopecky P, Matias M, et al.** Comparison of three protocols for tight glycemic control in cardiac surgery patients. *Diabetes Care*. 2009;32:757–761.
19. **Hovorka R, Kremen J, Blaha J, et al.** Blood glucose control by a model predictive control algorithm with variable sampling rate versus a routine glucose management protocol in cardiac surgery patients: a randomized controlled trial. *J Clin Endocrinol Metab*. 2007;92:2960–2964.
20. **Ouattara A, Lecomte P, Le Manach Y, et al.** Poor intraoperative

blood glucose control is associated with a worsened hospital outcome after cardiac surgery in diabetic patients. *Anesthesiology*. 2005;103:687–694.

21. **Koskenkari JK, Kaukoranta PK, Kiviluoma KT, Raatikainen MJ, Ohtonen PP, Ala-Kokko TI.** Metabolic and hemodynamic effects of high-dose insulin treatment in aortic valve and coronary surgery. *Ann Thorac Surg*. 2005;80:511–517.
22. **Dandona P, Aljada A, Mohanty P, Ghanim H, Hamouda W, Assian E, Ahmad S.** Insulin inhibits intranuclear nuclear factor κB and stimulates IκB in mononuclear cells in obese subjects: evidence for an anti-inflammatory effect? *J Clin Endocrinol Metab*. 2001;86:3257–3265.
23. **Akhtar S, Barash PG, Inzucchi SE.** Scientific principles and clinical implications of perioperative glucose regulation and control. *Anesth Analg*. 2010;110:478–497.
24. **Donatelli F, Cavagna P, Di Dedda G, et al.** Correlation between pre-operative metabolic syndrome and persistent blood glucose elevation during cardiac surgery in non-diabetic patients. *Acta Anaesthesiol Scand*. 2008;52:1103–1110.
25. **Van den Berghe G, Wouters PJ, Bouillon R, et al.** Outcome benefit of intensive insulin therapy in the critically ill: insulin dose versus glycemic control. *Crit Care Med*. 2003;31:359–366.
26. **Meyfroidt G, Keenan DM, Wang X, Wouters PJ, Veldhuis JD, Van den Berghe G.** Dynamic characteristics of blood glucose time series during the course of critical illness: effects of intensive insulin therapy and relative association with mortality. *Crit Care Med*. 2010;38:1021–1029.
27. **Falciglia M, Freyberg RW, Almenoff PL, D'Alessio DA, Render ML.** Hyperglycemia-related mortality in critically ill patients varies with admission diagnosis. *Crit Care Med*. 2009;37:3001–3009.
28. **Whitcomb BW, Pradhan EK, Pittas AG, Roghmann MC, Perencevich EN.** Impact of admission hyperglycemia on hospital mortality in various intensive care unit populations. *Crit Care Med*. 2005;33:2772–2777.
29. **Krinsley JS, Meyfroidt G, van den Berghe G, Egi M, Bellomo R.** The impact of premorbid diabetic status on the relationship between the three domains of glycemic control and mortality in critically ill patients. *Curr Opin Clin Nutr Metab Care*. 2012;15:151–160.
30. **Haga KK, McClymont KL, Clarke S, et al.** The effect of tight glycaemic control, during and after cardiac surgery, on patient mortality and morbidity: a systematic review and meta-analysis. *J Cardiothorac Surg*. 2011;6:3.

Anal Bioanal Chem
DOI 10.1007/s00216-015-9133-0

RESEARCH PAPER

Metabolomic profiling of urinary changes in mice with monosodium glutamate-induced obesity

Helena Pelantová [1,2] · Simona Bártová [1,3] · Jiří Anýž [4] · Martina Holubová [5] · Blanka Železná [5] · Lenka Maletínská [5] · Daniel Novák [4] · Zdena Lacinová [6] · Miroslav Šulc [1] · Martin Haluzík [6] · Marek Kuzma [1]

Received: 29 July 2015 / Revised: 25 September 2015 / Accepted: 19 October 2015

Abstract Obesity with related complications represents a widespread health problem. The etiopathogenesis of obesity is often studied using numerous rodent models. The mouse model of monosodium glutamate (MSG)-induced obesity was exploited as a model of obesity combined with insulin resistance. The aim of this work was to characterize the metabolic status of MSG mice by NMR-based metabolomics in combination with relevant biochemical and hormonal parameters. NMR analysis of urine at 2, 6, and 9 months revealed altered metabolism of nicotinamide and polyamines, attenuated excretion of major urinary proteins, increased levels of phenylacetylglycine and allantoin, and decreased concentrations of methylamine in urine of MSG-treated mice. Altered levels of creatine, citrate, succinate, and acetate were observed at 2 months of age and approached the values of control mice with aging. The development of obesity and insulin resistance in 6-month-old MSG mice was also accompanied by decreased mRNA expressions of adiponectin, lipogenetic and lipolytic enzymes and peroxisome proliferator-activated receptor-gamma in fat while mRNA expressions of lipogenetic enzymes in the liver were enhanced. At the age of 9 months, biochemical parameters of MSG mice were normalized to the values of the controls. This fact pointed to a limited predictive value of biochemical data up to age of 6 months as NMR metabolomics confirmed altered urine metabolic composition even at 9 months.

Keywords Mouse model · Monosodium glutamate (MSG) induced obesity · Diabetes · NMR · Metabolomics · Urine

Electronic supplementary material The online version of this article (doi:10.1007/s00216-015-9133-0) contains supplementary material, which is available to authorized users.

✉ Marek Kuzma
kuzma@biomed.cas.cz

[1] Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20 Prague 4, Czech Republic

[2] Department of Analytical Chemistry, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic

[3] Department of Analytical Chemistry, University of Chemistry and Technology Prague, Technická 5, 166 28 Prague 6, Czech Republic

[4] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Prague 6, Czech Republic

[5] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic

[6] 3rd Medical Department, 1st Faculty of Medicine, Charles University and General Faculty Hospital in Prague, U nemocnice 1, 128 08 Prague 2, Czech Republic

Introduction

Obesity is one of the most prevalent health problems commonly associated with an increased risk of type 2 diabetes mellitus (T2DM), arterial hypertension, cardiovascular disease, and several types of cancer [1]. The elucidation of the pathophysiology of obesity and related complications and their successful treatment require detailed characterization of involved biochemical pathways to dissect optimal nodes for therapeutic intervention. Metabolomic strategy based on NMR spectroscopy followed by multivariate statistical analysis is a perfect tool for understanding such a complex pathophysiological condition as obesity at a detailed metabolic level [2–4].

Numerous rodent models have been developed and used extensively to explore the etiopathogenesis of obesity and its complications [5, 6]. The main strategies can be based on

Published online: 17 November 2015

transgenic [7] or knockout [8] animals and on the obesity induced by a high-fat diet [9–12] or chemically [13–15] in mice without genetic modifications.

Monosodium glutamate (MSG)-induced obesity resulting from subcutaneous injections of MSG to neonatal rodents represents a useful chemically induced rodent model of obesity and insulin resistance in mice and rats. The MSG administration induces specific lesions in the arcuate nucleus (ARC) of the hypothalamus with the damage of most ARC neurons [16] together with ARC area shrinkage, third ventricle widening, and thinning of median eminence [17]. In contrast to ARC, the main food intake regulating center of the hypothalamus, other nuclei in the hypothalamus remain intact after MSG administration [18, 19]. MSG-treated rodents have disturbed light/dark eating cycle owing to impaired vision [20, 21]. They develop obesity with an increased adiposity at sustained body weight [22] resulting rather from a lower metabolic rate than from elevated food intake. Recently, we reported that mice with MSG obesity showed substantially attenuated food intake after 24-h fasting compared to their lean controls [23]. Other features of the MSG rat and mice model similar to human obesity include hyperinsulinemia and insulin resistance [24]. On the other hand, glucose levels in MSG obese male mice remain normal despite marked hyperinsulinemia [25] making MSG mice a suitable model of obesity combined with insulin resistance without overt diabetes which is a very common combination in patients worldwide.

Although several authors followed the biochemical alterations in MSG obese mice [25–27] in detail during past years, none of them deals with this model in terms of metabolomics. In the present study, we show a detailed characterization of the metabolic status of MSG mice using a combination of biometric, biochemical, and hormonal parameters with exploration of the NMR-based metabolome of urine.

Material and methods

Experimental animals

All animal experiments followed the ethical guidelines for animal experiments and the Act of the Czech Republic No. 246/1992 and were approved by the Committee for Experiments with Laboratory Animals of the Academy of Sciences of the Czech Republic. NMRI mice (Charles River, Sulzfeld, Germany) were housed at 23 °C with a daily cycle of 12 h light and dark (light from 6 am). Mice had free access to water and were fed standard chow diet (Mlýn Kocanda, Jesenice, Czech Republic) containing 25, 9 and 66 % calories as protein, fat and carbohydrate, its energy content was 3.4 kcal/g.

For MSG-induced obesity, newborn male NMRI mice were subcutaneously (SC) injected with L-glutamic acid sodium salt hydrate (Sigma, St. Louis, USA) (4 mg/g body weight) from postnatal day 2–8 as described previously [23]. Body weight was monitored once weekly up to the age of 9 months. Urine for NMR-based metabolomics was collected from MSG and control mice (10 animals per group) at the age 2, 6, and 9 months. For this purpose, the mice were housed in individual metabolic cages (Tecniplast, Italy) without access to food and 24 h urine was collected. The samples were stored at −80 °C until NMR analysis.

Overnight-fasted MSG obese mice and their controls at the age 2, 6 and 9 months (n=10) were sacrificed by decapitation starting at 8:00 a.m. The trunk blood was collected, and the plasma was separated and stored at −80 °C. The white adipose tissue (WAT, i.e., subcutaneous—SCAT and intraperitoneal—IPAT), brown adipose tissue (BAT), and the liver of all mice were dissected, weighed, and flash-frozen in liquid nitrogen and stored at −70 °C until RNA extraction. The rate of adiposity was expressed as the fat-to-body weight ratio (ratio of the total adipose tissue weight to the total body weight).

Biochemical experiments

Glucose tolerance test

An intraperitoneal glucose tolerance test (IPGTT) was performed in 6-month-old overnight fasted MSG mice and their controls starting at 8:00 a.m. by IP injected glucose (2 g/kg). Blood glucose was then measured at 0, 15, 30, 60, 90, 120 and 180 min following injection using a Glucocard glucometer (Arkray, Kyoto, Japan).

Determination of hormonal and biochemical parameters

Blood plasma insulin concentrations were measured by RIA assay (Millipore, St. Charles, MI, USA and Linco Research, St. Charles, MI, USA, respectively), and leptin concentrations were determined by ELISA assay (Millipore, St. Charles, MI, USA). Glucose levels were measured using a Glucocard glucometer. All measurements were carried out according to the protocols recommended by the manufacturers.

Determination of mRNA expression

mRNA expressions were determined in 6-month-old mice. Samples of IPAT, SCAT, and liver were processed as described in [28]. Determination of the mRNA expression of genes of interest (acetyl-CoA carboxylase 1 (ACACA)) and fatty acid synthase (FASN) in liver, IPAT and SCAT; major urinary protein-1 (MUP1) in the liver, lipoprotein lipase (LPL), fatty acid binding protein 4 (FABP-4), adiponectin, leptin, peroxisome proliferator-activated receptor alfa and gamma (PPAR-α, PPAR-γ) in IPAT and SCAT, PPAR-α, PPAR-γ, and uncoupling protein 1 (UCP-1) in BAT,

nicotinamide *N*-methyltransferase (NNMT) in SCAT, IPAT, BAT, and liver was performed using an ABI PRISM 7500 instrument (Applied Biosystems, Foster City, CA, USA). The expression of beta-2-microglobulin (B2M) was used to compensate for variations in input RNA amounts and the efficiency of reverse transcription, and the modified formula $2^{-\Delta Ct}$ was used to calculate the relative gene expression. For NNMT mRNA expressions, additional controls of glyceraldehyde-3 phosphate dehydrogenase (GADPH) and beta-glucuronidase (GUSB) were also applied.

Statistical analyses of biochemical parameters and mRNA expressions

The data are presented as the means±SEM for the number of animals indicated in the figures and tables. The data were analyzed by a *t* test, as stated in the figure and table legends, using the Graph-Pad Software (San Diego, CA, USA); $p<0.05$ was considered statistically significant.

Determination of proteins in urine

Protein concentrations were quantified using bicinchoninic acid (Sigma) with bovine serum albumin (Sigma) as a standard in the microscale mode [29]. In the next step, a native polyacrylamide gel electrophoresis was performed for qualitative description of protein in the analyzed samples. The same volume of urine containing a protein mixture was premixed for each sample with five-times concentrated electrophoretic sample buffer and separated on 12 % basic native PAGE in the 90 mM Tris–borate buffer system and 1 mM EDTA (pH 8.0) using 200 V/15 cm for 1 h at 4 °C. The protein bands were visualized using silver staining and bovine serum albumin (Sigma) was used as a molecular weight standard.

NMR spectroscopy

NMR experiments

Before NMR experiments urine samples were thawed at room temperature and centrifuged at 12,000 rpm for 5 min. The 200 μl of supernatant was diluted with 340 μl of H_2O and mixed with 60 μl of phosphate buffer (1.5 M KH_2PO_4 in D_2O containing 2 mM NaN_3 and 0.1 % trimethylsilylpropionic acid (TSP), pH 7.4) and transferred into 5 mm NMR tubes. NMR data were acquired on a 600 MHz Bruker Avance III spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a 5 mm TCI cryogenic probe head. All experiments were performed at 300 K. Automatic tuning and matching, shimming and adjusting of 90° pulse length were performed for each sample.

One-dimensional nuclear Overhauser enhancement spectroscopy (1D-NOESY) was initially acquired using Bruker pulse sequence noesygppr1d with the following parameters: presaturation in relaxation delay (4 s) and mixing time (100 ms) by 25 Hz saturation pulse centered on the water resonance, number of scans (NS)=32, 64 k of data points (TD), spectral width (SW) 20 ppm. Due to the strong background of broad signals the Carr–Purcell–Meiboom–Gill (CPMG) pulse sequence with water presaturation (Bruker pulse sequence cpmgpr1d) was then performed with NS=64, TD=64 k, SW=20 ppm, presaturation in relaxation delay 4 s, echo time 0.3 ms, loop for T2 filter 126. A short J-resolved experiment with presaturation (NS=2, SW=16 ppm, TD=8 k, number of increments=40, SW=78.125 Hz in the indirect dimension, relaxation delay 2 s) was executed for better metabolite identification. Additional 2D ^{1}H-^{1}H correlation spectroscopy (COSY) and ^{1}H-^{13}C heteronuclear multiple-quantum correlation (HMQC) experiments (Bruker pulse sequences cosygpprqf and hmqcphpr) were performed for selected samples.

Data pre-processing

Acquired data were processed with Topspin 3.2 software (Bruker BioSpin, Rheinstetten, Germany). Line broadening of 0.3 Hz for the 1D-NOESY and CPMG experiment was applied. Spectra were automatic phased, baseline corrected and referenced to TSP (0.00 ppm). To reduce the dimensionality of the data matrix and to eliminate the chemical shift variation, all NMR spectra were uniformly binned to 0.04 ppm intervals. Regions corresponding to water (4.68–4.90 ppm), urea (5.65–5.90 ppm) and TSP (−0.12–0.12 ppm) were removed prior to analysis. The normalization to the total spectral area was performed on each spectrum to compensate the different dilution of original urine samples.

Multivariate statistical analysis of NMR data

For the purpose of further analyses all measured spectra were exported from TopSpin as comma-separated values text files. Data were then collected using the MATLAB programming language [30] and arranged into the data matrix. The standard projection to latent structure discriminant analysis models (PLS-DA) were used to find metabolites altered in examined groups according to experiments' settings [31]. For each model, the spectra were centered by subtraction of mean value and Pareto scaled. The leave-one-out validation was used to assess the quality of various models and to choose the correct number of latent variables with respect to the performance of model on the validation data set. The fraction of explained variation R^2 and the fraction of predicted variation Q^2 were also computed to provide more detailed evaluation of trained models. The interpretation of PLS-DA models was performed by the examination of loading plots and S plots. The interesting bins in relation to the response were pointed out by

introducing limits to denominated values in S plots (correlation to the response greater than 0.5 and the loading coefficient greater than 0.05). Comparison of significant bins with the proton spectra revealed that several important signals were split between two bins. Four data matrixes mutually shifted by 0.01 ppm were therefore calculated and each signal was correctly centered into bin. In the next step, bins identified by PLS-DA models were subjected to the standard Student's two sample *t* test and the *p* values were computed.

Metabolite identification and quantification

Metabolites were identified by matching the ^{1}H NMR spectra with Chenomx NMR Suite 7.6 (Chenomx Inc., Edmonton, AB, Canada), BBIOREFCODE 2-0-3 (Bruker BioSpin GmbH, Rheinstetten, Germany) databases, and with published assignments. The identified metabolites were confirmed using COSY and HMQC experiment. Due to the free access of mice to water during the sample collection, urine samples differed in dilution and thus a relative quantification of individual metabolites was performed using the integral intensity of corresponding bins normalized to the total spectral area.

Results

Biochemical parameters

MSG mice and their age-matched controls were examined for fat and body weight and metabolic parameters at 2, 6, and 9 months of age as shown in Table 1. At 2 months of age, the body weight of MSG mice did not differ significantly from their age-matched controls, but their weight of total WAT was significantly higher compared to their age-matched controls. At 6 months of age, both the body weight and weight of total WAT were significantly increased in MSG mice compared to their age-matched controls. On the other hand, 9-month-old MSG mice were significantly heavier than their age-matched controls; however, their weight of total WAT was similar to the controls. Fasting glucose was only significantly increased in MSG mice at the age of 2 months. The levels of insulin and leptin were significantly increased in 2- and 6-month-old MSG mice compared to their age-matched controls. Only leptin, but not insulin level, was higher in 9-month-old MSG mice compared to the age-matched controls (Table 1). Even though MSG mice at the age of 2 and 6 months already revealed significant fat accumulation resulting in increased leptin level and conditions typical for pre-diabetes with a mild increase in blood glucose and increased insulin levels, at the age of 9 months their glucose and insulin levels did not significantly differ from their age-matched controls. Interestingly, plasma levels of free fatty acids and triglycerides did not differ significantly between MSG mice and their age-matched controls at 2, 6 or 9 months. In an intraperitoneal glucose tolerance test (IPGTT) performed in 6-month-old MSG mice and their age-matched NMRI controls, neither the glucose levels nor the area under the curve (AUC) exhibited any significant difference between MSG and control groups (data not shown).

Liver and adipose tissue of mRNA expressions in 6-month-old MSG mice

Obesity and insulin resistance impact changes in mRNA expressions of enzymes related to lipid metabolism in subcutaneous (SCAT) and visceral adipose tissue (IPAT) and liver. In adipose tissue the expression of leptin, the main adipokine produced by adipose tissue, is generally directly proportional to the mass of tissue. In MSG mice, a significantly increased leptin mRNA expression in SCAT (Fig. 1) but not IPAT was detected compared to controls. As the SCAT mass was double the IPAT mass in MSG mice, an increased circulating leptin level in MSG mice resulted then pre-eminently from the increased mRNA expression in SCAT.

Lipogenesis occurs in liver and adipose tissue, both SCAT and IPAT. In MSG mice, mRNA expression of ACACA, the rate-limiting step of fatty acid synthesis, was quite opposite in liver and fat. The ACACA mRNA expression was

Table 1 Metabolic parameters of 2-, 6-, and 9-month-old MSG mice and their age-matched controls

Mice	Weight [g]	White adipose tissue [% weight]	Glucose [mmol/l]	Insulin [ng/ml]	Leptin [ng/ml]
Control 2 months	40.38±0.94	4.19±0.42	6.68±0.44	0.98±0,15	1.92±0.41
MSG 2 months	42.42±0,56	11.53±0.62 ***	8.47±0.38 **	3.61±0,59 ***	28.75±3.95 ***
control 6 months	46. 53±1.08	4.48±0.37	6.55±0.39	0.97±0,20	4.02±0.56
MSG 6 months	53.08±2.11 *	9.08±0.89 ***	6.48±0.42	1.57±0,19 *	23.10±4.06 ***
control 9 months	56.54±1.27	7.22±0.60	6.87±0.35	2.44±0.63	16. 65±2.50
MSG 9 months	61.53±1.89 *	6.10±0.62	6.67±0.73	3.96±0.41	33.25±4.69 **

Data are mean±SEM, $n=10$ animals per group. Significance is *$p<0.05$, **$p<0.01$ and ***$p<0.001$ vs. corresponding control group using unpaired *t* test

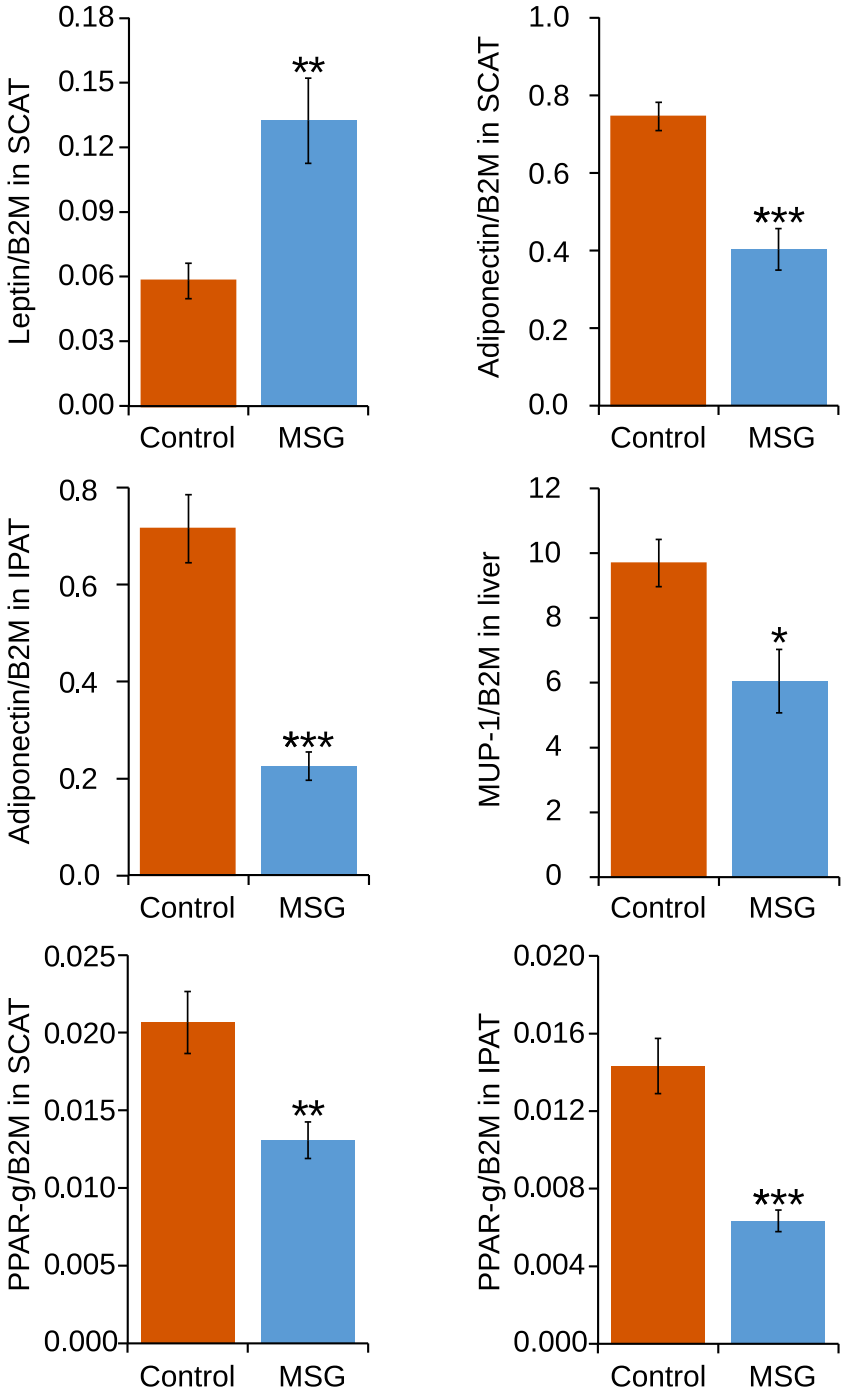


Fig. 1 Selected mRNA expressions for both control and MSG groups at 6 months of age: leptin in SCAT, adiponectin in SCAT and IPAT, MUP1 in liver, PPAR-γ in SCAT and IPAT (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$)

significantly enhanced in liver, but significantly attenuated in both SCAT and IPAT in MSG mice compared to controls (Fig. S1 in the Electronic Supplementary Material). The mRNA expression of FASN, the next enzyme catalyzing fatty acid synthesis, was analogously decreased in both SCAT and IPAT (Fig. S1), but was similar in liver of MSG mice regarding controls. Adipose tissue is the site of both lipogenesis and lipolysis. Lipolysis in adipose tissue was assessed by the mRNA expression of FABP-4 that mediates fatty acid uptake, transport and metabolism in cytoplasm of the adipocytes and by LPL mRNA expression, which were both lower in both SCAT and IPAT of MSG mice compared to controls. As mRNA expressions of both lipogenetic enzymes ACACA and FASN and both lipolytic substances FABP-4 and LPL in SCAT and IPAT were lower in MSG mice compared to controls (Fig. 1), it seems that lipid metabolism in adipose tissue of MSG mice was generally attenuated. mRNA expressions of PPAR-α in SCAT and IPAT were also lower in MSG mice compared to controls. No difference in BAT mRNA UCP-1 indicates a similar thermogenesis in both MSG and control mice.

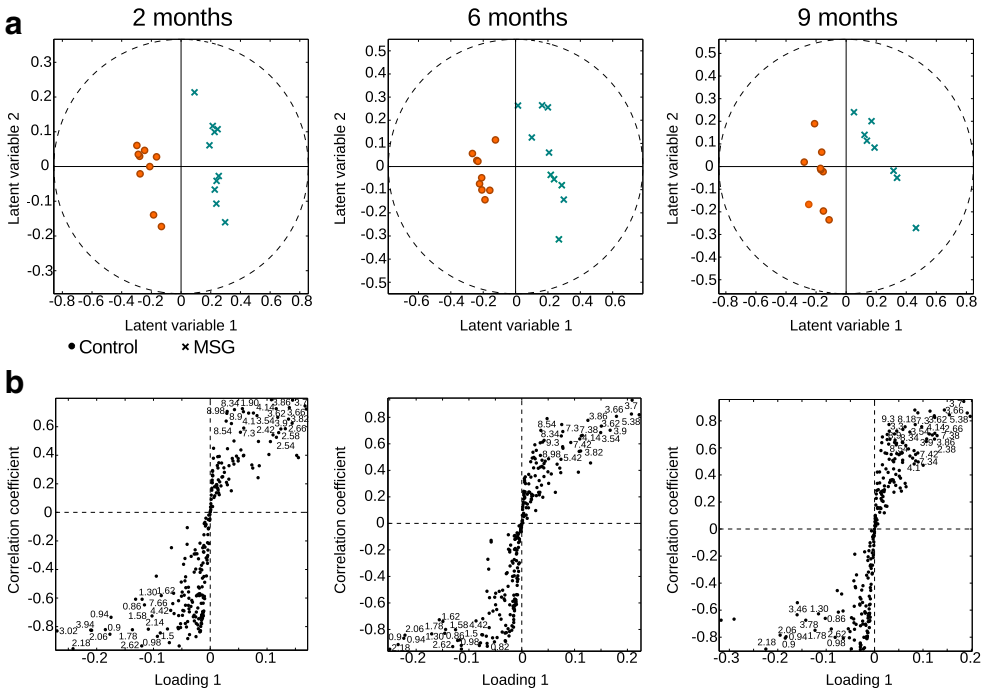


Fig. 2 Scores plots (**a**) and corresponding s-plots (**b**) of PLS-DA models for 2-, 6- and 9-month old mice

A significantly decreased mRNA expression of adiponectin, a marker inversely proportional to a diabetes state, in both subcutaneous and visceral fat of MSG mice (Fig. 1) pointed to a diabetes state.

In this study, the NNMT mRNA expression was not enhanced in SCAT, IPAT, and BAT. NNMT expression in liver (using B2M, GADPH, and GUSB as compensation) reveals an insignificant increase for MSG mice compared to controls

Table 2 Relative changes of significant metabolites in 2-, 6- and 9-month-old MSG mice compared to their age-matched controls (Student's *t* test was used to indicate the significance of change: *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$)

Metabolite	δ_H [a]	2 months	6 months	9 months
1-Methylnicotinamide (MNA)	9.29 (s)	↑↑***	↑*	↑**
Trigonelline	9.13 (s)	↓↓**	↓*	↓*
Nicotinamide *N*-oxide	8.76 (m)	↓**	↓**	↓*
N-methyl-4-pyridone-3-carboxamide (4-PY)	8.55 (d)	↑↑***	↑↑**	↑↑**
N-methyl-2-pyridone-5-carboxamide (2-PY)	8.34 (d)	↑↑***	↑*	↑**
Phenylacetylglycine (PAG)	7.43 (m)	↑	↑**	↑**
Allantoin	5.39 (s)	↑	↑***	↑***
Creatine	3.94 (s)	↓***	x	x
Citrate	2.68 (d)	↑↑**	x	x
Methylamine	2.61 (s)	↓***	↓***	↓***
Succinate	2.41 (s)	↑*	x	x
N-isovalerylglycine (IVG)	2.18 (d)	↓↓***	↓↓***	↓↓***
Unassigned acetyls	2.06 (s)	↓***	↓***	↓***
Acetate	1.93 (s)	↑↑*	x	x
Putrescine	1.78 (m)	↓***	↓***	↓***
Unassigned acids	0.90 (m)	↓***	↓***	↓**
Broad protein resonances	0.74 (br m)	↓↓***	↓↓***	↓↓***

[a] Representative bin used for the quantitation

↑ or ↓Relative change higher than±20 %

↑↑ or ↓↓Relative change higher than±40 %

x relative change less than±20 %

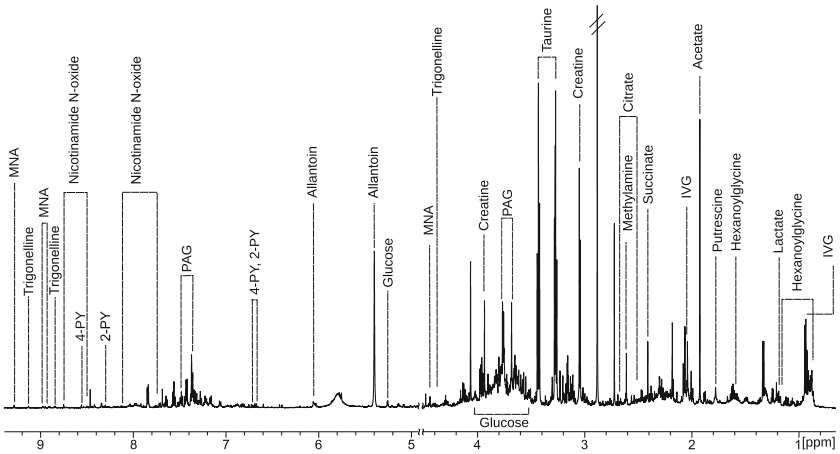


Fig. 3 Representative CPMG spectrum of a 2-month-old MSG mouse with assignment of particular signals of important metabolites

as a common trend in all three experiments. Nevertheless, it indicates correlation with an increased mRNA expression of lipogenetic enzymes ACACA and FASN in liver, but not in fat tissues of MSG mice compared to controls.

The mRNA liver expression of another diabetic marker—major urinary protein-1 (MUP1) [32, 33]—in MSG mice was found by one-third less than in controls (Fig. 1). It again points to a diabetic state at normoglycemic levels in 6-month-old MSG mice.

NMR-based metabolomics

As the impact of obesity and insulin resistance on the metabolism is not fully understood, a non-targeted analysis of metabolic changes and thus a determination of potential biomarkers is beneficial. We therefore performed NMR-based metabolomics of urine to follow changes in the complex set of end products of metabolism. ^{1}H NMR spectra of urine samples from MSG-treated (MSG) and age-matched untreated (control) mice were analyzed after 2, 6, and 9 months of age. Visual inspection of 1D-NOESY spectra revealed a huge amount of broad signals predominantly in the control group. Similar fact complicated evaluation of 1D-NOESY spectra in metabolomic analysis of *db/db* mice [34]. As the signals of small metabolites were distorted by the intense background of high molecular weight (HMW) compounds, we decided to perform additional CPMG experiments to attenuate broad signals (even though their residues were still detected) and use CPMG spectra for all subsequent analyses.

In concordance with the experiment's settings, three similar models were trained. All the PLS-DA models relate the class membership of mice to the MSG-treated group and control group to the spectra measured at the age of 2, 6, and 9 months. For each model, the spectra were mean-centered and Pareto scaled. The leave-one-out validation was carried out to assess the prediction quality of the model and to choose the number of latent variables according to the prediction accuracy criterion. The results of leave-one-out validation are presented in Table S1 in the Electronic Supplementary Material. The PLS-DA score plots for 2-, 6- and 9-month-old mice together with the corresponding S-plots are displayed in Fig. 2. The significant bins were identified as described in the experimental part; relative changes of their intensities are summarized in Table 2. Figure 3 displays representative CPMG spectrum of an MSG mouse with identification of important metabolites.

The following differences in metabolites relative to controls were identified for 2-month-old MSG mice: elevated 1-methylnicotinamide (MNA), *N*-methyl-4-pyridone-3-carboxamide (4-PY), *N*-methyl-2-pyridone-5-carboxamide (2-PY), citrate, succinate, acetate, and decreased trigonelline, nicotinamide *N*-oxide, creatine, methylamine, *N*-isovalerylglycine (IVG), putrescine, unassigned acetyls and aliphatic acids, HMW compounds.

