Mobile Data Acquisition towards Contextual Risk Assessment for Better Disease Management in Diabetes

Cosmin Cernazanu-Glavan*, Diana Lungeanu**, Stefan Holban*
* Politehnica University of Timisoara, Department of Computers, Timisoara, Romania
** Victor Babes University of Medicine and Pharmacy Timisoara, Department of Functional Sciences/Medical Informatics and Biostatistics, Timisoara, Romania
cosmin.cernazanu@cs.upt.ro, dlungeanu@umft.ro, stefan.holban@cs.upt.ro

Abstract—**Diabetes is one of the serious chronic medical conditions for which the patient's personal involvement into an adequate management of the disease is essential. This paper presents a framework to complement existing approaches in diabetes care through information technology instruments. Existing tools have been targeted at controlling the blood glucose level and generating alerts against hypo/hyper-glycaemia events, individualizing risk prediction based on epidemiological knowledge, or improving patient's education and motivation. We propose analyzing contextual data from patients, i.e. using data collected through mobile devices, on a daily basis, to uncover the implications of the life style and individual behaviour towards the risk for diabetic complications. The paper presents the rationale and the framework for this approach, rather than results from an accomplished work.**

I. Introduction

Within the context of life expectancy increase at a European level, the health status of the general population is in close relation to the successful management of the chronic medical conditions, especially those with a soaring prevalence like the diabetes.

Diabetes mellitus is a global epidemic and a huge public health problem in most countries. More than 300 million people or 6.4% of the world's adult population are living with diabetes. The diabetic population is expected to exceed 438 million by 2030, accounting for 7.8% of the adult population. Diabetes treatment costs have become a global burden: global health care expenditure to treat and prevent diabetes and related complications was roughly $376 billion USD in 2010. The International Diabetes Federation(IDF) predicts this number to exceed $490 billion by 2030.

Managing diabetes mellitus should involve a team consisting of the patient, medical specialist, personal physician and family members. In real life, although under specialized care, most patients get specialized advice only during the medical appointments, so they make their own decisions regarding daily therapeutic interventions. Moreover, too often, the medical experts receive sparse and low quality data regarding the daily life of their patients; therefore, on occasions, they might use unreliable information for making important decisions. During the last decade, e-Health systems proved to bring successful solutions and the role of the smart mobile devices has been constantly increasing, leading to the concept of m-Health [1-4] as a helpful tool in managing chronic diseases (by facilitating a seamless access to expert medical care together with effective self-managing instruments).

Although only partly related to the lifestyle, diabetes has debilitating and life threatening complications for which the prognosis is closely linked to successfully approaching the disease management as a combination of multiple factors, i.e. blood glucose concentration (BGC) in the context of nutrition, physical activity, or lifestyle in general [4-9]. Therefore, being able to make sense of contextual data originating in different sources became a must in controlling the health parameters [9] and personalizing the care recommendations [8, 10, 11]. Many projects world-wide have been focusing on data mining and knowledge extraction for either predicting the risk for long term complications [12], or uncovering the contextual information hidden in the anamnesis procedure and life style [13, 14].

II. Contextual Risk Assessment of Life Style Implications in Diabetes

A. *Development of contextual risk assessment tools*

We propose the development of a contextual risk assessment toolbox (Figure 1), employing instruments from intelligent data analysis to process data collected periodically from the cohort of patients, e.g. on a daily basis by connecting the mobile device to the central database. The data collected through the mobile device would be: blood glucose concentration (BGC) values (preferably by using continuous glucose monitoring systems, at least during the development phase, or directly connecting to the glucose meter device); nutritional values for each meal; physical activity read through the device accelerometer; pulse read through the device built-in camera; spatial locations and the time spent on each of them, read from the device built-in GPS. All the values would be registered as time dependant series.

Additional demographic and static data are considered for contextual processing, e.g. age, sex, physical constitution, co-morbidities, weather conditions, stress indicators. This information is either available in the central database and editable by request (e.g. co-morbidities), or automatically taken from specialized web servers (e.g. weather conditions).

