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

Data Mining for Blood Glucose Prediction and Knowledge Discovery in Diabetic Patients: The METABO Diabetes Modeling and Management System

Eleni Georga, Vasilios Protopappas, *Member, IEEE,* Alejandra Guillen, Giuseppe Fico, Diego Ardigo, Maria Teresa Arredondo, Themis P. Exarchos, *Student Member, IEEE*, Demosthenes Polyzos, and Dimitrios I. Fotiadis, *Senior Member, IEEE*

Abstract—**METABO is a diabetes monitoring and management system which aims at recording and interpreting patient's context, as well as, at providing decision support to both the patient and the doctor. The METABO system consists of (a) a Patient's Mobile Device (PMD), (b) different types of unobtrusive biosensors, (c) a Central Subsystem (CS) located remotely at the hospital and (d) the Control Panel (CP) from which physicians can follow-up their patients and gain also access to the CS. METABO provides a multi-parametric monitoring system which facilitates the efficient and systematic recording of dietary, physical activity, medication and medical information (continuous and discontinuous glucose measurements). Based on all recorded contextual information, data mining schemes that run in the PMD are responsible to model patients' metabolism, predict hypo/hyper-glycaemic events, and provide the patient with short and long-term alerts. In addition, all past and recently-recorded data are analyzed to extract patterns of behavior, discover new knowledge and provide explanations to the physician through the CP. Advanced tools in the CP allow the physician to prescribe personalized treatment plans and frequently quantify patient's adherence to treatment.**

I. Introduction

Diabetes Mellitus is a chronic metabolic disease characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. The chronic hyperglycemia of diabetes is associated with long-term complications such as neuropathies, nephropathies and retinopathies [1]. The vast majority of cases of diabetes fall into two broad categories, Type 1 Diabetes Mellitus (T1DM) and Type 2 Diabetes Mellitus (T2DM). The cause of T1DM is an absolute deficiency of insulin secretion which leads to an uncontrolled increase of blood glucose levels. T2DM is a much more prevalent category. The cause is a combination of resistance to insulin action and an inadequate compensatory insulin secretory response.

Recently, a number of computer-based expert systems for supporting decision making in diabetes management have been proposed [2]-[4]. One of the most popular mathematical models based on compartmental structure is the AIDA model [5]. Compartment models (CMs) are commonly used for modeling complex systems with dynamics such as physiological systems. Since the physiological interactions regarding diabetes are multidimensional, nonlinear, stochastic, time-variant, and patient specific, artificial neural networks (ANNs) [6], [7] have also been used very successfully. However, most of previous works used a combined model of them or other mathematical models. Mougiakakou et al. [8] presented simulation models, based on combination of CMs with recurrent ANNs, for short-term prediction of glucose-insulin metabolism of Type 1 diabetes patients. A neural network algorithm based on physiological parameters such as heart rate, QT interval and skin impedance has been developed for the detection of hypoglycemia episodes in children with Type 1 diabetes mellitus [9]. Furthermore, the performance of a number of ANNs variations, including CMs, time-series neural-network, and recurrent ANNs for predicting BGL is compared in [10], while Sparacino et al. in [11] presented low-complexity prediction strategies based on continuous blood glucose measurements.

The aim of this paper is to present the METABO diabetes management system and elaborate on the data mining schemes that are used (a) for modeling patient's metabolism and also (b) for extracting patient-specific knowledge. In METABO, modeling is used to predict blood glucose values and provide decision-support to the patient, whereas knowledge extraction aims at quantifying and explaining patient's performance.

Manuscript received April 7, 2009. This work is part funded by the European Union, Project METABO "Controlling Metabolic Diseases Related to Metabolic Disorders", FP7-ICT-2007–1–216270.

E. Georga, T. P. Exarchos and D. I. Fotiadis are with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, Greece, GR 45110 (email: egeorga@cs.uoi.gr, exarchos@cc.uoi.gr, fotiadis@cs.uoi.gr).