The same approach applied for 6- and 9-month-old models revealed similar sets of changes in the following metabolites relative to age-matched controls: elevated MNA, 4-PY, 2-PY, PAG, allantoin, and decreased trigonelline, nicotinamide *N*-

Table 3 Relative changes of significant metabolites in MSG and control mice at 9 months of age expressed as a ratio of 9/2 months concentrations (Student's *t*-test was used to indicate the significance of change: $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$)

Metabolite	δ_H [a]	MSG	Controls
1-Methylnicotinamide (MNA)	9.29 (s)	x	$\uparrow\uparrow^{**}$
Creatine	3.94 (s)	x	$\downarrow^{***}$
Taurine	3.43 (t)	x	$\uparrow^{**}$
Citrate	2.68 (d)	$\downarrow\downarrow^{*}$	$\downarrow$
2-Oxoglutarate (2-OG)	2.45 (t)	$\downarrow^{*}$	$\downarrow^{**}$
Succinate	2.41 (s)	$\downarrow^{*}$	x
Broad protein resonances	0.74 (br m)	$\downarrow\downarrow^{**}$	$\downarrow^{***}$

[a] Representative bin used for the quantitation

$\uparrow$ or $\downarrow$ Relative change higher than ±20 %

$\uparrow\uparrow$ or $\downarrow\downarrow$ Relative change higher than ±40 %

x relative change less than ±20 %

oxide, creatine, methylamine, IVG, putrescine, unassigned acetyls and aliphatic acids, and HMW compounds.

To reveal the origin of HMW compounds responsible for broad signals in spectra, the protein determination by bicinchoninic acid was performed on selected samples of three MSG and three control animals at the age of 2, 6, and 9 months. These experiments confirmed a substantially reduced level of excreted proteins in urine of MSG-treated mice compared to their age-matched controls; the ratio of MSG/control protein concentration was 66 % at the age of 2 months, 70 % at 6 months, and 53 % at 9 months. The results of protein determination are summarized in Table S3 in the Electronic Supplementary Material. Subsequent native polyacrylamide gel electrophoresis (Fig. S2 in the Electronic Supplementary Material) also shows a reduction of protein amount in three MSG samples compared to three controls and the visualized trends of protein amount approximately correspond to the determined protein concentrations. The qualitative differences of MUP isoforms pattern were not found in the analyzed samples.

The impact of aging on metabolite concentration was also studied by the comparison of 9-month-old and 2-month-old animals in both the MSG and control groups. Two additional models relating proton NMR urine spectra to the age of mice for the MSG-treated group and control group was trained. The results of leave-one-out validation of aging models are presented in Table S2 in the Electronic Supplementary Material. PLS-DA score plots for the untreated (control) and MSG-treated (MSG) model are depicted in Figure S3 in the Electronic Supplementary Material.

The 9-month-old control group had significantly elevated levels of MNA with taurine and decreased concentration of 2-oxoglutarate (2-OG), proteins, and creatine compared to the same animals at 2 months of age. The analogous comparison of 9- and 2-month-old MSG mice displayed significantly reduced concentration of citrate, 2-OG, proteins, succinate and proteins for the older animals. Concentration differences of the above-mentioned metabolites depending on the age are displayed for both the control and MSG groups in Table 3.

Discussion

Obesity and its related complications represent a highly prevalent condition of complex metabolic derangements often resulting in long-term metabolic and cardiovascular complications and premature death. Here we used NMR-based metabolomics of urine of MSG mice—a model of obesity and insulin resistance—to identify novel metabolites associated with obesity phenotype. We have also explored the influence of aging on the above described parameters. The evaluation of biochemical and hormonal parameters confirmed that both 2- and 6-month-old MSG mice developed obesity with hyperleptinemia and hyperinsulinemia. This phenotype is closely similar to obese patients at a high risk of development of late complications of obesity, making this mouse model beneficial for the study of human obesity.

Since the development of obesity and insulin resistance is reflected in mRNA expressions of enzymes participating in lipid metabolism, the changes in mRNA expressions induced by MSG obesity in SCAT, IPAT, BAT, and liver were also explored. Impaired insulin sensitivity and metabolic derangements in 6-month-old mice were indicated by the significantly decreased mRNA expressions of adiponectin in both subcutaneous and visceral fat, suggesting the prediabetic state. Increased leptin mRNA expression together with higher levels of leptin in serum, observed in our study, were also reported in the MSG model of CD-1 mice [27], probably reflecting a leptin resistance as described in numerous model of obesity.

As mRNA expressions of both lipogenetic and lipolytic enzymes in SCAT and IPAT were lower in MSG mice compared to controls, it seems that lipid metabolism in adipose tissue of MSG mice was generally attenuated. This state seems to be explained by the fact that PPAR-γ mRNA expression was attenuated in both SCAT and IPAT and also in BAT of MSG mice compared to controls. PPAR-γ is considered a general transcription factor of all ACACA, FASN, FABP-4, and LPL, and of adiponectin, for which all of the mRNA transcriptions were attenuated in fat tissue of MSG mice.

After the characterization of the MSG model using biochemical parameters and mRNA expressions was defined, our work was focused on tracking changes in the metabolic composition of urine as detected by NMR metabolomics.

The first interesting finding of the NMR-based study was higher levels of MNA and its oxidation products 2-PY and 4-PY in urine of MSG obese mice relative to control mice. MNA and its metabolites primarily originate from liver and their increase in urine was demonstrated as a positive biomarker of peroxisome proliferation [35, 36], which is linked with inflammation, obesity, and metabolic syndrome [37, 38]. On the other hand, urine of MSG mice contained lower levels of nicotinamide *N*-oxide, the oxidation product of nicotinamide, and trigonelline (*N*-methylnicotinate), the methylation product of nicotinic acid. Our NMR metabolomic data therefore suggest that nicotinamide in MSG obese mice is rather subjected to methylation than oxidation. The increase in MNA and 2-PY and the decrease in trigonelline was reported in T2DM patients, *db/db* mice and *fa/fa* rats [34].

Nicotinamide is methylated by *S*-adenosylmethionine (SAM) under NNMT (nicotinamide *N*-methyltransferase) catalysis. SAM further provides propylamine for polyamine metabolism. Contrary to nicotinamide, the utilization of SAM for polyamine metabolism seems to be attenuated in MSG mice as we observed decreased levels of polyamine putrescine. Putrescine, the major urinary excretion product of male mice, can

be synthesized from ornithine by ornithine decarboxylase (ODC) and/or by polyamine oxidase from acetylspermidine. Our NMR data indicate substantially reduced excretion of putrescine in MSG mice, which could be attributable to the blunted catabolism of polyamines. Polyamines play an important role in glucose metabolism and control of adipose tissue expansion. The influence of spermidine/spermine *N*1-acetyltransferase (SSAT) on adipose tissue formation in mice was reported [39]. It was shown that SSAT overexpression in mice resulted in an increased metabolic rate, reduced fat and increased insulin sensitivity. Furthermore, reduced polyamine metabolism via the inhibition of ODC by 2-difluoromethylornithine elevated ATP in adipose tissue and had substantial influence on the SSAT mouse phenotype. Similarly, it was demonstrated that activated polyamine metabolism is associated with reduced WAT and decreased concentration of acetyl-CoA and malonyl-CoA in WAT [40]. Taken together, these data support the hypothesis that activated polyamine catabolism has a strong influence on white adipose tissue and glucose metabolism control. Our results in obese MSG mice indicate that the reduction of polyamine metabolism might be related to the accumulation of WAT and obesity development as well as to the disturbed endocrine function and insulin resistance in the adipose tissue as documented by increased leptin and decreased adiponectin mRNA expression.

In previous studies, adipose tissue NNMT expression correlated with adiposity in mice of various strains with diet-induced obesity [41]. NNMT expression was increased in WAT and liver of both genetically and diet-induced obese and diabetic mice and NNMT inhibition increased SAM and NAD^+ levels as well as ODC and SSAT expression and activity in adipose tissue and polyamine urine excretion [42]. It was suggested that inhibition of NNMT could serve as a protection against diet-induced obesity via increasing polyamine flux [42]. Here, we found a non-significant tendency to increased mRNA NNMT expression in liver but not in adipose tissue of MSG mice compared to controls. We suggest that this trend might be due to increased mRNA expression of lipogenetic enzymes in liver and an attenuated expression of not only lipogenetic and lipolytic enzymes in WAT, but also of PPAR-γ, the general regulator of their transcription.

Another interesting fact, observed in NMR experiments and also confirmed by determination by bicinchonic acid, was a high amount of proteins in urine of control but not MSG mice. In contrast to human urine which is under normal conditions free of proteins, normal mouse urine contains a large amount of proteins belonging to a family of major urinary proteins (MUPs). MUPs are mainly synthesized in the liver, circulate in the bloodstream and are easily excreted into urine owing to their low molecular weight [33]. Their role in the organism is still not fully elucidated. It is known that MUPs bind volatile pheromone compounds and thus participate in chemical communication as urine scent marks [43]. The impact of MUPs on energy metabolism in mice was intensively studied in mouse models of genetic and high-fat (HF) diet-induced obesity and diabetes [32, 44]. The expression of MUP1 and its circulating concentrations were substantially reduced in *db/db* mice and dietary obese mice compared with their lean controls. Overexpression of MUP1 significantly attenuated hyperglycemia and glucose intolerance in both models [32]. MUP1 administration in *db/db* mice increased energy expenditure and improved glucose tolerance [44]. Additionally, blood and urine levels of MUP5 were regulated by dietary restriction of mice [45]. Markedly lowered liver mRNA expression and levels of excreted urinary proteins observed in MSG mice in our study are in good agreement with the above-mentioned results and can be attributed to obesity and insulin resistance resulting from MSG obesity. Collectively, these data suggest that altered MUPs levels in mice may play a role in the regulation of glucose metabolism and energy homeostasis.

Increased urine levels of phenylacetylglycine (PAG) and decreased urinary levels of methylamine (MA) were also observed in MSG-treated mice. Both of these metabolites are affected by gut microbiota that has been previously described to be markedly altered in obesity/T2DM [46]. The levels of PAG and MA were lowered in wild-type mice on an HF diet compared to a standard diet during the study of resistance to diet-induced obesity in AHNAK knockout mice [8]. On the other hand, elevated concentrations of urinary PAG were observed in rat model of HF diet-induced obesity [47]. It was speculated that overproduction of PAG in animals with high body weight gain is related to the obesity via gut bacteria. Since the gut microbiota has not been explored in our study, the exact mechanism for increased PAG levels in MSG mice remains to be further elucidated.

Increased levels of acetate, citrate, and succinate, markers of Krebs cycle, in the urine of MSG mice were most probably the result of metabolic distress under anaerobic metabolism. Interestingly, these changes were most pronounced in the 2-month-old MSG mice, while they were less apparent in older animals (Fig. S4 in the Electronic Supplementary Material). Reduced concentration of IVG in MSG obese mice are in accordance with decrease of IVG observed in human model of obesity [48].

Moreover, increased levels of allantoin in urine of MSG mice were observed, which is in agreement with investigation in T2DM patients [49]. Studies have linked allantoin as a biomarker of oxidative stress [50] resulting often in inflammation [51]. Thus, the increased allantoin levels in MSG mice may indicate increased oxidative stress and subclinical inflammation resulting from metabolic derangements [52]. However, allantoin in mice can be produced not only by chemical oxidation of uric acid but also via enzymatic reaction with urate oxidase [53] and therefore the mechanism responsible

for increased allantoin level in MSG mice is less clear. Allantoin levels in urine were elevated in mouse model of Alzheimer's disease [54] and in captopril-treated atherosclerotic mice [55]. Furthermore, allantoin was also described as a biomarker of gut inflammation in a mouse model of Crohn's disease [56]. In another study, allantoin was decreased in genetically obese and diabetic *db/db* mice [34]. The significance of allantoin was explored in HF diet-fed mice [57]. Interestingly, the chronic administration of allantoin significantly reduced the body weight of obese mice and lowered HF diet-induced hyperleptinemia suggesting that it plays a complex role in the regulation of an energy homeostasis.

Increased level of glucose in MSG mice urine reached the normal value of the control group at 9 months of age. Although this change is not significant due to the large variance of MSG group data, detected trend correlates well with the normalization of plasma glucose at 6 months.

In addition to the influence of obesity phenotype, our study has also revealed the age-dependent changes of individual metabolites in both control and MSG groups. Reduced concentrations of MUPs, citrate and 2-oxoglutarate were detected in both control and MSG groups with aging. While the increase of MNA, taurine and decrease of creatine were observed in control animals, the levels of these metabolites remained almost unaltered in MSG group. On the contrary, the reduction of succinate was only detected in MSG mice, but not in controls. A significantly higher level of MNA also discriminated old male BALB/c mice in the study of aging [58]. Decreased citrate and 2-OG contributed to characterization of older diabetic mice and rats, with these trends not detected in the control group [34]. Figure S4 clearly shows the development of individual metabolite levels during aging for both MSG and control mice. It is obvious that the initial difference in concentrations of creatine, citrate, succinate, and acetate between the MSG and control groups, which we detected at 2 months of age, substantially diminished in 6- and 9-month-old animals.

Conclusions

In conclusion, our study has revealed numerous differences in the urine metabolome of obese MSG mice relative to control animals with a further distinct influence of aging in both groups.

Detected alterations in nicotinamide metabolism and MUP excretion were also supported by corresponding trends in mRNA expression of NNMT and MUP1. The importance of these changes and its applicability to human pathophysiology of obesity and its complications needs further elucidation. Several biochemical parameters, indicating the development of obesity and insulin resistance in MSG mice, were normalized at the age of 9 months. This fact pointed to a limited predictive value of biochemical parameters, as NMR metabolomics confirmed the altered urine metabolic composition of MSG mice even at 9 months.

Acknowledgments We are indebted to H. Vysušilová for her excellent technical assistance.

Compliance with ethical standards

Funding The research was financially supported by the Czech Science Foundation (Grant No. GA13-14105S), grant for long-term conceptual development of the Institute of Microbiology (RVO: 61388971) and Institute of Organic Chemistry and Biochemistry (RVO: 61388963). The project was conducted within the "Prague Infrastructure for Structure Biology and Metabolomics" which has been built up by financial support of the Operational Program Prague – Competitiveness (Project No.: CZ.2.16/3.1.00/24023). The data analyses were partially supported by the Grant Agency of the Czech Technical University in Prague No. SGS13/203/OHK3/3T/13. We would also like to acknowledge project LO1509 of the Ministry of Education, Youth and Sports of the Czech Republic.

Conflict of interest Helena Pelantová, Simona Bártová, Jiří Anýž, Martina Holubová, Blanka Železná, Lenka Maletínská, Daniel Novák, Zdena Lacinová, Miroslav Šulc, Martin Haluzík, and Marek Kuzma declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted.

References

1. Qatanani M, Lazar MA (2007) Mechanisms of obesity-associated insulin resistance: many choices on the menu. Genes Dev 21:1443–1455. doi:10.1101/gad.1550907
2. Chen P, Liu J (2007) Metabonomics and diabetes mellitus. Adv Ther 24:1036–1045. doi:10.1007/BF02877709
3. Faber JH, Malmodin D, Toft H, Maher AD, Crockford D, Holmes E, Nicholson JK, Dumas ME, Baunsgaard D (2007) Metabonomics in diabetes research. J Diabetes Sci Technol 1:549–557
4. Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB (2009) Metabolomics applied to diabetes research: moving from information to knowledge. Diabetes 58:2429–2443. doi:10.2337/db09-0580
5. Lutz TA, Woods SC (2012) Overview of animal models of obesity. Curr Protoc Pharmacol Chapter 5:Unit5.61. doi:10.1002/0471141755.ph0561s58
6. Carroll L, Voisey J, van Daal A (2004) Mouse models of obesity. Clin Dermatol 22:345–349. doi:10.1016/j.clindermatol.2004.01.004
7. Katsuda Y, Ohta T, Shinohara M, Bin T, Yamada T (2013) Diabetic mouse models. Open J Anim Sci 03:334–342. doi:10.4236/ojas.2013.34050
8. Kim IY, Jung J, Jang M, Ahn YG, Shin JH, Choi JW, Sohn MR, Shin SM, Kang D-G, Lee H-S, Bae YS, Ryu DH, Seong JK, Hwang G-S (2010) 1H NMR-based metabolomic study on resistance to diet-induced obesity in AHNAK knock-out mice. Biochem Biophys Res Commun 403:428–434. doi:10.1016/j.bbrc.2010.11.048
9. Boulangé CL, Claus SP, Chou CJ, Collino S, Montoliu I, Kochhar S, Holmes E, Rezzi S, Nicholson JK, Dumas ME, Martin F-PJ

(2013) Early metabolic adaptation in C57BL/6 mice resistant to high fat diet induced weight gain involves an activation of mitochondrial oxidative pathways. J Proteome Res 12:1956–1968. doi:10.1021/pr400051s

10. Shearer J, Duggan G, Weljie A, Hittel DS, Wasserman DH, Vogel HJ (2008) Metabolomic profiling of dietary-induced insulin resistance in the high fat-fed C57BL/6J mouse. Diabetes Obes Metab 10:950–958. doi:10.1111/j.1463-1326.2007.00837.x
11. Duggan GE, Hittel DS, Sensen CW, Weljie AM, Vogel HJ, Shearer J (2011) Metabolomic response to exercise training in lean and diet-induced obese mice. J Appl Physiol 110:1311–1318. doi:10.1152/japplphysiol.00701.2010
12. Jung J, Kim IY, Kim YN, Kim J, Shin JH, Jang Z, Lee H-S, Hwang G-S, Seong JK (2012) H-1 NMR-based metabolite profiling of diet-induced obesity in a mouse mode. BMB Rep 45:419–424. doi:10.5483/BMBRep.2012.45.7.248
13. Graham ML, Janecek JL, Kittredge JA, Hering BJ, Schuurman H-J (2011) The streptozotocin-induced diabetic nude mouse model: differences between animals from different sources. Comp Med 61:356–360
14. Bunyan J, Murrell EA, Shah PP (2008) The induction of obesity in rodents by means of monosodium glutamate. Br J Nutr 35:25. doi:10.1079/BJN19760005
15. Nagata M, Suzuki W, Iizuka S, Tabuchi M, Maruyama H, Takeda S, Aburada M, Miyamoto K (2006) Type 2 diabetes mellitus in obese mouse model induced by monosodium glutamate. Exp Anim 55:109–115
16. Olney JW (1969) Brain lesions, obesity, and other disturbances in mice treated with monosodium glutamate. Science 164:719–721. doi:10.1126/science.164.3880.719
17. Broberger C, De Lecea L, Sutcliffe JG, Hökfelt T (1998) Hypocretin/orexin- and melanin-concentrating hormone-expressing cells form distinct populations in the rodent lateral hypothalamus: relationship to the neuropeptide Y and agouti gene-related protein systems. J Comp Neurol 402:460–474. doi:10.1002/(SICI)1096-9861(19981228)402:4<460::AID-CNE3>3.0.CO;2-S
18. Elefteriou F, Takeda S, Liu X, Armstrong D, Karsenty G (2003) Monosodium glutamate-sensitive hypothalamic neurons contribute to the control of bone mass. Endocrinology 144:3842–3847. doi:10.1210/en.2003-0369
19. Zelezná B, Maixnerová J, Matysková R, Haugvicová R, Blokesová D, Maletínská L (2009) Anorexigenic effect of cholecystokinin is lost but that of CART (Cocaine and Amphetamine Regulated Transcript) peptide is preserved in monosodium glutamate obese mice. Physiol Res 58:717–723
20. Edelstein K, Pfaus JG, Rusak B, Amir S (1995) Neonatal monosodium glutamate treatment prevents effects of constant light on circadian temperature rhythms of adult rats. Brain Res 675:135–142. doi:10.1016/0006-8993(95)00048-U
21. Mistlberger RE, Antle MC (1999) Neonatal monosodium glutamate alters circadian organization of feeding, food anticipatory activity and photic masking in the rat. Brain Res 842:73–83. doi:10.1016/S0006-8993(99)01836-3
22. Djazayery A, Miller DS, Stock MJ (1979) Energy balances in obese mice. Ann Nutr Metab 23:357–367. doi:10.1159/000176281
23. Maletínská L, Toma RS, Pirnik Z, Kiss A, Slaninová J, Haluzík M, Zelezná B (2006) Effect of cholecystokinin on feeding is attenuated in monosodium glutamate obese mice. Regul Pept 136:58–63. doi:10.1016/j.regpep.2006.04.020
24. Remke H, Wilsdorf A, Müller F (1988) Development of hypothalamic obesity in growing rats. Exp Pathol 33:223–232. doi:10.1016/S0232-1513(88)80076-8
25. Matyšková R, Maletínská L, Maixnerová J, Pirník Z, Kiss A, Železná B (2008) Comparison of the obesity phenotypes related to monosodium glutamate effect on arcuate nucleus and/or the high fat diet feeding in C57BL/6 and NMRI mice. Physiol Res 57:727–734
26. Hernández-Bautista RJ, Alarcón-Aguilar FJ, Del C, Escobar-Villanueva M, Almanza-Pérez JC, Merino-Aguilar H, Fainstein MK, López-Diazguerrero NE (2014) Biochemical alterations during the obese-aging process in female and male monosodium glutamate (MSG)-treated mice. Int J Mol Sci 15:11473–11494. doi:10.3390/ijms150711473
27. Roman-Ramos R, Almanza-Perez JC, Garcia-Macedo R, Blancas-Flores G, Fortis-Barrera A, Jasso EI, Garcia-Lorenzana M, Campos-Sepulveda AE, Cruz M, Alarcon-Aguilar FJ (2011) Monosodium glutamate neonatal intoxication associated with obesity in adult stage is characterized by chronic inflammation and increased mRNA expression of peroxisome proliferator-activated receptors in mice. Basic Clin Pharmacol Toxicol 108:406–413. doi:10.1111/j.1742-7843.2011.00671.x
28. Maletínská L, Matyšková R, Maixnerová J, Sýkora D, Pýchová M, Spolcová A, Blechová M, Drápalová J, Lacinová Z, Haluzík M, Zelezná B (2011) The Peptidic GHS-R antagonist [D-Lys(3)]GHRP-6 markedly improves adiposity and related metabolic abnormalities in a mouse model of postmenopausal obesity. Mol Cell Endocrinol 343:55–62. doi:10.1016/j.mce.2011.06.006
29. Wiechelman KJ, Braun RD, Fitzpatrick JD (1988) Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation. Anal Biochem 175:231–237
30. (2013) MATLAB version 8.1.0.604 (R2013a). Natick, Massachusetts: The MathWorks Inc
31. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130. doi:10.1016/S0169-7439(01)00155-1
32. Zhou Y, Jiang L, Rui L (2009) Identification of MUP1 as a regulator for glucose and lipid metabolism in mice. J Biol Chem 284:11152–11159. doi:10.1074/jbc.M900754200
33. Zhou Y, Rui L (2010) Major urinary protein regulation of chemical communication and nutrient metabolism. In: Litwack G (ed) Vitamins and hormones 1st ed., Vol. 83. Elsevier, Amsterdam, pp 151–163. doi:10.1016/S0083-6729(10)83006-7
34. Salek RM, Maguire ML, Bentley E, Rubtsov DV, Hough T, Cheeseman M, Nunez D, Sweatman BC, Haselden JN, Cox RD, Connor SC, Griffin JL (2007) A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. Physiol Genomics 29:99–108. doi:10.1152/physiolgenomics.00194.2006
35. Ringeissen S, Connor SC, Brown HR, Sweatman BC, Hodson MP, Kenny SP, Haworth RI, McGill P, Price MA, Aylott MC, Nunez DJ, Haselden JN, Waterfield CJ (2003) Potential urinary and plasma biomarkers of peroxisome proliferation in the rat: identification of N-methylnicotinamide and N-methyl-4-pyridone-3-carboxamide by 1H nuclear magnetic resonance and high performance liquid chromatography. Biomarkers 8:240–271. doi:10.1080/1354750031000149124
36. Delaney J, Hodson MP, Thakkar H, Connor SC, Sweatman BC, Kenny SP, McGill PJ, Holder JC, Hutton KA, Haselden JN, Waterfield CJ (2005) Tryptophan-NAD+ pathway metabolites as putative biomarkers and predictors of peroxisome proliferation. Arch Toxicol 79:208–223. doi:10.1007/s00204-004-0625-5
37. Kersten S (2002) Peroxisome proliferator activated receptors and obesity. Eur J Pharmacol 440:223–234. doi:10.1016/S0014-2999(02)01431-0
38. Stienstra R, Mandard S, Patsouris D, Maass C, Kersten S, Müller M (2007) Peroxisome proliferator-activated receptor alpha protects against obesity-induced hepatic inflammation. Endocrinology 148:2753–2763. doi:10.1210/en.2007-0014
39. Pirinen E, Kuulasmaa T, Pietilä M, Heikkinen S, Tusa M, Itkonen P, Boman S, Skommer J, Virkamäki A, Hohtola E, Kettunen M, Fatrai S, Kansanen E, Koota S, Niiranen K, Parkkinen J, Levonen A-L, Ylä-Herttuala S, Hiltunen JK, Alhonen L, Smith U, Jänne J, Laakso

M (2007) Enhanced polyamine catabolism alters homeostatic control of white adipose tissue mass, energy expenditure, and glucose metabolism. Mol Cell Biol 27:4953–4967. doi:10.1128/MCB.02034-06

40. Jell J, Merali S, Hensen ML, Mazurchuk R, Spernyak JA, Diegelman P, Kisiel ND, Barrero C, Deeb KK, Alhonen L, Patel MS, Porter CW (2007) Genetically altered expression of spermidine/spermine N1-acetyltransferase affects fat metabolism in mice via acetyl-CoA. J Biol Chem 282:8404–8413. doi:10.1074/jbc.M610265200
41. Grubb SC, Maddatu TP, Bult CJ, Bogue MA (2009) Mouse phenome database. Nucleic Acids Res 37:D720–D730. doi:10.1093/nar/gkn778
42. Kraus D, Yang Q, Kong D, Banks AS, Zhang L, Rodgers JT, Pirinen E, Pulinilkunnil TC, Gong F, Wang Y, Cen Y, Sauve AA, Asara JM, Peroni OD, Monia BP, Bhanot S, Alhonen L, Puigserver P, Kahn BB (2014) Nicotinamide N-methyltransferase knockdown protects against diet-induced obesity. Nature 508:258–262. doi:10.1038/nature13198
43. Beynon RJ, Hurst JL (2004) Urinary proteins and the modulation of chemical scents in mice and rats. Peptides 25:1553–1563. doi:10.1016/j.peptides.2003.12.025
44. Hui X, Zhu W, Wang Y, Lam KSL, Zhang J, Wu D, Kraegen EW, Li Y, Xu A (2009) Major urinary protein-1 increases energy expenditure and improves glucose intolerance through enhancing mitochondrial function in skeletal muscle of diabetic mice. J Biol Chem 284:14050–14057. doi:10.1074/jbc.M109.001107
45. Giller K, Huebbe P, Doering F, Pallauf K, Rimbach G (2013) Major urinary protein 5, a scent communication protein, is regulated by dietary restriction and subsequent re-feeding in mice. Proc Biol Sci 280:20130101. doi:10.1098/rspb.2013.0101
46. Holmes E, Li JV, Athanasiou T, Ashrafian H, Nicholson JK (2011) Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. Trends Microbiol 19:349–359. doi:10.1016/j.tim.2011.05.006
47. Kim S-H, Yang S-O, Kim H-S, Kim Y, Park T, Choi H-K (2009) 1H-nuclear magnetic resonance spectroscopy-based metabolic assessment in a rat model of obesity induced by a high-fat diet. Anal Bioanal Chem 395:1117–1124. doi:10.1007/s00216-009-3054-8
48. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, Gallup D, Ilkayeva O, Wenner BR, Yancy WE, Musante G, Surwit R, Millington DS, Butler MD, Laura P (2013) A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab 9:311–326. doi:10.1016/j.cmet.2009.02.002.A
49. Chung W-Y, Benzie IFF (2013) Plasma allantoin measurement by isocratic liquid chromatography with tandem mass spectrometry: method evaluation and application in oxidative stress biomonitoring. Clin Chim Acta 424:237–244. doi:10.1016/j.cca.2013.06.015
50. Kand'ár R, Záková P (2008) Allantoin as a marker of oxidative stress in human erythrocytes. Clin Chem Lab Med 46:1270–1274. doi:10.1515/CCLM.2008.244
51. Yardim-Akaydin S, Sepici A, Ozkan Y, Simşek B, Sepici V. Evaluation of allantoin levels as a new marker of oxidative stress in Behçet's disease. Scand J Rheumatol 35:61–64. doi:10.1080/03009740510026878
52. Esser N, Legrand-Poels S, Piette J, Scheen AJ, Paquot N (2014) Inflammation as a link between obesity, metabolic syndrome and type 2 diabetes. Diabetes Res Clin Pract. doi:10.1016/j.diabres.2014.04.006
53. Serkova N, Fuller TF, Klawitter J, Freise CE, Niemann CU (2005) H-NMR-based metabolic signatures of mild and severe ischemia/reperfusion injury in rat kidney transplants. Kidney Int 67:1142–1151. doi:10.1111/j.1523-1755.2005.00181.x
54. Fukuhara K, Ohno A, Ota Y, Senoo Y, Maekawa K, Okuda H, Kurihara M, Okuno A, Niida S, Saito Y, Takikawa O (2013) NMR-based metabolomics of urine in a mouse model of Alzheimer's disease: identification of oxidative stress biomarkers. J Clin Biochem Nutr 52:133–138. doi:10.3164/jcbn.12-118
55. Leo GC, Darrow AL (2009) NMR-based metabolomics of urine for the atherosclerotic mouse model using apolipoprotein-E deficient mice. Magn Reson Chem 47(Suppl 1):S20–S25. doi:10.1002/mrc.2470
56. Dryland PA, Love DR, Walker MF, Dommels Y, Rowan D, Roy NC, Helsby N, Browning BL, Zhu S, Copp BR, Ferguson LR (2008) Allantoin as a biomarker of inflammation in an inflammatory bowel disease mouse model: NMR analysis of urine. Open Bioact Compd J 1:1–6
57. Chung H-H, Lee KS, Cheng J-T (2013) Decrease of obesity by allantoin via imidazoline I 1 -receptor activation in high fat diet-fed mice. Evid Based Complement Alternat Med 2013:589309. doi:10.1155/2013/589309
58. Calvani R, Brasili E, Praticò G, Capuani G, Tomassini A, Marini F, Sciubba F, Finamore A, Roselli M, Marzetti E, Miccheli A (2014) Fecal and urinary NMR-based metabolomics unveil an aging signature in mice. Exp Gerontol 49:5–11. doi:10.1016/j.exger.2013.10.010

Journal of Pharmaceutical and Biomedical Analysis 115 (2015) 225–235

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Strategy for NMR metabolomic analysis of urine in mouse models of obesity— from sample collection to interpretation of acquired data

Helena Pelantová[a,b], Martina Bugáňová[a,f], Jiří Anýž[c], Blanka Železná[d], Lenka Maletínská[d], Daniel Novák[c], Martin Haluzík[e], Marek Kuzma[a,*]

[a] Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20 Prague 4, Czech Republic
[b] Department of Analytical Chemistry, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic
[c] Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Prague 6, Czech Republic
[d] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic
[e] 3rd Medical Department, 1st Faculty of Medicine, Charles University and General Faculty Hospital in Prague, U nemocnice 1, 128 08 Prague 2, Czech Republic
[f] Faculty of Chemical Technology, University of Chemistry and Technology Prague, Technická 5, 166 28 Prague 6, Czech Republic

ARTICLE INFO

Article history:
Received 14 April 2015
Received in revised form 26 June 2015
Accepted 29 June 2015
Available online 17 July 2015

Keywords:
NMR metabolomics
Mouse
Obesity
Urine

ABSTRACT

The mouse model of monosodium glutamate induced obesity was used to examine and consequently optimize the strategy for analysis of urine samples by NMR spectroscopy. A set of nineteen easily detectable metabolites typical in obesity-related studies was selected. The impact of urine collection protocol, choice of ^{1}H NMR pulse sequence, and finally the impact of the normalization method on the detected concentration of selected metabolites were investigated. We demonstrated the crucial effect of food intake and diurnal rhythms resulting in the choice of a 24-hour fasting collection protocol as the most convenient for tracking obesity-induced increased sensitivity to fasting. It was shown that the Carr-Purcell-Meiboom-Gill (CPMG) experiment is a better alternative to one-dimensional nuclear Overhauser enhancement spectroscopy (1D-NOESY) for NMR analysis of mouse urine due to its ability to filter undesirable signals of proteins naturally present in rodent urine. Normalization to total spectral area provided comparable outcomes as did normalization to creatinine or probabilistic quotient normalization in the CPMG-based model. The optimized approach was found to be beneficial mainly for low abundant metabolites rarely monitored due to their overlap by strong protein signals.

1. Introduction

Type 2 diabetes mellitus (T2DM) arises as a consequence of pathologically damaged utilization of glucose with substantial impact on other biochemical processes. Evidently, its incidence in recent years on a worldwide scale has had strong growing tendency [1,2]. Since metabolic pathways throughout the body are mutually interconnected, the course of diabetes is frequently coupled with visceral fat accumulation and obesity resulting in an increased risk of simultaneous cardiovascular diseases [3,4]. Metabolomics can not only help to understand the biochemical signature of diabetes and concomitant problems but can also contribute to early diagnostics development [5,6].

Rodents, particularly mice, have been extensively reported as commonly used models for metabolomics studies [7–11]. Referring to the diabetic issue, special attention has to be devoted to mouse models of obesity including genetically treated individuals as well as mice with chemically or diet induced obesity [12–16]. The investigation of biofluids, mainly urine and blood plasma, in obese models allows the qualitative and quantitative determination of metabolites, the presence and dynamic fluctuations of which are directly linked to the degree of metabolic damage, and thus may serve as an early disease marker. Urine is used with the advantage of an extremely diverse biological mixture loaded with information reflecting the current physiological stage of living organisms at the defined time point. It contains a huge set of structurally different metabolites and other water-soluble compounds, the concentrations of which vary with genetic and environmental conditions, e.g. drugs or food intake. Additionally, it can even be easily obtained using non-invasive

* Corresponding author.

http://dx.doi.org/10.1016/j.jpba.2015.06.036
0731-7085/

techniques, typically by 24-hour collection in metabolic cages for rodent models.

