Proceedings of the
2nd International Workshop on

Artificial Intelligence for Diabetes

held in conjunction with the
16th Conference on Artificial Intelligence in Medicine (AIME)

Vienna, Austria
24th June 2017

Volume Editors

Pau Herrero
Beatriz López
Clare Martin

http://tinyurl.com/aime-aid2017

Committees

Organising Committee

- Pau Herrero - Imperial College London, UK
- Beatriz López - University of Girona, Spain
- Clare Martin - Oxford Brookes University, UK

Program Committee

- Aldea, Arantza - Oxford Brookes University, UK
- Armengol, Eva - Artificial Intelligence Research Institute-CSIC, Spain
- Bellazzi, Ricardo - Università degli Studi di Pavia, Italy
- Bondia, Jorge - Universidad Politécnica de Valencia, Spain
- Bromuri, Stefano - Open University, Netherlands
- Georgiou, Pantelis - Imperial College London, UK
- Kafali, Ozgur - North Carolina State University, USA
- Klonoff, David C. - University of California, USA
- Lustrek, Mitja - Jozef Stefan Institute, Slovenia
- Marling, Cindy - Ohio University, USA
- Neerincx, Mark - Delft University of Technology, NL
- Oliver, Nick - Imperial College London, UK
- Torrent-Fontbona, Ferran - University of Girona, Spain
- Schumacher, Michael I. - University of Applied Sciences Western Switzerland
- Stathis, Kostas - University of London, UK
- Valls, Aida - Universitat Rovira i Virgili, Spain

Acknowledgments and Sponsors

Thanks to Cristian Roman and Pablo Gay for the support received in the management of the Workshop as well as the edition of the proceedings.

This workshop has received funding from the EU Horizon 2020 research and innovation programme under grant agreement No 689810.

Editor's Preface

Following the success of the first edition of the workshop on Artificial Intelligence for Diabetes (AID) held in The Hague, on 2016, under the umbrella of the ECAI conference, we are very proud to present its second edition at the AIME conference.

The workshop starts with an invited talk, which will provide an overview of the utilisation of Artificial Intelligence (AI) in diabetes with a particular focus on gestational diabetes. It continues with regular presentations carried out by people from up to six different countries, including UK, Spain, Italy, USA, Switzerland and Colombia. The afternoon is dedicated to strengthening the research relationships within the community.

On this occasion, the workshop organization takes a step forward to boost the research community by presenting the website ai4diabetes.org. The purpose of this website is to share the results and advances in the management of diabetes using AI techniques. This will enhance progress in the field by establishing synergies among researchers, as well as being a reference point for various stakeholders including users' associations and companies. We therefore warmly invite all researchers working in this field to join.

We hope that you enjoy the workshop and join ai4diabetes.org.

The Organizing Committee

Pau Herrero (Imperial College of London, UK)
Beatriz López (University of Girona, Spain)
Clare Martin (Oxford Brookes University, UK)

Vienna, Austria June 24th, 2017

Table of Contents

Keynote

Artificial Intelligence in Diabetes Care: The Challenge of Supporting Patients in Their Daily Living

M. Elena Hernando

Profesora Titular / Associate Professor
Grupo de Bioingenieria y Telemedicina, ETSI Telecomunicacion, Universidad Politecnica de Madrid, Avda Complutense 30, 28040 Madrid, Spain

Abstract. Diabetes is a complex disease where healthcare professionals and patients have to work together to achieve a good metabolic control. Patients play an active role in their own care and need to have the knowledge to make decisions adapted to their daily living. In such scenario, AI applications make possible to support patients decisions at any scenario of their daily living and additionally they open the door to react at time scales smaller than the programmed face-to-face visits times. In the past decade, the diabetes management paradigm is being transformed by the combination of continuous glucose monitoring and insulin pump data. Additionally, current wearable technologies allow to monitor other physiological parameters (i.e. heart rate, sleep quality, physical activity, etc) that have a close relationship with metabolic control. Nowadays, the challenge is to create proactive AI systems, integrated in the healthcare information systems, accessible from patients devices and able to collect relevant multi-parametric data from patients in a seamless way. Aspects as user interfaces design and user modelling play a crucial role in the success of users interaction with AI systems. This talk will discuss the lessons learned in previous experiences and projects and will present new opportunities for AI in diabetes.

Automatic Adjustment of Basal Insulin Infusion Rates in Type 1 Diabetes using Run-to-Run Control and Case-Based Reasoning

Pau Herrero[1], Peter Pesl[1], Monika Reddy[2], Nick Oliver[2] and Pantelis Georgiou[1]

[1]Centre for Bio-Inspired Technology, Imperial College London, London, SW7 2AZ, UK
[2]Charing Cross Hospital, Imperial College Healthcare NHS Trust, London, W2 1NY, UK

pherrero@imperial.ac.uk

Abstract. People with type 1 diabetes mellitus rely on a basal-bolus insulin regimen to roughly emulate how a non-diabetic person's body delivers insulin. Adjusting such regime is a challenging process usually conducted by an expert clinical. Despite several guidelines exist for such purpose, they are usually impractical and fall short in achieving optimal glycemic outcomes. Therefore, there is a need for more automated and efficient strategies to adjust such regime. This paper presents, and *in silico* validates, a novel technique to automatically adapt the basal insulin profile of a person with person with type 1 diabetes. The presented technique, which is based on Run-to-Run control and Case-Based Reasoning, overcomes some of the limitations of previously proposed approaches and has been proved to be robust in front of realistic intra-day variability. Over a period of 5 weeks on 10 virtual adult subjects, a significant reduction on the percentage of time in hyperglycemia (<70mg/dl) (from 14.3±5.6 to 1.6±1.7, $p< 0.01$), without a significant increase on the percentage of time in hypoglycemia (>180mg/dl) (from 10.2±5.9 to 1.6±1.7, $p=0.1$), was achieved.

Keywords: Diabetes management, insulin therapy, artificial intelligence, case-based reasoning, run-to-run control.

1 Introduction

Type 1 diabetes mellitus (T1DM) is an autoimmune condition characterized by elevated blood glucose levels due to the lack of endogenous insulin production [1]. People with T1DM require exogenous insulin delivery to regulate glucose. Current therapies for T1DM management include the administration of multiple daily injections (MDI) or continuous subcutaneous insulin infusion (CSII) with pumps.

A basal-bolus insulin regimen involves taking a longer acting form of insulin to keep blood glucose levels stable through periods of fasting and separate injections of shorter acting insulin to prevent rises in blood glucose levels resulting from meals. Such regimen attempts to roughly emulate how a non-diabetic person's body delivers insulin. People with T1DM using insulin pumps are able to pre-program different infusion rates

along the day, hence they are able to achieve a finer control [2]. Fig. 1 shows an example of the plasma insulin profile of a typical MDI basal-bolus regimen.

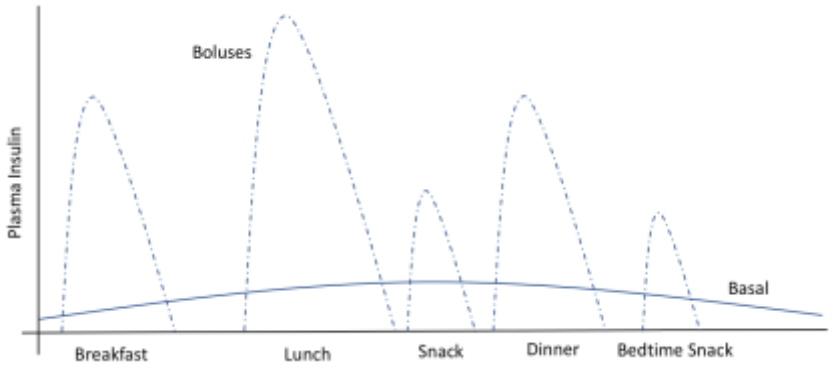

Fig. 1. Plasma insulin profile of a typical basal-bolus regimen. Dashed lines correspond to meal bolus and solid line to the basal insulin injection.