V. Protopappas and D. Polyzos are with the Department of Mechanical Engineering and Aeronautics, University of Patras, Greece, GR 26500 (email: vprotop@mech.upatras.gr).

M. A. Guillén is with the Department of Service Solutions from CRDM in Medtronic Spain, Madrid, ES 28050 (email: alejandra.guillen@medtronic.com).

G. Fico is with the Lifestyle Supporting Technologies Group, Technical University of Madrid, ES 28040 (email: gfico@lst.tfo.upm.es).

D. Ardigo is with the Department of Internal Medicine and Biomedical Sciences, University of Parma, Parma, Italy, IT 43100 (email: d.ardigo@libero.it).

Maria Teresa Arredondo is with the Lifestyle Supporting Technologies Group, Technical University of Madrid, ES 28040 (email: mta@lst.tfo.upm.es).

II. The METABO System

A. System Description

METABO provides a comprehensive system running in every-day life environments for continuous and multi-parametric monitoring of the metabolic status of patients with diabetes disease and for the provision of advanced and personalized decision-support to both the patient and the treating physicians. The METABO system aims at:

- the integration of traditional data with lifestyle information in a continuous and consistent way using suitable sensors
- the development of advanced personal metabolic models using data mining techniques
- the prevention of life-threatening situations and provision of decision support to the patient
- the discovery of patient-specific knowledge and the creation of adaptive healthcare pathways which will enable the personalization of treatment.

B. METABO Architecture

The METABO system consists of three main components, namely: the Patient Mobile Device, the Mobile Control Panel and the Central System. The METABO architecture is presented in Fig. 1. The components of the architecture are:

- Patient Mobile Device (PMD): It is a portable device (e.g. a PDA) connected to peripheral sensors and devices (such as continuous and discontinuous glucose meters, physical activity sensors, etc) and contains tools and interfaces for manual data input, data analysis, visualization and feedback.
- Central System (CS): It hosts all patients' databases and manages communication with experts' Control Panel and patients' PMDs.
- Control Panel (CP): It is a complete data management environment that allows the experts evaluate patient's performance and prescribe personalized treatment.

C. Comprehensive Monitoring

The lifestyle monitoring system developed in METABO is responsible to analyze data related with patients' lifestyle which may either directly affect glucose levels (e.g. insulin and food intake, exercise) or influence glucose control and diabetes management in the long-term (e.g. healthy diet, changes in physical activity profile). Specifically, food intake is manually recorded at various levels of detail through suitable tools and interfaces on the PMD, such as food atlas, personalized food lists, simple CHO counting. Drug intake monitoring allows for notification of diabetes-related medications (type, dose and timing) and also considers drugs not specifically targeting at diabetes but directly affecting glucose metabolism (e.g. corticosteroids). Physical activity monitoring concerns the recording and analysis of simple daily activities from pedometers and also of high-intensity exercise from more specialized sensors. In specific cases, the patient can also manually select the type, duration and intensity of exercise activities through suitable interfaces on the PMD. Moreover, regular glycaemic levels both in continuous and discontinuous mode are collected and graphically visualized. The patient's profile enables the maintenance, synchronization and visualization of patient's demographic and clinical data relevant to diabetes management. Also, an agenda is used for permitting the patient to record events as well as for notification of reminders and appointments.

At the same time, feedback is provided to the patients including alerts and recommendations for assisting them in taking decisions, actions as well as educational and motivational material for being compliant with the treatment plans. The system also serves as an alternative communication channel for providing enhanced interaction between the patient and the doctors.