NMR spectroscopy belongs to the most used analytical platforms employed in metabolomics, giving a large amount of reproducible information on chemical substances of various structures [17–19]. Several pulse sequences are widely recommended for NMR-based applications; detailed experimental protocols were prepared depending on the origin of the sample [20,21]. 1D-NOESY with presaturation became the almost exclusive method for urine samples, bringing such benefits as a good quality of solvent suppression, robustness and easy automation [22]. Apparently, it is the most dominant pulse sequence in the field of urine-based metabolomics. The presence of high molecular weight (HMW) compounds is not expected in human urine and proteinuria is considered a signal of renal damage. On the contrary the occurrence of urinary proteins in rodents is absolutely natural. These mouse proteins, commonly referred to as major urinary proteins (MUPs), are carriers of the important signaling roles and participate in the regulation of glucose and lipid metabolism [23]. Protein signals cover the whole range of the measured spectrum and can interfere with signals of interest, preventing their correct quantitation. The CPMG pulse sequence [24,25], editing signal intensities on the basis of relaxation time, can be the method of choice in such case. Unwanted protein signals are suppressed due to their different relaxation time, thus there is no need for sample pre-processing. This experiment is well established for plasma and serum samples in metabolomics, but it is almost always omitted in NMR-based mouse urine study.

The *J*-resolved experiment (JRES) [26], in which splitting of signals due to *J*-coupling is separated from the effect of chemical shift, is the only routinely performed 2D technique in metabolomic analysis [27]. The problem with the signal overlap may be partially overcome in 1D projection (pJRES) of spectrum where each signal corresponds to a singlet, while information about the multiplicity is preserved in the second dimension. Although improved JRES pulse sequences were designed for precise metabolite quantification [28,29], only fast JRES experiment is usually incorporated into standard protocol of high-throughput metabolomic profiling as the source of additional information for metabolite assignment.

Acquired spectra should be processed by phase and baseline correction, normalization and binning prior to the statistical analysis. The normalization is of great importance especially for urine spectra because it compensates the variability of sample dilution. Few methods of normalization are available. The normalization to the total spectral area (TSA) is the standard normalization method in NMR metabolomics [30]. Its prerequisite that total spectral integral is directly proportional to the sample dilution fails in the case when several samples contain a high concentration of a single metabolite. This major drawback of TSA can be overcome by the probabilistic quotient normalization (PQN) [31] where the inter-sample variation is estimated by computing quotients between given test spectrum and reference median spectrum. Very common routine method is the normalization to creatinine which is used mainly in human studies [32–35], yet its application in rodent models has been reported as well [36,37]. As long as the excretory system is working properly, creatinine with its relatively stable concentration can serve as a perfect standard for corrections caused by different urine volumes. However, the metabolic alterations linked to diabetes and obesity increase the risk of kidney disease, and thus the creatinine clearance has to be appraised very cautiously.

For the purpose of our study the mouse model of chemically induced obesity was employed [16,38–40]. The monosodium glutamate (MSG) administered to newborn animals induces specific lesions in the arcuate nucleus resulting in the development of obesity combined with insulin resistance [41]. Metabolic changes should be manifested in early stages as fluctuation, particularly in metabolites in biofluids, and thus could be determined by appropriate NMR procedures.

The main aim of the presented work was to examine if commonly used standard procedures are convenient also for mouse models of induced obesity. Our attention was simultaneously focused on the impact of different experimental factors on metabolite concentrations detected in urine of obese mice at several levels. We studied how variations of collection protocol can influence detected concentrations of metabolites. Moreover, we evaluated spectra acquired using three pulse sequences (1D-NOESY, CPMG, JRES) and finally the effect of three different normalization procedures on resulting data was compared.

2. Materials and methods

2.1. Experimental animals

All animal experiments followed the ethical guidelines for animal experiments and the Act of the Czech Republic No. 246/1992 and were approved by the Committee for Experiments with Laboratory Animals of the Academy of Sciences of the Czech Republic.

The study was performed on an NMRI model of mice (Charles River, Sulzfeld, Germany). L-glutamic acid sodium salt hydrate (Sigma, St. Louis, USA) was injected to newborn male mice in a single subcutaneous dose (4 mg per g of body weight) postnatal day 2–8 as described previously [42]. The mice were maintained at 23 °C with 12–12 h of light-dark regime (light from 6 AM) with free access to water. The chow diet (Mlýn Kocanda, Jesenice, Czech Republic) contained 25, 9 and 66% calories counting as protein, fat and carbohydrate respectively. The total energy content was 3.4 kcal/g.

2.2. Urine collection

Urine for NMR-based metabolomics was collected from both 2-month old MSG-treated and control mice (10 animals per group) using three different collection protocols. The morning urine was collected directly through gentle pressure at abdomen (protocol A). During the 24-hour urine collection the mice were housed in individual metabolic cages (Tecniplast, Italy) without access to food (protocol B) or with access to food ad libitum (protocol C). Samples were collected with the addition of NaN_3 and stored at −80 °C until NMR analysis.

2.3. NMR experiments

The urine sample was thawed at room temperature directly before the NMR experiment and centrifuged at 12,000 rpm for 5 min. The 200 μl of supernatant was diluted with 340 μl of H_2O, mixed with 60 μl of phosphate buffer (1.5 M KH_2PO_4 in D_2O containing 2 mM NaN_3 and 0.1 % trimethylsilyl propionic acid (TSP), pH 7.4) and transferred into 5 mm NMR tube. Directly collected morning urine (protocol A) and one aliquot of sample from the B protocol were measured in 3 mm tubes filled with a mixture of 50 μl supernatant and 150 μl diluted phosphate buffer. NMR data were acquired on a 600 MHz Bruker Avance III spectrometer (Bruker BioSpin, Rheinstetten, Germany) equipped with a 5 mm TCI cryogenic probe head. All experiments were performed using standard manufacturers' software (Topspin 3.2, Bruker BioSpin, Rheinstetten, Germany) at 300 K with automatic tuning and matching, shimming and adjusting of 90° pulse length for each sample.

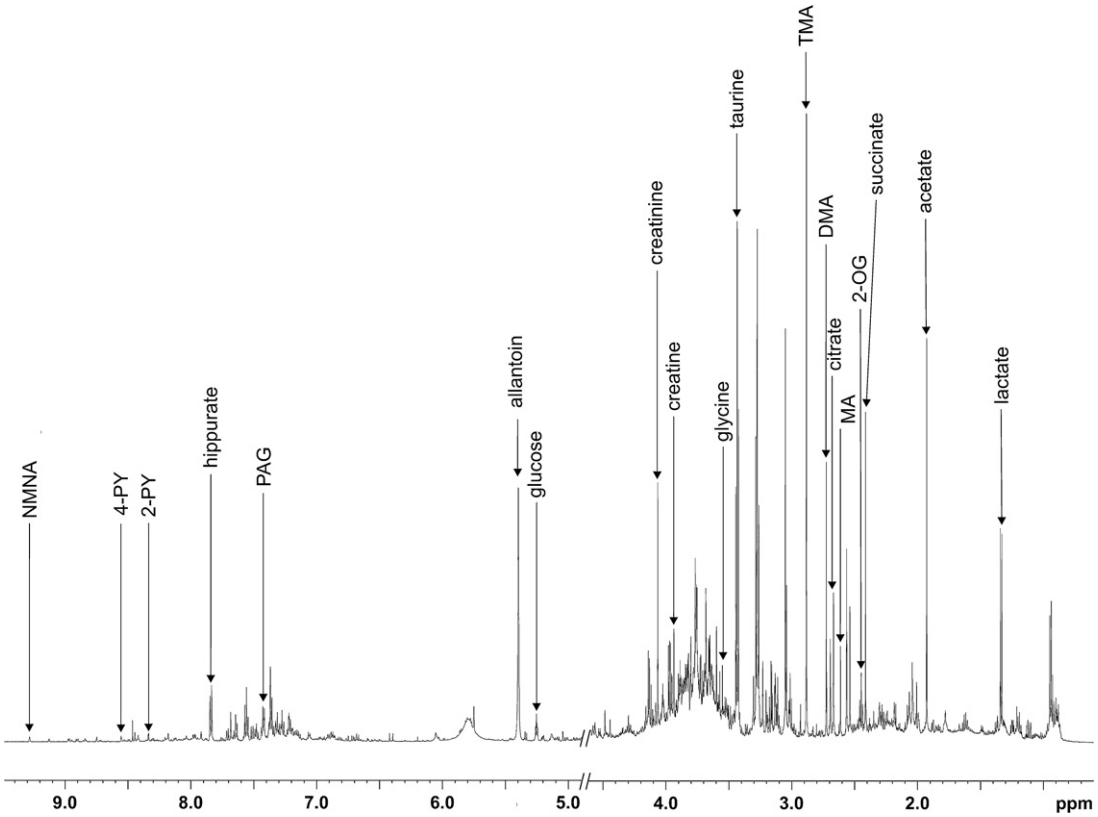


Fig. 1. Representative CPMG spectrum of MSG mouse with the assignment of particular signals of selected metabolites.

1D-NOESY spectra were acquired using Bruker pulse sequence noesygppr1d with following acquisition parameters: number of scans (NS) = 32 (NS = 64 for 3 mm tubes), 64k of data points (TD), spectral width (SW) of 20 ppm. Water resonance was suppressed during relaxation delay (4 s) and mixing time (100 ms) by 25 Hz (γB_1) saturation pulse centered on the water resonance. The CPMG spectra with presaturation were obtained by Bruker pulse sequence cpmgpr1d, which was performed with NS = 64, TD = 64k, SW = 20 ppm, relaxation delay 4 s with presaturation pulse centered on the water signal, echo time 0.3 ms, loop for T2 filter 126. A short 2D J-resolved (Bruker pulse sequence jresgpprqf) experiment with presaturation (NS = 2, SW = 16 ppm, TD = 8k, number of increments = 40, SW = 78.125 Hz in the indirect dimension, relaxation delay 2 s with presaturation pulse centered on the water signal) was executed for better metabolite identification and as an additional source of data for multivariate analysis.

Acquired data were processed by Topspin 3.2 (Bruker BioSpin, Rheinstetten, Germany) software. Line broadening of 0.3 Hz for 1D-NOESY, CPMG, and 1D projection of the *J*-resolved experiment was applied. Spectra were automatically phased, baseline corrected and referenced to the peak of TSP (0.00 ppm).

2.4. Data analysis

Metabolites were identified by matching the 1D spectra (NOESY or CPMG) with databases Chenomx NMR Suite 7.6 (Chenomx Inc., Edmonton, AB, Canada), BBIOREFCODE 2-0-3 (Bruker BioSpin GmbH, Rheinstetten, Germany) and with previously published assignments.

All NMR spectra were uniformly binned to 0.04 ppm intervals; the regions corresponding to water (4.68–4.90 ppm), urea (5.65–5.90 ppm) and TSP (−0.12–0.12 ppm) were consequently excluded. The normalization to the total spectral area, probabilistic quotient normalization or normalization to the creatinine signal at 4.06 ppm was performed to compensate the different dilution of original urine samples; all methods will be discussed below. To avoid splitting of important signals between adjacent bins and distortion of bin intensities, four data matrices mutually shifted by 0.01 ppm were calculated and each signal under study was correctly centered into an appropriate bin.

Principal component analysis (PCA) was performed using standard algorithm based on singular value decomposition [43], data were mean centered and Pareto scaled before the analysis. Further statistical analysis was performed by the standard partial least square discriminant analysis (PLS-DA), as in the case of PCA data were mean centered and Pareto scaled. Both above mentioned procedures utilized the MATLAB program [44]. The leave-one-out validation was used to assess the quality of different models and to choose the correct number of latent variables with respect to the performance of the model on the validation data set. The fraction of explained variation R^2 and the fraction of predicted variation Q^2 were also computed to provide more detailed evaluation of trained models. Selected subset of metabolites was then analyzed in more detail using results of PLS models. The standard Student's two sample t-test for ratio of means was carried out using the R environment for statistical computing [45] and appropriate package mratios [46] to validate the significance of metabolites concentration ratios for selected bins described in the next chapter.

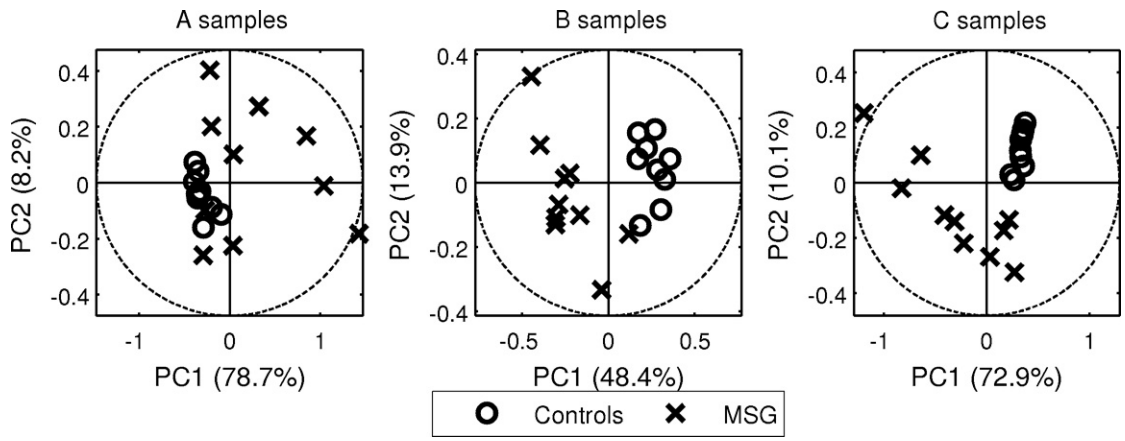


Fig. 2. The comparison of PCA scores plots for urine collection protocol A, B, and C (explained variance of the first two PCs is indicated in the figure).

3. Results and discussion

The NMR-based analysis of urine samples allowed us to identify more than forty metabolites. Nineteen metabolites were selected for subsequent analysis (the assignment of the corresponding signals is depicted in Fig. 1). The metabolites selection was driven by several factors. From analytical point of view, suitable metabolites could be easily and reliably detected using well-defined multiplets with chemical shifts covering the whole range of a proton spectrum. In addition, the concentration of the most and the least abundant metabolites in this set differs in two orders of magnitude, allowing us to monitor possible effects on compounds presented in very different amounts. In terms of biochemistry, altered levels of selected metabolites should be related to diabetes or obesity [47–49]. Each metabolite in this study was represented by one signal centered into a bin with 0.04 ppm width. The metabolite concentration was expressed as the corresponding bin intensity normalized to the total spectral area, to creatinine or using probabilistic quotient normalization (see Section 3.3.). All prospective effects of different experimental factors were then studied on the same set of metabolites. Selected metabolites with representative signals are listed in Table 1.

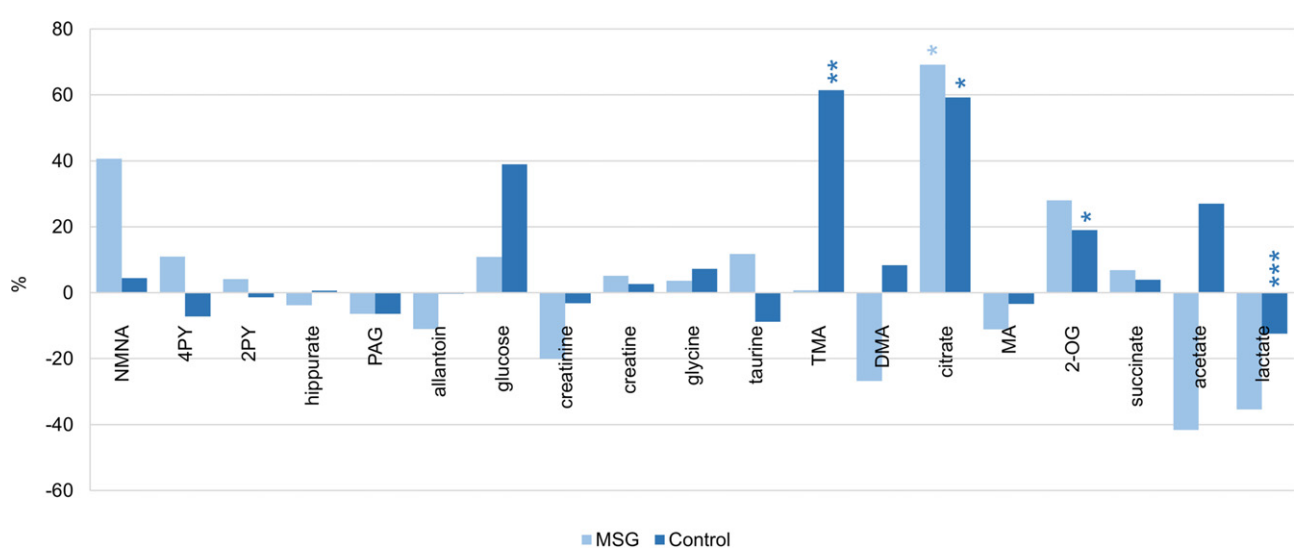


Fig. 3. The relative impact of diurnal variations (expressed as the ratio of concentrations detected using protocols A and C) – evaluated independently for both MSG and control mice. Student's *t* test for ratios was used to indicate the significance of change in metabolites' levels between groups: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

Table 1
The list of metabolites selected for this study with representative signals.

Selected signal [ppm]	Metabolite (with abbreviation)
9.29 (s)	1*N*-Methylnicotinamide (NMNA)
8.55 (d)	*N*-Methyl-4-pyridone-3-carboxamide (4-PY)
8.34 (d)	*N*-Methyl-2-pyridone-5-carboxamide (2-PY)
7.84 (dd)	Hippurate
7.46 (t)	Phenylacetylglycine (PAG)
5.39 (s)	Allantoin
5.25 (d)	Glucose
4.06 (s)	Creatinine
3.93 (s)	Creatine
3.54 (s)	Glycine
3.43 (t)	Taurine
2.88 (s)	Trimethylamine (TMA)
2.72 (s)	Dimethylamine (DMA)
2.68 (d)	Citrate
2.61 (s)	Methylamine (MA)
2.45 (t)	2-Oxoglutarate (2-OG)
2.41 (s)	Succinate
1.92 (s)	Acetate
1.34 (d)	Lactate

3.1. Impact of urine collection protocol

Urine collection protocol was the first experimental factor under study. Three different protocols described in the experimental part were compared.

Ideally, sampling of urine should not affect metabolic processes. From this point of view the best choice could be the urine collection in metabolic cages with free access to food and water (protocol C), but due to a substantial risk of urine contamination by food particles it is hard to employ. The direct urine collection (protocol A) is rather problematic as the volume of urine is usually very small and considerably variable. The urine composition is not negatively influenced by fasting, but the levels of metabolites can reflect diurnal variations. Furthermore, animals are not unified relative to the food intake and food pattern. The most common and widely recommended method applied for rodent models is 24-hour collection in metabolic cages without access to food (protocol B). It should be borne in mind that the all-day fasting can considerably change the metabolic profile of urine of small animals and thus affect the results of obesity-related study.

Small volumes of directly collected morning urine (protocol A) had to be measured in NMR tubes with a 3 mm diameter. To assure that resulting data are directly comparable (especially when measurements were carried in a 5 mm probe), urine samples of protocol B were prepared in both 3 and 5 mm tubes. The statistical evaluation confirmed almost negligible effect of the tube diameter on final data (data not shown).

Standard 1D-NOESY experiments were acquired for all samples (e.g. three different samples A, B, C per animal). Visual inspection of PCA score plots (see Fig. 2) indicates that the best group separation was achieved by the commonly used protocol B of 24-hour collection of fasting urine. The analysis of normalized concentrations of individual metabolites was performed for deeper insight into the impact of particular urine collection protocols.

If we consider protocol C as an approximate picture of 24-hour-long metabolism, the ratio of particular metabolite concentrations in the samples from protocols A (morning collection) and C (24-hour urine with access to food) may reflect potential changes of selected metabolites induced by circadian rhythms (see Fig. 3).

Elevated levels of glucose, trimethylamine (TMA), citrate, 2-oxoglutarate (2-OG), and acetate and reduced level of lactate were observed in the morning urine (protocol A) of the control group. The morning urine of the MSG group displayed the higher concentration of 1*N*-methylnicotinamide (NMNA), citrate, 2-OG and lowered concentration of dimethylamine (DMA), acetate and lactate. However, many changes mainly in the MSG group were not considered as significant and are not distinguishable from inherent variation of metabolites levels among individual mice. Observed trends in both groups nevertheless correspond to those previously described in literature. Gavaghan et al. [50] also studied diurnal variations for two phenotypically normal mouse strains, where increased concentrations of TMA, DMA, citrate, 2-OG, and succinate in the morning urine were detected compared to the afternoon urine collection. Williams et al. [51] detected increases in taurine, creatinine, allantoin, and 2-OG in the evening urine samples of Zucker fatty rats. Bollard et al. [52] reported analogous diurnal variations among morning and evening urine collection in Sprague–Dawley rats: elevated levels of citrate, glycine, creatine, and glucose in morning urine and raised concentrations of taurine, creatinine, and hippurate in the evening.

The free access to food during the all-day urine collection is the only difference between protocols B and C. Therefore, the ratios of metabolite concentrations detected in samples B and C show the influence of food intake on urine composition. Fig. 4 displays the different effect of fasting on particular metabolites independently for both groups of animals. MSG mice became less tolerant to the fasting compared to the control NMRI group, which is apparent from the significant variations of almost all observed metabolites. The restricted access to food primarily increased concentrations of NMNA, *N*-methyl-2-pyridone-5-carboxamide (2PY), *N*-methyl-4-pyridone-3-carboxamide (4PY) and phenylacetylglycine (PAG) and decreased the glucose level in MSG treated animals. The changed sensitivity to fasting is probably the main reason for the better separation of the MSG and control group using the B protocol, which was observed in PCA analysis (see Fig. 2).

The different response on restricted food intake between the examined and control animals may overlap or distort the model-related metabolic differences. We consider it a major drawback of the commonly used fasting protocol B. We assume that whenever the model and control animals are not from the same strain, the effect of fasting should be taken into account and evaluated separately for both groups. On the other hand, in models of chemically or diet induced obesity both the examined and control groups only differ in the treatment inducing the obesity (e.g. MSG injection, high-fat diet). All detected alterations during fasting can therefore be directly attributed to the metabolic changes caused by obesity and insulin resistance development, and thus can be utilized as markers within the model under study. For instance, the elevated concentrations of NMNA and its oxidation products 2PY and 4PY observed under pre-diabetes condition were negligible during standard food intake and might be easily omitted using protocol A or C only.

3.2. Impact of pulse sequence

Originally, the 1D-NOESY pulse sequence was performed on urine samples from fasting animals (protocol B) as recommended in standard protocol [20]. Intense broad resonances attributed mainly to major urinary proteins were detected in the spectra. Therefore we also performed the CPMG experiment, which is usually applied for plasma or serum samples, to suppress the undesirable influence of high molecular weight compounds. Furthermore, JRES spectra were acquired and their 1D projections (pJRES) were also subjected to data analysis. We compared how the pulse sequence employed for data acquisition can affect the quality of a spectrum, and in this manner influence the detected metabolite concentrations and resulting statistical model.

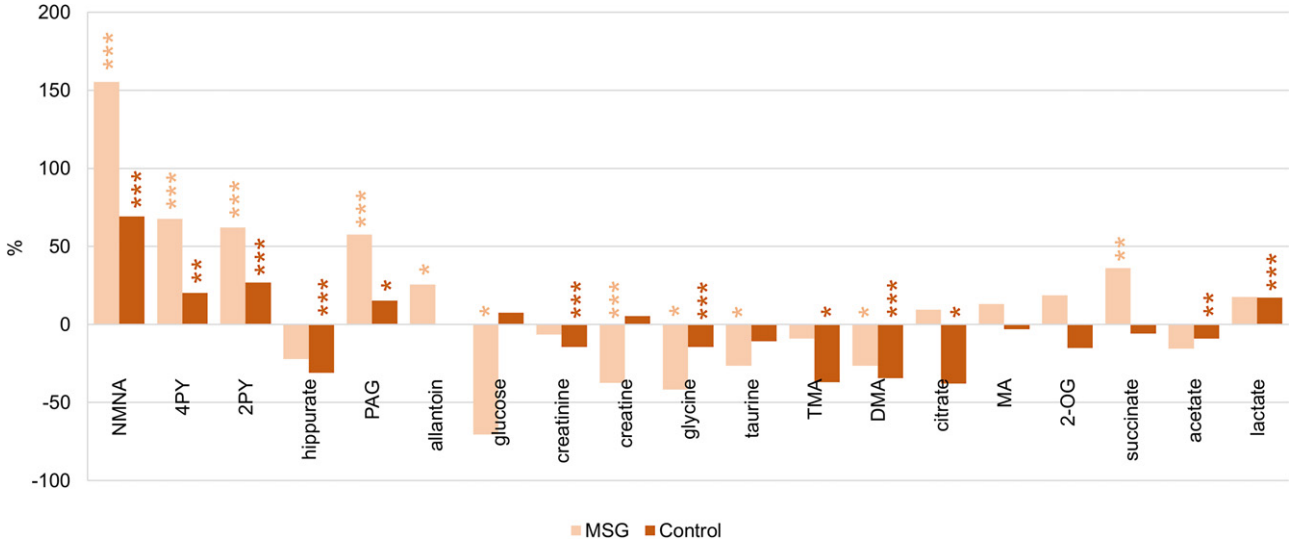


Fig. 4. The relative impact of fasting during urine collection (expressed as the ratio of concentrations detected using protocols B and C) – evaluated independently for both MSG and control mice. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$.

Data obtained by 1D-NOESY, CPMG, and pJRES were initially analyzed by the PLS-DA method. Three PLS-DA models were trained and validated. The visual inspection of scores plots (Fig. 5) shows slightly worse group separation in 1D-NOESY than the CPMG and pJRES model, which is reflected in the characteristics of respective PLS models. In comparison with PLS models for CPMG and pJRES spectra the 1D-NOESY model needs, as a result of validation procedure, markedly more latent variables. The CPMG and pJRES models are both sufficiently accurate with only one latent variable, the 1D-NOESY model needs three latent variables. These distinctly worse characteristics propagate into other performance measures such as the fraction of explained variation R^2 and the fraction of predicted variation Q^2 (see legend to Fig. 5), which are both, similarly to the number of latent variables, smaller in the 1D-NOESY model than in the other models.

Although all models generally seem to be appropriate for group discrimination, differences among 1D-NOESY, CPMG, and pJRES spectra are apparent in Fig. 6.

The impact of pulse sequence is well illustrated on the expanded aromatic part of the proton spectra. The strong protein background in the 1D-NOESY spectrum (blue line) overlaps less intensive signals preventing their quantification. On the other hand, pJRES (green line) provides spectra with a flat baseline and only a small signal overlap, but low abundant metabolites resonating around 9 ppm are not detected due to a low signal-to-noise (S/N) ratio in the fast single-scan *J*-resolved experiment optimized for the

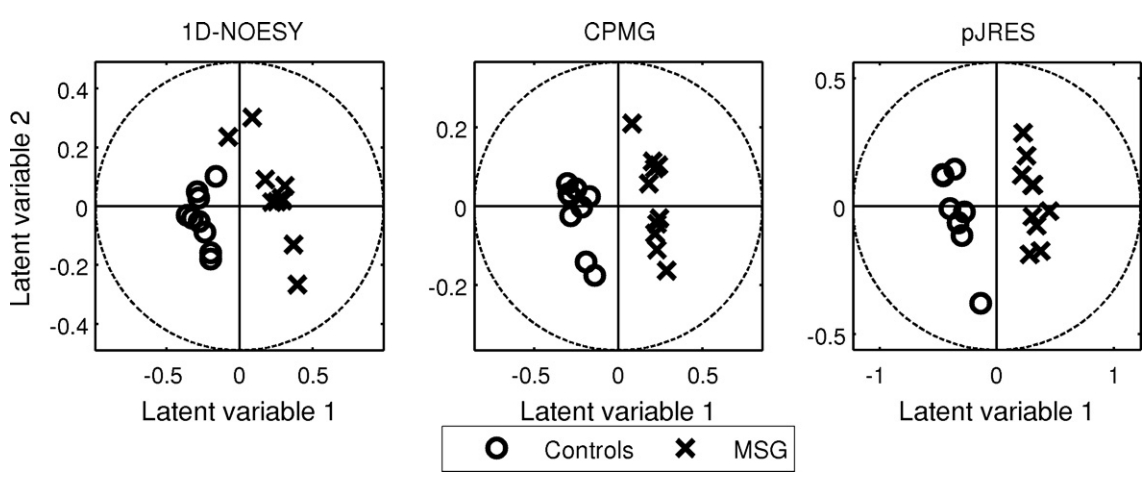


Fig. 5. Scores plots of PLS-DA models based on 1D-NOESY, CPMG and pJRES experiments. The values of fraction of explained variation R^2 and fraction of predicted variation Q^2 for individual models are: $R^2=0.97$ and $Q^2=0.87$ for 1D-NOESY, $R^2=0.98$ and $Q^2=0.93$ for CPMG, and $R^2=0.97$ and $Q^2=0.92$ for pJRES.

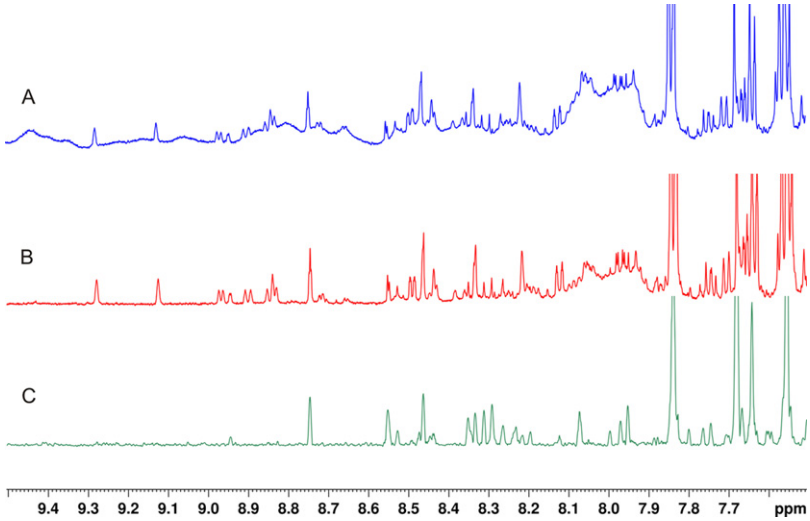


Fig. 6. Comparison of 1D-NOESY (A), CPMG (B), and pJRES (C) spectra for the representative sample of control mouse – expansion of aromatic region with less intense metabolites.

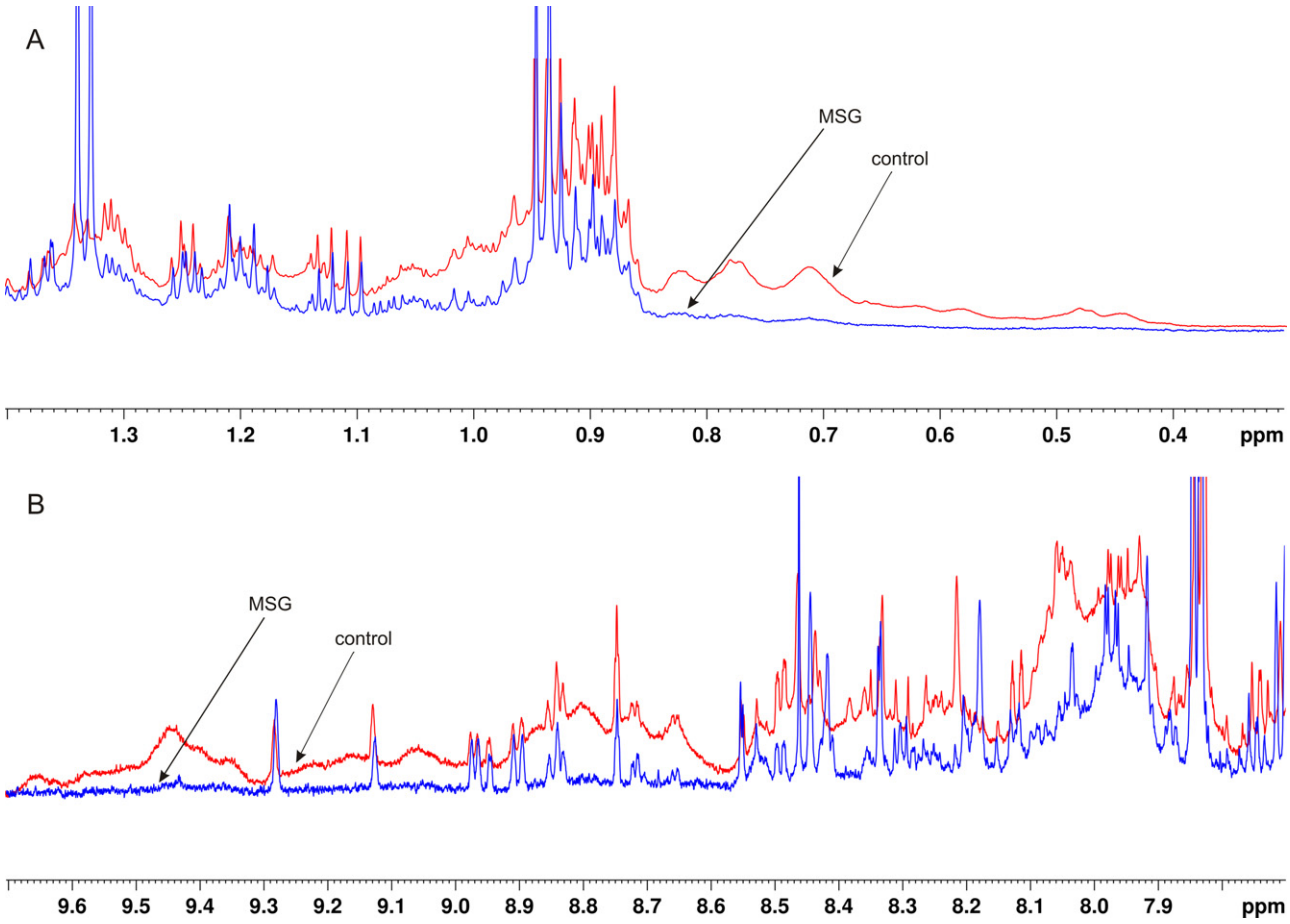


Fig. 7. Comparison of 1D-NOESY spectra for the representative sample of the control and MSG mouse – expansion of right (A) and left (B) region of spectrum.