Several medical guidelines exist for adjusting insulin doses basal-bolus regime [3][4], which usually require a cumbersome process of trial and error conducted by an expert clinical. In addition, this regimen is not static and changes over the time due circadian variation in hormone levels (e.g. dawn phenomenon), physical exercise, psychological stress, recurrent illness, or long-term changes in insulin sensitivity due to lifestyle (e.g. obesity) [1]. Therefore, there is a need for more automated and efficient strategies to adjust such regime.

2 Related Work

In order to automatically adjust the basal insulin regimen of a pump, Palerm et al. proposed a Run-to-Run control algorithm that requires six capillary blood glucose measurements along the day at specific times [5]. However, this approach might be impractical in a real-life scenario due to the numerous capillary measurements required.

In the context of an artificial pancreas (i.e. hybrid closed-loop control), a Run-to-Run approach which adapts the basal insulin delivery during the night and the carbohydrate-to-insulin ratio during the day, based on some performance indices calculated from subcutaneous continuous glucose sensor data has been proposed by Toffanin et al. [6].

It is important to note that, apart from considering intra-day changes in insulin sensitivity due to circadian variations of hormonal levels, none of these approaches take into consideration the variability on insulin requirements due to other factors such as physical exercise, alcohol, stress or menstrual cycle.

3 Methods

In this paper, we present a novel technique to automatically adjust the basal insulin profile of a person with type 1 diabetes. For this purpose, a Run-to-Run algorithm [7] incorporating a new control law, which avoids some of the limitations of previously proposed techniques, is introduced (e.g. number of capillary measurements). Then, Case-Based Reasoning [8] is employed to account for intra-subject insulin sensitivity variability. This is done by storing in a subject-specific case-base representing scenarios with significantly different insulin requirements (e.g. dinner after exercise vs. dinner after watching a movie) and therefore, requiring a different basal insulin dosing. Then, a Run-to-Run algorithm is applied to each one of the cases in the case-base in order to adapt the solutions (i.e. basal rate) if considered to be suboptimal.

3.1 Case-Based Reasoning

Case-Based Reasoning (CBR) is an artificial intelligence problem solving framework that solves a newly encountered problem, based on the information obtained from previously solved problems and stored as cases in a case-base [8]. A case is defined by

$$Case_i = \{Problem_i, Solution_i, Outcome_i\}, \qquad (2)$$

where $Problem_i$ is the description of the problem to be solved (e.g. controlling basal glucose); $Solution_i$ is the solution to $Problem_i$ (e.g. basal insulin rate); and $Outcome_i$ is the outcome resulting of applying $Solution_i$ to $Problem_i$ (e.g. glucose outcome).

CBR is usually described in four steps: *Retrieve* the most similar cases to the problem to be solved from the case-base; *Reuse* the solutions of retrieved cases; *Revise* the outcome of the applied solution to the new problem; and *Retain* the new problem if its solution is considered useful for solving future problems. **Figure 4** show the four steps of the CBR cycle (*Retrieve, Reuse, Revise, Retain*) applied to the problem of basal insulin dosing.

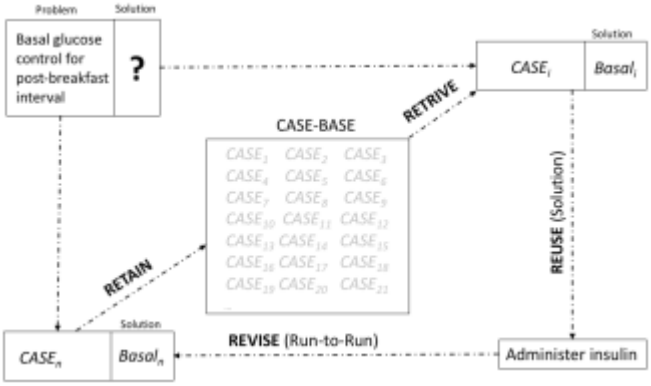

Figure 4. CBR cycle (*Retrieve, Reuse, Revise, Retain*) applied to the problem of basal insulin dosing

In this work, cases are stored in a subject-specific case-base representing scenarios with significantly different insulin requirements (e.g. dinner after exercise vs. dinner after watching a movie) and therefore, requiring different basal insulin rates.

Cases are retrieved from the case-base by computing the Euclidian distance between the current problem and all the cases in the case-base and by selecting the case with the shorter distance.

If the closest retrieved case is still very distant from the current scenario, its solution can be reused (Reuse step) by applying a set of simple rules to guarantee that the applied solution is safe (e.g. reduce basal rate by 30%) and a new case is created for the new scenario. If the retrieved case is equal to the current scenario, then no new case is created.

In order to perform the Revision step, the R2R algorithm presented below is employed, which adapt the solution (i.e. basal rate) of the retrieved, or newly created, case when the glucose outcome is considered sub-optimal based on the analysis of the postprandial CGM measurements.

3.2 Run-To-Run control

Run-To-Run (R2R) is a control methodology designed to exploit repetitiveness in the process that is being controlled [7]. Insulin dosing, and in particular basal insulin delivery, has a repetitive nature. Therefore, R2R control can be used to exploit such characteristic.

The proposed R2R algorithm is based on the hypothesis that the time of the day-time when a person is closest to the fasting condition is right before the meals. Therefore, the pre-meal glucose measurement is used to adjust the basal insulin rate of the previous post-prandial period. Regarding the night-time period, this is split in two parts, from dinner time to 6 hours after dinner and from that time to breakfast time. Note that, unlike the six glucose measurements required by the R2R algorithm by Palerm et al. [5], the proposed algorithm only requires four.

The control law used to adjust the basal rates is defined by

$$Basal_i^{k+1} = Basal_i^k + K \cdot (G_i - G_T), \qquad (1)$$

where sub-index i indicates the time interval, the super-index k indicates the iteration, K is a tuning gain, G_i is the pre-meal glucose value for corresponding interval, and G_T is the glucose target.

It is important to remark that previously proposed R2R algorithms for basal insulin adjustment [5][7] are able deal with intra-day variability due circadian variation, but are not able to deal with inter-day variability due to other factors such as exercise, alcohol, stress, and menstrual cycle. For this purpose, the utilization of Case-Based Reasoning [8] has been proposed in this work.

4 Experimentation and Results

4.1 *In Silico* Evaluation under Intra-Day Variability

The latest version of the UVa-Padova T1DM simulator (v3.2) [9] was used to evaluate the proposed to automatically adapt basal insulin. 10 adult subjects were used for this purpose. Basal insulin infusion rates were initialised the one provided by the the simulator. Meal boluses were computed using a standard bolus calculator [Schmidt 2014] using the default parameters provided by the simulator.

A six-week scenario was selected in order to leave enough time to the basal adaptation mechanism to converge. The selected daily pattern of carbohydrate dose intake was 7am (60g), 1pm (100g) and 7pm (80g).

Intra-day variability and uncertainty were introduced in the simulator as described by Herrero and colleagues [10]. Note that despite all the variability introduced in the simulator, only four different cases were required within the CBR algorithm (i.e. post-breakfast, post-lunch and post-dinner, and night time).

The following glycemic metrics, which are widely accepted by the artificial pancreas community to evaluate glucose controllers, were selected for comparison purposes: mean blood glucose (BG in mg/dl); percentage time in glucose target range [70,180] mg/dl (% in T); percentage time below target (i.e. hypoglycemia) (% < T); percentage time above target (i.e. hyperglycemia) (% > T);

4.2 Glycemic Outcomes

Table 1 shows the results corresponding to the 10 virtual adults after one week of no adaptation and at week 6 after 5 weeks of basal adaptation. **Fig. 2** shows a weekly evolution of three of the evaluated glycemic metrics

Table 1. Glycemic outcomes corresponding to the 10 virtual adult subjects.