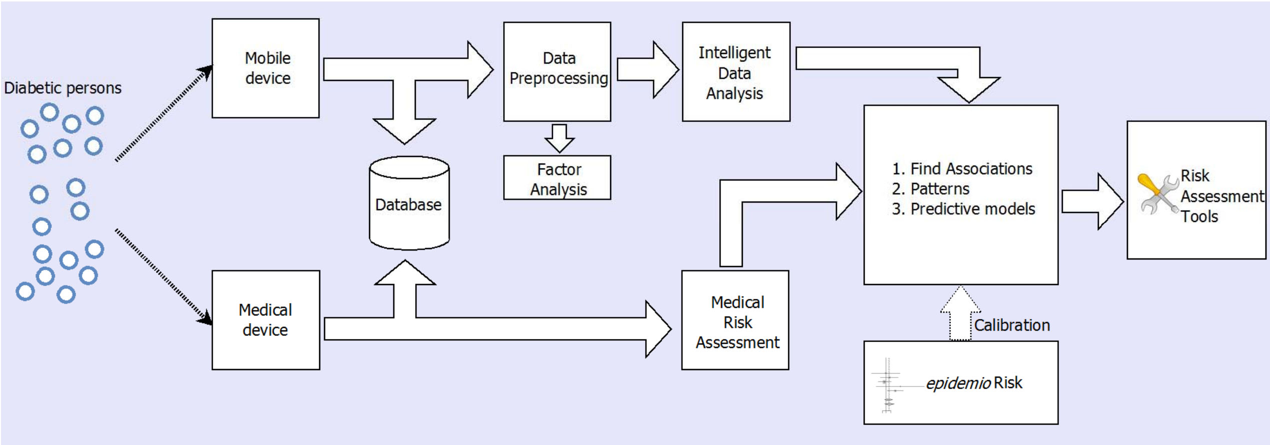

Figure 1. Strategy and stages in developing the tools for contextual risk assessment. The developed algorithms and models undergo a medical validation stage at the point of care and a further step-wise calibration using risk assessment instruments already in use.

All raw data are stored in a central data base, for further reference during the medical risk assessment appointments periodically scheduled and for the validation stage.

Before going through the intelligent data analysis, the collected data undergo a pre-processing phase. Representation of information as vectors of features/characteristics is the keystone of the whole concept and offers flexibility across both the different levels of processing and the different time stages of knowledge extraction. The feature vectors include descriptive statistics (e.g. specific mean and dispersion values calculated for days and months) and newly generated information (e.g. de degree of life organization regarding the meals' time and their nutritional balance, or the regularity of physical activities, the variation range in BGC values and the frequency of hyper- or hypo-glycaemia episodes, etc.)

Monitoring data have already been exploited for knowledge extraction, e.g. Campos-Delgado et al [15] regulated the BGC values with a fuzzy-based controller incorporating information from the monitoring process. Magni and Bellazzi [16] built a stochastic model able to extract BGC values and assess the variability in time dependant series. Other approaches used expert systems [17], cascade learning system [18] or even evolutionary approaches [19] to extract useful information from monitoring medical data.

Our approach is not only more comprehensive (i.e. including medical and non-medical data), but also much simpler and more flexible, with scope for further improvements. Thus, we can apply the same method to evaluate the risk for multiple diabetic complications.

In addition, a factor analysis is performed to ascertain the importance of each parameter in the feature vectors. Starting from pre-processed data and based on the medical risk assessment outcome, by applying knowledge discovery techniques, the risk assessment tools and the risk groups are built. Moreover, the time dynamics for each individual feature vector is considered and subsequent anomaly detection algorithms are applied.

A nice software tool for visualizing and stratifying high dimensional patient data was developed by Harle et al [20]. Another approach consisted in creating a "black box" model of a Support Vector Machine [21], and then used the model to reach certain diagnostic decision in diabetes mellitus, as well.