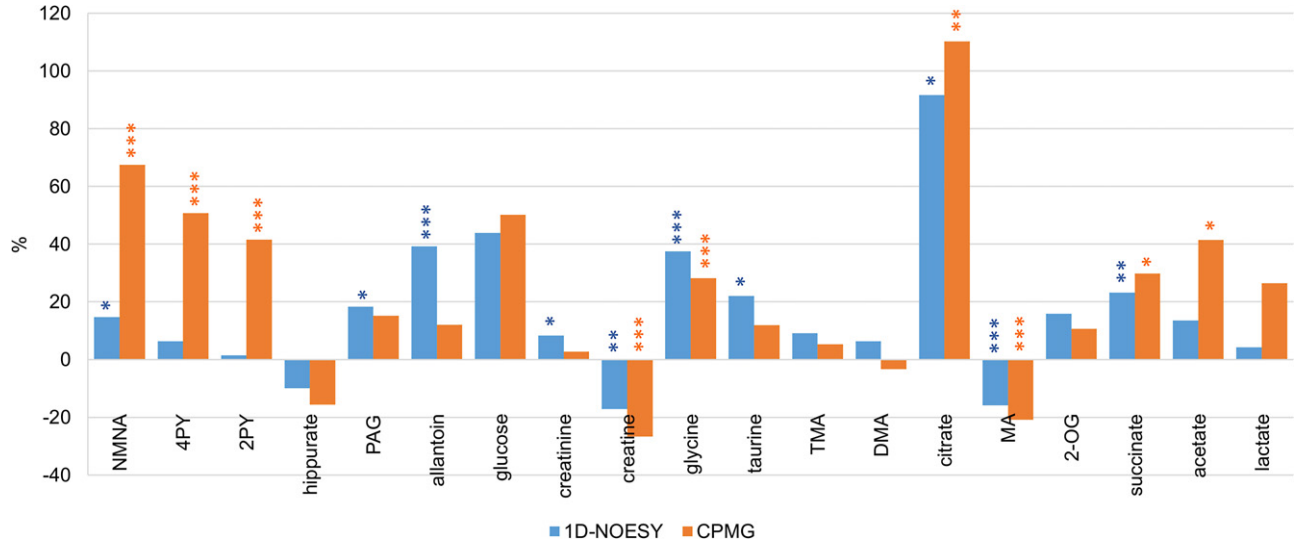


Fig. 8. Relative changes of metabolite concentrations (expressed as MSG/control concentration ratios) – the comparison of the 1D-NOESY and CPMG based model. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

high-throughput mode. The CPMG spectrum (red line) displays well resolved signals with a satisfactory S/N ratio and nearly flat baseline.

Although the use of mouse models is very common in NMR metabolomic studies of obesity and diabetes, we noticed only a few papers where CPMG experiment was performed instead of 1D-NOESY to eliminate protein signals in urine [6,53]. During metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human [5] high levels of proteins in *db/db* mouse model, which authors explained by proteinuria, also negatively affected 1D-NOESY spectra mainly in the aliphatic and aromatic region. The results acquired by CPMG were reported only sporadically. Moreover, they have never been thoroughly discussed and compared with the widely used and generally accepted 1D-NOESY sequence for NMR metabolomics in urine. For this reason, we focused attention on the comparison of metabolite levels detected by 1D-NOESY and CPMG experiments in mouse urine samples, naturally containing proteins.

Apart from the signal overlap, the substantial difficulty of 1D-NOESY spectra stems from the fact that the amount of excreted protein is significantly higher in control than in MSG mouse urine (see Fig. 7). The average decrease of protein signals intensity read at 9.45 ppm was 26%, and at 0.80 ppm even 40%. Changed levels of major urinary proteins (MUPs) relative to the altered regulation of glucose metabolism were observed in several studies [54–56]. Reduced concentration of MUPs was, for instance, described in *db/db* mice and dietary obese mice compared to their lean controls. It means that the above mentioned problems with 1D-NOESY spectra may be common for other mouse obesity models as well.

The changes of individual metabolites were calculated as the ratio of MSG and control bin intensities to evaluate outputs of different pulse sequences. The significant discrepancy between data acquired by 1D-NOESY and CPMG pulse sequences is depicted in Fig. 8. The model based on 1D-NOESY data shows substantially smaller changes in levels of NMNA, 4PY, 2PY, acetate, and lactate which are caused by the overlap of their peaks with broad protein resonances in corresponding spectral region. Because the amount of proteins excreted by MSG mice is reduced compared to control animals, the contribution of protein signals to bin intensities decreases in MSG samples. Thus, the real increase of particular metabolites can be compensated by the lowering protein signals in the NOESY-based model. It is crucial mainly for low abundant compounds. For instance, the important increase of nicotin amide metabolites (NMNA, 2PY, 4PY) would be regarded insignificant using standard 1D-NOESY experiments.

3.3. Impact of normalization method

Data normalization is a necessary procedure for urine samples to compensate deviations arising from the different sample dilution. The three main normalization methods applied in NMR-based metabolomics were compared: normalization to the total spectral area (TSA), probabilistic quotient normalization (PQN) and normalization to creatinine. Even though the last method is often used as a standard in biomedical research, its application may be limited either by impaired renal function or by signal overlap of creatinine peak in NMR spectrum resulting in its difficult quantitation. Fig. 9 reports the results of these three methods on MSG-treated vs. control concentration ratios for CPMG and 1D-NOESY pulse sequences.

All examined methods of normalization provided quite consistent results for the CPMG-based model. In the case of experimental problems with creatinine quantification both TSA and PQN methods thus can be applied. Average deviation among data obtained by normalization to TSA and PQN compared with creatinine normalization was 6% and 2%, resp. On the other hand, concentration changes derived from 1D-NOESY spectra normalized to TSA demonstrated positive deviation of about 12% from those normalized to creatinine. Due to the lower protein excretion in MSG mice the total area of their 1D-NOESY spectra is smaller than the TSA of control samples, and it results in higher values of the MSG/control ratio. Generally, the higher the decrease of protein level in model than control samples, the greater the values of detected metabolite changes normalized to TSA than creatinine, and vice versa. However, PQN was able to satisfactorily suppress effect of protein background in 1D-NOESY spectra and with approximately 3% difference from values obtained by normalization to creatinine became suitable normalization method regardless of the employed pulse sequence.

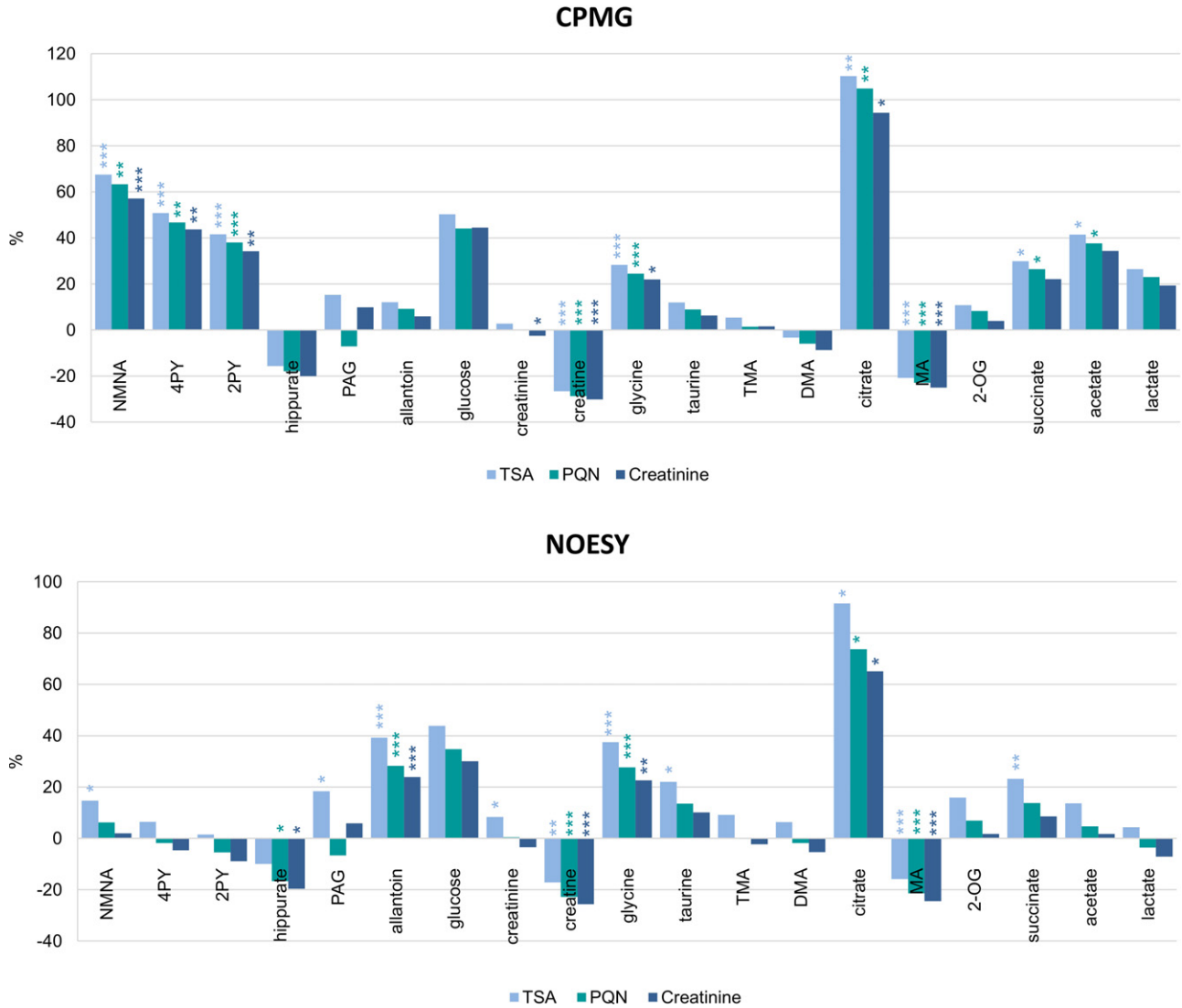


Fig. 9. Relative changes of metabolite concentrations (expressed as MSG/control concentration ratios) using normalization to TSA, PQN and normalization to creatinine for the 1D-NOESY and CPMG based model. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

4. Conclusions

We examined the impact of various experimental factors on the concentrations of nineteen selected metabolites. Based on the results, we can conclude that the effect of fasting and diurnal rhythms is reflected in urine collection protocol and affects metabolite levels. The 24-hour urine collection without access to food was determined to be advantageous in the models of diet- or chemically induced obesity. In this protocol even metabolic changes provoked by fasting can be utilized for model description.

The significant amount of proteins naturally occurring in urine samples complicates detection of small metabolites in 1D-NOESY spectra. However, the main issue is not urinary proteins as such, but the substantial change of their concentration related to the obesity development. For this reason the CPMG pulse sequence was suggested as the method of choice for urine-based metabolomics in mouse models. Satisfactory attenuation of broad protein resonances and flat baseline in resulting spectra can lead to improved metabolite quantitation whether using multivariate statistical methods or a matching with spectral databases. In addition, it allows also evaluation of low abundant metabolites which are not usually monitored.

Application of commonly used normalization to TSA provides analogous results with creatinine normalization only if the comparable total spectral area of all samples can be expected. This condition was satisfied by CPMG but not by 1D-NOESY spectra where significant changes of protein concentration among model and control animals in obesity-related mouse models play important role and substantially lower an accuracy of standard TSA normalization method. Although this problem can be overcome introducing PQN, the consistency of data normalized by all three methods also favors CPMG before 1D-NOESY pulse sequence in above mentioned animal models.

Since metabolomics requires an individual and tailored approach according to studied issues, we can propose the 24-hour fasting urine collection in combination with the CPMG experiment as the appropriate strategy for obesity-related studies on mouse models.

Acknowledgements

The research was financially supported by the Grant Agency of the Czech Republic (Grant No. GA13-14105S), grant for long-term conceptual development of the Institute of Organic Chemistry and Biochemistry (RVO: 61388963). The project was conducted within the "Prague Infrastructure for Structure Biology and Metabolomics" which has been built up by financial support of the Operational Program Prague – Competitiveness (Project No.: CZ.2.16/3.1.00/24023). The data analyses were partially supported by the Grant Agency of the Czech Technical University in Prague No. SGS13/203/OHK3/3T/13. We would like to acknowledge also the project LO1509 of the Ministry of Education, Youth and Sports of the Czech Republic. We are indebted to H. Vysušilová for her excellent technical assistance.

References

[1] G. Danaei, M.M. Finucane, Y. Lu, G.M. Singh, M.J. Cowan, C.J. Paciorek, et al., National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2. 7 million participants, Lancet 378 (2011) 31–40.
[2] S. Wild, G. Roglic, A. Green, R. Sicree, H. King, Global prevalence of diabetes – Estimates for the year 2000 and projections for 2030, Diabetes Care 27 (2004) 1047–1053.
[3] M. Takahara, I. Shimomura, Metabolic syndrome and lifestyle modification, Rev. Endocrinol. Metab. Disord. 15 (2014) 317–327.
[4] W. Kannel, T. Dawber, A. Kagan, N. Revotskie, J. Stokes, Factors of risk in the development of coronary heart disease – six year follow-up experience the Framingham Study, Ann. Intern. Med. 55 (1961) 33–50.
[5] R.M. Salek, M.L. Maguire, E. Bentley, D.V. Rubtsov, T. Hough, M. Cheeseman, et al., A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human, Physiol. Genomics 29 (2007) 99–108.
[6] S.C. Connor, M.K. Hansen, A. Corner, R.F. Smith, T.E. Ryan, Integration of metabolomics and transcriptomics data to aid biomarker discovery in type 2 diabetes, Mol. Biosyst. 6 (2010) 909–921.
[7] J.L. Griffin, Understanding mouse models of disease through metabolomics, Curr. Opin. Chem. Biol. 10 (2006) 309–315.
[8] K. Kanasaki, D. Koya, Biology of obesity: lessons from animal models of obesity, J. Biomed. Biotechnol. 2011 (2011), article ID 197636.
[9] A.A. Butler, R.D. Cone, Knockout models resulting in the development of obesity, Trends Genet. 17 (2001) S50–S54.
[10] L. Carroll, J. Voisey, A. van Daal, Mouse models of obesity, Clin. Dermatol. 22 (2004) 345–349.
[11] T.A. Lutz, S.C. Woods, Overview of animal models of obesity, Curr. Protoc. Pharmacol. (2012), Chapter 5 Unit 5.61.
[12] R. Buettner, J. Schölmerich, L.C. Bollheimer, High-fat diets: modeling the metabolic disorders of human obesity in rodents, Obesity 15 (2007) 798–808.
[13] J.-Y. Jung, I.-Y. Kim, Y.-N. Kim, J.-S. Kim, J.-H. Shin, Z.-H. Jang, et al., 1H NMR-based metabolite profiling of diet-induced obesity in a mouse model, BMB Rep. 45 (2012) 419–424.
[14] I.Y. Kim, J. Jung, M. Jang, Y.G. Ahn, J.H. Shin, J.W. Choi, et al., 1H NMR-based metabolomic study on resistance to diet-induced obesity in AHNAK knock-out mice, Biochem. Biophys. Res. Commun. 403 (2010) 428–434.
[15] J. Shearer, G. Duggan, A. Weljie, D.S. Hittel, D.H. Wasserman, H.J. Vogel, Metabolomic profiling of dietary-induced insulin resistance in the high fat-fed C57BL/6J mouse, Diabetes. Obes. Metab 10 (2008) 950–958.
[16] M. Nagata, W. Suzuki, S. Iizuka, M. Tabuchi, H. Maruyama, S. Takeda, et al., Type 2 diabetes mellitus in obese mouse model induced by monosodium glutamate, Exp. Anim. 55 (2006) 109–115.
[17] A. Smolinska, L. Blanchet, L.M.C. Buydens, S.S. Wijmenga, NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review, Anal. Chim. Acta 750 (2012) 82–97.
[18] I.F. Duarte, S.O. Diaz, A.M. Gil, NMR metabolomics of human blood and urine in disease research, J. Pharm. Biomed. Anal. 93 (2014) 17–26.
[19] A.-H.M. Emwas, R.M. Salek, J.L. Griffin, J. Merzaban, NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations, Metabolomics 9 (2013) 1048–1072.
[20] O. Beckonert, H.C. Keun, T.M.D. Ebbels, J. Bundy, E. Holmes, J.C. Lindon, et al., Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts, Nat. Protoc. 2 (2007) 2692–2703.
[21] A.-H. Emwas, C. Luchinat, P. Turano, L. Tenori, R. Roy, R.M. Salek, et al., Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review, Metabolomics (2014), http://dx.doi.org/10.1007/s11306-014-0746-7
[22] R.T. Mckay, How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: an examination of the pulse sequence components and evolution, Concepts Magn. Reson. Part A 38A (2011) 197–220.
[23] Y. Zhou, L. Rui, Major urinary protein regulation of chemical communication and nutrient metabolism, in: G. Litwack (Ed.), Vitamins and Hormones 83, Elsevier Inc., Amsterdam, 2010, pp. 151–163.
[24] H. Carr, E. Purcell, Effects of diffusion on free precession in nuclear magnetic resonance experiments, Phys. Rev. 94 (1954) 630–638.
[25] S. Meiboom, D. Gill, Modified spin-echo method for measuring nuclear relaxation times, Rev. Sci. Instrum. 29 (1958) 688–691.
[26] W.P. Aue, J. Karhan, R.R. Ernst, Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy, J. Chem. Phys. 64 (1976) 4226–4227.
[27] C. Ludwig, M.R. Viant, Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox, Phytochem. Anal. 21 (2010) 22–32.
[28] M.J. Thrippleton, R.A.E. Edden, J. Keeler, Suppression of strong coupling artefacts in J-spectra, J. Magn. Reson. 174 (2005) 97–109.
[29] H.M. Parsons, C. Ludwig, M.R. Viant, Line-shape analysis of J-resolved NMR spectra: application to metabolomics and quantification of intensity errors from signal processing and high signal congestion, Magn. Reson. Chem. 47 (Suppl 1) (2009) S86–S95.
[30] A. Ross, G. Schlotterbeck, F. Dieterle, H. Senn, NMR spectroscopy techniques for application to metabonomics, in: J.C. Lindon, J.K. Nicholson, E. Holmes (Eds.), Handb. Metabonomics Metabolomics, Elsevier Science B.V., Amsterdam, 2007, pp. 55–112.
[31] F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics, Anal. Chem. 78 (2006) 4281–4290.
[32] C. Zuppi, I. Messana, F. Forni, C. Rossi, L. Pennacchietti, F. Ferrari, et al., 1H NMR spectra of normal urines: reference ranges of the major metabolites, Clin. Chim. Acta. 265 (1997) 85–97.
[33] S. Bouatra, F. Aziat, R. Mandal, A.C. Guo, M.R. Wilson, C. Knox, et al., The human urine metabolome, PLoS One 8 (2013) e73076.
[34] E.J. Saude, D. Adamko, B.H. Rowe, T. Marrie, B.D. Sykes, Variation of metabolites in normal human urine, Metabolomics 3 (2007) 439–451.
[35] E. Saude, B. Sykes, Urine stability for metabolomic studies: effects of preparation and storage, Metabolomics 3 (2007) 19–27.
[36] R. Calvani, E. Brasili, G. Praticò, G. Capuani, A. Tomassini, F. Marini, et al., Fecal and urinary NMR-based metabolomics unveil an aging signature in mice, Exp. Gerontol. 49 (2014) 5–11.
[37] D. Portilla, S. Li, K.K. Nagothu, J. Megyesi, B. Kaissling, L. Schnackenberg, et al., Metabolomic study of cisplatin-induced nephrotoxicity, Kidney Int. 69 (2006) 2194–2204.
[38] J.W. Olney, Brain lesions, obesity, and other disturbances in mice treated with monosodium glutamate, Science 164 (1969) 719–721.
[39] R. Matyšková, L. Maletínská, J. Maixnerová, Z. Pirník, A. Kiss, Comparison of the obesity phenotypes related to monosodium glutamate effect on arcuate nucleus and/or the high fat diet feeding in C57BL/6 and NMRI mice, Physiol. Behav. 57 (2008) 727–734.
[40] R.J. Hernández-Bautista, F.J. Alarcón-Aguilar, M.C. Del Escobar-Villanueva, J.C. Almanza-Pérez, H. Merino-Aguilar, M.K. Fainstein, et al., Biochemical alterations during the obese-aging process in female and male monosodium glutamate (MSG)-treated mice, Int. J. Mol. Sci. 15 (2014) 11473–11494.
[41] A. Djazayery, D.S. Miller, M.J. Stock, Energy balances in obese mice, Ann. Nutr. Metab. 23 (1979) 357–367.
[42] L. Maletínská, R.S. Toma, Z. Pirnik, A. Kiss, J. Slaninová, M. Haluzík, et al., Effect of cholecystokinin on feeding is attenuated in monosodium glutamate obese mice, Regul. Pept. 136 (2006) 58–63.
[43] I.T. Jolliffe, Principal Component Analysis, xxix, 2nd ed., Springer, New York, 2002, pp. 487, ISBN 0387954422.
[44] MATLAB version 8.1.0.604 (R2013a). Natick, Massachusetts: The MathWorks Inc., 2013.
[45] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2015 http://www.R-project.org/
[46] G.D. Djira, M. Hasler, D. Gerhard, F. Schaarschmidt (2012). mratios: Inferences for ratios of coefficients in the general linear model. R package version 1.3.17. http://CRAN.R-project.org/package=mratios
[47] B. Xie, M.J. Waters, H.J. Schirra, Investigating potential mechanisms of obesity by metabolomics, J. Biomed. Biotechnol. 2012 (2012), article ID 805683.
[48] I.R. Lanza, S. Zhang, L.E. Ward, H. Karakelides, D. Raftery, K.S. Nair, Quantitative metabolomics by H NMR and LC–MS/MS confirms altered metabolic pathways in diabetes, PLoS One 5 (2010) e10538.
[49] I. Messana, F. Forni, F. Ferrari, C. Rossi, B. Giardina, C. Zuppi, Proton nuclear magnetic resonance spectral profiles of urine in type II diabetic patients, Clin. Chem. 44 (1998) 1529–1534.
[50] C.L. Gavaghan, I.D. Wilson, J.K. Nicholson, Physiological variation in metabolic phenotyping and functional genomic studies: use of orthogonal signal correction and PLS-DA, FEBS Lett. 530 (2002) 191–196.
[51] R.E. Williams, E.M. Lenz, J.A. Evans, I.D. Wilson, J.H. Granger, R.S. Plumb, et al., A combined 1H NMR and HPLC–MS-based metabonomic study of urine from obese (fa/fa) Zucker and normal Wistar-derived rats, J. Pharm. Biomed. Anal. 38 (2005) 465–471.
[52] M.E. Bollard, E. Holmes, J.C. Lindon, S.C. Mitchell, D. Branstetter, W. Zhang, et al., Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high-resolution 1H NMR spectroscopy of

urine and pattern recognition, Anal. Biochem. 295 (2001) 194–202.
[53] G.C. Leo, A.L. Darrow, NMR-based metabolomics of urine for the atherosclerotic mouse model using apolipoprotein-E deficient mice, Magn. Reson. Chem. 47 (2009) S20–S25, Suppl 1.
[54] Y. Zhou, L. Jiang, L. Rui, Identification of MUP1 as a regulator for glucose and lipid metabolism in mice, J. Biol. Chem. 284 (2009) 11152–11159.
[55] X. Hui, W. Zhu, Y. Wang, K.S.L. Lam, J. Zhang, D. Wu, et al., Major urinary protein-1 increases energy expenditure and improves glucose intolerance through enhancing mitochondrial function in skeletal muscle of diabetic mice, J. Biol. Chem. 284 (2009) 14050–14057.
[56] K. Giller, P. Huebbe, F. Doering, K. Pallauf, G. Rimbach, Major urinary protein 5, a scent communication protein, is regulated by dietary restriction and subsequent re-feeding in mice, Proc. Biol. Sci 280 (2013), article ID 20130101

Handbook of Research on ICTs for Human-Centered Healthcare and Social Care Services

Maria Manuela Cruz-Cunha
Polytechnic Insitute of Cavado and Ave, Portugal

Isabel Maria Miranda
Polytechnic Insitute of Cavado and Ave, Portugal

Patricia Gonçalves
Polytechnic Institute of Cávado and Ave, Portugal

Medical Information Science
REFERENCE

Managing Director: Lindsay Johnston
Editorial Director: Joel Gamon
Book Production Manager: Jennifer Yoder
Publishing Systems Analyst: Adrienne Freeland
Development Editor: Myla Merkel
Assistant Acquisitions Editor: Kayla Wolfe
Typesetter: Christina Henning
Cover Design: Jason Mull

Published in the United States of America by
Medical Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Handbook of research on ICTs for human-centered healthcare and social care services / Maria Manuela Cruz-Cunha, Isabel Maria Miranda, and Patricia Goncalves, editors.
volumes cm
Includes bibliographical references and index.
Summary: "This book is a comprehensive collection of relevant research on technology and its developments of ICTs in healthcare and social services, focusing on the emerging trends in the social and healthcare sectors such as social networks, security of ICTs, and advisory services"-- Provided by publisher.
ISBN 978-1-4666-3986-7 (hardcover) -- ISBN 978-1-4666-3987-4 (ebook) -- ISBN 978-1-4666-3988-1 (print & perpetual access) 1. Medical informatics--Handbooks, manuals, etc. 2. Information storage and retrieval systems--Medicine--Handbooks, manuals, etc. 3. Medical care--Computer networks--Handbooks, manuals, etc. I. Cruz-Cunha, Maria Manuela, 1964- editor of compilation. II. Miranda, Isabel Maria, 1954- editor of compilation. III. Goncalves, Patricia, 1974- editor of compilation. IV. Title: Handbook of research on information and communication technologies for human-centered healthcare and social care services.
R858.H3228 2013
610.285--dc23
2012048438

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter 14
Does IT Bring Hope for Wellbeing?

Daniel Novák
Czech Technical University in Prague, Czech Republic

Olga Štěpánková
Czech Technical University in Prague, Czech Republic

Sébastien Rousseaux
Wallon Project and SPES Project, Belgium

Massimo Busuoli
ENEA, Belgium

Marco Carulli
ENEA, Belgium

Gianluca D'Agosta
ENEA, Italy

Teresa Gallelli
University of Bologna, Italy

Miroslav Uller
Czech Technical University in Prague, Czech Republic

Miloš Mráz
Charles University, Czech Republic

Martin Haluzík
Charles University in Prague, Czech Republic

Paul Richter
Newcastle University, UK

Mike Martin
Newcastle University, UK

Rob Wilson
Newcastle University, UK

Marzia Bettocchi
Bologna Municipality, Italy

Tiziana Mucci
National Council of Research, Italy

Emanuela Pipitone
University of Modena and Reggio Emilia, Italy

Claudio Lamberti
University of Bologna, Italy

Annunziata Siena
Processes Innovation Unit of Local Health Authority, Italy

Carlo Descovich
Bologna Health Authority, Italy

Boris Dvořáček
DEEP, Czech Republic

Christine Petioky
Vienna Social Fund, Austria

František Babič
Technical University of Košice, Slovak Republic

Marek Lenart
University of Vienna, Austria

Alexander Wöhrer
University of Vienna, Austria

DOI: 10.4018/978-1-4666-3986-7.ch014

ABSTRACT

The first part of this chapter reviews the design, implementation, and customer experience with the OLDES SW tele-care platform developed within the EU project Older people's e-services at home. The OLDES solution has been successfully tested at two different locations: in Italy with the participation of a group of 100 seniors (including 10 senior citizens suffering from heart disease), and in the Czech Republic, with the involvement of a group of 10 diabetic patients. The suggested OLDES approach proved to be an effective solution for municipalities, hospitals, and their contact centres for providing health and social services. The project partners therefore decided to develop a second generation of the system called SPES (Support to Patients through E-Service Solutions), which started in April 2011. The SPES project aims at transferring the original approach and results achieved in implementing the OLDES focusing on new target problem domains: dementia, mobility-challenged persons, respiratory problems, and social exclusion.

1. INTRODUCTION

We live in times when the age structure of whole continents has been changing significantly Department of Economic and Social Affairs of the United Nations Secretariat, 2010), and the percentage of the population that is aging is raising concern for governments and also for health insurance systems. Much hope is now placed in new assistive technologies that could offer reasonable help by developing reliable, low-cost technical tools with intuitive control that are ready to improve the effectiveness, interoperability and compatibility of national systems of health and social care. This has become a key task for ensuring sustainable delivery of social and economic benefits throughout the European territory in the short-to-medium term. A large number of recent projects have tried to deal with disease management for the elderly, including e.g. Confidence (Gonzalez & Vlez, 2008), Attentianet (Attentianet, n.d.), Enable (Enable, n.d.), K4Care (K4Care, n.d.), SENSACTION-AAL (Sensaction-AAL, n.d.). INTEL is even developing the dedicated Health Guide PHS6000 platform (Intel, n.d.). However, one of the major problems of IT-assisted ambient technologies remains the motivation of patients, and their daily involvement in disease management systems. This applies not only to the patient him/herself (Stroetmann, Husing, Kubitschke, & Stroetmann, 2002) but also to all the stakeholders that form the care-chain in support of older persons. This problem has been addressed by the EU project OLDES (045282): "Older People's e-services at home" (www.oldes.eu), which has been co-funded under the EU IST Programme in the period from 2007-2010. This paper reviews our experience and pilot results obtained in the course of the project.

The main goal of the OLDES project, which was carried out by a consortium of 11 partners, has been to create and test new technological solutions to improve the quality of life of older people, by developing a very low-cost, easy-to-use care platform designed to make life easier for older people in their own homes. The leading idea of the project was to offer a uniform technological solution which would provide all features required by the providers of social care systems and also by those catering for various tele-health application scenarios. The feasibility of the OLDES concept has been evaluated for patients with Congestive Heart Failure (CHF) and for patients with type 2 diabetes mellitus.

The system is based on established Internet and Tele-care communication standards, and combines social computing (user entertainment services provided through easy-to-access thematic

channels and special interest forums supported by animators) with health-care facilities. A federated solution of this kind offers the following advantages (Martin, Ponsard, Walsh, Baines, Rousseaux, Rinaldi, & Tamburriello, 2008):

- It offers a chance to break the social isolation that results from the decreased mobility of older people, through its tele-accompany service.
- The social services can motivate the users of the OLDES system towards a healthy life style by offering space to discussion groups on this topic, and by providing relevant audio-visual content in a well understandable form.
- The content is created by social workers or volunteers, who are often themselves seniors, and who thus participate in the life of the online community in a novel way.
- Selected vital parameters are monitored using tele-monitoring devices, and are stored safely in a database run by a call-centre. When an emergency situation is identified (e.g. as a result of an automated evaluation of recently acquired physiological data), the call centre staff is ready to react. The data can also be accessed at any time through Web interface by medical staff caring for the patient.
- The solution can quickly be adapted to the current health status of the service-user, without requiring any additional training: if the user is used to communicating with the system when consuming its social service, he/she will find it quite natural to use the same communication platform, using the same control philosophy, when necessary, for a tele-health application.

One of the distinguishing features of the OLDES system is that its client (an older person) communicates with the system using a remote controller only—no keyboard is involved. OLDES attempts in this way to overcome the problem of technophobia, which often has a negative impact e.g. on the perception of eHealth services among potential senior users (Hardicker & Grant, 2010; Stroetmann, Husing, Kubitschke, & Stroetmann, 2002): the computer-based solution is masked by a TV screen and its remote controller, which has been specially adopted for this purpose.

Section 4.3 will highlight the main results obtained in the OLDES project. Section 4.5 will provide further evidence that computer-assisted interactive IT has great potential for improving the quality of care for patients (Jackson, Bolen, Brancati, Batts-Turner, & Gary, 2006; Inglis, Clark, McAlister, Ball, Lewinter, Cullington, Stewart, & Cleland, 2010).

Finally, there will be a brief description of the follow-up project, the SPES Programme (Support Patients through E-services Solutions), funded by the European Regional Development Fund (ERDF). This project aims to provide dedicated new solutions for other health and social conditions (e.g. asthma, cognitive impairments), and will test them carefully in four pilot runs between 2011-2013.

2. DESIGN: ASSUMPTIONS AND PRINCIPLES

One of the most significant problems that actors such as municipalities, health-care service providers and other related service-providers have to face in providing e-Care systems is caused by the wide range of users and the need to tailor the system to each user through affordable and accessible low-cost services.

An example of the scenario building process is shown in Figure 1: it positions OLDES services in the more general context of traditionally offered services, and it reviews the existing processes by which applicants for services have to be assessed. Most elderly people are not in perfect health – they suffer from various health problems that signifi-

cantly influence their perception of wellbeing. In most cases, the target person feels fit enough to remain at home on his/her own. However, there are certain periods when an older person does not feel well, and he/she needs to be admitted to hospital. After the treatment, when her/his health status improves, the person can return to her/his routine activities at home. It is clear that if we are able to prevent these relapses leading to hospitalization the person's wellbeing will be very positively influenced. Thus, OLDES can be perceived as a tool for offering a new set of options for GPs and for their patients.