D. Personal Metabolic Modeling & Decision Support

Metabolic modeling focuses on capturing the glucose

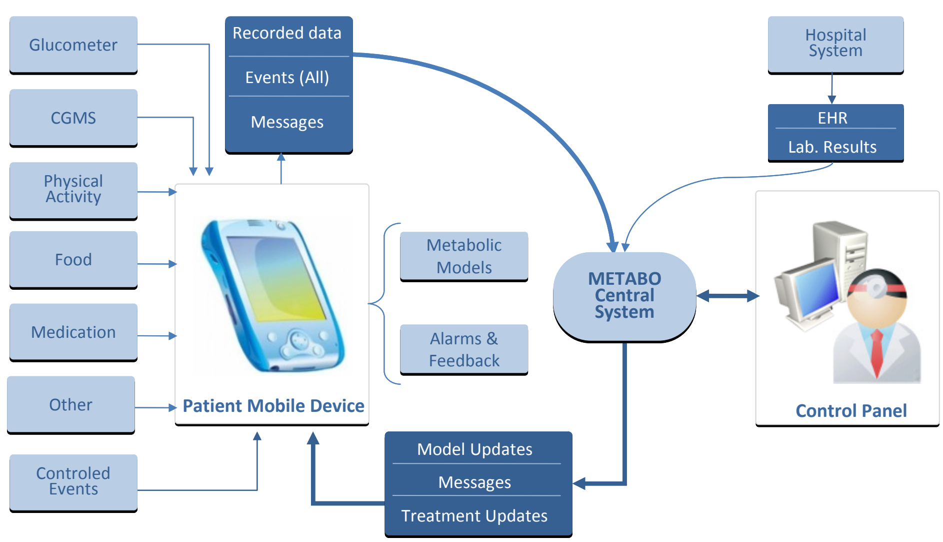

Fig. 1. METABO Architecture.

response to patient's lifestyle and treatment with the aim of predicting blood glucose values and providing immediate alerting and decision support to the patient. In METABO, metabolic modeling is achieved not only by considering the recent glycaemic profile of the patient for analyzing the insulin-glucose interaction but also by exploiting all available contextual information (effect of activity, diet, etc.). In this context, a multi-level scheme is implemented in which mathematic models that describe the main metabolic processes are combined with data mining techniques able to predict glycaemic response.

E. Personalized Treatment

Personalized treatment is realized in terms of clinical pathway workflows. Initially, the "standard clinical pathway" [12] is considered as the default pathway for all patients and contains the basic steps and activities to be followed. During the entrance of a new patient to METABO, he/she is classified and assigned to a specific scenario through suitable classification rules. At this starting point, the physician prescribes to the patient a healthcare pathway which is specific for the scenario that the patient belongs to. This pathway includes some predefined specific steps to be added to the standard clinical pathway. After the initialization, phase in which scenario-specific pathways have been prescribed, the patient is followed by clinicians through the METABO system and his/her healthcare pathway can be further tailored according to the personalized knowledge that has been discovered from all recorded information. In this way, the physician can design personalized pathways by modifying thresholds in rules or inserting/deleting steps in the workflow.

III. METABOLIC MODELING

The architecture of metabolic modelling is presented in Figure 2. The involved architectural modules are:

- Feature Extractor: Calculates and interprets the monitored data to diabetes-related features.
- Personal Metabolic Model: It is composed of two main functional submodules; the physiological model and the glucose predictive model able to provide predictions of the patient's blood glucose profile over a specific time horizon.
- What-If Advisor: Generates personalized advices on actions to be taken or to be avoided according to pre-defined "what-if" queries selected by the patient. The result is mainly triggered by the glucose value predicted by the Personal Metabolic Model according to the selected scenario.
- METABO clinical decision support (CDSS) Model: Generates short-term and long-term alerts on hypo/hyper-glycaemic events and provides recommendations to the patient (e.g,. correction of bolus insulin, extra fingerstick testing, etc.).

Compartmental models are used to compute the absorption function of exogenous glucose and subcutaneous injected insulin exploiting all recorded dietary and insulin-medication related information, respectively. The outputs of the compartmental models along with other measured data are used as input to the predictive model which provides an estimation of future blood glucose levels. In order to model the blood glucose metabolism of a diabetic patient, we consider data mining techniques for regression, as well as time-series analysis methods. It should be noted that different predictive models are expected to be developed for patients that belong to different categories.