Although basically similar to other solutions in medical data analysis, our toolbox would offer personalization for each patient, but not in a constraining manner. The personalization process would be achieved over time and with a minimum intervention from the patient. This approach would help a better understanding of the process itself and offer an easy customization for each diabetes complication risk.

At the point of care, medical risk assessment is used for periodically validating the contextual risk models, while a step-wise calibration is carried out employing models based on epidemiological data, already validated and in current medical practice. These epidemiology-based models, as in [22], combine epidemiological data and risk factors from published medical reports to build comprehensive risk models based on current medical knowledge.

On the other hand, we propose uncovering hidden contextual relationships in everyday life, which is a novel approach, complementary to the current medical solutions for addressing the risk assessment. This insight into patient's life potentiates the medical knowledge and offers the important advantage of providing a wealth of reliable data, based on which reliable models can be built.

B. *Planned roadmap and milestones*

Developing the toolbox for contextual analysis would be a sub-study of a larger validation study aimed at assessing risk for long-term complications in diabetes, by employing epidemiological state of the art. Such an approach from two perspectives and at different time granulation, would offer important advantages: (a) making use of existing know-how, methodology and experience in model validation; (b) saving resources (i.e. time and effort) for parallel validation on the same cohort of patients; (c) providing increased power and reliability, by using the same cohort of subjects. In conclusion, it would be beneficial both to patients and to the doctors.

The first milestone is the development of the software application for context data acquisition on smart mobile devices and storing in a central database.

We envision developing a two-level management system for the diabetic disease, i.e. a front-end patientcentred mobile application (front-end level) having risk assessment instruments (back-end level) behind. The mobile access and personalized dialogue is meant to have a motivational impact on the diabetic patient, while giving day-to-day assistance.

The second step consists of data collection. Data would be collected from a sub-group of the main cohort of patients recruited in the epidemiological study. We aim at having 60 patients in this sample, selected based on their representativeness for both types of diabetes. We expect to have a drop-out rate less than 15%, therefore having at least 50 subjects with complete data, which should suffice for the reliability target.

There might be patients who would be reluctant to have a mobile device on them during their current activities and provide personal data on a daily basis. This might impair the representativeness of the patients' cohort from whom we collect the data. However, by explaining to the patients the importance of the study, relying on the increasing number of population who can afford smart mobile devices and on the incentives provided by the motivating front-end mobile application, we expect to be able to have a representative cohort of diabetic patients.

C. *Personalized contextual risk assessment*

After having developed the risk assessment toolbox, the instruments would be put to work, i.e. providing individualized feed-back on a daily basis. The data flow for patient-specific knowledge discovery and personalized risk assessment based on contextual data is shown in Figure 2.