The process leading to the creation of the most suitable model for cooperation among all the stakeholders (social services, medical institutions, charities, and volunteers) has been described previously (McLoughlin, Maniatopulos, Wilson, & Martin, 2009). This paper focuses on the selection of communication tools for linking the service user and the OLDES system. A vast number of information sources can be used to monitor the social and health condition of an older person with the aim of identifying early signs of deterioration of his/her health status (Rinaldi, Martin, & Gaddi, 2011). For this purpose, we can apply first of all physiological values provided by regular measurements as planned by the medical doctor (See Table 1 for a review of the devices selected for the OLDES project). Some complementary information can be obtained from the usage of home devices, data from interactions with the ambient intelligent environment (Maly, Curin, Kleindienst, & Slavik, 2008) or with the user interface (behaviour data), data from an analysis of the interaction with a professional carer (social data), as well as more static data, e.g. anamnesis (clinical information).

2.1. The Structure of the OLDES System

The OLDES architecture consists of two main parts, namely the *Local Hub* and the *Central Hub,* as depicted in Figure 2. The Local Hub collects physiological data measured by sensors, and enables users to communicate together or with a tele-accompany operator through Voice-Over-IP.

Figure 1. The main parties entering the stage influencing the wellbeing of a partially dependent person (e.g. target elderly). SSA - Social Security Administration, GP - General Practitioner, UVG - "Unità di Valutazione Geriatrica"

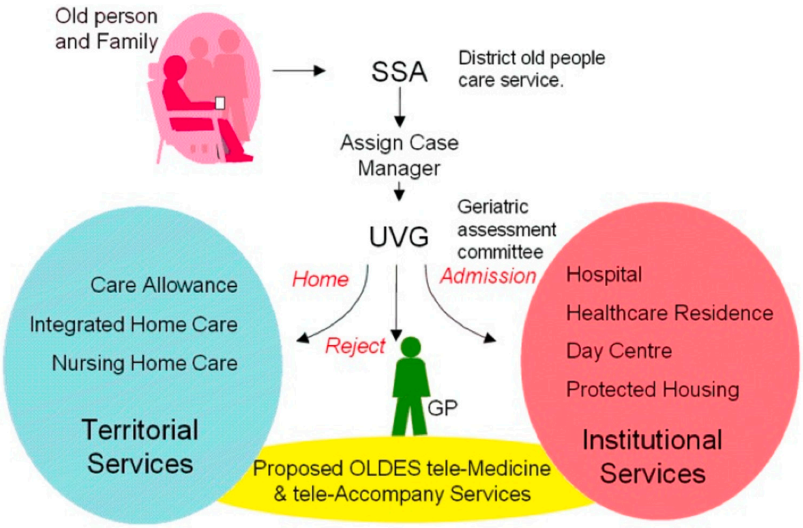
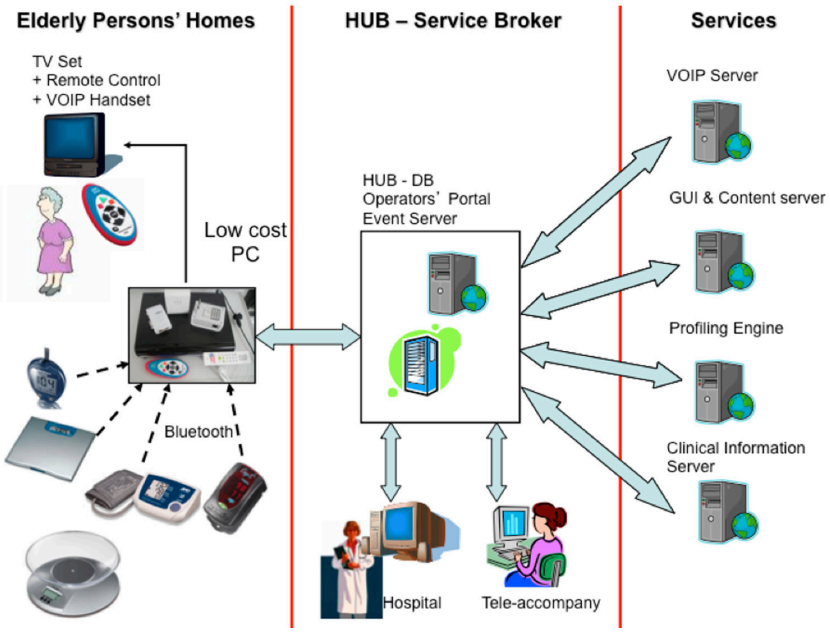


Table 1. Devices used in the OLDES project

DEVICE TYPE	PICTURE	PRODUCER	MODEL	COMMUNI-CATION
Weighing scale		A&D Medical	UC-321PBT	Bluetooth
Blood pressure meter		A&D Medical	UA-767PBT	Bluetooth
Food weighing scale		Salter-housewares	Salter 1016	Bluetooth
Pulse oximeter		Nonin	Onyx II 9560	Bluetooth
ECG		Zephyr Technology	BioHarness BT	Bluetooth
Glucometer		Lifescan	OneTouch Ultra	Bluetooth

Figure 2. OLDES architecture



The information is collected by a low-cost laptop and sent via a secured Internet connection to the Central Hub. Each older person has his/her own health agenda stored in the Central Hub.

The central HUB provides a unique and secure connection point to the OLDES central platform. It enables physicians, GPs, and tele-accompany members to access the system through easy-to-use and secure Web portals. Medical personnel, social workers and volunteers in the call centre are always available to provide support to the client, whenever he/she needs it. The service servers implement the services provided to the stakeholders connected to the system: Voice-Over-IP, the Graphical User Interface (GUI) generation engine and content server, the profiling engine, and medical data storage and analysis. A healthy lifestyle is promoted through appropriate scheduling of the thematic entertainment channels.

The OLDES project considers cost efficiency to be a very important issue - all implemented software in the Local and Central Hub is therefore open-source. Open-source components have been used and integrated to achieve some of the OLDES supported functionalities.

2.2. Local HUB: Home System

1. **Low-Cost Computer:** OLDES software is portable, and does not rely on specific hardware. It has been tested and validated on various PC-based systems. For the pilot, an ASUS EeeBox was used and deployed in the older people's homes.
2. **Remote Control:** OLDES has substituted the classical keyboard of the standard notebook or laptop with a remote controller in order to simplify interaction with the system. The main idea of the Infra-Red remote control was to ensure functionality combining features offered by teletext (familiar from standard TV sets) and by a joystick-like navigation system described next. The Infrared Remote Control consists of two devices: 1) the Remote Control keypad with an IR transmitter 2) the USB Infra-red Receiver. The Remote Control keypad is a final consumer product manufactured by American vendor Fobis Technologies.
3. **The USB Infra-Red Receiver:** Is a specific USB dongle with two main competencies: a) receive, demodulate and decode an infra-red signal b) communicate over USB and emulate a regular keyboard's key presses (See Figure 6a).
4. **Medical Devices:** Used for capturing medical data were selected according to several criteria: they have to be highly reliable, minimally invasive and easy to use for an older person. This means that they have to ensure measurements with a sufficient level of accuracy without the need to calibrate them before each measurement. The following devices and sensors have been tested in the design and development of the OLDES platform: a sphygmomanometer for blood pressure measurement, a belt for electrocardiogram (ECG) monitoring, a pulse oximeter for monitoring the oxygen saturation of the patient's blood, a LifeScan glucometer for measurements of glucose levels, together with scales for body weight measurement and an interactive scale for daily food intake records (Table 1). In addition, some sensors for ambient monitoring (e.g. temperature and humidity in the patient's home) have been included in order to check the conditions in the patient's environment, since elevated temperatures can seriously impact the health of elderly people in summer periods (See Figure 8b). The selected medical devices were very well accepted by the clients, because these devices offer a easy-to-use user interface, and they require only a few actions from the user to operate properly. The data resulting from the measurements was transmitted to the

PC through wireless connections (mainly through Bluetooth), which limited inputs or technical interactions of the human user as much as possible.

5. **Supportive Home Appliances:** The diabetes pilot project enables users to control their diet, namely their input of carbohydrates, and to keep a suitable food diary. For this purpose, a kitchen scale is wirelessly connected to the computer. The user selects the type of food that is currently on the scales, using the remote controller to find out its nutrition values. The values are automatically stored in the computer memory.
6. **Health Monitoring Software:** The main functionalities of this software tool chain collect health data measurements from various medical devices. They communicate using various technologies (Bluetooth, USB, RS232), provide a preliminary analysis of the data, and send the measurements to the OLDES central HUB through a secured encrypted Internet connection. These operations do not require any human interaction.
7. **Graphical User Interface Client:** It is important to ensure that the intended users are able (and willing) to access the offered functionalities in a simple manner - this is the purpose of the GUI (in the case of OLDES controlled by the remote controller only). The GUI framework is implemented as a standard client-server approach. GUI client, which runs on a low-cost computer, connects to the GUI service server to fetch resources and dialog pages (See Figure 3). Based on standard Web technologies, the client simply uses a Web browser running in full screen mode. It is accessible not only using OLDES software, but can be displayed using any Firefox Web browser. It communicates via AJAX with the server. In contrast to traditional Web applications, the clients look and behave similarly to desktop applications based on a user-interface toolkit such as Java AWT or Swing. The GUI client is written in Javascript and uses SVG (Scalable Vector Graphics) for rendering GUI pages and controls.

2.3. Central HUB

1. **Graphical User Interface Server**: The server part is implemented in PHP, and is responsible for handling client requests and generating GUI pages. The GUI pages are generated dynamically by the server on the basis of data and configurations stored in a database. Thanks to this feature, the look&feel of GUI is customisable for each user. Based on the user profile, menus and functionalities can be added for each user independently. The pages that are generated are sent as XML to the client, where the XML is parsed and the GUI controls are visualized instantiated. The GUI server also contains basic support for content management, representing the content as a set of content nodes. Examples of supported content types include the welcome screen, menu, article and contact list (See Figure 6).
2. **Voice-Over-IP Server**: The OLDES platform offers communication services adapted to the needs of an older person and based on modern Voice-Over-IP communication technologies. The open-source Asterisk PBX server (www.asterisk.org) was selected to implement this functionality. Asterisk is an advanced and stable VOIP server which offers plenty of functionalities. More specifically, Asterisk integrates a conferencing module, which has been used in OLDES to create and evaluate a dynamic discussion group system for older peoples. A specific Web-based interface was developed to manage the VOIP server functionalities and to enable social workers to create and manage the conference rooms in an easy and adapted way.

3. **Clinical Information Server**: This server receives, stores and analyses the medical data sent by the local hubs. An advanced data base, called Electronic Health and Social Record, was designed and implemented to store all the data. An intelligent medical data analysis tool has been integrated. It combines a support vector machine and a fuzzy logic engine for analysing the medical data and generating alerts and alarms. These signals are used to support the decision process of the physicians.
4. **Central Server:** The Hub works as the connection point between the home systems, the carers as clinicians or e-Care operators, and the service providers and maintainers. While the older clients access the Hub services through a low-cost computer, the other stakeholders can access OLDES functions through a unique Web portal customised according to their roles and needs. The data transfer from the home system software to the central server uses the XML-RPC (Remote Procedure Calls) Web service protocol.

3. EVOLUTION OF THE SYSTEM

3.1. Platform Design

Given the lack of familiarity of seniors with technological devices, the usability and accessibility-related aspects of e-Health systems have always represented an important dimension in the implementation of tele-assistance platforms. During the last decade, much attention has been devoted to problems of usability and how it can be improved (Stone, Jarrett, Woodroffe, & Minocha, 2005). A natural way to approach this goal is by improving the usability of the systems through customizable solutions potentially replicable in different care situations. Extensive experience indicates that the design of interactive systems cannot follow the same strategy as that applied in other fields of engineering design (Mayhew, 1999), where the initial design decisions can be carried out analytically without relying on a prototype. We designed the OLDES personal healthcare system interface by iterating paper and software prototyping design. Moreover, in the case of OLDES, the classical keyboard has been substituted by a remote controller, carefully selected and specially modified for the project, in order to ease interaction with the system. A user-centred approach has been applied to test the usability of the proposed tool and its suitability for the capacities of older people (Fisk, Rogers, Charness, Czaja, & Sharit, 2004), adapting the configuration of the GUI (Graphical User Interface) on the basis of test activities carried out with older people (Macik, Hanzl, Klima, & Slovacek, 2009).

To manage the risk of a mismatch in the interface between the technical domain and the user domain, the project has applied the broad framework of user-centered design that shaped the dimensions of technology acceptance. Particular attention was paid to four principles of human-centred design that are listed in the relevant ISO13407 standard:

- Active involvement of users and a clear understanding of user and task requirements.
- Appropriate allocation of function between users and technology.
- Iteration of design solutions.
- Multi-disciplinary design.

3.2. Computer without a Keyboard

The term Tangible User Interface (TUI) refers to a system that gives physical form to digital information and/or computational functions. One of the major motivating arguments for TUIs is that they utilize our spare capacity for interaction, in terms of the degrees-of-freedom and sensitivity of our hands, our bimanual skill, fine-grained mo-

tor control, peripheral vision and spatial memory (Dubuc & Edge, 2004). Whilst TUIs themselves have not been widely applied in the field of assistive technology, there are countless assistive technologies already in use which make use of the physical form. The successful adoption of these artefacts shows that older adults are comfortable with using assistive technologies in a familiar physical form, and thus the transition from low-tech assistive devices to hi-tech devices may be relatively smooth.

A crucial task was to design a TUI for navigating in the OLDES menu that is as simple as possible, and that takes into account the capabilities and constraints of elderly people. Some ideas from the USERFit handbook (Poulson, Ashby, & Richardson, n.d.) have proved to be very inspiring, and led to the design of remote control that combines tele-text functionality from a standard TV set with a joystick-like navigation system.

3.2.1. TUI Evolution

The initial suggestion depicted in Figure 4a was based on the concept of a standard TV remote controller. It consisted of:

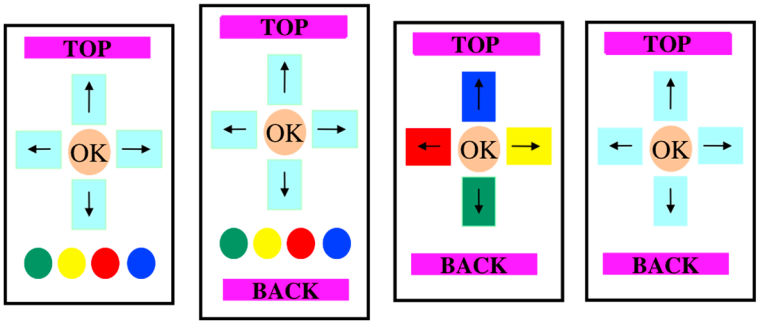
1. A top bar-button that can ensure a jump to the top screen (depending on the given OLDES setting).
2. Four joystick buttons that enable navigation in menus.
3. Coloured teletext-like buttons that serve mainly for quick access to important points in the menu.

After careful testing with a few volunteers, a "back button" was added in order to support back-tracking by coming back to the previous screen (See Figure 4b). Laboratory tests revealed that the similarity of performance between the OK button and the teletext buttons is misleading for some elderly users, since both types of buttons can be understood as confirmation. For this reason, the

Figure 3. Example of a graphical user interface: (a) the introductory screen in the Bologna pilot project. (start: today, always, my appointment, call to, how to, entertainment); (b) weather forecast with rain probability prediction

teletext buttons were replaced by coloured joystick buttons shown in Figure 4 c. During screen design for the diabetes pilot project, no need for quick access (hot keys) was identified. This allowed further simplification of the final TUI design (See Figure 4d). A commercial weemote®dV (Weemote, n.d.) remote controller for Set-Top devices featuring Interactive TV offered just the functionalities required by our last prototype (See Figure 4d). This product has been adopted by the OLDES system after a minor modification.

Figure 4. Schematic illustration of various remote control candidates



3.2.2. Inclusive Design

The user interface has to meet the requirements of its users and also the requirements for the task to be performed. It is important not to overwhelm the user with too much detail, while providing her/him with all necessary information. We have applied paradigm of user-centered design. To find the proper trade-off, we first applied paper prototyping (Snyder, 2003), with the clear goal of designing a suitable user interface for older persons suffering from diabetes. The user-interface guided the user through all the necessary steps that she/he experienced when trying to accomplish tasks, e.g. measurements of blood glucose level, weight and blood pressure, or the choice of appropriate food in compliance with the recommended amount of carbohydrates. The experience gained through paper-prototyping testing had a clear influence on the design of the final software prototype.

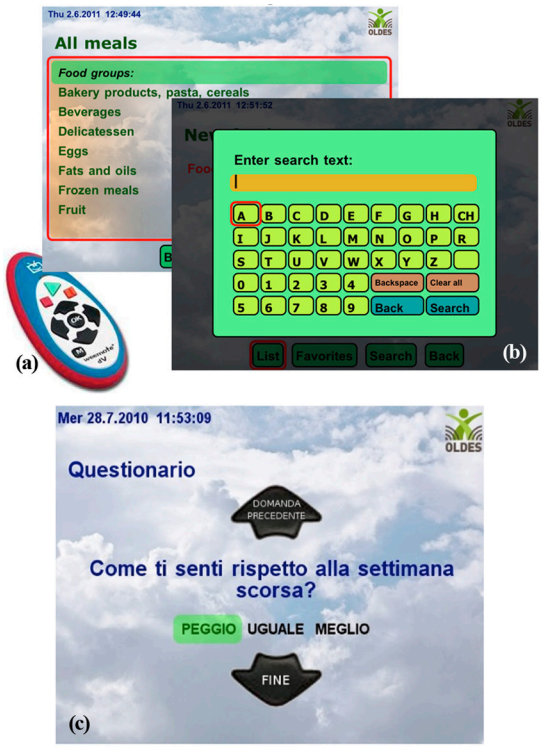
The user designed the user interface for her/himself according what she/he considered suitable. Several mock-ups had been prepared in advance as main menu and sub-menus, icons, titles of the menus. If the user suggested an issue, and no prepared components were ready, coloured papers were used to substitute for the issues as menus, the glycemia value, or a picture - see Figure 5a. Secondly, a software prototyping approach was applied. During the software prototype test sessions in the usability lab, the user was seated in the testing room along with the moderator, who guided in the usability test. The moderator tried to intervene as little as possible – he/she was under instructions to give clues and push the user ahead only in the event that the user was apparently lost or confused. All the user's actions were monitored in the observer room by means of two video cameras (front view and rear view) and an audio recording (See Figure 5b. Even inexperienced computer users were able to accomplish the set of defined tasks within a 30 minutes period. Experienced users took 10-15 minutes to accomplish the task.

4. FIELD TESTING

Two different pilots were implemented to test and validate the platform that had been created. One pilot, in Bologna, was targeted at 100 older persons, of whom 10 subjects were affected by Congestive Heart Failure (CHF), and the second pilot was tested in Prague, focusing on patients with diabetes mellitus. All users of the system (service users, clients) were adequately informed of the aims of the study, and gave written informed consent to participate in the study. In addition, it was made clear to the patients that their inclusion in our study would not modify the usual clinical follow-up. The homes of all the service users were equipped with all the necessary HW, namely the OLDES platform, including both the

Figure 5. Prototyping sessions: (a) paper prototyping, (b) software prototyping

Figure 6. Graphical and tangible user interface with the weemote®dV controller: (a) Weemote®dV controller used for controlling the graphical user interface; (b) Prague pilot user interface – food item selection; (c) Bologna pilot user interface – example of a questionnaire (how did you feel last week – better, no change, worse)



tele-accompany system and the tele-assistance system, the sensors for remote monitoring of physiological parameters such as body weight, blood pressure, blood oxygen saturation and heart rate by one-lead ECG. Patients and caregivers were trained in how to handle the devices, the PC and the platform. The enrolment and training activity, carried out by local services providers, was an important aspect both of the implementation and of the design of the system, and of the use of the computer and the medical devices by end users.

4.1. The OLDES Tele-Accompany Offers

The main intention of the project was to offer to the user various means of contact and communication, and also the following services:

- **VoIP Phone Calls:** Within the network of service users. This technology was used not only to provide contact with the dedicated carer in the social care centre. Each user could make free calls to other users belonging to the same network.
- **Multimedia Content:** (Text, audio, video) Was created and targeted to the specific audiences.
- **Participation in Discussion Groups:** On specific targeted contents. Various discussion groups were established in which the users could participate in moderated voice chats animated by a moderator.

- **Personal Agenda Management:** Each user could have an individualized reminder of events that she/he did not want to miss (planned visits, anniversaries, activities, favourite TV programs, etc.).

All of these services were accessible from the OLDES start screen (See Figure 3a), and the user was able to select from them according to her/his taste and mood.

4.2. Description of the OLDES Tele-Health Pilot Studies

The Bologna pilot study tested the OLDES platform on a cohort of 100 older persons. In addition to the tele-accompany services provided for the whole range of participants, 10 of them suffering from CHF were included in the cardiovascular pilot study, where the tele-accompany part was complemented by remote monitoring through sensors measuring selected physiological parameters (body weight, blood pressure, blood oxygen saturation, ECG). The 90 elderly participants in the test of the OLDES tele-accompany services platform were selected by the Social Services Unit of the Savena District (Municipality of Bologna) together with an ad hoc tele-accompany working group composed of "technical" representatives of the local partners. The ten cardiac patients were selected among the out-patients of the Cardiology Division of Bellaria Hospital in Bologna. To achieve the target number of 100 active participants in the tests, it proved necessary to recruit a significantly higher number of potential service users: there were 53 withdrawals from the trial out of 132 persons recruited in the lifetime of the OLDES project. The main reasons for the withdrawals were related to technical problems that prevented the OLDES system being installed in the home of the recruited person. The most serious problems (that beyond the control of the project partners) could only be identified when a technician made the first personal visit to the home of the potential service user. The main problems were that a broadband connection was not available in the locality (30% of cases), or that the home TV set could not be connected to the PC provided by the project e.g. because there was no SCART input (21%).

The pilot study in Bologna tested the OLDES telemedicine platform for 10 subjects with CHF. The patients were followed up by the Heart Failure Unit, Cardiology Division, Bellaria Hospital, Bologna. Of these ten patients who were initially selected and agreed to participate in the study, two subsequently backed away from the trial. At the time when the system was being installed, one of the others had a TV set malfunction, and this patient was therefore not recruited. In the case of two more patients, the installation took too much time, so that their results could not be included in the final assessment. Thus, the pilot study has collected data from 5 patients suffering from congestive heart failure. All patients met the inclusion criteria for the Telemedicine Pilot Study: (1) known diagnosis of congestive heart failure, (2) age > 65 years, (3) clinical and therapeutic stabilization for at least 3 weeks.

The telemedicine system tested in Bologna also included the digital administration of a clinical questionnaire designed on the basis of clinical questions to which patients are usually asked to provide answers during their outpatient visits. The television set was used for friendly visualization of the clinical questionnaires. The patients used the weemote®dV remote control switch to select their answers. The answers to the questionnaires were collected by the server, and were displayed on the workstation of the cardiologist. Using this collected information, the cardiologist was able to monitor the patient's lifestyle, and ascertain whether it was appropriate to her/his state of illness. The administration of a digital clinical questionnaire was the most innovative aspect of the telemedicine system tested in Bologna (See Figure 7).

Figure 7. Example of the food diary (food intake monitoring)

cz | en | it Doctor Doctor | logout

OLDES

Web Portal for Doctors

ADD PATIENT | REMOVE PATIENT'S DATA | PATIENTS | FOOD MENU

EDIT | GENERAL | ANAMNESIS | FOOD MENU | GLYCEMIA | BLOOD PRESSURE | WEIGHT | ENVIRONMENT

Jan 2, 2010 12:34:54 PM - Feb 2, 2010 12:34:54 PM << >> month | week | day | 8-hours

Time	Menu	Composite	Name of food	Energy [kJ]	Protein	Fat	Sacharides	Weight [g]	Unit	Description
Daily summary - 1/8/10				9707.0	66.3	20.8	184.5	-	8.52	
1/8/10	Menu summary -			4116.0	0.0	0.0	0.0	-	0.73	
1/8/10 3:33 PM		weighted	Smažený vepřový řízek, bram. Salát	4116.0	0.0	0.0	0.0	73.0	0.73	
1/8/10	Menu summary -			2184.0	19.1	5.6000004	64.4	-	4.28	
1/8/10 3:49 PM		weighted	Mandarinky	197.0	0.9	0.3	10.6	74.0	1.48	
1/8/10 3:49 PM		weighted	Inst. polévka hrachová (Vitana)	1420.0	18.2	5.3	53.8	238.0	2.38	
1/8/10 3:49 PM		weighted	Rohlík, houska (4-5dkg)	567.0	0.0	0.0	0.0	42.0	0.42	
1/8/10	Menu summary -			1630.0	23.7	4.8	61.5	-	0.74	
1/8/10 4:03 PM		weighted	Vita káva	1630.0	23.7	4.8	61.5	74.0	0.74	
1/8/10	Menu summary -			1777.0	23.5	10.4	58.600002	-	2.77	
1/8/10 5:55 PM		weighted	Minerálka light	59.0	0.0	0.0	3.4	174.0	1.74	
1/8/10 5:55 PM		weighted	Chléb žitný Vita	1063.0	5.9	1.4	53.8	62.0	0.62	
1/8/10 5:55 PM		weighted	Šunka dušená	655.0	17.6	9.0	1.4	41.0	0.41	
Daily summary - 1/9/10				3630.0	53.46	26.779999	77.25	-	18.16	
1/9/10	Menu summary -			2283.0	36.2	15.4	62.0	-	3.47	
1/9/10 7:50 AM		weighted	Minerálka light	59.0	0.0	0.0	3.4	5.0	0.05	
1/9/10 7:50 AM		weighted	Minerálka light	59.0	0.0	0.0	3.4	187.0	1.87	
1/9/10 7:50 AM		weighted	Chléb žitný Vita	1063.0	5.9	1.4	53.8	46.0	0.46	
1/9/10 7:50 AM		weighted	Eidam 30% t.v.s.	1102.0	30.3	14.0	1.4	109.0	1.09	
1/9/10	Menu summary -			110.0	0.0	0.0	6.6	-	8.02	
1/9/10 7:59 AM		weighted	Čaj expreso Tea	110.0	0.0	0.0	6.6	802.0	8.02	
1/9/10	Menu summary -			1186.0	16.16	11.08	7.65	-	6.17	
1/9/10 1:19 PM		weighted	Vývar z drůbežího masa	808.0	16.16	11.08	7.65	515.0	5.15	

The Prague pilot study focused on older persons suffering from Type 2 diabetes mellitus (T2DM). Diabetes mellitus is a chronic metabolic disorder characterized by increased glucose levels due to insulin secretion deficiency and decreased sensitivity of the peripheral tissues to insulin effects, which can ultimately lead to the development of acute (hypoglycemia, hyperglycemia) and long-term complications (cardiovascular, cerebrovascular, renal, ophthalmic). In this context, the project offered the users the possibility to control their diet, allowing the carer or the physician to monitor selected physiological functions and glucose excursions online. Diet is a critical aspect of diabetic compensation, since patients are required to strictly adhere to predefined diet composition and energy food intake. Diet adherence is a major problem in most patients with T2DM, and in the elderly in particular. A programmable interactive scale incorporated into the OLDES platform is able to calculate the exact amount of energy and nutrients encompassed in a particular portion of food. The platform's memory stores data on calories, saccharides, proteins and lipids ingested throughout the day and compares it to the maximum daily doses recommended for the particular person. This process has been implemented through the use of wireless scales installed in patients' homes, where the food is weighed and assessed with regard to its nutrition values. As cardiovascular complications are the most common cause of death in diabetic patients, blood pressure measurement, i.e. the simplest cardiovascular parameter, was also adopted into the system. Patients were asked to measure their body weight once a day, and their blood pressure and glucose level three times per day – the results are stored by the system and can be visualized at

Figure 8. Data for one of the patients in the Prague pilot study (male, age: 73 years, monitoring time: 10 days) as provided to the physician. (a) One excessive peak (stress provoked), some minor peaks, average, systolic values slightly above target. (b) ambient parameters: temperature and humidity

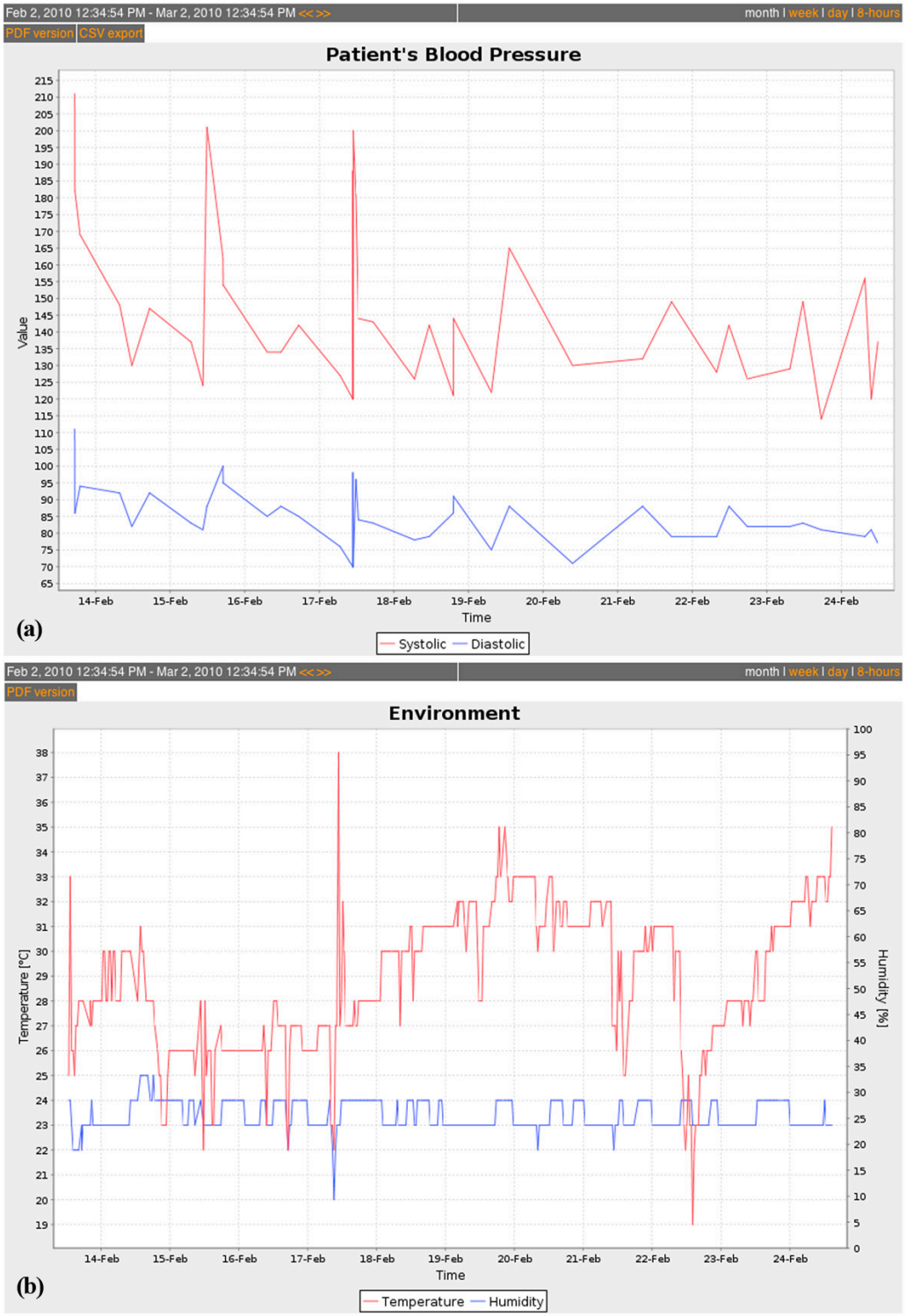


Figure 9. Glucose and blood pressure development

various scales, compare Figure 8 and Figure 9. After the data that has been acquired is sent into the OLDES platform, the system evaluates them using an appropriate algorithm. The system can raise medical alarms if necessary, or can point out some extreme values to the physician, who can make appropriate recommendations.

The data provided valuable insights into the diabetic and blood pressure compensation of each patient, and into their various eating habits. Lists of ingested meals, exactly weighted and with the precise amount of calories and nutrients that they contain, are generally much more helpful to a diabetologist or dietician for identifying major dietetic problems than the largely inaccurate meal lists standardly written down by the patient her/himself. In this way, the patients participating in the pilot can be shown their most common dietetic mistakes – ranging from high energy, saccharide, and fat intake and wrong selection of meals to an insufficient number of meals throughout the day. The blood glucose profiles made it possible to unravel any dangerous patterns in glucose excursions 24 hours a day, as shown in Figure 9a, where trends toward morning and evening hypoglycemia are displayed. This observation resulted in appropriate changes in insulin dosage based on the medical doctor's recommendation: "Reduce night NPH insulin from 12 to 10. Reduce lunch preprandial insulin from 10 to 8. And do not forget to take a second supper at the time when you inject the night insulin dose!"

Blood pressure profiles accessible online also provided an impressive opportunity to identify patients with dangerous high or low pressure peaks, and to react accordingly. Let us mention the case of another diabetic patient with no history of hypertension taking two drugs for lowering blood pressure for other indications: alpha-blockers primarily for prostatic hyperplasia and diuretics for renal insufficiency. Figure 9b shows one excess blood pressure value throughout the monitoring period, with an otherwise acceptable profile, although in the presence of type 2 diabetes mellitus slightly lower blood pressure levels are recommended. On the basis of this observation, the medical doctor recommended: "Continue blood pressure monitoring for 1-2 weeks more; in the case of persistently elevated blood pressure levels consider adding other antihypertensive medicaments – ACE inhibitors or Ca blockers."

4.3. Evaluation and Results

4.3.1. Tele-Accompany

In the pilot project evaluation, we paid particular attention to the users' evaluation of the OLDES home system installations, their feelings and expectations. These elements were recorded through 2 *ad hoc* questionnaires filled in by a facilitator. The first questionnaire (See Appendix 1) evaluates the installation process of the system, and the second (See Appendix 2) wraps up the experience of the user during her/his active participation in the project. 86 questionnaires were completed in Bologna, 58 of them before the system had been installed and 28 after the system had been installed. The average ratings (0-10 low/high satisfaction with and interest in the technology) were high for the questions on knowledge of the technology (8.5), curiosity (9.3), and asking questions (9.5). However, there were low ratings on indifference (1.2), i.e. they were not indifferent, and on concern (2.8), though the level of concern grew a little after installation of the system. Interest in the technology is on a moderate level, slightly little lower before installation of the system (4.8) and growing to almost 6 after installation. The findings from the first questionnaires show high average ratings for the question asking about satisfaction with the OLDES system (7.9), the question about the positive role of OLDES in terms of social contact and leisure time (7.8), and the users' evaluation of the assistance received from the call centre operators (7.9). The evaluation model also included observations, focus groups and interviews with representative users.