Week	% in T	% < T	% > T	BG
#1	75.5±6.7	14.3±5.6	10.2±5.9	118.2±8.6
#6	84.9±7.6	1.6±1.7	13.4±7.2	137.2±9.0
p-value	< 0.01	< 0.01	0.1	< 0.01

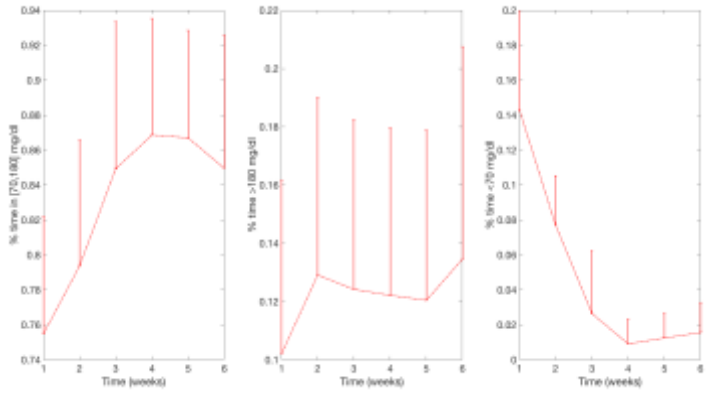

Fig. 2. Weekly evolution of the glycemic metrics. Error bars represent the standard deviation.

When analyzing the weekly evolution of the evaluated glycemic metrics, it can be observed that glycemic metrics take about 3 weeks to converge without significant oscillations towards a steady state value and remain fairly stable along the simulation.

In a real-life scenario, the convergence rate might take longer due to the consideration of more cases representing other scenarios such as exercise, alcohol consumption, hormone cycles or stress.

Clinical trials are required to validate the proposed technique in a real environment. Future work includes the evaluation of the presented technique with a meal bolus adaptation technique.

5 Conclusions

In a virtual T1DM population (10 adults) over a 6-week scenario with intra-day variability, the presented basal insulin adaptation technique significantly reduces hypoglycemia without a significant increase in hyperglycemia.

Acknowledgement

This project has received funding from European Union Horizon 2020 research and innovation programme under grant agreement No 689810 (PEPPER).

References

1. Daneman, D.: Type 1 diabetes. The Lancet 367(9513), 847-858 (2006).
2. Pickup, J.: Insulin pumps. Int J Clin Pract Suppl. 170, 16-9 (2011).

3. Walsh J., Roberts, R. Bailey, T.: Guidelines for insulin dosing in continuous subcutaneous insulin infusion using new formulas from a retrospective study of individuals with optimal glucose levels. J Diabetes Sci Technol.4:1174-1181 (2011).
4. Davidson, PC. Hebblewhite, HR. Steed, RD. Bode, BW.: Analysis of guidelines for basal bolus insulin dosing: basal insulin, correction factor, and carbohydrate-to-insulin ratio. Endocr Pract.14,1095-1101 (2008).
5. Palerm, CC. Zisser, H. Jovanovič, L. Doyle, FJ. A Run-to-Run Control Strategy to Adjust Basal Insulin Infusion Rates in Type 1 Diabetes. Journal of process control 18(3-4),258-265 (2008).
6. Toffanin, C. Visentin, R. Messori, M. Di Palma, F. Magni, L. Cobelli, C. Towards a Run-to-Run Adaptive Artificial Pancreas: In Silico Results. in *IEEE Transactions on Biomedical Engineering*, 99,1-1 (2017).
7. Wang, Y. Gaoc, F. Doyle III, FJ. Survey on iterative learning control, repetitive control, and run-to-run control. Journal of Process Control 19(10), 1589–1600 (2009).
8. Aamodt, A. Plaza, E. Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Commun 7(1), 39-59 (1994).
9. Kovatchev, B., Breton, M. Dalla-Man, C. Cobelli, C. In silico preclinical trials: a proof of concept in closed-loop control of Type 1 diabetes. J. Diabetes Sci. Technol 3(1),44–55 (2009).
10. Herrero, P. Pesl, P. Bondia, J. Reddy, M. Oliver, N. Georgiou, P. Toumazou, C. Method for automatic adjustment of an insulin bolus calculator: in silico robustness evaluation under intra-day variability. Comput Methods Programs Biomed. 2015;119(1):1-8.

A CBR-based bolus recommender system for type 1 diabetes

Ferran Torrent-Fontbona, Beatriz Lopez, and Alejandro Pozo-Alonso

eXiT research group, University of Girona, Girona 17003, Catalunya

Abstract. People with type 1 diabetes mellitus usually need to administer bolus insulin before each meal to keep the blood glucose level in the target glycaemic range. However, the factors involved in the calculation of the appropriate dose can change due to multiple factors and with an unknown relation. This may increase the error in the bolus calculation, and therefore, increase the chances of hypoglycaemia and hyperglycaemia. This paper proposes a bolus recommender system based on case based reasoning developed under project PEPPER, with the objective of recommending personalised and adaptive bolus doses. The system has been tested with *in silico* adults with UVA/PADOVA T1DM simulator. Results show that the use of the proposed bolus recommender system increases the percentage of time in the target glycaemic range.

Keywords: Diabetes, bolus recommender system, case based reasoning, patient empowerment

1 Introduction

Type 1 diabetes mellitus (T1DM) is a chronic disease, which requires people with it to check their blood glucose level and administer insulin to themselves to control and maintain blood glucose in the target range. People with T1DM usually use two types of insulin: bolus insulin, which is a fast acting insulin, and basal insulin, which is a slow acting insulin. However, the calculation of the needed amount of either bolus or basal at each time is not easy and the parameters to calculate it may change due to several factors. Then, T1DM people self-regulate, with the help of clinicians every several weeks or months, the needed doses.

This paper proposes a Case Based Reasoning (CBR) [1] approach to estimate the parameters for bolus calculation and recommend bolus doses. CBR is a lazy learning methodology [5], which consists of using past experiences to solve future problems. CBR traditionally incorporates four main main steps [1]: retrieve, reuse (or adaptation), revise (or evaluation), and maintenance (or storage and management of the case base). These consist of (i) identifying prior similar experiences to the problem to be solved; (ii) adapting the solutions of prior experiences to find a solution to the new problem; (iii) evaluating the outcome of the proposed solution and repair it if necessary; and (iv) storing the current

experience (problem and solution) for further problems and manage the case base.

There are various commercial applications that facilitate the calculation of the amount of insulin, usually the bolus dose for a given basal. The author in [2] provides a wide review of the current applications for bolus calculation. However, these applications require the user to estimate the needed parameters to calculate the bolus, such as the insulin to carbohydrates ratio (ICR) besides the amount of carbohydrates. Moreover, these parameters may change due to different factors and without a known relation. This usually makes these applications very ineffective.

Nevertheless the literature presents approaches capable to iteratively adjust some of the parameters for bolus calculation. In this regard, D. Brown presents in [2] a bolus calculator approach based on Case Based Reasoning (CBR) [1], which considers carbohydrate intakes, preprandial blood glucose levels of a few previous meals in order to calculate the appropriate bolus. Conversely, this paper does not consider previous meals and the implementation of the CBR steps are different. Shashaj *et al.* propose in [6] a run-to-run algorithm [9] instead of a CBR, to iteratively adjust bolus calculator parameters. The point of using a CBR is to have different parameters' values for different situations. The authors in [3,4] combine a run-to-run algorithm and a CBR of prototypes to iteratively select and adjust the parameters of a bolus calculator depending on the context (time of day and physical activity). The revise step of CBR approach presented in this paper is based on the one presented in [3]. However, the proposed revise methodology presents some modifications (see Section 2.3). Moreover, the proposed retrieve, reuse and maintenance steps differ from [3], since the proposed approach is not a CBR of prototypes and, therefore, the size of case base can dynamically change according to the attributes of the cases.