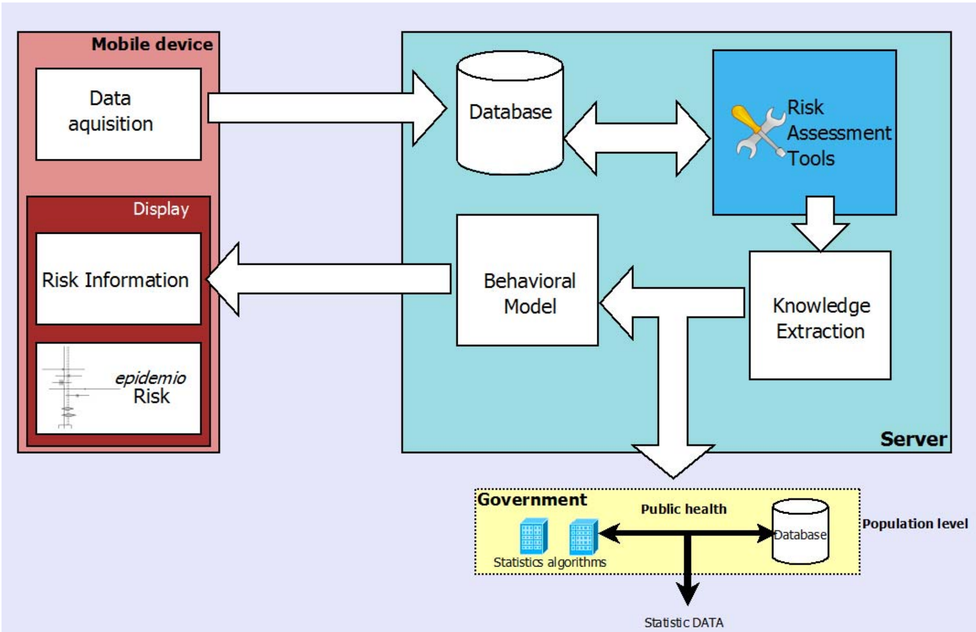

Figure 2. Personalized contextual risk assessment, based on individual data. In parallel, good quality and reliable data accrue in public health databases for further use

The individual data are collected through the mobile device and subsequently processed with the risk assessment tools, able to extract knowledge regarding each particular patient. This understanding of each individual profile is a two-fold gain:

(i) On the one hand, the results consist of an individualized model, i.e. based on patient's own feature vector reflecting its degree of similarity to the risk prototypes, time dynamics of feature vectors, variation and/or anomaly detected in individual behaviour, etc.

(ii) On the other hand, the extracted knowledge goes to a population level database, contributing to knowledge accumulation on the long term. These individual profiles' evolution over time constitute a valuable resource for further validation and data mining approaches.

The individualized behavioural model generates two types of feed-back: a patient-customized feed-back and a professional level profile of risk aimed at the medical specialist (providing an insight into the patient's risk profile in the context of his/her every day life, particular metabolic responses to life situations or life patterns, etc.) The message generated for the patient is comprehensive: individual risk in the context of the epidemiologically known, evidence-assessed medical risk, combined with the own life-style context.

The manner to communicate the information (displayed information, alarms, SMSs) depends on their importance, but it can be also configured by the user. In special cases (multiple hypo/hyper-glycaemia into one day), a third person can be notified.

A special task of the Knowledge Extraction module is the aggregation of information from previous time periods in order to find mistakes in user's behavior. Monitoring for a long time certain aspects from diabetic's life, it can be determined if he/she regularly makes mistakes concerning any aspects of the daily routine (the way he/she performs the physical exercises, wrong insulin doses, flawed approach or algorithm in meals' assessment and planning, etc).

As Figure 2 shows, the mobile application keeps its two-level architecture: a front-end customized interaction and a back-end level for data collecting. On a regular basis, e.g. twice a day, it synchronizes with the server application.

The intelligent instruments would provide the means for personalizing the medical care in general and the management of diabetes in particular. Moreover, the contextual risk models are adaptable both to the context itself and to the normal physiological changes in a diabetic's life, which might prove to be an important characteristic for youngsters suffering of type I diabetes.

We do not aim at replacing existing medical instruments and risk assessment methods, but rather complementing and enriching them.

D. Ethical issues

In order to assess the implications of the daily alimentary and physical behaviour of patients with diabetes, personal data are collected. All information collected is treated as confidential and solely used for the outcomes of the study.

The results of all performed tests are handled by qualified personnel, stored in a protected database, communicated only to the patient, and the name or other personal information data are available only for the on-site employees. Throughout the study, each patient is assigned a unique identifier number for correspondence between medical and behavioural data and a secure de-identification protocol is designed and submitted for ethics committee approval before any data collection begins. The personnel analyzing data have no access to identified or identifiable information.

The same ethical principles and de-identification protocol are applied while using the risk assessment tools for providing personalized feed-back to the diabetic patients.

III. DISCUSSION

Personal risk assessment for long and short term complications is a key factor in diabetes management. We propose analyzing contextual data about patients' life by employing mobile technology and developing appropriate tools for personalized risk assessment and monitoring.