The main findings can be summarized as follows: the users found the OLDES system a useful tool for socialization, for being updated and for finding information in everyday life. All of the content was considered important, with a preference expressed for healthcare and wellness topics. The opportunity to use OLDES to get involved and remain involved in local activities is also a fundamental aspect of the service, since lack or loss of participation is perceived by seniors as a basic form of social isolation. It is interesting to note that most the terms employed by the users for describing OLDES seem to perceive the device as a sort of TV channel, possibly because they could not find any other simple but more accurate definition. Consequently, the passivity that typifies TV consumption is the most immediate attitude towards this new medium. Thus it is necessary to stimulate users to utilize the system to more interactively, in order to exploit the socialization potentials of the OLDES platform more fully. Nevertheless, all users had confidence in the project and wanted it to continue. They were confident that the technical difficulties would be resolved. All users hoped to keep receiving useful and updated content that they could consult daily.

4.3.2. Telemedicine

We will now review the perception of the OLDES system by its 10 users in the diabetes pilot study, as evaluated from two points of view, using the questionnaires shown in the Appendix: from the point of view (1) of the person providing technical and training support, and (2) of the opinions of the users. The first questionnaire, under the title "Installation Process Evaluation Form", was completed by the technician after delivering the OLDES system to the user. To deliver the system, the technician had to install all the equipment in the user's home and to provide training for the user. The user could also seek help from the technician while the pilot study was operating.

We present the results for 8 users only, because the remaining two users withdrew from the pilot study after one day (two ladies in their 80s). The technician evaluation results of only eight patients is therefore included. In case of nominal/verbal answers, the scale corresponds to scores from 1 to 10: whereby 1 is the lowest evaluation and 10 is the highest. The original questionnaires are available in Appendix 1 and Appendix 2, and the results are summarized in Appendix 3.

This survey was based on the opinions of 8 users. In order to obtain some statistically meaningful results we point out that it would be necessary to have (1) a minimum pilot duration of one month, and (2) 25-30 users involved in the study. The average age of the users was 75.3 years. In general, the OLDES system was perceived very positively. The first user scored the OLDES system as neutral. One of the reasons for this comparatively low score may be that this was the first pilot study, in the course of which several technical bugs had to be resolved. It is no surprise that this process lowered the reliability of the system in eyes of the user, and caused discomfort for the user. An apparently promising sign from the point of the potential for future deployment is the difference between the answers provided by the users for two questions:

a) How do you appreciate the concept of the OLDES system?

b) How are you satisfied with the OLDES system in general?

It seems that the OLDES experience convinced the users that such a system can have a positive effect: While the average answer to the question a) was 7.1 (See Figure 10a), the average rating for question b) was 8.7 (See Figure 10b).

The users rather quickly got used to navigating in the graphical user interface using a remote controller, and considered it mainly easy to use. However, some users complained about the stiffness of the buttons of the remote controller. The

Figure 10. Evaluation results for OLDES system: (a) how do you appreciate the OLDES system? (b) generally, how satisfied are you with the OLDES system?

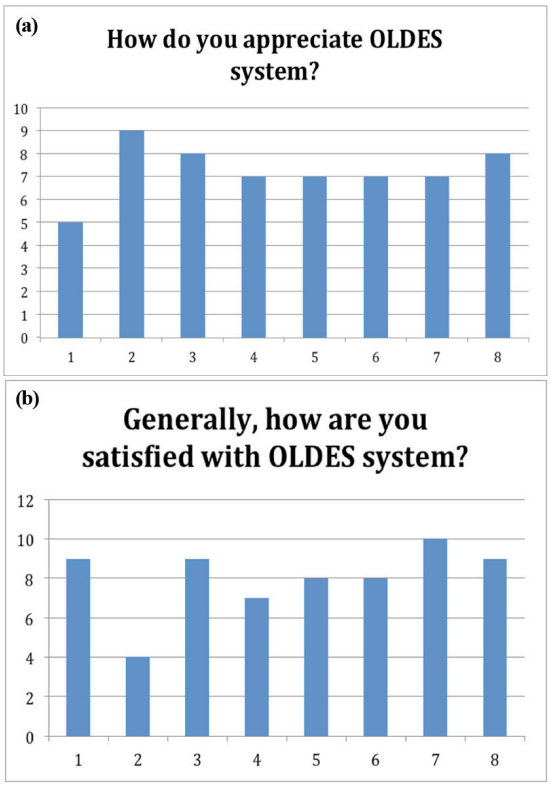
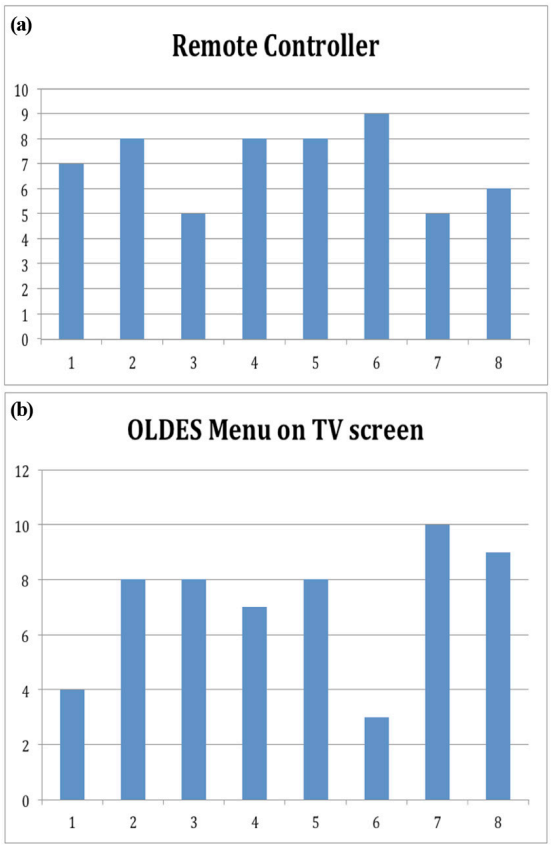


Figure 11. Results of the evaluation of the tangible (a) and graphical user interfaces (b)



average score was 7.0 for remote controller ergonomics (See Figure 11a) and 7.1 for usability of the graphical user interface (See Figure 11b). The only case where dissatisfaction was expressed was with the food selection menu, where the user can choose among more than 2000 food items presented in the OLDES database. To find a more user-friendly solution to the food selection menu problem will be a challenging task for future research. It is interesting to note that most of the users considered as very beneficial the fact that the OLDES system provides a means for inspecting the review of their own data collected in the course of the previous. Surprisingly, fewer than 50% of the users expressed interest in having this collection of data shown in graphical form.

The success of the health monitoring pilot studies from clinical point of view is indicated by evidence that OLDES provides clinicians in the pilot sites with:

- Early indications of potential problems that allow for immediate intervention.
- Reliable information that enables decisions can be made quickly and accurately.

This evidence was gathered from clinicians through interviews and observations in the two pilot sites. It is not possible to quantify this evidence.

Clinical participants from the pilot sites proposed a set of specific health status measures that can be identified by physicians for the patients

participating in the health monitoring pilot studies. In the case of the pilot in Prague for diabetic patients, the set of measures aims to:

- Improve glycaemic control and diabetic compensation parameters, in order to reduce the rate and the severity of long-term diabetic complications and the need for hospitalization.
- Improve the safety of diabetic therapy by reducing the rate of hypoglycaemia, especially during the night.
- Improve blood pressure control to reduce cardiovascular complications.

In the health monitoring pilot (Bologna), anticipated improvements for older persons affected by CHF are:

- Reducing need for hospital re-admission for heart failure.
- Improving the safety of the patient, and the time that the patient can stay in his/her own home.

In Bologna, in the telemedicine pilot study, the preliminary data collected at the beginning and at the end of the study through ad hoc online questionnaires produced the following main results:

1. The OLDES platform is sufficiently flexible and easy to use.
2. The user acceptance and perceived utility and quality of the equipment was also good, even for older patients. Patients needing assistance from caregivers showed lower accceptance.

4.4. Lessons Learned

The technology used by the system must be fully reliable in order to gain the trust and confidence of its users. The OLDES system is a complex technology chain, the reliability of which is determined by its weakest link. Surprisingly, the weakest link proved to be *Internet access*. Even in such large towns as Bologna and Prague there proved to be some locations where the quality of the signal was so unstable that functionality of the system varied dramatically from day to day, and from session to session. This kind of behaviour can greatly undermine confidence in the system, and can result in disappointment. It is therefore most important to check the local quality of Internet access in advance before recruiting potential users. The quality of the Internet signal also needs to be taken into account in the questionnaires used for evaluating the system.

The next problem turned to be the *pushbuttons of the remote controller*. This was perceived by the users as a situation when the system does not react to their actions. This behaviour was reported by several clients, and in most cases it proved due to a discrepancy between the soft touch of the user and the stiffness of the pushbuttons. Technical problems of this kind need to be discovered during the installation phase, and the operator should have some alternative means for dealing with them.

For the control system, it would be good to have the option of working with a touch screen. This will be considered in the follow-up SPES project, see later elaboration. The stability of the platform is also a crucial consideration in the testing phase. Reliable hardware has to be carefully selected prior to implementation of the pilot study. When weighing considerations of stability and cost, it is necessary to have in mind that low cost hardware can lead to low-quality performance. In addition, the use of open-source software can turn out in the long term to be more expensive than the use of a commercial software product. It took us a considerable amount of time to implement voice-IP using the Asterisk system, and also to develop the graphical user interface. Servicing costs must be also taken into account.

Some Psychological Traits: What is taken for granted by a trainer can be perceived by the target

user as something completely new. A sufficient amount of time, a cheerful attitude and true respect and sympathy for the user's difficulties are therefore essential ingredients for making the user like and utilize the OLDES system once it has been installed. A successful interaction aimed at correctly explaining the OLDES platform should take at least 20 minutes, over and above the time required for the installation itself. There is certainly a role for volunteers in explaining how to use the OLDES system.

When *training* the older people to make use of technology, it is often forgotten that a user who may be completely unable to perform straightforward IT operations is nevertheless a senior human being whose greater experience of life needs to be valued, and whose intelligence, culture, powers of observation and sense of humour have necessarily been diminished by age. Users recognize very quickly and clearly when something has gone wrong. When there is a problem, it is important to distinguish between a true technical glitch and a misunderstanding by the user, and to deal with the two kinds of problem accordingly. The user will immediately notice if the operator is insincere (for instance, if she/he tries to conceal technological problems). A good tip for avoiding this kind of risk is always to stress the continuity of care between the current new project that they are participating in and the local health system they they already know and trust. If some technological problems occur, the best solution is to explain them to the user in plain language, showing honest appreciation for their patience and for their willingness to participate in an experiment that will benefit from their help. An effective installation phase should not take too much time, and all the devices should work properly in order not to discourage the potential users.

In all cases, the tests showed that the OLDES system was much more successful when the senior could rely on *support from a relative or neighbour* who offered to help her/him to master the functions of the system. However, relatives who failed to encourage the senior also made a great difference. There were even cases where the family persuaded a potential user not to participate. The social and family situation of the user need to be carefully evaluated, both before she/he is included in the programme and during the installation phase. The use of technology is problematic for this target group, and any independent condition that creates stress can also have effects on the willingness of a senior to learn. However, senior users who could rely on a stable marital situation or on the presence of family members seemed positively curious about the IT challenge, regardless of their age or difficulties.

Topics Popular Among the Users: The Bologna team was surprised to find how much the seniors appreciated the idea of watching videos on the platform. Many users complained about the lack of interesting programs on modern TV channels and radios: uploading clips from old films, songs or sports classics was a highly appreciated initiative, especially when the upload was in response to a specific request. The Bologna team also uploaded and offered to its clients some videos that had been shot at the workshops held with the system users, i.e. with clients of the OLDES system. Most users were unused to seeing their own image on a screen. They were extremely entertained by this option, and were also more excited about participating in subsequent workshops, which offered opportunities for new socialization.

The Prague users of the health part of the project were asked what additional functionalities they would like to have in the next version of the OLDES system. All of them suggested a panic button, and 75% wanted some kind emergency warning system to be built-in (it could draw attention to a range of household events, e.g. a gas leakage or a fire threat).

4.5. Results

4.5.1. Quality of Life: Social Aspects

There is emerging evidence that use of the OLDES system has some positive impact on various dimensions of the users' quality of life. Evidence of this nature has been captured in various forms. For example, the questionnaire-based data collection process undertaken during the OLDES workshops concluded that the users find the system useful for communicating with other people, for fostering socialization and for combatting social isolation. OLDES users also reported feeling a greater sense of security due to the option of calling the CUP2000 call centre, and feeling more informed about local events and issues. In some cases, having greater access to information leads to increased opportunities for users to attend local events. Given the very limited opportunity to collect data on users' experience of being an OLDES system users over a period of time, it is problematic to make any claims about users feeling part of a care community.

Some of the qualitative data collected towards the end of the project suggest that participating in the social processes of the project itself (in addition to the direct effect of being a OLDES system user) has yielded quality of life benefits for some users, particularly for those attached to voluntary associations who have taken part in various OLDES-related activities, and who have generally collaborated closely with the project team.

4.5.2. Cost Saving

A number of cost saving outcomes and improved productivity for service providers were anticipated by the partners at the outset of the project. These included:

- Doctors will spend less time visiting patients who do not need visits.
- Early intervention may prevent hospitalization.
- A carer can monitor more patients than before because of monitoring assistance. and because the monitoring software is able to raise an alarm.

At the same time, it was anticipated that new costs might be incurred, e.g. due to false alarms generated by health monitoring.

The OLDES system has verified the technical feasibility of the telemedicine dimensions of the system system, and that the hardware and software component installed at the users' homes have been accepted by the patients.

A greater number of participants will be required in order to develop the cost-effectiveness of tele-monitoring.

5. CONCLUSION AND FURTHER PLANS

Participating in the OLDES project stimulated older persons to use ICT instruments and solutions, enriched their skills and provided an opportunity for them to interact in the modern world of communications. The project has demonstrated the feasibility of daily use of a low-cost home access point that makes a range of services available, supports social interaction and efficiently manages some problems related to chronic diseases (diabetes and chronic heart failure) in cooperation between the social services and the medical services by monitoring physiological parameters. Experience gained from the pilot runs in several dozens of households has indicated the following directions for further research and development:

1. The OLDES system limited the *control interface* to the remote controller (See Figure

6). This device was very well accepted by some users, though some complained about it. It seems that a touch screen could accommodate the requirements of those who had problems with the remote controller. Ideally, it should be possible to choose the means of control and the user interface to suit the needs of different users.

2. The users in the OLDES system interacted with several medical devices. Some of them required very a simple interaction, e.g. pressing button to start the measurement or to transfer the data. Despite the easy one-step action, some users experienced difficulties resulting in data losses. In future, medical devices should operate automatically, without user control or interaction, in order to ensure *regular monitoring* of a patient's health status.
3. The OLDES system was designed for use in the homes of its target users. However, there are some wealthier target groups whose users would benefit considerably from a mobile application. Can the OLDES solution be generalized to a *mobile setting*?
4. The multimedia content proved to be most popular feature for most OLDES tele-accompany users. In order to retain their interest, the content had to be regularly updated. This turned out to be a rather time-consuming and human resource-consuming activity. To attract users towards the other more sophisticated services offered by the system, it was necessary to encourage and invite them repeatedly. It remains an open question what technical means could be used to simplify the updating of social content, and to *improve the social inclusion* of potential users with very limited IT skills. It is clear that the ultimate solution cannot rely on extra work by the professional support staff.

The follow-up project, SPES, will aim to offer some new solutions for these issues, and it will test them carefully in four pilot runs between 2011-2013. The original OLDES system is currently being adapted for use by mobile users, and with alternative means of control. The new SPES system will be tested on new types of health conditions (e.g. asthma, dementia). In addition, SPES will pay special attention to the problems at the interface between technology and social inclusion, as mentioned previously in item 4.

We will now review the aims of in the planned pilot studies that will be carried out ensured in new locations. The studies will work with each user for a period of at least several weeks.

The *Ferrara pilot* (Italy) is targeted at patients affected by chronic respiratory failure, requiring long-term oxygen therapy and non-invasive mechanical ventilation, who are already under observation from pneumologists for periodic clinical controls. The aim is to provide patients suffering from *breathing problems* with a system that can remotely monitor their health status using a saturimeter and several other non-invasive medical devices, in order to enable more efficient use of medical aids and more efficient organization of the the healthcare service provided by medical staff. A complementary medical Web application will enable clinicians and all other stakeholders to access patients' data and report clinical considerations.

At the same time, this pilot study will also test the use of an alternative control system: the patient will be provided with a touch-screen netbook, which will replace the home gateway provided for the OLDES project. The touch-screen netbook will allow the system to:

- Collect data locally from medical devices.
- Store the data until it can be sent to the central HUB.
- Properly and safely send data to the central HUB.
- Show the user some basic functionalities (e.g. a list of measurements correctly sent).
- Provide an agenda.
- Administer questionnaires.

The communication network is investigating the possibility of using ADSL where already available (i.e. when the user already has an ADSL line) and, in all the other cases, a UMTS/GSM connection.

The *Vienna pilot study* (Austria) will seek tailor-made solutions for people with dementia who face health risks and who are prone to accidents because they tend to get lost in the street, etc. These people often have to move to inpatient care, although they would prefer to continue to have self-determined lives in their own homes. Their families are overburdened by constantly having to care for their disoriented relatives. In addition, demented people do not receive stimulation and motivation appropriate to their needs. If people, for example, confuse. day and night, summer and winter, and breakfast and supper, they need structured and comprehensible input. The overall objectives of the Vienna pilot study will be to enhance the quality of life and the self-determination of older persons with dementia, as the SPES system will make it possible to supervise them without restricting their movement. This should lead to a decrease in the number of admissions to nursing homes due to dementia. In addition, applications will be tested to support reality orientation and memory training for users by offering them visual and audio programmes on screens. In general, the use of this form of technology should raise the degree of security for disoriented persons both in the public space and in everyday life.

The SPES solution aims to incorporate as many as possible of the technical means currently available on the market for the target users. These technical means range from simple products such as automatic pill dispensers and telephones with the photographs of people frequently called, to more complex tracking devices, using Global Positioning Systems (GPS), which also assist in locating people with dementia when they wander off (Maurice & Nugent, 2010).

An emergency alarm system for day-care centres, flat-sharing communities and private homes will be designed and tested. The solution will include environment sensors, e.g. at the entrances and exits of private apartments, so that adequate warnings will be given before the senior leaves home (e.g. a 'talking key' that says 'Please take me with you' or a stove that says 'Please turn me off'). The resulting system should lead to more security for the senior, reducing her/his anxiety and supporting her/his autonomy. The intention is to enrich the everyday lives of seniors.

For relatives, it should provide relief from their burden, reduce stress, and improve relationships within families. The pilot should support their work, as carers can concentrate better on their true tasks of assisting the protagonists instead of constantly supervising them. The results should meet legal responsibility requirements. For caregivers, the pilot should represent a new professional experience, broadening their methodological approaches and lessen their burden of responsibility.

Another user group requiring mobile application will be that of the *Boskovice pilot* (Czech Republic) that will be focused on 40 mobility-impaired persons followed by a non-governmental organization called DEEP. These persons are often illiterate in the field of assistive devices and monitoring services including e-health concept. Therefore, as the first step, courses and appropriate training will be provided in a range of registered social services motivating seniors and disabled persons in the field of communication services, education, social and professional integration. DEEP team will select 40 persons that will take part in the SPES pilot testing of services based on monitoring through sensors and detectors. The location is, predominantly, the city of Brno in the Czech Republic. It will be necessary to i) inform, ii) motivate, and iii) to teach people how to use appropriate tools and sensors. DEEP is anticipating addressing at least 200 persons from which 40 will be chosen for pilot testing purpose using motivation/health questionnaires. Based on the analysis and the specific needs of their households, the system users will be provided with the individual-specific peripherals such as identification bracelets for disoriented people,

identification of dangerous inclinations, local remote control of movement, supervision of household activities (on/off household appliances, windows and door opening etc). Each of these peripherals currently relies on GSM and GPRS data stream based on IP gateway and SMS. This ensures two-way transmission, possibility of the remote setting, and emergency operations from the control centre via the GSM gateway.

The *Kosice pilot* (Slovakia) will be devoted to the problems related to the social aspects of technology development as described in the item 4. The main aim will be to ensure means that can improve social inclusion of older people through suitable ICT solutions designed and developed within SPES project. The older persons will be able to enrich their daily routine by various communication and social features for spending leisure or mediating psychological support. The offered solution will be based on detailed analyses of existing opportunities, requirements and expectations from target group of seniors who will be offered a chance to remain in contact with their relatives, neighbors or some other people with similar hobbies or health problems. The pilot locality will contain small flats or senior house under responsibility of Municipality of Košice. The user-friendly environment will offer interesting functionalities for stimulation of interactions between or with older persons as discussion forum, personal or group instant messaging, emails and virtual meetings. These basic features could be extended with opportunities of dating, exchanging information or photo sharing. The social services will be mediated by appropriate Windows-based platform, designed and implemented directly to Kosice pilot requirements. Suitable methods and approaches from domains of Data mining and Social Network Analyses will be applied to evaluate collected historical data in order to customize the user's environment (list of people with similar hobbies, list of people with similar daily habits, etc.) or to identify interesting and useful information for future improvement. Technical University of Košice will be responsible for relevant software modules and technical support during pilot realization.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support provided by the following grants: STREP IST-2005-045282-AAL OLDES (Older People's e-services at home), SPES (Contract Number 3CE286P2): Support Patients through E-services Solutions - CENTRAL EUROPE Programme co-financed by the ERDF, and by Ministry of Education, Youth, and Sport of the Czech Republic with grant No. MSM6840770012 titled "Transdisciplinary Research in Biomedical Engineering 2."

REFERENCES

K4Care. (n.d.) *Knowledge-based homecare eservices for an ageing Europe*. Retrieved November 14, 2011 from http://www.k4care.net

Attentianet. (n.d.). *Advanced tele-assistance convergent network for chronic disease patients and elders project*. Retrieved November 14, 2011 from http://www.attentianet.eu/

Dubuc, L., & Edge, D. (2004). TUIs to ease: Tangible user interfaces in assistive technology. In *Proceedings of the 2nd Cambridge Workshop on Universal Access and Assistive Technology*. IEEE.

Enable. (n.d.). *Ambient assisted living solutions based on advanced ICT technologies*. Retrieved November 14, 2011 from http://www.enable-project.eu/

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2004). *Designing for elderly adults: Principles and creative human factors approaches*. London, UK: CRC Press. doi:10.1201/9781420023862.

Gonzalez & Vlez. (2008). Confidence: Support to elderly peoples independent living. In Proceedings of Psykologi. Psykologi.

Hardicker, N. R., & Grant, M. J. (2010). Factors that influence public engagement with eHealth: A literature review. *International Journal of Medical Informatics*, *80*, 1–12. doi:10.1016/j.ijmedinf.2010.10.017 PMID:21112244.

Inglis, S. C., Clark, R. A., McAlister, F. A., Ball, J., Lewinter, C., & Cullington, D. et al. (2010). Structured telephonesupport or telemonitoring programmes for patients with chronic heart failure. *The Cochrane Library*, 8.

Intel. (n.d.). *Health guide PHS6000*. Retrieved November 14, 2011 from www.intel.com/healthcare/telehealth

Jackson, C. L., Bolen, S., Brancati, F. L., Batts-Turner, M. L., & Gary, T. L. (2006). A systematic review of interactive computer-assisted technology in diabetes care interactive information technology in diabetes care. *Journal of General Internal Medicine*, *21*(2), 201–202. doi:10.1007/s11606-006-0261-2 PMID:16606384.

Macík, M., Hanzl, M., Klíma, F., & Slováček, M. (2009). Digital household user interfaces for elderlys. In *Ambient Intelligence Perspectives II, Selcted Papers from the Second International Ambient Intelligence Forum*. Amsterdam: IOS Press.

Malý, I., Cuřin, J., Kleindienst, J., & Slavík, P. (2008). Creation and visualization of user behavior in ambient intelligent environment. In *Proceedings of the 12th International Conference on Information Visualisation (IV08) - International Symposium of Human-Computer Interaction*. Los Alamitos, CA: IEEE Computer Society.

Martin, M., Ponsard, C., Walsh, S., Baines, S., Rousseaux, S., Rinaldi, G., & Tamburriello, F. (2008). OLDES: Designing a low-cost, easy-to-use e-care system together with the stakeholders. In *Proceedings of the International Conference on Computers Helping People with Special Needs*. IEEE.

Maurice, D., & Nugent, D. (2010). *Supporting people with dementia using pervasive health technologies*. Berlin: Springer.

Mayhew, D. J. (1999). *The usability engineering lifecycle: A practitioner's handbook for user interface design*. San Francisco: Morgan Kaufmann.

McLoughlin, I., Maniatopulos, G., Wilson, R., & Martin, M. (2009). How to die before you get old? *Public Management Review*, *11*(6), 857–880. doi:10.1080/14719030903319002.

Poulson, D., Ashby, M., & Richardson, S. (n.d.). USERfit: *A practical handbook on user-centred design for rehabilitation and assistive technology*. Retrieved November 14, 2011 from http://www.iris-design4all.org/

Rinaldi, G., Martin, M., & Gaddi, A. (2011). Establishing an infrastructure for telecare: Combining the socio-technical and the clinical in handbook of digital homecare: Successes and failures. Communications in Medical and Care Compunetics, 3..

Sensaction-AAL. (n.d.). *Sensing and action to support mobility in ambient assisted living*. Retrieved November 14, 2011 from http://www.sensaction-aal.eu/

Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces*. San Francisco: Morgan Kaufmann.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. San Francisco: Morgan Kaufmann.

Stroetmann, V. N., Hüsing, T., Kubitschke, L., & Stroetmann, K. A. (2002). The attitudes, expectations and needs of elderly people in relation to e-health applications: Results from a European survey. *Journal of Telemedicine and Telecare*, *2*, 82–83. doi:10.1258/135763302320302154 PMID:12217149.

The Department of Economic and Social Affairs of the United Nations Secretariat. (2010). *The 2009 revision and world urbanization prospects*. Retrieved November 14, 2011 from http://esa.un.org/unpd/wup/doc_highlights.htm

Weemote. (n.d.). *Weemote®dV remote controller*. Retrieved November 14, 2011 from http://www.weemote.com/products.html

ADDITIONAL READING

Chan, M., Campo, E., Laval, E., & Estève, D. (2002). Validation of a remote monitoring system for the elderly: Application to mobility measurements. *Technology and Health Care*, *10*(5), 391–399. PMID:12368559.

Hancock, G. A., Reynolds, T., Woods, B., Thornicroft, G., & Orrell, M. (2003). The needs of older people with mental health problems according to the user, the carer, and the staff. *International Journal of Geriatric Psychiatry*, *18*, 803–811. doi:10.1002/gps.924 PMID:12949848.

Isaia, G., Bo, Nobili, G., Cappa, G., Mondino, S., Pilon, S., & Massaia, M. (2009). Costs of the in-home patients affected by dementia. *Archives of Gerontology and Geriatrics*, *49*(1), 147–151. doi:10.1016/j.archger.2009.09.024 PMID:19836628.

Orpwood, R., Gibbs, C., Adlam, T., Faulkner, R., & Meegahawatte, D. (2005). The design of smart homes for people with dementia – User interface aspects. *Universal Access in the Information Society*, *4*(2), 156–164. doi:10.1007/s10209-005-0120-7.

Park, D. C. (1992). Applied cognitive aging research. In Craik &Salthouse, (Eds.), The Handbook of Aging and Cognition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reinersmann, A., Lauriks, S., van der Roest, H. G., Meiland, F. J. M., Davies, R. J., & Moelaert, F. et al. (2007). Review of ICT-based services for identified unmet needs in people with dementia. *Ageing Research Reviews*, *6*(3), 223–246. doi:10.1016/j.arr.2007.07.002 PMID:17869590.

KEY TERMS AND DEFINITIONS

Ambient Assitive Living: Is set of tools how to help to ensure health, safety, and well-being of older people.

Food-Intake or Consumption: Is sum of food consumed by a person. Control of food intake is very important for managing blood glucose levels in diabetes mellitus patients.

Social Services: Generally consist of counseling and psychotherapy, human services management, social welfare policy analysis, policy and practice development, community organizing, international, social and community development, advocacy, teaching, and social and political research. In our approach, the social services addressed communication among participants, broadcasting entertainment content over Internet or invitations to culture events.

Tangible User Interface (TUI): Is a user interface in which a person interacts with digital information through the physical environment. In the OLDES projectt remote controller was used while in case of the SPES project touch screen was applied.

User-Centered Design: Is the process of designing a tool, such as a Website's or application's user interface, from the perspective of how it will be understood and used by a human user.

APPENDIX 1

INSTALLATION PROCESS EVALUATION FORM QUESTIONNAIRE

Evaluating the Installation Process and the User´s First Impression

To be filled in by an OLDES representative, not a user.

Id_____ Telemedicine user

Visit: a to install the OLDES system b _____ days after the installation of the OLDES system

Evaluation following the visit referred to in option a

- What are the user´s expectations?
- Why has the user agreed to take part in this study?

Observations made during the installation:

- Does the elderly person show an interest in technology? ☐ No ☐ Yes
- Does he/she have any knowledge of technology? ☐ No ☐ Yes
- Is he/she apprehensive? ☐ No ☐ Yes
- Is he/she curious? ☐ No ☐ Yes
- Is he/she indifferent? ☐ No☐ Yes
- Does he/she have any questions? ☐ No ☐ Yes
- Does the user seek some support (e.g. in everyday tasks, food delivery)? ☐ No ☐ Yes
- Does he/she expect a financial reward? ☐ No ☐ Yes

Evaluation following the visit referred to in option b

- How many days has the person been using the OLDES system? _____
- Has the user been recording his/her results on paper (for later checks)? ☐ No ☐ Yes
- Is the user satisfied with the system? ☐ No ☐ Yes
- Has the system boosted the user´s self-confidence? ☐ No ☐ Yes
- Does the user find the system useful? ☐ No ☐ Yes

Circle reasons why the OLDES system seems to benefit the user e.g. the system makes him/her feel more secure; the user is familiar with similar services and is interested in them; the OLDES provider is based near the user´s home or is territorially ´accessible´.

APPENDIX 2

USER EVALUATION OF OLDES SYSTEM

Questionnaire Evaluating End-User Satisfaction with the Platform for Oldes System Telecommunication Services

Id_____ Telemedicine user ___ Face-to-face interview ___ Filled in by the user

1. Sex ☐ M ☐ F
2. Age _____
3. Who do you live with?
   ☐ I live on my own.
   ☐ Partner (spouse)
   ☐ Carer
   ☐ Son/daughter
   ☐ Relative/family member not stated above
   ☐ Person not stated above

4. Who helps you with everyday duties (shopping, preparing food, etc.)?
   ☐ Nobody
   ☐ Partner (spouse)
   ☐ Carer
   ☐ Son/daughter
   ☐ Relative/family member not stated above
   ☐ Person not stated above

5. What is your overall impression of the OLDES system?

Mark the score you would give to the system on a scale from 1 (very poor = the lowest score) to 10 (an excellent system) with X (Box 1).

Box 1.

Points	1 (poor)	2	3	4	5	6	7	8	9	10 (excellent)
Evaluation										

6. Have you found it easy or difficult to use the following tools used by the OLDES system?

USING THE REMOTE CONTROL

Box 2.

Points	1 (difficult)	2	3	4	5	6	7	8	9	10 (easy)
Evaluation										

USING THE ON-SCREEN MENU OF THE OLDES SYSTEM

Box 3.

Points	1 (difficult)	2	3	4	5	6	7	8	9	10 (easy)
Evaluation										

USING ´THE PRINTED GUIDE - OLDES SYSTEM MANUAL´

Box 4.

Points	1 (difficult)	2	3	4	5	6	7	8	9	10 (easy)
Evaluation										

7. Select further services the OLDES system should also offer:
   - ☐ Button (on the remote control) to summon help
   - ☐ Sensor to detect emergencies in the home, e.g. a gas leak, a water tap running inadvertently, a fire
   - ☐ Phone calls made free of charge
   - ☐ Information on frequently used phone numbers
   - ☐ Up-to-date information on activities in the vicinity of your home (e.g. opening hours, surgery hours, a timetable of favourite tram lines…). State pieces of information you are interested in.
   - ☐ Medication administration reminders
   - ☐ Switching lights off/on operated centrally (in selected rooms)
8. What other medical appliances do you wish to include (in the system)?
9. What would you want the OLDES system platform to include in the future?

Would you like to receive a summary of particular recorded data? ☐ YES ☐ NO

If yes, select categories you are interested in (Box 5).

Box 5.

Recorded data	Blood pressure	Weight	Daily calorie intake	Glycaemia	Room temperature	Room humidity levels

Select which period this summary should cover:

☐ Day
☐ Week
☐ Month

Circle your chosen way of presenting the requested data (See Table 2).

Table 2.

Chart

Table

Date	Weight	Glycaemia
1.3.2010	68	5.6
2.3.2010	69	7.3
3.3.2010	68	5.9
4.3.2010	70	8.3
5.3.2010	69	4.8
6.3.2010	68	5.2
7.3.2010	69	6.1

10. Have you ever experienced problems with any of the OLDES system components?
    ☐ YES, with the on-screen menu of the OLDES system.
    ☐ YES, with the OLDES system remote control.
    ☐ NO

11. If you experienced some problems with the OLDES system, who did you ask for help?
    ☐ The technician who installed the OLDES system.
    ☐ Friends or relatives
    ☐ Somebody else (please specify) _____
    ☐ Nobody

12. Score the comprehensibility of the instructions you received during the installation of the OLDES system (Box 6).

Box 6.