The metabolic models and the corresponding CDSS are intended to be installed and run in the Personal Mobile Device every 20-30 min during steady-state and 5-10 min after meal. The personalization of the whole system is achieved through a training process which will be periodically carried out in the Central System.

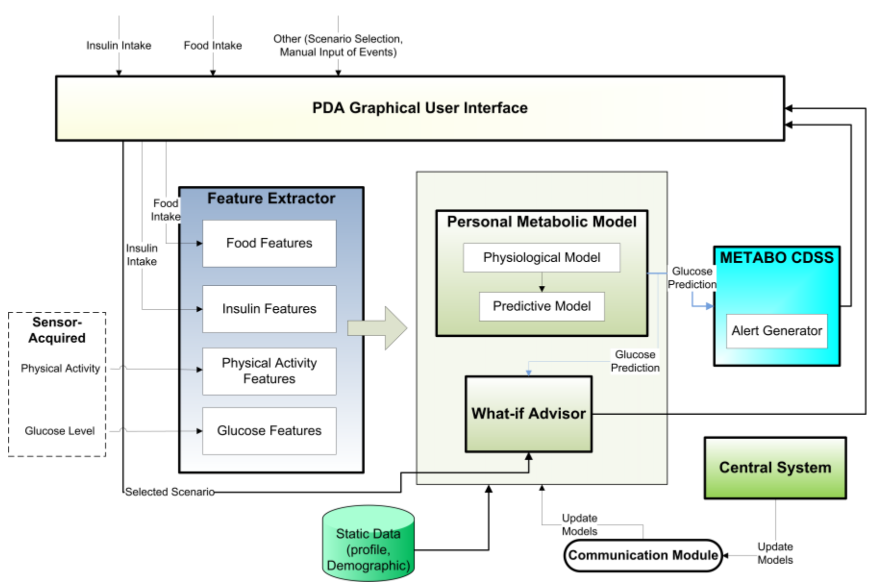

Fig. 2. Architecture of personalized modeling and decision support.

IV. Knowledge Extraction

Knowledge extraction is performed through suitable data mining tools on the Mobile Control Panel that assist the physician in understanding and quantitatively interpreting patient's behavior and evaluate the response to current treatment.

The detection of changes or the assessment of regularities within patient's lifestyle and medical data contribute to the interpretation of his metabolic status. Therefore, physicians are provided with suitable tools for specifying the parameters for which patterns are recognized and the corresponding attributes that establish them. Furthermore, association analysis is used to discover interesting associations among lifestyle parameters, intake of medication, glycaemia and health status collected either for a specific patient or for groups of patients. Intelligent clustering techniques are employed aiming at defining groups of patients who have high-dimensional similarities hidden within all contextual data. Equally important, the physicians have the capability of treating patients by taking advantage of any established knowledge that is available for the identified clusters. Finally, physicians can assess the effectiveness of previous treatments and forecast the performance/response of a patient to new treatments or modifications through suitable classification schemes that assign a patient to a class of medical status.

Generally, with data mining schemes the physicians can gain a deeper understanding of patient's behavior as well as achieve personalized treatment.

V. Clinical Validation

The evaluation of the METABO system will be performed in three phases. The first phase is an observational study which aims at collecting synchronized real-life data for model training and testing. In the second phase, METABO will be tested in clinical practice for usability and acceptability issues, as well as for training the medical team and the patients on its use. Finally, the third phase involves multi-center clinical trials in order to evaluate the clinical effectiveness of METABO and thus to improve metabolic control compared with existing clinical practise. The approximate duration of the three phases is 2 weeks per patient, 2-4 weeks and 6 months, respectively. Selected sub-pilots will be included in all phases with the aim to collect data in specific subpopulations or specific clinical situations. These sub-pilots will mainly investigate the effect food intake, physical exercise and comorbidities on patients' glucose metabolism.