A behavioural and contextual model is essential both for understanding the person's life style and for appropriately customizing the output to be displayed on the mobile device. In order to successfully customizing the user interaction interface (to convey a motivating message), additional information is taken into account (e.g. age, computer skills, personal preferences, level of education, etc.).

Emerging technologies for mobile monitoring and disease management are already offered [23,24] and they are well received by the patients. However, our approach comes with an innovative perspective over the individualized assessment of risk in the context of every day life analysis. We also expect uncovering new information and associations between one's health condition and style of life. While the life expectancy increases, prevention and life style awareness are constant issues in our modern society [25, 26].

Moreover, the population level of information processing is an invaluable long term benefit of the systematic, direct and complete data collection through mobile devices: the data are correct and reliable as a sound base for health policies, while the validated risk assessment methods and tools provide important instruments for further healthcare interventions. They facilitate a comprehensive approach in health policies and public health interventions, taking into account more than statistical and epidemiological evidence employed at present.

IV. CONCLUSIONS

Diabetes usually develops into long-term complications, e.g. retinopathy, chronic kidney disease, neuropathy, chronic cardiac disease, each dramatically impairing the patient's life. The proposed framework is applicable for any of these complications or for combinations of them, as well. This is an important advantage of our approach, i.e. it goes beyond classical controlled studies and is able to offer a holistic assessment of risk, both in developing the toolbox and in subsequent individual contextual risk assessment.

At the same time, the presented framework can be easily extended to processing contextual information for other health related issues, e.g. overweight or cardiac diseases. Although already supported by the classical epidemiological evidence, most healthcare intervention programmes would gain power and sustainability making use of the mobile data collection and adequate contextual interpretation.

REFERENCES

[1] H. Alemdar, and C. Ersoy, “Wireless sensor networks for healthcare: A survey,” *Computer Networks*, vol. 54, no.15, pp. 2688–2710, 2010.

[2] A. Pantelopoulos, and N.G. Bourbakis, “A survey on wearable sensor-based systems for health monitoring and prognosis,” *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 40, no. 1, pp. 1–12, 2010.

[3] S. Krishna, S.A. Boren, and A.E. Balas, “Healthcare via cell phones: a systematic review,” *Telemedicine and e-Health*, vol. 15, no. 3, pp. 231–240, 2009.

[4] B. Holtz, and C. Lauckner, “Diabetes management via mobile phones: a systematic review,” *Telemedicine and e-Health*, vol. 18, no. 3, pp. 175–184, 2012.

[5] American Diabetes Association, “Nutrition recommendations and interventions for diabetes,” Diabetes Care, vol. 31, no. 1, pp. 61–78, Feb. 2008.

[6] American Diabetes Association, “Physical activity/exercise and diabetes,” Diabetes Care, vol. 27, no. 1, pp. 58–62, Jan. 2004.

[7] American Diabetes Association, “Standards of medical care in diabetes,” Diabetes Care, vol. 34, no. 1, pp. 11–61, Jan. 2011.

[8] T. Broens, A. Van Halteren, M. Van Sinderen, and K. Wac, “Towards an application framework for context-aware m-health applications,” *International Journal of Internet Protocol Technology*, vol. 2, no. 2, pp. 109–116, 2007.

[9] L. Grandinetti, and O. Pisacane, “Web based prediction for diabetes treatment,” *Future Generation Computer Systems*, vol. 27, no. 2, pp. 139–147, 2011.

[10] J. Hong, E.H. Suh, J. Kim, and S. Kim, “Context-aware system for proactive personalized service based on context history,” *Expert Systems with Applications*, vol. 36, no. 4, pp. 7448–7457, 2009.

[11] A.J. Jara, M.A. Zamora, and A.F. Skarmeta, “An internet of things---based personal device for diabetes therapy management in ambient assisted living (AAL),” *Personal and Ubiquitous Computing*, vol. 15, no. 4, pp. 431–440, 2011.