Points	**1 (poor)**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 (excellent)**
Evaluation										

13. Do you find the OLDES system manual useful (Box 7)?

Box 7. Score the usefulness accordingly

Points	**1 (poor)**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 (excellent)**
Evaluation										

14. What is your overall satisfaction with the OLDES system (Box 8)?

Box 8.

Points	**1 (poor)**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10 (excellent)**
Evaluation										

15. Why are you satisfied/not satisfied with the OLDES system?
16. What improvements to the OLDES system would you recommend?

APPENDIX 3

EVALUATION RESULTS OF OLDES SYSTEM

Table 3.

User Evaluation of OLDES System									
ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**AVG**
Sex	M	M	F	F	M	**M**	F	F	
Age	82	85	73	74	84	**55**	80	69	75,3
Household Sharing	With Partner	With Partner	With Partner	Alone	With Daughter	With Partner	With Daughter	With Relative	
Who helps you with daily duties (shopping, cooking, etc.)	Partner	Partner	Partner	Nobody	Daughter	Partner	Daugher	Relative	
How do you appreciate OLDES system	5	9	8	7	7	7	7	8	7,3
Do you consider easy/difficult using OLDES component?									
Remote Controller	7	8	5	8	8	9	5	6	7,0
OLDES Menu on TV screen	4	8	8	7	8	3	10	9	7,1
Using paper manuals	6	8	3	6	8	7	4	3	5,6
Choose further functionalities which OLDES system should offer?									
First aid/ emergency button	X	X	X	X	X	X	X	X	
Sensor for emergency household event detection (gas leak, fire, etc.)		X	X	X	X		X	X	

continued on following page

Table 3. Continued

User Evaluation of OLDES System									
ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**AVG**
Free phone calls			X	X			X	X	
List of recent calls				X			X	X	
Actual information (timetables, doctor schedule, shop schedule, etc.)				timetable				wheather, shopping timetable	
Pills reminder					X	x		X	
Central light control (switch off/on)		X			X		X	X	
Which further medical devices do you recommend?			More dietic items		Sleep quality, daily activity & snaps, memory training games		Heart beat, ECG, camera monitoring	ECG, telephone calls	
What else should OLDES system contain?	Blood pressure, glycemia, weight	Blood pressure, glycemia	Blood pressure, glycemia, weight, calorie intake	Blood pressure, glycemia, weight, calorie intake	Blood pressure, weight, calorie intake	Blood pressure, glycemia, weight	Blood pressure, weight, calorie intake	Blood pressure, weight, temperature, glycemia	
Choose time span of your interest	Week	Week	Week	Month	Week, month	Week		Day, month	
Choose type of presentation		Table	Table	Graph	Table	Table	Graph	Graph	
Did you experience problems with any OLDES component	Yes – remote controller	Yes - menu	Yes - menu	Yes – internet	Yes – controllder	Yes – menu	Yes – menu, controller	Yes – controller	
Who helped you in case of problem?	Technician	Technician	Technician	technician	Relatives	Technician	Relatives	Technician, relatives	
Was provided training clear?	10	10	9	9	8	8	10	9	9,1
Is OLDES system useful?	10	8	9	8	9	6	4	5	7,4
Generally, how are you satisfied with OLDES system?	9	4	9	7	8	8	10	9	8,0

31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Diabetes management in OLDES project

D. Novák, M. Uller, S. Rousseaux, M. Mráz, J. Smrž, O. Štepanková, M. Haluzík, M. Busuoli

***Abstract*— EU project OLDES (Older People's e-services at home) develops easy to use and low cost ICT platform in order to offer a better quality of life to elderly people directly in their homes through innovative systems of tele-accompany, tele-assistance and tele-medicine. The elderly are able to access the services and send relevant medical data from their home by being connected to the central server via a low cost PC which is based on Negroponte paradigm. The OLDES platform interface uses television screens controlled through a remote control customized for the elderly. The feasibility of OLDES project is evaluated by the pilot study concentrating on compensation of diabetic patients. Compensation of diabetes is achieved by monitoring glucose glycemia level, blood pressure and weight. Moreover, the patient feeds into OLDES system daily consumption of food using interactive food scales and obtains advice if necessary.**

I. INTRODUCTION

The consequence of increasing life expectancy and decreasing birth rates is that population is becoming increasingly older [1]. Several projects deals with elderly disease management as CONFIDENCE [2], ATTENTIANE[3], ENABLE[4], K4Care[5]. Very similar platform to OLDES is developing INTEL under name Health Guide PHS6000 [6].

The aim of the project is to offer new technological solutions to improve the quality of life of older people in their homes. Thanks to an advanced and innovative technological platform, OLDES provides user entertainment services, through easy-to access thematic channels and special interest forums supported by animators, health care facilities based on established Internet and Tele-care communication standards. However, one of the major problems in IT-assisted ambient technologies is patient motivation and daily involvement in disease management systems [7]. OLDES promotes healthy life style by providing audio-visual content which helps in patients' social inclusion and their motivation. The content is created by social workers or volunteers which are often senior themselves. Information about constant monitoring of vital parameters and health state is surveyed in a call centre by GPs.

The following reviews suggest that computer-assisted interactive IT is an important tool and has great potential to improve diabetes care [8],[7].

This paper presents OLDES technical solution and the implementation of OLDES system project in management of patients suffering from diabetes.

D. Novák, O. Štepanková, M. Uller, J. Smrž, is with Department of Cybernetics, Czech Technical University in Prague, Czech Republic, xnovakd1@labe.felk.cvut.cz, S. Rousseaux is with Centre dExcellence en Technologies de lInformation et de la Communication, Belgium, M. Mráz and M.Haluzík is with Third Department of Medicine, Charles University, Czech Republic, M. Busuoli is with Ente per le Nuove Tecnologie, l'Energia e l'Ambiente-ENEA, Brussels, Belgium

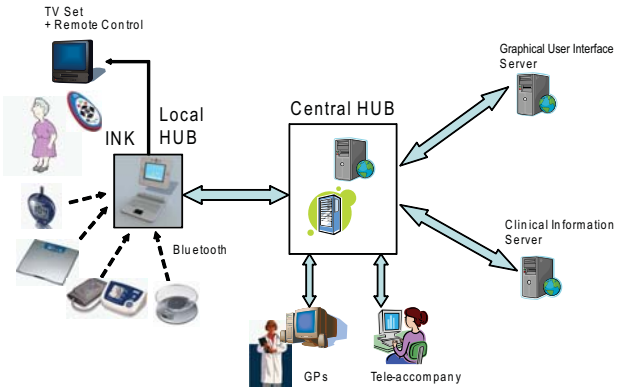


Fig. 1. Oldes architecture

The OLDES architecture consists of two main parts as depicted in Fig.1: *Local Hub* collects physiological data measured by sensors and communication channels implemented by Voice-IP. The information is collected by low-cost INK Laptop and sent via a secured ADSL link to *Central Hub*. Each senior has a health agenda stored in Central Hub. The senior continuously receives social-medical support by medical doctors, social workers and volunteers situated in the call center. Healthy life style is promoted trough an adequate scheduling of thematic entertainment channels. The feasibility of OLDES concept is evaluated under cardiomyopathy and diabetes pilot projects.

OLDES implements the following functionalities:

- Low Cost computer-based system: The elderly persons are provided with a low cost computer based system (the INK computer) which works as the access point for OLDES functionalities and services. The computer is connected to a classic television set which displays all the information provided by the platform by a graphical user interface (GUI).
- Adapted graphical user interface: The graphical user interface is especially designed to meet elderly person usability requirements. Its content is fully dynamic and it is based on web technologies. To select the options and access contents, elderly have an easy to use remote control.
- Tele-health monitoring system: The elderly persons are provided with communicant Bluetooth medical devices. The INK computer installed in elderly person houses automatically collects the data measured by these devices and sends them to a central repository in a secured way.
- Web portal: A web portal which provides interfaces for system administrators, GPs and professionals, discus-

sion groups animators and tele-accompany members.

II. Methodology

Cost efficiency is a major goal of the OLDES project therefore all implemented software in Local and Central Hub is open-source.

A. Local HUB:Home system

This section describes the hardware and software components of the home system located on the left side of Fig. 1.

On the hardware side the home system is composed of: i) a computer based set-top box adapted to elderly persons needs and called the INK computer, ii) an easy-to-use remote control and its receiver iii) an audio handset, iv)a set of medical devices communicating with the INK through Bluetooth or USB interface.

1) INK computer: The INK computer is an x86-based computer built on a VIA C7-M platform produced by Canadian company INK-media. This computer is based on Negroponte's principle: it has low cost ($200US), fanless and diskless operation with no mechanical moving parts, and full-featured while maintaining a small form factor suitable for use as a set-top box. The INK computer operates with a Linux Ubuntu operating system running OLDES software tools.

2) Remote Control: OLDES substituted classical keyboard with a remote controller in order to ease interaction with the system. The main idea of the Infra-Red remote control is to ensure functionality combining features offered by tele-tex (familiar from the standard TV sets) and by a joystick-like navigation system. The Infra-red Remote Control consists of two devices: 1) Remote Control keypad with IR transmitter 2) USB Infra-red Receiver. The first device is a final consumer product Weemote dV manufactured by American vendor Fobis Technologies. The second device is a specific USB dongle with two main competences: a) receive, demodulate and decode infra-red signal b) communicate over USB and emulate a regular keyboard's key presses - see Fig. 2.

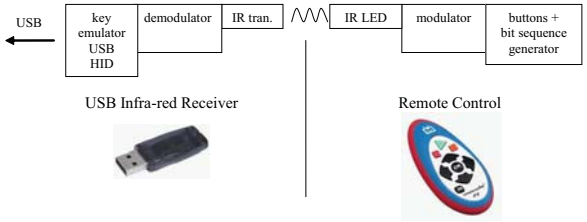


Fig. 2. Remote Controller

USB IR Receiver is implemented in Atmel's ATmega8 microcontroler. IR waveform is sampled by built-in UART peripheral. USB protocol is completely handled by software. The receiver is powered from the bus. No additional software is needed since the USB HID keyboard is a cross-platform spread standard.

3) Medical Devices: We have defined for diabetes pilot project list of patients parameters we have to monitor using medical devices listed in Table I.

TABLE I
Physiological values gathered during diabetes project

Measurement	Producer	Communication
glucose level	Lifescan OneTouch Ultra	Bluetooth
blood pressure	A&D UA-767	Bluetooth
personal weight	A&D UC-321	Bluetooth
food intake	Salter 1016	USB

The selected medical devices are easy to use by an old person and as less invasive as possible. This signifies that the medical devices are based on an easy to use user interface which requires only a few actions from the elderly person to operate properly. Lifescan glucometer doesn't provide an integrated Bluetooth communication interface. To ease device usage by the elderly persons and to ease its integration to OLDES system we developed an smart Bluetooth module, called WAND, which is able to communicate with the glucometer, collect the measurement in its memory and send theses data to the INK computer wirelessly. It has a very simple user interface composed of a simple button. When a measurement is taken, the user has just to press this button to send the data to the INK through Bluetooth.

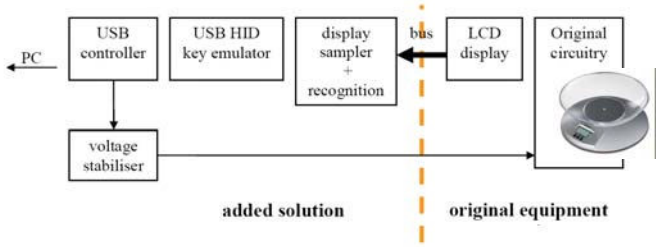


Fig. 3. Kitchen scales implementation

4) Supportive Home Appliances: The diabetes pilot project offers the users a possibility to control their diet, namely the input of their sacharids. The user selects the type of food which is currently on the scales using the remote controller to find out its nutrition values. A kitchen wireless scales based on Salter commercial scales were constructed. Primarily, the scale has no means of transferring measured data to any other electronic device. A custom solution that enables the scale to communicate with a personal computer is based on unidirectional USB connection.

Fig. 3 shows the data flow from scale electronics to the personal computer. Path is fully digital, thus no additional error of measurement occurs. Since the LCD display is used as a gateway to the original circuitry, value presented to user over USB is always the same as the value on the LCD. The accuracy of measurement guaranteed by the manufacturer is not influenced by added electronics.

5) Health Monitoring Software: The high level functionalities of this software tool chain are to collect health data

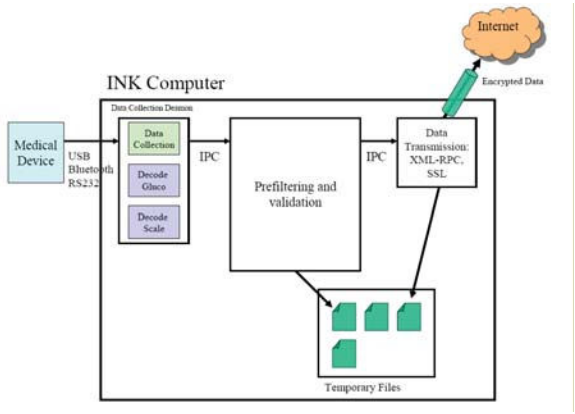


Fig. 4. Health Monitoring Software Architecture

measurements from various medical devices communicating using different technologies (Bluetooth, USB, RS232,), to pre-analyse these data and to send the measurements to OLDES central HUB through a secured encrypted internet connection. These operations don't require any human interaction.

Health Monitoring System is composed of several independent software modules communicating together through inter-process communication channels (IPC) - see Fig. 4. We have efficiently separated the data collection software modules from the data processing and data transmission modules.

Health Monitoring System architecture is basically composed of 3 main software modules:

- Data collection software which is device specific. Indeed, the different medical devices we selected communicate using not compatible data protocols, we need to analyse and decode the received data string. This is done by separate software function symbolised in Fig. 4 by boxes Decode Gluco and Decode Scale. There are as much decoding functions as different medical devices. The decoded data are encoded in a unified protocol and sent to the Prefiltering and Validation software module through an inter-process communication channel (IPC)
- Prefiltering and Validation software. This software module receives the data collected by the data collection software through an IPC and i) analyse the collected measurement data to avoid sending wrong data to the central, ii) ensure INK internal data security by copying the received measurement in INK non-volatile memory to avoid losing measurements if the internet connexion is down and the INK crashes.
- Data Transmission software. This software module sends the measurements to the central system through an SSL encrypted internet connection. It implements the XML-RPC (Remote Procedure Call) Webservice protocol which is a standardised protocol enabling software to transfer structured XML-based data over an HTTP communication.

6) Graphical User Interface Client: The aim of the OLDES project is to provide a complex of new interactive IT tools supporting independent life of seniors. Since all the envisaged tools count upon active involvement of their users, it is clear that the process cannot be reduced to development of novel IT functionalities. It is most important to ensure that the intended users are able (and willing) to access the offered functionalities in a simple way - this is the purpose of the graphical user interface (GUI). The interface is controlled by the remote controller only. The GUI framework is implemented as a standard client-server approach. GUI client, which runs on INKs, connect to a central GUI server to fetch resources and dialog pages. The client is running in web browser and via AJAX communicates with the server. In contrast to traditional web applications, the clients look and behave similarly to desktop applications based on an user-interface toolkit like Java AWT or Swing -see Fig. 5 for examples. The GUI client is written in Javascript and uses SVG (Scalable Vector Graphics) for rendering GUI pages and controls. The look&feel of GUI is configurable by means of skins.

B. Central HUB: call center

This section describes the hardware and software components of the call center located in Fig. 1.

1) Graphical User Interface Server: The server part is implemented in PHP and is responsible for handling client requests and generating GUI pages. Generated pages are sent as XML to the client, where the XML is parsed and the GUI controls are instantiated. The GUI server contains also basic support for content management, representing the content as set of content nodes. The examples of supported content types are welcome screen, menu, article and contact list - see Fig. 5. Each content type is defined in its own PHP generator template; the designed API makes it simple to add new content types.

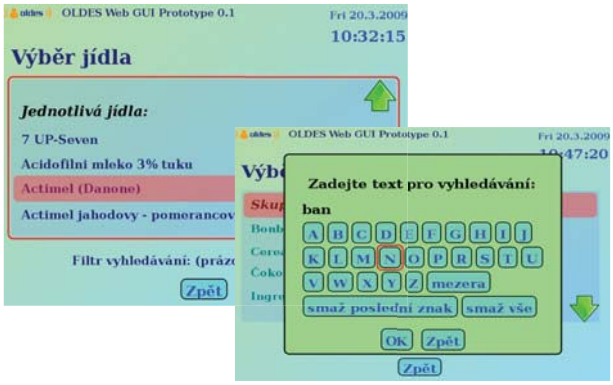


Fig. 5. Food selection using order list or virtual keyboard

2) Clinical Information Server: This server implements an XMLRPC-based webservice which collects the measurement data sent by the INK computers. When data are received, this webservice inserts them in a centralised data base and sends back an acknowledgement to the INK. This webservice is implemented using standard and open-source

technologies like PHP and Apache2 server. All necessary information gathered during clinical pilot is accessible via web clinical portal. Physicians can mainly inspect physiological values and food intake in case of diabetes pilot project - see Fig. 7. The server is implemented using Java Tomcat Servlet container and Apache Struts 1 web application framework.

III. Results and Discussion

A. Diabetes Pilot Project

The main goal of diabetes's pilot is to achieve better compensation of diabetes in hard-to-compensate patients by flexible individualized approach to insulin dose adjustment. The system is developed system for monitoring of physiological functions and self-diagnostics of diabetes in form of advisory system for a patient who is "feeling bad"- this situation is often caused by changes of arterial blood pressure (bad compensated hypertension, hypotension caused by high dose of antihypertensives).

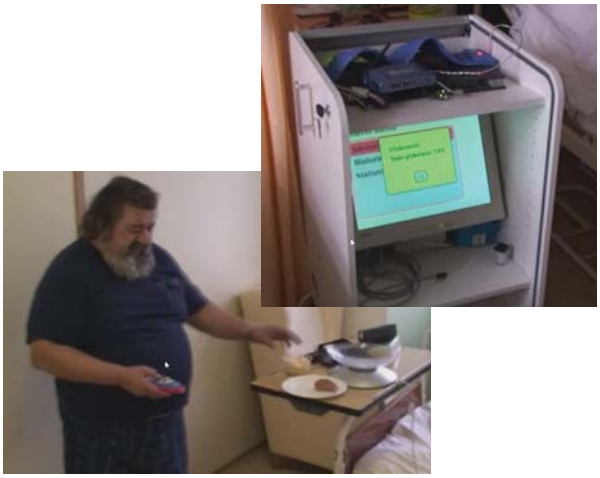


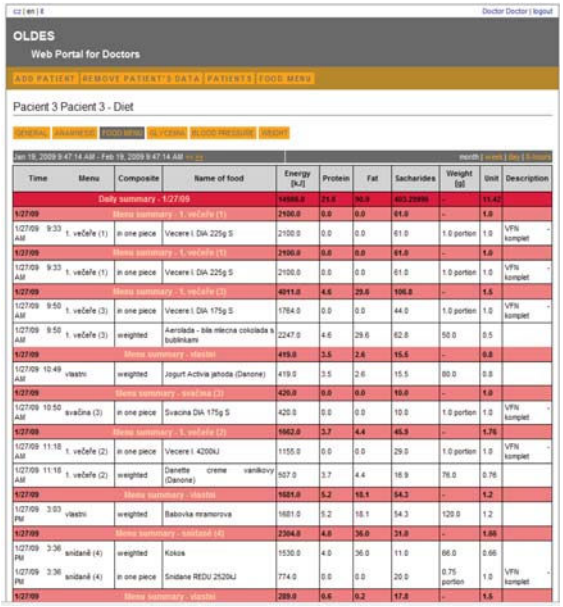
Fig. 6. Example of food intake. Patient must select food item before food scaling using the remote controller

Patients measure its weight once a day, blood pressure three times per day, glucose level three times per day. After the data are transmitted via low cost INK computer to the OLDES central node, the physician can make recommendations of further procedures as acute intervention, eventually modification of chronic medication at home conditions without necessity of stay in the hospital, etc. Another part of OLDES is based on diabetic diet with restricted amount of sacharides and movement activities.

These patients can use interactive scale connected to a computer database of sacharides amount in frequent foodstuff - see Fig. 6. Automatic computation of total daily consumption simplifies patient's control of sacharides intake. The resulting information should be exported to a dietitian who can suggest recommendations for modification of patient's diet - see Fig.7.

IV. Conclusions and Future Works

The expected results of the project and of the diabetes pilots can be resumed as the following: i) providing an open



OLDES

Web Portal for Doctors

Pacient 3 Pacient 3 - Diet

Time	Menu	Composite	Name of food	Energy [kJ]	Protein	Fat	Sacharides	Weight [g]	Unit	Description
Daily summary - 1/27/09										
1/27/09	Menu summary - 1. večeře (1)			2100.0	0.0	0.0	61.0	-	1.0	
1/27/09 9:33 AM	1. večeře (1)	in one piece	Vecere I. DIA 225g S	2100.0	0.0	0.0	61.0	1.0 porton	1.0	VFN komplet
1/27/09	Menu summary - 1. večeře (1)			2100.0	0.0	0.0	61.0	-	1.0	
1/27/09 9:33 AM	1. večeře (1)	in one piece	Vecere I. DIA 225g S	2100.0	0.0	0.0	61.0	1.0 porton	1.0	VFN komplet
1/27/09	Menu summary - 1. večeře (3)			4011.0	4.6	29.6	106.0	-	1.5	
1/27/09 9:50 AM	1. večeře (3)	in one piece	Vecere I. DIA 175g S	1764.0	0.0	0.0	44.0	1.0 porton	1.0	VFN komplet
1/27/09 9:50 AM	1. večeře (3)	weighted	Aerolada - bila mlecna cokolada s bublinkami	2247.0	4.6	29.6	62.0	50.0	0.5	
1/27/09	Menu summary - vlastní			419.0	3.5	2.6	15.5	-	0.8	
1/27/09 10:49 AM	vlastní	weighted	Jogurt Activia jahoda (Danone)	419.0	3.5	2.6	15.5	80.0	0.8	
1/27/09	Menu summary - svačina (3)			420.0	0.0	0.0	10.0	-	1.0	
1/27/09 10:50 AM	svačina (3)	in one piece	Svacina DIA 175g S	420.0	0.0	0.0	10.0	1.0 porton	1.0	VFN komplet
1/27/09	Menu summary - 1. večeře (2)			1062.0	3.7	4.4	45.9	-	1.75	
1/27/09 11:18 AM	1. večeře (2)	in one piece	Vecere I. 4200kJ	1155.0	0.0	0.0	29.0	1.0 porton	1.0	VFN komplet
1/27/09 11:18 AM	1. večeře (2)	weighted	Danette creme vanilkovy (Danone)	507.0	3.7	4.4	16.9	76.0	0.76	
1/27/09	Menu summary - vlastní			1681.0	5.2	18.1	54.3	-	1.2	
1/27/09 3:03 PM	vlastní	weighted	Babovka mramorova	1681.0	5.2	18.1	54.3	120.0	1.2	
1/27/09	Menu summary - snídaně (4)			2304.0	4.0	36.0	31.0	-	1.66	
1/27/09 3:36 PM	snídaně (4)	weighted	Kokos	1530.0	4.0	36.0	11.0	66.0	0.66	
1/27/09 3:36 PM	snídaně (4)	in one piece	Snidane REDU 2520kJ	774.0	0.0	0.0	20.0	0.75 porton	1.0	VFN komplet
1/27/09	Menu summary - vlastní			289.0	0.6	0.2	17.8	-	1.5	

Fig. 7. Example of food daily consumption

service oriented platform allowing a granular management of users and of their authorisations, and an easy connection of new services by public, private or non-profit service providers; ii) providing a set of protocols for accessing existing services; iii) provide a low cost home access point allowing to access the services and to efficiently manage chronic diseases as diabetes applying vital physiological signals monitoring.

Acknowledgment

The project was supported by STREP IST-2005-045282-AAL OLDES (Older People's e-services at home) and by Ministry of Education, Youth and Sport of the Czech Republic with the grant No. MSM6840770012 entitled "Transdisciplinary Research in Biomedical Engineering II".

References

[1] P. D. of the Department of Economic and W. P. P. Social Affairs of the United Nations Secretariat, "The 2008 revision and world urbanization prospects," 2008.

[2] N. Gonzlez and I. Vlez, "Confidence: Support to older peoples independent living," in *In proceedings of Psykologi*, 2008.

[3] ATTENTIANET, "Advanced teleassistance convergent network for chronic disease patients and elders project." [Online]. Available: http://www.attentianet.eu/

[4] ENABLE, "Ambient assisted living solutions based on advanced ict technologies." [Online]. Available: http://www.enable-project.eu/

[5] K4Care, "Knowledge-based homecare eservices for an ageing europe." [Online]. Available: http://www.k4care.net

[6] Intel, "Health guide phs6000." [Online]. Available: www.intel.com/healthcare/telehealth

[7] N. S. et al., "The attitudes, expectations and needs of elderly people in relation to e-health applications: results from a european survey," *Journal of Telemedicine and Telecare*.

[8] C. J. et al., "A systematic review of interactive computer-assisted technology in diabetes care interactive information technology in diabetes care," *Journal General Internal Medicine*.

Chapter 5

Analysis of event and actigraph time series in schizophrenia and bipolar disorder

Analysis of actigraph parameters for relapse prediction in bipolar disorder: a feasibility study

D. Novák, F. Albert, F. Španiel

Abstract— **The paper presents a framework for early identification of prodromal syndromes od mania or depression in bipolar disorder. The framework may mitigate relapses and improve patient functioning. The methodology consists of long-term actigraphy monitoring and simplified self-assessment tool to determine manic or depression events. Eight patients were involved in the feasibility study, spanning period of 150 months, resulting in 17 relapses and 3 hospitalizations in total. We concluded that the most promising parameter extracted from actigraphy recording is a circadian rhythm's interdaily stability. Using developed trend analysis applied on interdaily stability parameter, we achieved sensitivity and specificity about 65, resp. 68. We hypothesized that this performance is both mainly due to missing values in data and due to small amount of relapses.**

I. INTRODUCTION

Bipolar disorder is an episodic and recurrent condition that is frequently disabling daily life. Bipolar Disorder is associated with an increased suicide risk [1]. Annual insurance payments were greater for medical services for persons with bipolar disorder than for patients with other behavioral healthcare diagnoses. Primary goals of the treatment of bipolar disorder are to stabilize the patient and prevent the recurrence of episodes [2]. The early identification of prodromal symptoms of mania or depression may allow rapid interventions that mitigate relapses and improve life quality of patient. Dysregulation of the sleep wake cycle and other circadian rhythms may underlie the pathophysiology of bipolar disorder [3]. One possible way how to quantify circadian rhythms and sleep patterns is use of actigraph technology. The advantage of actigraphy over traditional polysomnography is that actigraphy can conveniently record continuously for 24-hours a day for days, weeks or even longer. This paper is focused on feasibility study of analysing relapse events (bipolar, mania) using long-term actigraphy monitoring and a self-assessment of patient's mood.

II. METHODOLOGY

A. Experimental Set-up

Two types of data were monitored: objective data represented by movement activity recordings by use of actigraphy and subjective data acquired by questionnaire. Movement activity was recorded using Actiwatch device (Camntech, Cambridge, UK). Actiwatches were configured to record movements in 2-minute epochs. Patients worn the device on the non-dominant wrist, the most common site for activity-based measurement. Actiwatch data were downloaded quartly using the ActiReader. Regular self-assessment of patient's mood was performed by a questionnaire. Questions are depicted in Table I. The values of each answer varied between 0 to 9. A score of 0 represents very faint feeling to the concrete question. A score of 9 represents very strong feeling to the concrete question. Technically, the questionnaire was sent by SMS message once a week. Additionally, during quarterly ambulatory visit, patient was asked to fill Young Mania Rating Scale (YMRS) and Hamilton Rating Scale for Depression (HRSD) [2] to clinically determine a possible relapse. Missing data exclusion criteria were based on Someren [4] who recommended that periods over 60 minutes without movement, even during sleep, are extremely unlikely to occur naturally.

TABLE I
SELF-ASSESMENT QUESTIONNAIRE

ID	Question
1	I feel like I am able to do anything
2	I really feel well inside
3	It seems to me that I will not succeed at anything
4	I am depressed
5	I feel full of energy
6	I feel speed up
7	I have racing thoughts
8	I am overly active
9	I am restless
10	I am impulsive
11	My moods change a lot
12	I feel like people are out to get me
13	I feel like the world is against me
14	I feel iritated
15	I feel argumentative
16	I get distracted easily
17	I cannot concentrate well
18	I sleep well

B. Data Set-up

Inclusion criteria for the study were a primary diagnosis of BD-I with no significant comorbidity, not working shiftwork, absence of a physical disability that may interfere with recording of ambulatory wrist movement, and a currently stable clinical state. In total, eight patients were monitored within 2006-2013 period. Two patients (n.2 and n.3) were withdrawn from the study due to systematic no-cooperation. All patients except of patient n.2 were medicated (alternatives of lithium). The basic characteristics are summarized in Table II. Data from questionnaire for patient n.7 and 8 were not available (NA). The legend is the following: Dur-

D. Novák, F. Albert is with Department of Cybernetics, Czech Technical University in Prague, Czech Republic, xnovakd1@labe.felk.cvut.cz, F. Španiel is with Prague Psychiatric Centre, Czech Republic.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

involvement of a patient in the study in months, #SMS-number of questionnaires filled (sent by SMS), number of events (mania or/and depression), #Vis-number of ambulatory visits, #Hos-number of hospitalizations, Mis-percentage of missing actigraphy data.

TABLE II
BASIC STUDY STATISTICS

ID	Sex	Age	Dur	#SMS	#Event	#Vis	#Hos	Mis
1	F	41	80	340	14	19	2	26
4	M	34	6	16	0	7	0	10
5	M	42	5	12	0	5	0	0
6	F	37	14	23	3	6	1	4
7	F	31	7	4	NA	NA	NA	53
8	F	27	37	NA	NA	NA	NA	14

C. Sleep period detection

Sleep disturbance is recognized as an essential aspect of affective illness. Therefore, it is important to detect sleeps' periods and to extract parameters from the sleep period in actigraph recording. First, the actigraph signal was smoothed by median filter $N = 20$. Secondly, three windows of different length were defined ($l_1 = 10, l_2 = 120, l_3 = 30$ samples). In the next step, an algorithm was searching $l_1 - 1$ non-zero samples from the left of ith-sample and l_1 zero samples to the right. All such as samples were stored into array P. Consequently, in the array P the algorithm was detecting such as index where l_2 samples (to the right) contain 75% zero values. This sample determines the beginning of the sleep. The detection of the end of sleep is similar. The end of sleep is determined by the first value between sleep-beginning tags, where l_3 samples are non zero.

D. Feature calculation

Standard parameters based on published algorithms for measures on actigraphy data were calculated. The extracted parameters can be divided into three parts: Night-Sleep analysis and Day-Awake analysis [5] and Circadian Rhythm Analysis [4]. Regarding Sleep analysis, the following parameters were extracted:

Assumed Sleep T_S,

$$T_S = t_{s1} - t_{s2} \tag{1}$$

where t_{s1} is wake time and t_{s2} time of fall asleep.

Actual awake time T_A,

$$f\left(x\right)\begin{cases} 0: & \sum_{y-t_{s1}}^{t_{s2}} a(x+y)\cdot w_A y \leq w_A \\ 1: & \sum_{y-t_{s1}}^{t_{s2}} a(x+y)\cdot w_A y > w_A \end{cases} \tag{2}$$

$$T_A = \sum_{t=t_{S1}}^{t_{S2}} f_A(a(t)), p_A = \frac{T_A}{T_S}\cdot 100 \tag{3}$$

where a represents actigraphy data, T_A is number of samples that represent physical activity, p_A is share of activity time in percentage, w_A is awake distribution.

Actual sleep time T_{AS}:

$$T_{AS} = T_S - T_A, p_{AS} = \frac{T_{AS}}{T_S}\cdot 100 \tag{4}$$

where T_{AS} is time of sleep without physical activity, p_{AS} is percentage share

Number of sleep bouts n_S:

$$f_{nS}\left(x\right)\begin{cases} 0: & w_S(a(x-1)) = 0 \vee w_S(a(x)) = 1 \\ 1: & w_S(a(x-1)) = 1 \wedge w_S(a(x)) = 0 \end{cases} \tag{5}$$

$$n_S = \sum_{t=t_{S1}}^{t_{S2}} f_{nS}(data(t)) \tag{6}$$

where n_s is number of segments with activity lower than a threshold and w_s is sleep distribution, Definition of number of wake bouts n_A is similar.

Sleep efficiency p_{ef}:

$$p_{ef} = \frac{T_S - T_A}{T_S} \tag{7}$$

Additionally the rest of sleep parameters were similarly calculated accordingly to (1) - (5), namely: Number of minutes immobile T_I, number of segments without activity n_I, number of segments that last one minute or less n_{I1}, fragmentation index I_f, total activity score S_A and minutes marked as active T_M.

Considering Day-Awake analysis, we detected periods of day-inactivity $NAPs$ bounded by two constants: $NAP_{min} = 10min$ and $NAP_{max} = 90min$. Furthermore, number of NAP per each day $NAP_{eachday}$ and total length of NAPs NAP_{total} were extracted. Next, total daytime activity was operationalised using the M10 variable which is defined in the next paragraph.

Another group of parameters is derived from Circadian Rhythm Analysis established by Someren [4]: IS (Interdaily Stability), IV (Intra-Daily Variability), L5 (provides the average activity level for the sequence of the least five active hours) and M10 (indication of the phase of the most ten active hours).