Following this line of research, PEPPER (Patient Empowerment through Predictive PERsonalised decision support) project has the objective of providing a personalised adaptive decision support system for bolus dosing that combines multiple data sources. This paper presents a CBR-based bolus recommender system for T1DM and analyses its performance with UVA/PADOVA T1DM simulator [7].

2 CBR-based bolus recommender system

The CBR-based bolus recommender system presented in this paper has the objective of recommending an appropriate bolus dose to people with T1DM before a meal. In order to do so, the CBR considers a set of attributes that describe the situation or case, and the insulin to carbohydrates ratio (ICR) of the user as the solution of the case. Then, this ICR is used to calculate the appropriate bolus dose for a given amount of carbohydrates. The remainder of the section explains the implementation of the CBR steps for bolus recommendation.

2.1 Retrieve

The retrieve step is responsible for selecting similar cases to the query (or new) case. Retrieve methods usually calculate the distance between the query case and those in the case base and, then, select the closest ones.

The proposed CBR methodology considers the ICR of the user as the solution. Therefore, the considered attributes must be variables that can modify the ICR of a user. These variables include time of day (for intra-day variability), physical activity, stress, hours of sleep, alcohol ingestion, ambient temperature, etc. However, there are contextual factors that usually have a great impact on the ICR, e.g. menstruation and digestive illness. As a consequence the proposed retrieve methodology consists of two steps: context reasoning and selection. Context reasoning consists of choosing the appropriate contextual case base and selection consists of choosing the closest cases (in the corresponding contextual case base) to the query case. Note that this retrieve methodology implies that the CBR system manages not one, but several case bases.

2.2 Reuse

The reuse step consists of adapting the solutions of the retrieved cases to the query case. This paper proposes a weighted average of the retrieved ICRs using the distance between the retrieved cases and the query case.

Once the ICR of the query case is derived, the bolus recommendation is calculated as follows:

$$B = \frac{CHO}{ICR} + \frac{G_c - G_{sp}}{ISF} - IOB \tag{1}$$

where CHO is the amount of carbohydrates of the meal, G_c is the blood glucose, G_{sp} is the standard blood glucose level, ISF is the insulin sensitivity factor and IOB is the remaining active insulin (insulin on board). ISF is calculated as stated in [8] using Equation (2), where W is the weight of the user in kg.

$$ISF = \frac{341.94 ICR}{W} \tag{2}$$

2.3 Revise

After the user administers a bolus (e.g. the recommended bolus) and has the meal, the postprandial phase starts. The proposed revise process is based on the idea proposed in [3], which relies on the assumption that an additional bolus is necessary to bring the minimum glucose value, G_{min}, in the glycaemic range. The value G_{min} is calculated as expressed in Equation (3) as the minimum glucose value measured by the continuous glucose monitor $G_{cgm(t)}$ between t_1 time after the meal time t_m and t_2 time after t_m with $t_1 < t_2$, e.g. $t_1 = 2h$ and $t_2 = 6h$.

$$G_{min} = \min_{t \in \{t_m + t_1, t_m + t_2\}} \{G_{cgm}(t)\} \tag{3}$$

Given the minimum postprandial blood glucose, the ICR is corrected if G_{min} is not in the glycaemic range ($[G_l, G_h]$). Then, the corrected ICR (ICR_c) is calculated according to Equation (4), where ICR_a is the previous ICR used in Equation (1) and α is the learning rate (e.g. 0.5). However, conversely to [3], the revise equation incorporates the learning rate α to smooth ICR changes.

$$ICR_c = (1-\alpha)ICR_a + \alpha \frac{CHO + \frac{G_c - G_{sp}}{341.94/W}}{B + IOB + \frac{G_{min} - G_{sp}}{ISF}} \tag{4}$$

2.4 Maintenance

The ICR can change over time. An example is the intra-day variability, which causes some periodicity in the ICR. This is solved by retrieving the appropriate cases of the case base. However, the ICR could change over time without an apparent periodicity because the physiology of the patient changes (age, body weight, etc.). Therefore, the CBR system has to deal with the concept drift problem. This paper proposes to deal with this problem in the maintenance step by replacing old cases with new cases if these are similar enough. Therefore, the proposed retain process is when there is a candidate query case to be stored in a context case base, check if there is another case within a distance lower than a particular threshold. In such a case, delete the old case and store the new one.

3 Results

The proposed CBR-based bolus recommended system has been tested on ten adult *in silico* subjects using the UVA/PADOVA T1DM simulator [7]. However, since the simulator does not incorporate intra-day variability of the insulin sensitivity, this was artificially introduced as proposed in [3].

The performance of the proposed CBR bolus recommender system has been compared in terms of time with blood glucose in target range (70 mg/dl to 180 mg/dl) with the use of a bolus calculator using Equation (1) and default *in silico* subjects' parameters.

Table 1 shows the average and standard deviation over twenty simulations of the percentage of time that subjects had blood glucose level in target range in 90-days simulations. It shows that *in silico* subjects using the CBR-based bolus recommender system increase their time in the glycaemic range and, in average, reduce the standard deviation, meaning that they have a more stable blood glucose level.

Since the simulation used only considers intra-day variability, the case base of the recommender system of each subject was formed by only four cases, the one used to initialise the case base, and another one for each meal (breakfast, lunch and dinner).

Table 1. Percentage time in target of eleven in silico adults using the CBR-based bolus recommender system or a bolus calculator. In bold face those significantly greater according to Wilcoxon tests, p-value = 0.05.

	CBR	Bolus calculator
Subject 1	79.49 ± 2.02	78.33 ± 5.20
Subject 2	$\mathbf{93.43 \pm 2.07}$	91.15 ± 1.26
Subject 3	$\mathbf{74.95 \pm 2.33}$	68.87 ± 4.36
Subject 4	84.76 ± 5.56	79.02 ± 10.08
Subject 5	$\mathbf{88.84 \pm 4.29}$	80.60 ± 7.00
Subject 6	81.91 ± 1.52	81.18 ± 2.02
Subject 7	69.68 ± 6.98	64.10 ± 6.91
Subject 8	$\mathbf{86.87 \pm 1.26}$	82.61 ± 5.82
Subject 9	$\mathbf{90.18 \pm 2.63}$	83.92 ± 3.58
Subject 10	$\mathbf{83.03 \pm 4.04}$	75.24 ± 7.98
Subject 11	$\mathbf{83.26 \pm 4.64}$	77.63 ± 5.39
Average	83.31 ± 3.39	78.51 ± 5.42

4 Conclusions

People with T1DM need to administer a bolus dose, usually before each meal. The calculation of the appropriate bolus is not easy and requires the estimation of a set of parameters, such as the insulin to carbohydrates ratio, that can change over time due to multiple factors. This paper presents a CBR-based bolus recommended system which has been developed under project PEPPER. The objective of the bolus recommender system is to provide personalised and adaptive bolus recommendations to the users (i.e. people with T1DM).

The system has been tested using eleven *in silico* adults with the the UVA/ PADOVA T1DM simulator. The results show that the proposed system increases the percentage of time of subjects' blood glucose in the glycaemic target range.

Despite the presented results, there is still research to do, especially in the management of missing data and in the maintenance of the case base.