VI. Conclusions

We have presented an innovative system for the monitoring and management of diabetes. The system collects a diverse set of patient data for comprehensive monitoring of lifestyle and disease-related data. All recorded data are analysed and modeled to create personalized models of patients' glucose metabolism and to provide decision support to the patient to self-manage the disease in his/her daily life. In this paper, we paid special attention to the description of metabolic modelling and knowledge extraction.

The techniques and methods developed within METABO will help towards the prediction of clinically critical events in both hypoglycaemia and hyperglycaemia. Moreover, appropriate recommendations are provided to the patients, tailored to their needs, in order to stabilize their diabetic profile. The patient is able to receive personalized advices on actions to be taken or to be avoided according to pre-defined queries. On the other hand, techniques for knowledge extraction help physicians understand and explain patient's behavior and evaluate the response to current treatment or to prescribe a personalized healthcare pathway.

Currently, METABO is at the system development phase. In parallel, data are collected through observational studies for developing and testing metabolic models and the associated decision-support and knowledge extraction schemes.

References

[1] D. Takahashi, Y. Xiao, F. Hu, and M. Lewis, "A survey of insulin dependent diabetes-part I: therapies and devices", *J. of Telemedicine and Applications*, Article ID 639019, vol. 2008, pp. 1-15, 2008.

[2] E. D. Lehmann, "Application of Information Technology in Clinical Diabetes Care", *Medical Informatics*, vol. 21, pp. 255-378, 1996.

[3] E. D. Lehmann, "Application of Information Technology in Clinical Diabetes Care", *Medical Informatics*, vol. 22, pp. 1-120, 1997.

[4] D. Takahashi, Y. Xiao, F. Hu, and M. Lewis, "A survey of insulin dependent diabetes-part II: control methods", J. *of Telemedicine and Applications*, Article ID 739385, vol. 2008, pp. 1-14, 2008.

[5] E. D. Lehmann, and T. Deutch, "A physiological model of glucose-insulin interaction in type 1 diabetes mellitus", *J. Biomed. Eng.*, vol. 14, pp. 235-242, May, 1992.

[6] W. A. Sandham, D. J. Hamilton, A. Japp, and K. Patterson, "Neural network and neuro-fuzzy systems for improving diabetes therapy", in *Proc. 20th Annu. Conf. IEEE Engineering in Medicine and Biology Society*, Hong Kong, China, 1998, vol. 20, no. 3, pp. 1438–1441.

[7] Z. Trajanoski, and P. Wach, "Neural Predictive Controller for Insulin Delivery Using the Subcutaneous Route", *IEEE Trans. Biomedical Engineering*, vol. 45, no. 9, pp. 1122-1134, Sept., 1998.

[8] S. Mougiakakou, A. Prountzou, and K. Nikita, "A Real Time Simulation Model of Glucose-Insulin Metabolism for Type 1 Diabetes Patients", in *Proc. 27th Annu. Conf. IEEE Engineering in Medicine and Biology Society*, Shanghai, China, 2005, pp. 298-301.

[9] H.T. Nguyen, N. Ghevondian, and T. Jones, "Neural-network Detection of Hypoglycemic Episodes in Children with Type 1 Diabetes using Physiological Parameters", in *Proc. 28th Annu. Conf. IEEE Engineering in Medicine and Biology Society*, New York City, USA, 2006, pp. 6053-6056.

[10] V. Tresp, T. Briegel, and J. Moody, "Neural Network Models for the Blood Glucose Metabolism of a Diabetic", *IEEE Trans. Neural Networks*, vol. 10, no. 5, pp. 1204-1213, Sept., 1999.

[11] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, and C. Cobelli, "Glucose Concentration can be Predicted Ahead in Time from Continuous Glucose Monitoring Sensor Time-Series", *IEEE Trans. Biomedical Engineering*, vol. 54, no. 5, pp. 931-937, May, 2007.

[12] P. Barnes, H. Tindall, C. Baynes, D. Hicks, K. McAuley, and M. Shukla "Diabetes Care Pathway", NHS, version 3, April 2008.