[12] S.G. Mougiakakou, C.S. Bartsocas, E. Bozas, N. Chaniotakis, D. Iliopoulou, I. Kouris, S. Pavlopoulios, A. Prountzou, M. Skevofilakas, A. Tsoukalis, K. Varotsis, A. Vazeou, K.

Zarkogianni, and K.S. Nikita, “SMARTDIAB: a communication and information technology approach for the intelligent monitoring, management and follow-up of type 1 diabetes patients,” *Information Technology in Biomedicine, IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 622–633, 2010.

[13] J.M. Tomczak, and A. Gonczarek, “Decision rules extraction from data stream in the presence of changing context for diabetes treatment,” *Knowledge and Information Systems*, vol. 34, no. 3, pp. 521–546, 2013.

[14] METABO Controlling chronic diseases related to metabolic disorders, site:http://www.metabo-eu.org/.

[15] D. U. Campos-Delgado, M. Hernandez-Ordonez, R. Femat, and A. Gordillo-Moscoso, “Fuzzy-based controller for glucose regulation in type-1 diabetic patients by subcutaneous route,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 11, pp. 2201–2210, Nov. 2006.

[16] P. Magni, and R. Bellazzi, “A stochastic model to assess the variability of blood glucose time series in diabetic patients self-monitoring,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 977–985, Jun. 2006.

[17] K. Polat, and S. Gunes, “An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease,” *Digital Signal Processing*, vol. 17, no. 4, pp. 702–710, Jul. 2007.

[18] K. Polat, S. Gunes, and A. Arslan, “A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine,” *Expert System Application*, vol. 34, no. 1, pp. 482–487, Jan. 2008.

[19] X. Chang, and J. H. Lilly, “Evolutionary design of a fuzzy classifier from data,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 4, pp. 1894–1906, Aug. 2004.

[20] C.A. Harle, D.B. Neill, and R. Padman, “Information Visualization for Chronic Disease Risk Assessment,” *Intelligent Systems, IEEE*, vol. 27, no. 6, pp.81–85, 2012.

[21] N. Barakat, A.P. Bradley, and M.N.H. Barakat, “Intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus,” *Information Technology in Biomedicine, IEEE Transactions on*, vol.14, no.4, pp.1114–1120, July 2010.

[22] T. Aspelund, O. Thornórisdóttir, E. Olafsdottir, A. Gudmundsdottir, A.B. Einarsdóttir, J. Mehlsen, S. Einarsson, O. Pálsson, G. Einarsson, T. Bek, and E. Stefánsson, “Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy, ” *Diabetologia*, vol. 54, no. 10, pp. 2525–2532, 2011.

[23] J. Tran, R. Tran, and J.R. White, “Smartphone-based glucose monitors and applications in the management of diabetes: an overview of 10 salient "Apps" and a novel Smartphone-connected blood glucose monitor,” *Clinical Diabetes*, vol. 30, no. 4, pp. 173–178, 2012.

[24] A. Fioravanti, G. Fico, M.T. Arredonto, and J.P. Leuteritz, “A mobile feedback system for integrated e-health platforms to improved self-care and compliance of diabetes mellitus patients, ” *33rd Annual Int Conf of the IEEE EMBS*, pp. 3550–3553, 2011.

[25] S.J.W. Robroek, K.G. Reeuwijk, F.C. Hillier, C.L. Bambra, R.M. van Rijn, and A. Burdorf, “The contribution of overweight, obesity, and lack of physical activity to exit from paid employment: a meta-analysis,” *Scand J Work Environ Health*, vol. 39, no. 3, pp. 233–240, 2013.

[26] M. Hamer, M. Kivimaki, A. Steptoe, “Longitudinal patterns in physical activity and sedentary behaviour from mid-life to early old age: a substudy of the Whitehall II cohort,” *L Epidemiol Community Health*, vol. 66, pp. 1110–1115, 2012.