Acknowledgement

This project has received funding from the grant of the University of Girona 2016-2018 (MPCUdG2016) and the European Union Horizon 2020 research and innovation programme under grant agreement No. 689810, `www.pepper.eu.com/`, PEPPER. The work has been developed with the support of the research group SITES awarded with distinction by the Generalitat de Catalunya (SGR 2014-2016).

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications 7(1), 39–59 (1994)

2. Brown, D.: Temporal case-based reasoning for insulin decision support. Ph.D. thesis, Oxford Brookes University (2015)
3. Herrero, P., Pesl, P., Bondia, J., Reddy, M., Oliver, N., Georgiou, P., Toumazou, C.: Method for automatic adjustment of an insulin bolus calculator: In silico robustness evaluation under intra-day variability. Computer Methods and Programs in Biomedicine 119(1), 1–8 (2015)
4. Herrero, P., Pesl, P., Reddy, M., Oliver, N., Georgiou, P., Member, S.: Advanced Insulin Bolus Advisor based on Run-To-Run Control and Case-Based Reasoning. IEEE Journal of Biomedical and Health Informatics 19(3), 1087–1096 (2015)
5. López, B.: Case-Based Reasoning: A Concise Introduction. Synthesis Lectures on Artificial Intelligence and Machine Learning 7(1), 1–103 (2013)
6. Palerm, C.C., Zisser, H., Bevier, W.C., Jovanovic, L., Doyle, F.J.: Prandial Insulin Dosing Using Run-to-Run Control: Application of clinical data and medical expertise to define a suitable performance metric. Diabetes Care 30(5), 1131–1136 (2007)
7. Visentin, R., Dalla Man, C., Kovatchev, B., Cobelli, C.: The University of Virginia/Padova Type 1 Diabetes Simulator Matches the Glucose Traces of a Clinical Trial. Diabetes Technology & Therapeutics 16(7), 428–434 (2014)
8. Walsh, J., Roberts, R., Bailey, T.: Guidelines for Optimal Bolus Calculator Settings in Adults. Journal of Diabetes Science and Technology 5(1), 129–135 (2011)
9. Wang, Y., Gao, F., Doyle, F.J.: Survey on iterative learning control, repetitive control, and run-to-run control. Journal of Process Control 19(10), 1589–1600 (2009)

T2DM phenotypes recognition by Careflow Mining

Arianna Dagliati, PhD[2], Lucia Sacchi, PhD[1], Valentina Tibollo, MS[2], Giulia Cogni[2], MD, Luca Chiovato, PhD[2], John H. Holmes, PhD[3], Riccardo Bellazzi, PhD[1]
1 Univesity of Pavia, Pavia, Italy;
2 Istituti Clinici Scientifici Maugeri, Pavia, Italy,
3 University of Pennsylvania, USA

Abstract. In this work we describe the application of a Careflow Mining algorithm to detect the most frequent careflows from process events. The applied method enriches the detected patterns with clinical data to define temporal phenotypes across the studied population. We defined a novel approach for dynamically stratifying a population. Novel phenotypes are discovered from heterogeneous data of 424 Italian T2DM patients, and compared in terms of metabolic control and complications.

Keywords. Careflow mining, Longitudinal data, Electronic Phenotypes

Introduction

A major source of complexity in the management of Type 2 Diabetes (T2DM) patients arises from the large variability of events they undergo during disease progression. These events, such as hospital admissions, follow-up clinic visits, laboratory tests, and therapy changes, are stored in different data repositories using different formats and occurring in temporal sequences that represent the patient careflow [1]. Current data management technologies [2] are able to gather and merge data distributed in Electronic Health Records (EHRs) and in Administrative Data Warehouses (DW), and enable the access to large quantities of complex data that can be exploited for the management of chronic diseases. The application of longitudinal analysis and careflow discovery to these data enables the recognition of hidden temporal patterns, population stratification, and phenotypes definition. Careflow mining techniques can automatically detect the most frequent temporal patterns from routinely collected administrative and/or clinical data. Once identified, the careflows might be exploited to identify sub-groups of individuals in large cohorts of patients. Electronic phenotyping has been defined as the use of EHRs and secondary data to recognize a set of clinical conditions and characteristics that define a sub-cohort of patients, which is a computable phenotype [3], [4]. Careflow mining techniques can be used as a novel approach for electronic phenotyping. Examples of frameworks able to automatically extract phenotypes from EHR data are still limited and none of them is able to represent longitudinal data. While they combine different sources of EHR data, they tackle the complexity through rule-based approaches [5], or they rely on semi-supervised methods to extract patient narratives, but without explicitly representing temporal relations among events [6]. Recent applications [7] have shown the possibility of performing accurate predictions of hospital acquired complications as a function of healthcare system exposure using temporal clinical data. Yet few works are applied to characterize T2DM. Some cross-sectional approaches [8] demonstrate improved performance in using EHR data, when compared to basic covariates. In [9] authors apply an high-throughput clinical phenotyping algorithm to recognize T2DM cases and controls, and demonstrate the impact of insufficient longitudinal data on the accuracy

of an algorithm, thus suggesting careful consideration of temporal aspects in the design and execution of T2DM phenotyping algorithms. In our previous work [10], we define electronic temporal phenotyping by mining careflows, and we show how careflow mining (CFM) can be a useful instrument to identify the evolution of clinical states over time, across a population of breast cancer patients. In this work we present an application of the method implemented in [10] to analyze T2DM evolution, especially in terms of complications and assessment of metabolic control. Moreover, we describe two additional features of the algorithm: (i) a careflow similarity measure based on Jaccard similarity and (ii) the possibility of detecting AND splits in a careflow.

Methods

Cohort description and electronic phenotypes approach. To apply the CFM algorithm and detect novel T2DM phenotypes, we considered a cohort of 424 patients diagnosed with T2DM after 2004, and currently treated at the Istituti Clinici Scientifici Maugeri (ICSM) hospital in Pavia, Italy. In this work, we focus on the analysis of the administrative data collected on this cohort by the Pavia Local Health Care Agency (ATS). In particular, on the basis of clinicians' interests, the algorithm was applied to the events patients underwent in the period between the diagnosis of T2DM and the first visit at ICSM, thus taking into account patients' encounters within health care facilities other than ICSM, or GPs follow-ups. As a first analysis step, we identified the sources of complexity for T2DM patients: (i) T2DM careflows can last for a long time (in our case for more than 10 years), (ii) the disease progression pace is slow and characterized by relatively frequent modifications (e.g changes in therapies and in the frequency of follow ups), (iii) sequences of events are very heterogeneous and influenced by multiple conditions. Phenotypes were extracted from all the procedures (ambulatory visit, follow ups and hospitalizations) executed during the considered observation window (from diagnosis to first visit at ICSM), previously preprocessed in order to obtain a consistent representation. After running the algorithm on these data, we focused on the characteristics of the extracted phenotypes to evaluate their clinical relevance. Clinical variables were used to enrich the phenotypes and to detect if transitions among events were due to meaningful medical episodes or if different careflows were associated witht different clinical responses. Qualitative representations of clinical time series of Glycated Hemoglobin (Hba1c) were used to characterize patients' sub-cohorts, and to understand if the mined paths are reliable drivers for phenotypes.

The CFM algorithm [10] takes into account the temporal nature of the data, explicitly including both process and clinical information in a two steps approach. The algorithm mines the most frequent patients' histories form process events, and illustrates the evolution of the disease enriching the histories with clinical information. The use of events temporal alignment to recognize similar profiles of patients adds a temporal dimension to the electronic phenotype. Its results are represented using a directed acyclic graph (DAG) where events are connected through arcs that represent temporal connections. T2DM disease progression can be illustrated through paths in the DAG.