E. Trend analysis

We suggest to find specific trends in extracted features before an event of depression or mania. Our idea is similar to shapelets which were defined by Keogh [6]. Informally, shapelets are time series subsequences which are in some sense maximally representative of a class. From point of temporal data mining, the class is here represented by depression or mania. However, due to low data resolution when each day is represented by one value of a feature, shapelets could not be applied. Therefore, an alternative concept based on extensive search approach and using simpler shapes is proposed. We call the basic shapes primitives. We assume that mania or depression event is marked by a presence of a primitive some time before the event.

The explanation of the alternative concept to Keogh optimization is straightforward. Firstly, the data sets must fulfill the following criteria: An event (mania od depression) gray

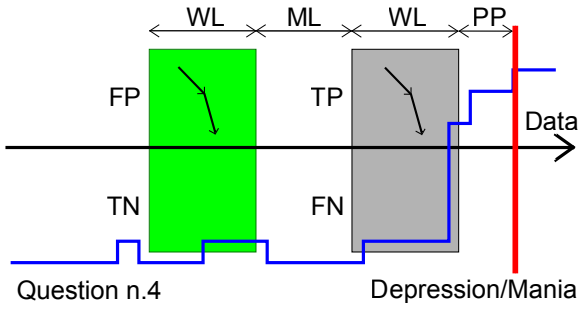


Fig. 1. Example of primitive [-1,-1] detected in data before event of a relapse (mania or depression). Legend: WL:window length, ML:minimum length, PP:prediction window, FP:false positive, FN:false negative, TP:true positive, TN:true negative

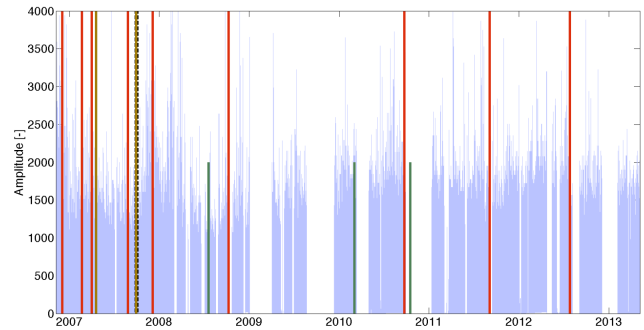


Fig. 2. Example of actigraphy data of patient n.1. Red line marks depression events, green line detects subjective depression events, orange markers indicate two hospitalizations and dashed black line visualizes start of the lithium treatment.

box in Fig. 1 has length of WL days, must precede the event by PP days and must be separated by ML days from a normal green box. The normal green box in Fig. 1 is of WL day long, is separated from the event green box by ML days and responses to question no.4 "I am depressed" must be below threshold 2. Parameter PP determines the minimum length in days before occurring relapse. To summarize, the green box determines periods when mood is stabilized while the gray box marks off depressive or mania events. Secondly, several basic primitives consisting of combination three $[p_1, p_2, p_3]$ or two $[p_1, p_2]$ samples (days) were established. For example, primitive [-1,1,0] represent decrease, increase and change of trend in a respective parameter. Finally, using extensive-search, the algorithm is searching such a primitive among set of $P = \{[p_1, p_2, p_3], [p_1, p_2]\}$ assuring that a corresponding ROC curve has the largest area under this curve. The ROC curve is defined by trade-off between sensitivity (Se) and specificity (Sp) values parametrized by WL value. To calculate Se and Sp, we defined false positive as appearance of a primitive (or primitives) in the whole Normal window, true negative as no-show of the primitive. Similarly, true positive is indicated by the primitive presence in the Event window, false negative as primitive missing in the Event window - see example in Fig. 5, where several primitives of [-1,-1] in the IS time series were detected.

III. Results and Discussion

Complete actigraphy record of patient n.1 over span of 6 years is depicted in Fig. 2. Events of depressions are visualized by vertical red lines, hospitalization are determined by orange lines. There are many missing values in the record, some of them are of considerable length. For example, note major drop-outs of more than 2 months between 2009-2011 period. Additionally, the patient was allowed not to wear the actigraph during vacations explaining regular drop-outs during summer periods.

The sleep detection algorithm detects precisely beginning and end of sleep periods as is shown in Fig. 3. Apparently, the user forgot to take an actigraph device after evening hygiene, no activity was detected from the night 28th till the morning 29th when the user started wearing the actigraph again.

The next phase was focused on data processing from sms questionnaire. The question 1 and 4 from Table I are shown along with depression (clinically confirmed during ambulatory visits and subjectively defined as events marked by question n.4 ¿5) and hospitalization events for patient n.1 in Fig. 4. It is important to determine the minimum number of questions which may reflect relapse onset. Therefore, we performed correlation analysis using gathered questionnaires across all patients with the aim of determining minimum set of independent questions.The following set of questions $Q = \{4, 6, 7, 8, 15\}$ was determined on 0.05 significance level. Several depression events visually correlate with value of question n.4 as can be seen in Fig. 4. Question n.1 is antagonistic to question n.4 confirming the consistency of responses in case of patient n.1. Similarly, correlation analysis was performed on a set of actigraphy parameters resulting in the following set of features $F = \{n_i, T_{AS}, T_M, IV, IS, M10, L5, NAP\}$.

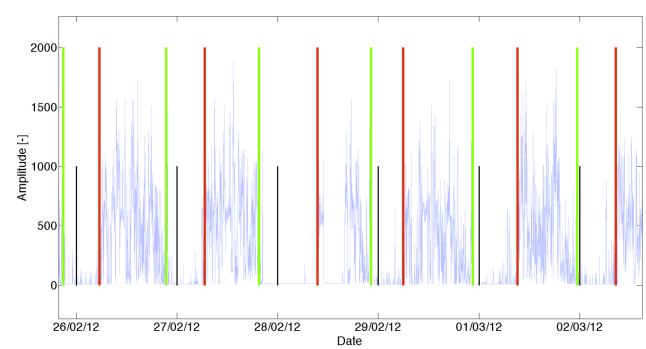


Fig. 3. Example of sleep detection in case of patient n.1. Red line marks start of sleep, green line detects send of sleep ,black marker indicates midnight.

Having independent features, we performed trend analysis using three and two points based primitives. Extensive search resulted in finding primitive [-1,-1] - see Fig. 5. This primitive applied on IS signal delivered the best performance in terms of maximization area under ROC. IS-Interdaily Stability quantifies the degree of regularity in the activity-rest pattern with a range of 0 to 1 where a value of 0 indicates a total lack of rhythm and a value of 1 indicates

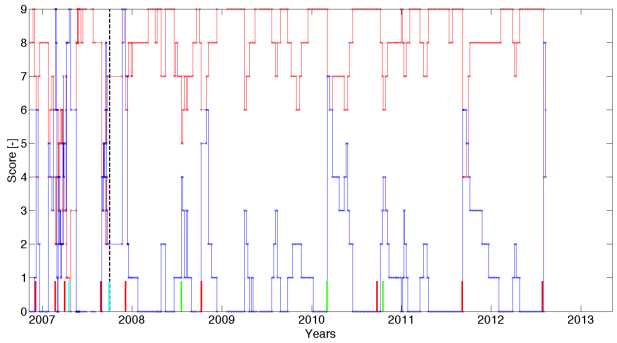


Fig. 4. Visualization of two antagonistic questions 1 (red) and 4 (blue) for patient n.1. Red line marks depression events, green line detects subjective depression events, orange markers indicate two hospitalizations and dashed black line visualizes start of the lithium treatment.

a perfectly stable rhythm. As can be observed, patient no.1 alternates along 0.5 value. An example of primitive detection is depicted in Fig. 5. Note the presence of two FP detection in green windows marked by asterisk and one FN detection marked by cross symbol.

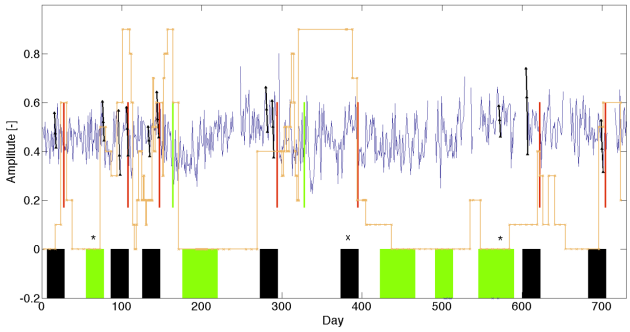


Fig. 5. Example of primitive detection [-1,-1] for patient no.1. Trend analysis parameters were set up to WL=22 and PP=0. Blue line is IS time series. Red line marks depression events, green line represent subjective depression events, black line with arrows visualize primitive [-1,-1]. Orange line is equal to question n.4 "I feel depressive".

Finally, comparison of two primitives [-1,-1] and [-1,0,1] applied on the IS parameter is depicted in Fig.6. The best trade off between Se and Sp values is about $Se = 65\%$, $Sp = 68\%$. The framework presented here determined the methodology for a bigger clinical trial. During this feasibility study we had to cope with low cooperability of patients resulting in total loss of 30% actigraphy data - see Table II. Furthermore, some patients did not send regularly subjective questions. Only three patients persisted in the study for a period longer than 12 months. Additionally, 17 depressions events were diagnosed while no mania event was presented in our data set. We postulate that online approach for actigraphy monitoring and implementation of alert system in case of missing data (actigraph records or questionnaire) is absolutely necessary for succesfull completion of long-term monitoring in such as ambitious project focusing on patients suffering from severe psychiatric disorder.

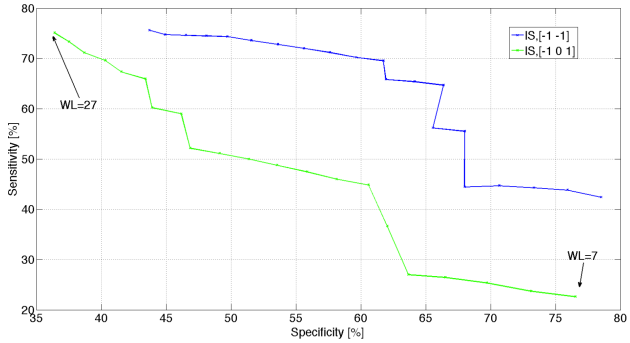


Fig. 6. ROC curve for all patient. ROC is paramatrized by WL parameter, $WL = 7, \ldots, 27$. Two primitives are depicted: [-1,-1]:blue line and [-1.0,1]:green line.

IV. CONCLUSIONS AND FUTURE WORKS

We presented the basic framework for early identification of syndromes of mania or depression. The framework may mitigate relapses and improve patient quality of life. The methodology consisted of long-term actigraphy monitoring and simplified self-assessment tool to determine manic or depression symptoms. We concluded that the most promising parameter exctracted from actigraph data is Interdaily Stability. Using proposed trend analysis and the IS parameter, we achieved sensitivity and specificity about 65%, resp. 68%. To increase performance, online monitoring of movement activity is essential. However, current actigraphic technology does not permit real-time monitoring of activity. This significantly limits its use for this purpose. Therefore, we designed own system which enables to monitor actigraphy in real time without recharging the battery for a period of one year (not covered in this paper) with similar technical parameters to existing offline commercial systems. We plan to use this solution in an ambitious clinical study involving more than 70 patients monitored for 12 months.

ACKNOWLEDGMENT

The project was supported by grant IGA NT14387 from Ministry of Health, Czech Republic.

REFERENCES

[1] R. J. Baldessarini, M. Pompili, and L. Tondo, "Suicide in bipolar disorder: Risks and management," *CNS Spectr*, vol. 11, no. 6, pp. 465–471, Jun 2006.

[2] A. P. Association, "Practice guideline for the treatment of patients with bipolar disorder (revision)," *Am J Psychiatry*, vol. 159, no. 4 Suppl, pp. 1–50, Apr 2002.

[3] M. Bauer, P. Grof, N. Rasgon, T. Bschor, T. Glenn, and P. C. Whybrow, "Temporal relation between sleep and mood in patients with bipolar disorder," *Bipolar Disord*, vol. 8, no. 2, pp. 160–167, Apr 2006.

[4] E. J. Van Someren, D. F. Swaab, C. C. Colenda, W. Cohen, W. V. McCall, and P. B. Rosenquist, "Bright light therapy: improved sensitivity to its effects on rest-activity rhythms in Alzheimer patients by application of nonparametric methods," *Chronobiol. Int.*, vol. 16, no. 4, pp. 505–518, Jul 1999.

[5] S. Ancoli-Israel, R. Cole, C. Alessi, M. Chambers, W. Moorcroft, and C. P. Pollak, "The role of actigraphy in the study of sleep and circadian rhythms," *Sleep*, vol. 26, no. 3, pp. 342–392, May 2003.

[6] L. Ye and E. Keogh, "Time series shapelets: A new primitive for data mining," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '09, 2009, pp. 947–956.

Available online at www.sciencedirect.com
ScienceDirect
Schizophrenia Research 98 (2008) 312–317
SCHIZOPHRENIA RESEARCH
www.elsevier.com/locate/schres

ITAREPS: Information Technology Aided Relapse Prevention Programme in Schizophrenia

Filip Španiel [a,d,e,*], Pavel Vohlídka [b], Jan Hrdlička [c], Jiří Kožený [a,d,e], Tomáš Novák [a,d,e], Lucie Motlová [a,d,e], Jan Čermák [a,d,e], Josef Bednařík [f], Daniel Novák [c], Cyril Höschl [a,d,e]

[a] *Prague Psychiatric Center, Ústavní 191 181 03 Prague 8, Czech Republic*
[b] *Eli Lilly and Company Ltd, Lilly House, Priestley Road Basingstoke, Hampshire, RG24 9NL, United Kingdom*
[c] *Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics, Karlovo namesti 13, 12135, Prague, Czech Republic*
[d] *3rd Faculty of Medicine, Charles University Prague, Czech Republic*
[e] *Center of Neuropsychiatric Studies, Czech Republic*
[f] *Eli Lilly CR s.r.o., Pobrezni 1a, 186 00 Prague 8, Czech Republic*

Received 15 May 2007; received in revised form 27 August 2007; accepted 6 September 2007
Available online 24 October 2007

Abstract

ITAREPS presents a mobile phone-based telemedicine solution for weekly remote patient monitoring and disease management in schizophrenia and psychotic disorders in general. The programme provides health professionals with home telemonitoring via a PC-to-phone SMS platform that identifies prodromal symptoms of relapse, to enable early intervention and prevent unnecessary hospitalizations. Its web-based interface offers the authorized physician a longitudinal analysis of the dynamics and development of possible prodromes.

This work presents preliminary findings from a one-year mirror-design follow-up evaluation of the programme's clinical effectiveness in 45 patients with psychotic illness. There was a statistically significant 60% decrease in the number of hospitalizations during the mean 283.3±111.9 days of participation in the ITAREPS, compared to the same time period before the ITAREPS entry (sign test, $p<0.004$). Variables significantly influencing the number of hospitalizations after the ITAREPS entry (medication compliance along with factors intrinsic to the ITAREPS, i.e. adherence to the programme and involvement of a family member) suggest a critical role of the programme in controlling the number of relapses and subsequent hospitalizations in psychosis.

Keywords: Psychotic disorders; Schizophrenia; Relapse prevention; Hospitalizations

1. Introduction

Schizophrenia is a major psychotic disorder that has devastating effects on the lives of patients and their caregivers.

The illness is accompanied by high rate of relapses and readmissions. An overview of studies investigating long term outcomes has shown that people with schizophrenia have a one-year relapse rate of 15 to 35%, rising to 80% at five years after the onset of first-episode, despite maintenance medication (Robinson et al., 1999). Achievement of full remission becomes less likely after each relapse (Wiersma et al., 1998). In addition, an increased number of relapses positively correlate with subsequent

* Corresponding author. Prague Psychiatric Center, Ústavní 191 181 03 Prague 8, Czech Republic. Tel.: +420 2 6600 3390; fax: +420 2 6600 3366.
E-mail address: spaniel@pcp.lf3.cuni.cz (F. Španiel).

0920-9964/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.schres.2007.09.005

social decline and reduced treatment response in patients with schizophrenia (Lieberman et al., 1996; Shepherd et al., 1989). Needless to say, these patients as well as their caregivers experience extremely high levels of burden related to the dramatic consequences of relapse and hospitalization.

Studies published so far suggest that the prediction of relapse appears to be an achievable goal for a substantial proportion of patients with schizophrenia. The successful prediction of relapse using new technologies e.g. telemedicine capabilities will most likely require a combination of components built into an integrated programme. These components may include monitoring of non-specific and specific prodromal symptoms, frequent assessments and an involvement of both patients and caregivers (Fitzgerald, 2001). Crisis interventions such as increasing the dose of antipsychotic medication, based on early detection within a relapse prevention programme, may reduce relapse and readmission rates compared to a treatment-as-usual group. Timely review of patient mental status with close to real-time feedback has been shown as a critical success factor in the management of psychosis (Herz et al., 2000).

The future wide availability of high bandwidth public wireless networks will give rise to new mobile health care (or M-Health) services. These approaches allow for home-based telemedicine ensuring the direct and prompt communication between the clinician and the patient or their caregiver.

The ITAREPS project (Information Technology Aided Relapse Prevention Programme in Schizophrenia) developed at the Prague Psychiatric Centre represents a step towards a highly customizable prodromal signs monitoring M-Health service platform.

1.1. The ITAREPS programme design

Participants enrolled in the ITAREPS programme (the patient and her/his family member) were instructed to complete a 10-item Early Warning Signs Questionnaire (EWSQ) by a Short Message Service (SMS) request sent weekly (on Thursdays) by an automated system to their mobile phones. Attendance of a family member at the ITRAEPS was highly recommended, albeit optional.

There were two versions of EWSQ, one for patients (EWSQ-10P) and the other for their family members (EWSQ-10FM). Full reporting on psychometric properties of EWSQ has been published elsewhere (Spaniel et al., 2007).

Nine of the EWSQ items covered the most common early non-specific warning sings of the relapse. The tenth item instructed patients and carers to list three specific symptoms which had individually preceded previous relapse or relapses in the given patient and which were not included under items 1–9 (Table 1). In case of more than one specific symptom occurrence, the most pronounced sign was selected and scored under item 10.

The EWSQ was designed to detect proportional worsening (or a new onset) of symptoms compared to the baseline (the previous week's completed questionnaire). Each item score ranged from 0 (no worsening or improvement of symptoms) to 4 (extreme worsening).

Table 1
Early warning sings questionnaire, patient and family member version

Item no.	EWSQ 10 Patient Version	EWSQ 10 Family Member Version
1	Has your sleep worsened since the last evaluation?	Change of the sleep pattern.
2	Has your appetite decreased since the last evaluation?	Marked behavioral changes.
3	Has your concentration, e.g., ability to read or watch TV, worsened since the last evaluation?	Social withdrawal.
4	Have you experienced fear, suspiciousness, or other uneasy feelings while being around other people since the last evaluation?	Deterioration in daily activities and functioning.
5	Have you experienced increased restlessness, agitation, or irritability since the last evaluation?	Deterioration in personal hygiene.
6	Have you noticed that something unusual or strange is happening around you since the last evaluation?	Loss of initiative, motivation.
7	Have you experienced loss of energy or interest since the last evaluation?	Eccentric thought content, marked preoccupation with strange ideas.
8	Has your capability to cope with everyday problems worsened since the last evaluation?	Marked poverty of speech and content of thoughts.
9	Have you experienced hearing other people's voices even when nobody was around since the last evaluation?	Irritability, restlessness, agitation, aggressivity
10	Have you noticed any other of your individual early warning signs since the last evaluation?	Have you noticed any other individual early warning signs since the last evaluation?

Completion of the questionnaire took approximately 5 min. Individual EWSQ scores were sent by participants back to the ITAREPS as a SMS message, presented as a string of ten digits (i.e. 0001202001). The information was then processed automatically. If the total score (and/or given score in prepsychotic sub items of the EWSQ) exceeded a given score threshold, an immediate ALERT was declared and the psychiatrist was notified of this by an e-mail message.

The e-mail ALERT message contained the code number of the patient and an Early Intervention Algorithm (EIA). The EIA was a key factor allowing for early crisis intervention. It consisted of five consecutive steps in which immediate phone contact with the patient and prompt evaluation of his/her current status was recommended.

The presence of early warning signs warranted an immediate 20% increase from baseline maintenance dose of antipsychotic within the next 24 h. The effectiveness of this particular intervention has been confirmed previously (Herz et al., 2000). To achieve the most appropriate 20% fractional dose increase, splitting or dividing of the tablets was to be done in the most suitable way according the relevant antipsychotic's Product Information. The fifth item of EIA emphasizes the importance of doctor–patient face-to-face contacts during the treatment, pointing on the fact that the program is not intended to reduce the number of visits.

Once an ALERT was declared, it continued for the next 3 week ALERT PERIOD, providing that the following 6 consecutive EWSQ scores showed no worsening of symptoms. If so, the ALERT PERIOD was withdrawn and the event announced to physician via e-mail along with recommendation concerning subsequent tapering down the medication to the pre-ALERT doses. During the ALERT PERIOD, participants were to return questionnaires twice-weekly upon SMS request (on Thursdays and Mondays).

In addition to that, more conservative score thresholds were adopted. If EWSQ scores exceeded those modified thresholds anytime during the ALERT PERIOD, an ALERT EMERGENCY was announced via e-mail. In such a case the ALERT PERIOD was extended for a further 3 weeks after each ALERT EMERGENCY message.

Admissions to the hospital and discharge dates were announced to the system by either and/or both patient and family member by SMS message. In order to record the event in the system, SMS had to follow a predefined form. This information was automatically integrated into line graphs accessible to clinicians. Information about hospital admissions were subsequently confirmed by the outpatient psychiatrist for the purpose of this clinical evaluation.

The patient data entering process was operated exclusively on the internet using standard browsers. In order to access the ITAREPS (www.itareps.com), participating psychiatrists needed a username and a password. Subsequently, his/her personal page was created in which his/her patients were registered and assigned a specific code number. The subject number and patient initials were recorded on a patient assignment log, which was kept exclusively by the outpatient psychiatrist.

During each personal web page visit, the physician could easily check the current medical status of his/her patient. Longitudinal score values (both from the patient and his/her carer) were available in a graphic form (line graphs) and in a detailed written description which was converted from the completed questionnaires returned as SMS messages from patient and his/her carer. In this way the psychiatrist could easily review the dynamics and development of possible prodromes.

All stored data contained no patient identifiers. The data were rendered unreadable to unauthorized persons. Data processing and security measures addressed all known security vulnerabilities including encryption of transmission and storage of information across the internet.

No specific training has been provided for ITAREPS participants. According to the clinical experience with ITAREPS program, ITAREPS User Manual distributed to the program participants provided a firm basis for appropriate management and system operation.

2. Methods

2.1. Subjects

A total of 45 patients and 39 family members were enrolled in the one-year, longitudinal, mirror-design evaluation of the ITAREPS clinical effectiveness since its implementation in clinical practice in the Czech Republic in September 2005. These subjects were recruited by their psychiatrists through 14 outpatient mental health facilities in the Czech Republic. Direct advertising in Czech peer-reviewed journal Psychiatrie was used for the purpose of out-patient facilities recruitment. No financial or other incentives were given to any participant to take part in this clinical evaluation.

Patients fulfilled ICD-10 criteria for diagnosis of schizophrenia, schizoaffective disorder, or acute polymorphic psychotic disorder with or without symptoms of schizophrenia, and were aged between 18 and 65 (Table 2).

Mean illness duration (time since occurrence of the first psychotic symptoms) was 6.9 (SD 4.8) years in the participants, resulting in 134 hospitalizations during a total of 256 patient/years. Thus, cumulative hospitalization

Table 2
The demographic and clinical characteristics of the patients

			Mean (SD)
Age	Male $N=27$ (60%)		31.3 (7.6) years
	Female $N=18$ (40%)		29.6 (5.8) years
CGI			2.6 (1.2)
			N
Diagnosis	Schizophrenia		29 (64.4%)
	Schizoaffective disorder		11 (24.4%)
	Acute polymorphic psychotic disorder with schizophrenia symptoms		4 (8.9%)
	Acute polymorphic psychotic disorder without schizophrenia symptoms		1 (2.3 %)
Antipsychotic medication	No		22 (48.8%)
	Yes	Atypical	19 (42.2%)
		Classical	4 (9%)

incidence in our patient group was 0.51 patient/years before the ITAREPS entry.

Because this follow-up used only clinical information without specific patient identifiers and the procedures required no deviation from standard clinical practice, informed consent was not obtained from participants. The protocol of ITAREPS programme was approved by the Ethics Committee of the Prague Psychiatric Center.

2.2. Measures

The following baseline measurements were obtained from outpatient psychiatrists: demographic data; diagnosis; illness history; CGI and current medication. Compliance was evaluated by therapists using a questionnaire with a 5-point scale (1 = always and 5 = never taking antipsychotic medication (Herz et al., 2000)). The adherence to the ITAREPS programme was measured as a percentage difference (PD) between the number of messages the patient/family member were supposed to send according to the ITAREPS programme and the number actually sent by both participants. Patients and family members were divided into two groups according to PD. Subjects with a PD of less than 30% undelivered SMS messages containing EWSQ scores were designated as "Cooperative" and "Uncooperative" if they had a PD of more than 30% SMS messages missing.

2.3. Statistics

Statistical analysis was performed using the sign test, the Mann–Whitney *U* test and the Fisher's exact test. The

Table 3
Relationship of patient variables to hospitalization status after the ITAREPS entry

Variable	Non-hospitalized during ITAREPS ($n=38$)		Hospitalized during ITAREPS ($n=7$)[a]		Difference between non-hospitalized and hospitalized during ITAREPS	
	Median	Minimum–maximum	Median	Minimum–maximum	Mann–Whitney *U* (*z* scores)	*p*
CGI	2.0	1–5	3.0	1–5	−1.064	0.3
Illness duration (days)	2178.5	51–3910	3489.0	2203–4940	−1.535	0.13
No. of hospitalizations before ITAREPS	2.0	0–13	6.0	1–15	−1.558	0.12
Medication compliance	1.5	1–3	2.0	1–5	−1.969	0.04[b]
No. of Alert states	0.0	0–26	0.0	0–1	−0.538	0.64
No. of days in ITAREPS	291.0	5–565	300.0	76–412	−0.047	0.97
		N		*N*		*p*[c]
Medicated/unmedicated		19/19		4/3		0.69
Inclusion of a family member Yes/No		35/3		4/3		0.05[b]
Patient cooperation Yes/No[d]		32/6		3/4		0.03[b]
Family member cooperation Yes/No[d]		21/17		1/6		0.09

[a] Two subjects were hospitalized twice during the follow-up period.
[b] Statistical significance at $p \leq 0.05$, exact significance, two-sided.
[c] Fisher's exact test.
[d] Cut off point of cooperativeness defined as more or less then 30% undelivered SMS messages.

criterion for statistical significance was set at $p \leq 0.05$ exact significance, two-sided to have adequate power to detect differences with the small number of subjects.

3. Results

In total 1540 EWSQs were completed and returned as a SMS message by patients along with 1020 EWSQs from family members during a total follow-up period of 34.7 patient/years.

The ITAREPS system announced the onset of early warning sings in 88 cases based on EWSQ scores sent by patients and in 47 cases based on family members' reports, during the follow-up period.

The overall patient drop-out rate during the evaluation period was 10%. The drop-out was defined as any discontinuation of participation in ITAREPS programme for more than 10 weeks.

There were 9 hospitalizations in the 45 patients participating in the ITAREPS programme (Table 3) for a mean of 283.3 ± 111.9 days follow-up, compared with 27 hospitalizations during the same length of time before entering the programme, resulting in a statistically significant 60% reduction in the number of hospitalizations during ITAREPS programme participation (sign test, $p < 0.004$). There was 100% decrease in the number of hospitalizations (from 13 to 0) in highly cooperative patients ($N = 21$, 47%, family member included, more than 70% SMSs returned by both participants) during the follow-up. Those subjects showed cumulative hospitalization incidence 0.42 patient/year before the ITAREPS entry.

In a second step, Mann–Whitney or Fisher's exact tests were performed to search for variables influencing the number of hospitalizations after ITAREPS enrolment (Table 2). No significant differences between hospitalized and non-hospitalized patients in the ITAREPS programme were found regarding the duration of illness, number of hospitalizations before entering ITAREPS, CGI ratings during initial visit, number of ALERT states during ITAREPS, medication status (on or off antipsychotic medication at the entry of the ITAREPS programme) or duration of the ITAREPS programme participation.

However, univariate comparisons showed that non-hospitalized patients were significantly more compliant with their medication, more adherent to the ITAREPS programme and had a family member significantly more frequently involved in the programme compared to subjects hospitalized after ITAREPS entry. A similar but non-significant trend related to higher family member adherence to the ITAREPS programme was found in non-hospitalized compared to hospitalized patients.

4. Discussion

The prevention of relapse and readmission has high priority in the treatment of patients with a psychotic disorder. There is an urgent need to develop effective relapse prevention programmes that are applicable in routine clinical practice.

The use of telemedicine capabilities to monitor chronically ill patients is becoming more clinically relevant and economically cost effective.

The novel M-health management package, ITAREPS, by means of a simple, interactive communication based on user–server interactivity between healthcare professional and project participants, can timely report prodromal symptoms of relapse and promote appropriate measures that can be taken for early pharmacological intervention.

This one-year mirror-design follow-up was carried out to evaluate the impact on outcome of early intervention enabled by the ITAREPS system in 45 patients with psychotic illness. Significantly fewer hospitalizations were shown in patients enrolled in the ITAREPS programme compared to the follow-up period before the programme entry. Among several variables that were considered to be potentially associated with the rate of hospitalizations after ITAREPS entry, only medication compliance, adherence to the ITAREPS programme and involvement of a family member in the project showed statistical significance. Those preliminary findings suggest that the parameters intrinsic to the ITAREPS programme, along with compliance, as a crucial factor in project efficiency resulted in a reduction of the number of relapses and subsequent hospitalizations in psychosis.

The results of this pilot clinical evaluation must be interpreted in light of its methodological limitations including absence of a control group and relatively small numbers of patients.

In addition, although a specific procedure was recommended according to Early intervention algorithm (EIA), the exact nature of the intervention during the Alert states could not be documented in this clinical evaluation. This is of particular importance in medication-free subjects ($n = 22$, 49%) in which antipsychotic readministration corresponds EIA during the ALERT state.

Therefore, a prospective, double-blind, randomized trial with the ITAREPS programme is warranted to confirm our preliminary findings.

5. Conclusions

The ITAREPS programme presents a novel approach towards relapse prevention which according to our

knowledge has not been used before. The programme offers affordable solution to face the challenges of frequent rehospitalizations and relapses in patients with schizophrenia and psychotic disorders in general. Besides this primary effect, one of the main goals of the program is implementing continuous quality improvement in the doctor–patient relationship resulting in the strengthening and positive interference with the rapport between physician and patient. Furthermore, ITAREPS user-friendly, web-based data capture system enables international large-scale naturalistic studies to be conducted, aimed at the evaluation of long-term outcomes, dynamic fluctuations of the course of psychotic illness and the rate of relapses related to the particular antipsychotic medication.

Role of funding source

Eli Lilly Czech Republic supported financially the project. This company had no further role in the collection, analysis and interpretation of data.

Contributors

Filip Španiel developed ITAREPS system, designed the study, wrote the protocol and manuscript. Pavel Vohlídka co-worked on study concept and design and drafting of the manuscript. Jan Hrdlička made substantial contributions to the manuscript in terms of data acquisition and analysis. Jiří Kožený undertook the statistical analysis and interpretation of the data. Lucie Motlová and Jan Čermák managed the literature searches, analyses and technical support of the study. Daniel Novák contributed to analysis and interpretation of the data. Josef Bednařík and Cyril Höschl were responsible for critical revision of the manuscript for intellectual content.

Conflict of interest

Pavel Vohlidka and Josef Bednařík are employees of Eli Lilly CR, s.r.o. which has provided a unrestricted grant for ITARERPS development. All other authors declare that they have no conflicts of interest.

Acknowledgement

ITAREPS has been developed in Prague Psychiatric Centre as a non-commercial, academic project, technically supported by Academia Medica Pragensis (Amepra). Amepra is an agent and ethical guarantor of relations between Eli Lilly Czech Republic, which supported the project, and the academic sphere involved in the project development and implementation.

ITAREPS is operating under the auspices of the Psychiatric Society CLS JEP of Czech Republic and Psychiatric Society SLS, Slovak Republic.

The authors wish to convey special thanks to Tracy J. Taylor for her comments in the writing of the manuscript.

References

Fitzgerald, P.B., 2001. The role of early warning symptoms in the detection and prevention of relapse in schizophrenia. Aust. N. Z. J. Psychiatry 35 (6), 758–764.

Herz, M.I., Lamberti, J.S., Mintz, J., Scott, R., O'Dell, S.P., McCartan, L., Nix, G., 2000. A program for relapse prevention in schizophrenia — a controlled study. Arch. Gen. Psychiatry 57 (3), 277–283.

Lieberman, J.A., Koreen, A.R., Chakos, M., Sheitman, B., Woerner, M., Alvir, J.M.J., Bilder, R., 1996. Factors influencing treatment response and outcome of first-episode schizophrenia: implications for understanding the pathophysiology of schizophrenia. J. Clin. Psychiatry 57 (Suppl 9), 5–9.

Robinson, D., Woerner, M.G., Alvir, J.M.J., Bilder, R., Goldman, R., Geisler, S., Koreen, A., Sheitman, B., Chakos, M., Mayerhoff, D., Lieberman, J.A., 1999. Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. Arch. Gen. Psychiatry 56 (3), 241–247.

Shepherd, M., Watt, D., Falloon, I., Smeeton, N., 1989. The natural-history of schizophrenia — a 5-year follow-up-study of outcome and prediction in a representative sample of schizophrenics. Psychol. Med. Monogr. Suppl. 15, 1–46.

Spaniel, F., Novak, T., Motlova, L., Hrdlicka, J., Höschl, C., 2007. Information technology aided relapse prevention program in schizophrenia (ITAREPS): reliability and validity of the early warning signs questionnaire. Psychiatrie 11 (3), 157–159.

Wiersma, D., Nienhuis, F.J., Slooff, C.J., Giel, R., 1998. Natural course of schizophrenic disorders: a 15-year followup of a Dutch incidence cohort. Schizophr. Bull. 24 (1), 75–85.