AND events. The results of the CFM algorithm are DAGs, where events occur in sequences, and parallel events are not directly mined during the discovery process. Process mining techniques, like the alpha algorithm [11], extract parallel routings if sequences like <B, C> and <C, B> are detected. We applied detection strategy inspired to the same relations among events, with further restrictions on the moment when these sequences occur. Once the histories are mined, the algorithm searches for sequences of events in the form <B, C> and <C, B> that occur exactly at the same point of the history. When the condition is detected the two sequences are merged and represented together as B AND C. The number of patients undergoing the events merged into the AND, and in the following event, are summed. Once the AND detection is performed a possible loss of information can occur, especially in terms of temporal and clinical characteristics of the cohorts. However, this feature could be very useful in the case of large data variability: form the electronic phenotyping point of view, the number of sub cohorts could can reduced when AND events are recognized.
Similarity measure. Mining careflows from process data allows precise descriptions of patients conditions and treatments [12]. Differently from other approaches, where the temporal similarity is exploited during clinical workflows mining [13], the careflows extracted from our algorithm represent sequences that are identical in timing and sequence of events' occurrence. We therefore decided to use a similarity measure that allows the comparison of temporal phenotypes after the algorithm execution. To this end, we exploited the Jaccard similarity coefficient [12]. The Jaccard index is commonly used in bioinformatics to compare genes and metabolic pathways [14], [15]. It is defined as the size of the intersection divided by the size of the union of the sample sets. In our application, the Jaccard similarity coefficient is computed for each pair of mined phenotypes, on the basis of the whole set of events patients underwent, not only the mined ones. The result is a symmetric NxN matrix, where N is the number of the mined frequent histories.

Results

In this work, the clinical events that we considered are ambulatory procedures, inpatient hospitalizations, short procedure unit (SPU) visits, and drug purchases. Each data stream was pre-processed as follows: (i) Ambulatory Procedures were grouped on the basis of the taxonomy used at ASL that indicates the ward where procedures have been performed; (ii) Hospitalization and SPU Procedures ICD9-CM codes were mapped into the first levels of the Clinical Classifications Software (CCS) for ICD-9-CM [16]; (iii) Drug Purchases, classified on the basis of ATC levels, were filtered in order to take into account only treatments for diabetes and further aggregated on the basis of clinicians' suggestions, (iv) T2DM Diagnosis, and First Visit at ICSM are extracted from the hospital EHR and used as censoring events. The most frequent careflows mined by the algorithm are shown in Figure 1. In this figure, it is possible to identify eight phenotypes ($\Phi 1$, .., $\Phi 9$), $\Phi 7$ was excluded as associated with less than 2 patients. The patients in $\Phi 1$ exit from the careflow immediately after diagnosis. Other sub-cohorts had one or more laboratory exams before the first visit at the hospital ($\Phi 2$) or a generic specialist visit into another clinical center ($\Phi 5$). Patients identified by the careflows $\Phi 3$, $\Phi 8$, $\Phi 9$ experienced one or more radiology procedures after the

diagnosis. Patients associated with Φ6 follow either the careflow "Diagnosis → Laboratory Exam → Radiology → End" or the careflow "Diagnosis → Radiology → Laboratory Exam → End". The event "Laboratory Exam **AND** Radiology" is detected and represented.

To assess the informative value of the careflow mined from process data, Hba1c values were selected as the most meaningful biomarkers of T2DM control. When compared with Kruskal-Wallis chi-squared test, Hba1c values results are significantly different (p-value << 0.01) between different phenotypes. Φ3 shows the highest value of Hba1c in the observed period. Thus, we assumed that Φ3 patients are more complicated, and clinicians spend more efforts in trying to stabilize their metabolic control. A first characteristic of these patients is that they have the longest period between the diagnosis and the first hospital visit (3.4 years, compared to 1.9 years in the rest of the cohort). Also drug purchasing information indicates that these patients receive more complex and complete diabetic treatments once they receive hospital care: they start to be treated with Glp-1_Analogues, Insulin, Repaglinide and Alpha glucosidase inhibitors. We also studied T2DM complications onset in the different groups. Patients in Φ3 show an interesting pattern: they experience more Macrovascular events (Acute Myocardial Infarction, Occlusion of Carotid Artery, Stroke) than any other phenotype, while these events arise after a longer period (on average 891.25 days after the first visit, compared with the mean of the rest of the population that is 758 days). Moreover, when Φ3 was grouped with the most similar phenotypes on the basis of the Jaccard similarity (Φ5, Φ6), the Kaplan-Mayer survival curve for Macrovascular events suggests a faster worsening of the conditions of these patients. Although the computation of the total sample size needed to reach a significance of 0.05 and a power of 0.8 is of 1,023 subjects, compared to the 351 available in the data set.

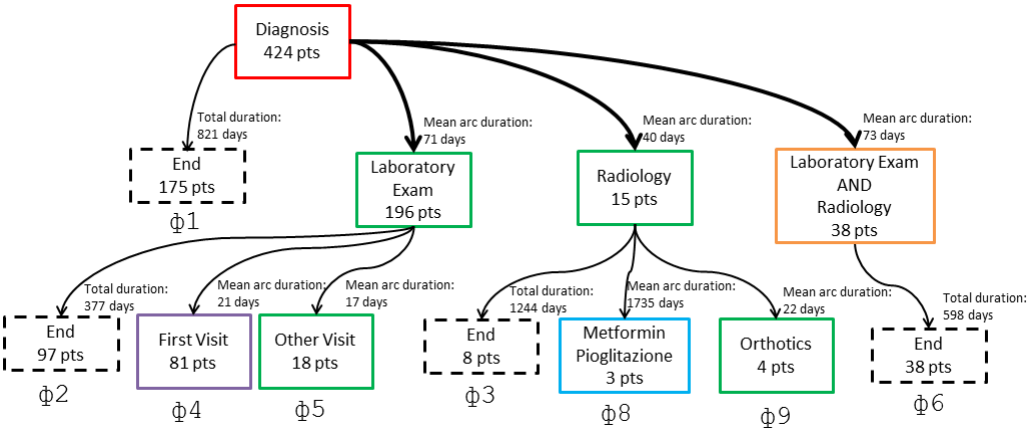

Fig. 1 The mined careflow.

Discussion

In this work we described how patients' management processes can be better investigated through longitudinal heterogeneous data. Carfelow mining approaches illustrate the evolution of the disease. Mining careflows through process data allows leveraging

on information already well-structured to be represented as event logs, avoiding costly procedures of preprocessing the entire corpus of clinical data. Process data trigger phenotypes definition through the segmentation of the entire population and search space reduction. The patterns mined are the substrate for the additional inferences, clinical data are used to enrich and compare specific pathways and prove the medical relevance of the automatically extracted phenotypes. The detected temporal patterns make possible to reconstruct clinical pathways and estimates metabolic control deviations, which might arise during the process of care. The presented analysis framework identifies interesting clusters of patients with similar care histories, allowing patient risk profiles to be reassessed accordingly.

Future efforts will be dedicated to improve the similarity measure, for example through the exploitation of Edit distance and OSS measures [17], especially for histories composed of larger number of events.

References

[1] S. Quaglini, M. Stefanelli, A. Cavallini, G. Micieli, C. Fassino, and C. Mossa, "Guideline-based careflow systems," *Artificial Intelligence in Medicine*, vol. 20, no. 1. pp. 5–22, 2000.

[2] I. S. Kohane, S. E. Churchill, and S. N. Murphy, "A translational engine at the national scale: informatics for integrating biology and the bedside," *J. Am. Med. Informatics Assoc.*, vol. 19, no. 2, pp. 181–185, 2012.

[3] R. L. Richesson *et al.*, "Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory.," *J. Am. Med. Inform. Assoc.*, vol. 20, no. e2, pp. e226-31, 2013.

[4] R. Rachel and S. Michelle, "Electronic Health Records-Based Phenotyping," *Rethinking Clinical Trials*, 2014. [Online]. Available: http://sites.duke.edu/rethinkingclinicaltrials/ehr-phenotyping/.

[5] G. N. Nadkarni *et al.*, "Development and validation of an electronic phenotyping algorithm for chronic kidney disease.," *AMIA Annu. Symp. Proc.*, vol. 2014, pp. 907–16, 2014.

[6] Y. Halpern, S. Horng, Y. Choi, and D. Sontag, "Electronic medical record phenotyping using the anchor and learn framework.," *J. Am. Med. Inform. Assoc.*, p. ocw011, 2016.

[7] J. L. Warner, P. Zhang, J. Liu, and G. Alterovitz, "Classification of hospital acquired complications using temporal clinical information from a large electronic health record," *J. Biomed. Inform.*, vol. 59, pp. 209–217, 2016.

[8] A. E. Anderson, W. T. Kerr, A. Thames, T. Li, J. Xiao, and M. S. Cohen, "Electronic health record phenotyping improves detection and screening of type 2 diabetes in the general United States population: A cross-sectional, unselected, retrospective study.," *J. Biomed. Inform.*, vol. 60, pp. 162–168, 2015.

[9] W. Q. Wei, C. L. Leibson, J. E. Ransom, A. N. Kho, and C. G. Chute, "The absence of longitudinal data limits the accuracy of high-throughput clinical phenotyping for identifying type 2 diabetes mellitus subjects," *Int. J. Med. Inform.*, vol. 82, no. 4, pp. 239–247, 2013.

[10] A. Dagliati, "Temporal electronic phenotyping by mining careflows of breast cancer patients.," *J Biomed Inf.*, vol. Feb, no. 66, pp. 136–147, 2017.

[11] W. Van Der Aalst, T. Weijters, and L. Maruster, "Workflow mining: Discovering process models from event logs," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 9, pp. 1128–1142, 2004.

[12] L. J. Frey, L. Lenert, and G. Lopez-Campos, "EHR Big Data Deep Phenotyping. Contribution of the IMIA Genomic Medicine Working Group.," *Yearb. Med. Inform.*, vol. 9, pp. 206–11, 2014.

[13] C. Combi, M. Gozzi, B. Oliboni, J. M. Juarez, and R. Marin, "Temporal similarity measures for querying clinical workflows," *Artif. Intell. Med.*, vol. 46, no. 1, pp. 37–54, 2009.

[14] S. Zhou *et al.*, "Diversity of Gut Microbiota Metabolic Pathways in 10 Pairs of Chinese Infant Twins," *PLoS One*, vol. 11, no. 9, p. e0161627, 2016.

[15] L.-Y. Wang *et al.*, "Time-dependent variation of pathways and networks in a 24-hour window after cerebral ischemia-reperfusion injury," *BMC Syst. Biol.*, vol. 9, no. 1, pp. 1–11, 2015.

[16] A Elixhauser, C. Steiner, and L. Palmer, "Clinical Classifications Software (CCS), 2014," *U.S. Agency Healthc. Res. Qual.*, no. November 2013, pp. 1–54, 2014.

[17] C. Gòmez-Alonso and A. Valls, "A similarity measure for sequences of categorical data based on the ordering of common elements," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5285 LNAI, pp. 134–145.

Drug Purchases Behavioral Patterns to Support Clinical Decisions in T2D Patients

Arianna Dagliati, PhD[2], Lucia Sacchi, PhD[1], Valentina Tibollo, MS[2], Giulia Cogni[2], MD, Luca Chiovato, MD, PhD[2], Riccardo Bellazzi, PhD[1]
1 Univesity of Pavia, Pavia, Italy;
2 Istituti Clinici Scientifici Maugeri, Pavia, Italy,

Introduction

Within the European funded project Mosaic, ten clinical and technical partners collaborate to improve Type 2 Diabetes (T2D) patients' characterization and to evaluate the risk of increasing disease complexity. The aim of the project was to implement a novel approach for the management of chronic T2D populations. Multiple databases from hospitals, local health care agencies and population studies were the project pillars. Clinical data from hospital information systems collected during routine practice, together with data recorded for billing purposes, provide an integrated research setting and enable a broader vision of individual patients' histories. One of the main project aims was the discovery of well-defined drug purchases patterns behaviors derived from administrative data streams. The developed method to retrieve drug purchases patterns was integrated as part of a Clinical Decision Support System (CDSS), structured as a Dashboard. The CDSS has the capability to improves patients' characterization and provide new insight in the disease evolution while combining these findings with clinical time series derived from EHRs. Dashboards are capable of merging information from different data sources and provide a visual summarized representation of key performance indicators. Their visual properties allow to provide summaries of a big volume of data through easy to read color coded graphical format in order to deliver an intuitive assessment of complex clinical conditions [1]. CDSS structured as dashboards are especially useful to represents temporal patterns and have been already used in pharmacotherapy and medication surveillance to enhance cognitive recognition of patterns in time [2]–[4].

Methods

The analysis of drug purchases data has been performed in the following steps:

1) Quantification of drug exposure. Each drug purchase is described by its Defined Daily Dose (DDD), which allows computing the expected number of therapy days related to that purchase. As the information about the actual drug dosages a patient should take was not available, drug purchases were used as a proxy for estimating drug intake. Some pharmaco-epidemiology studies [5]–[7] investigate the impact of drug adherence in the arising of T2D acute events and in glycemic control. These works successfully exploit Proportion of Days Covered (PDC) measure to assess patient exposure and compliance to specific drug. DDD values were used as proxy for the dosing regimen and PDC was exploited for the drug exposure analysis. PDC is calculated as the number of days with drug on hand divided by the number of days in the specified time interval. The PDC can be multiplied by 100 to yield to a percentage. Once the PDC values had been computed, it was possible to extract the data necessary to compute indicators and thresholds that show, for each patient, his/her behavior in purchasing drugs during the evolution of the disease.

2) Comparison of patient's behavior with a reference population. In order to assess if a patient does not follow the median behavior of the population we have computed a statistical test, i.e. the PDC value of the entire observation period of a patient has been compared with the population through a Wilcoxon test. For each patient, and for each of the drug purchased, the result of the procedure indicates if the patient behavior was in line with the rest of the population (when the p value < 0.05) or if he/she tend to purchase more or less of a certain drug.

3) Visual analytics for extracting interesting patterns. For each time interval where the PDC has been computed, a temporal abstraction (qualitative label) that indicates if the quantity of the drug purchased in the period is under or over a certain threshold has been graphically represented in the dashboard. In particular, the procedure to calculate the thresholds was tailored on the single patient for each of the purchased drug. The 33rd and 66th percentiles of the PDC was computed for each drug purchased, the value of PDC was compared to these thresholds and associated to a label that indicates a behavior that diverge form the patient's standards (over or under purchasing a drug). To give to physicians a general indication, when the PDC is in range, the label carries also the information about two fixed thresholds of 80% and 100%. The periods during which there are no prescriptions are labelled as interruptions.

Results

The visual analytics approach exploited to show in the CDSS the purchases of drugs during the disease evolution is shown in Figure 1 and Figure 2.

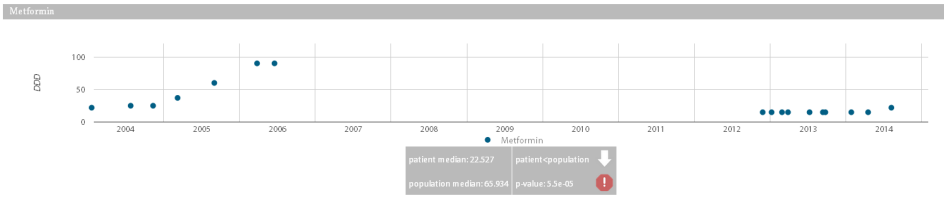

Figure 1. The graphs show drugs' purchases during the disease evolution. Each purchase is quantified with the DDD associated to each active principle. The grey box indicates if the patient purchased larger or smaller quantities of the drug (with the arrow pointing up or down) and if this difference is significant (with a red or green icon), when compared to other patients that are treated with the same drug.

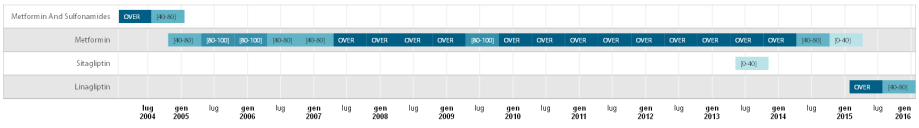

Figure 2. Therapy purchasing behavior graphs illustrate the PDC with a 6-month granularity. For each semester, the represented time line indicates through colors and labels the compliance of the patients to a certain therapy, on the basis of specific thresholds computed on the basis of his/her behavior. OVER labels indicate a value greater than the 66th percentile of the patients, other labels indicate the value range of the PDC for the considered period.

The system was evaluated near the Istituti Clinici Scientifici Maugeri hospital (Pavia, Italy) following a pre-post approach. The actions of 9 physicians were monitored for three months before (on 353 patients) and after (on 353 patients) the introduction of the CDSS. The comparison of the collected indicators in the pre and post phase shows a reduction of visits

duration, an increasing in the prescription of screening exams and in interventions on the physical activity.

Discussion

In this work we show how administrative data can be exploited as a population marker. The detection of drug consumption patterns allows stratifying the population on the basis of subjects' purchasing attitude. Thanks to the developed approach, physicians have the opportunity to use as a proxy of the complexity of clinical conditions the patterns of drug purchases. Both the procedures to derive the behavior indicators and the exposure thresholds produce results which summarize complex and rich information. The visual representation of overloaded information turns the complexity into an opportunity to gain insight, conclusions, and to support decision making.

References

[1] R. A. Simms, H. Ping, A. Yelland, A. J. Beringer, R. Fox, and T. J. Draycott, "Development of maternity dashboards across a UK health region; Current practice, continuing problems," *Eur. J. Obstet. Gynecol. Reprod. Biol.*, vol. 170, no. 1, pp. 119–124, 2013.

[2] D. R. Mould and M. C. Dubinsky, "Dashboard systems: Pharmacokinetic/pharmacodynamic mediated dose optimization for monoclonal antibodies," *J. Clin. Pharmacol.*, vol. 55, no. S3, pp. S51–S59, 2015.

[3] L. R. Waitman, I. E. Phillips, A. B. McCoy, I. Danciu, M. Robert, M. S. Halpenny, C. L. Nelsen, D. C. Johnson, J. M. Starmer, and J. F. Peterson, "Adopting real-time surveillance dashboards as a component of an enterprisewide medication safety strategy," *Jt. Comm. J. Qual. Patient Saf.*, vol. 37, no. 7, pp. 326–332, 2011.

[4] B. E. Dixon, A. M. Jabour, E. O. Kelly, and D. G. Marrero, "Improving Medication Adherence for Chronic Disease Using Integrated e-Technologies," *Medinfo 2013*, vol. 18, no. 7, p. 2013, 2013.

[5] F. M. de Vries, P. Denig, S. Vegter, H. J. Bos, M. J. Postma, and E. Hak, "Does a cardiovascular event change adherence to statin treatment in patients with type 2 diabetes? A matched cohort design," *Curr. Med. Res. Opin.*, vol. 31, no. 4, pp. 595–602, Apr. 2015.

[6] G. A. Nichols, A. G. Rosales, T. M. Kimes, K. Tunceli, K. Kurtyka, P. Mavros, and J. F. Steiner, "Impact on glycated haemoglobin of a biological response-based measure of medication adherence," *Diabetes. Obes. Metab.*, Apr. 2015.

[7] T. Simon-Tuval, N. Triki, G. Chodick, and D. Greenberg, "The association between adherence to cardiovascular medications and healthcare utilization," *Eur. J. Heal. Econ. HEPAC Heal. Econ. Prev. care*, Jun. 2015.

Negative results for the prediction of postprandial hypoglycemias from insulin intakes and carbohydrates: analysis and comparison with simulated data

Fabien Dubosson[1], Natalia Mordanyuk[2], Beatriz López[2], and Michael Schumacher[1]

[1] AISLab group, University of Applied Sciences Western Switzerland
[2] eXiT group, University of Girona

Abstract. Diabetic patients usually take insulin bolus right before eating a meal. A wrong dosage of insulin may lead to a hypoglycemia. Being able to anticipate such insulin-induced, postprandial hypoglycemias would enable warning of the patients about the risk associated with the quantity of insulin they are planning to take. In this work, we explore the feasibility of predicting these postprandial hypoglycemias by using information available at pre-meal time, such as glucose levels, planned insulin intakes and carbohydrates estimations. First, an experiment has been done on a dataset acquired on real patients, for which several classes of machine learning algorithms have been tried. The obtained results do not offer predictions that are useful enough to consider any usage in real-life applications. These kinds of datasets — acquired on real patients — suffer heavily from missing data and incorrect carbohydrates estimations though. In order to analyse the impact of these flaws on the obtained results, the same experiment has been run on a simulated dataset. Results support that even with the simulated dataset, which does not have missing data and which has precise carbohydrates intake, these features alone are not able to predict postprandial hypoglycemia. Therefore, improving the quality of patients annotations is not enough to solve the problem, and using these features without further features engineering does not offer good results.

Keywords: Hypoglycemia, Prediction, Insulin, Carbohydrates

1 Introduction

Type I diabetes patients usually take insulin bolus before starting a meal. The dosage of insulin depends on the quantity of carbohydrates they are going to ingest, and this quantity is usually estimated by the patients themselves. These estimations may be error prone, and taking too much insulin may lead to hypoglycemias which may be dangerous for patients lives in the long term.

In this work we explore the possibility of anticipating such postprandial hypoglycemias by using the information available juste before a meal: patient's glycemia levels, planned insulin intakes and carbohydrates estimations. Being able to do so would enable warning of the patients about the risk associated with the dosage of insulin they are planning to take.

A big part of the work going in the direction of hypoglycemias predictions focuses on live predictions from Continuous Glucose Monitoring (CGM) signals [1, 2]. Some are specially oriented toward patients with insulin pumps, as they try to detect when to stop basal insulin to prevent hypoglycemias [3, 4]. Our work, however, is targeting different audiences and objectives. The goal is to offer the possibility to anticipate hypoglycemias to patients not wearing any insulin pump or any continuous glucose monitoring device. This may assist patients to be confident about the planned insulin intake.

In [5], Reddy et al. presented a bolus calculator based on CGM device using Case-Based Reasoning (CBR) methods. They evaluated it while acquiring the dataset used in this study. Their work is closer to the objectives of this experiment than previously cited related works. However, their method is based on CGM devices as opposed to our work.

The two datasets are presented in Section 2. The Section 3 explains the methodology and presents the results, and finally results of this research are discussed in Section 4.

2 Datasets

This work is based on two datasets: a first one that have been acquired on type I diabetes patients, and a second one that have been generated by a simulator. These two datasets are presented in the following subsections.

2.1 Real-patients dataset

The database used in this study was provided by the Imperial College London [5]. The population consists of 10 patients (men and women), aged between 24 and 74 suffering from type I diabetes. The patients have been enrolled to a 6-weeks acquisition session, during which some of the patients' measures have been collected. In addition, the patients were provided a mobile application with an insulin bolus dose decision support system based on CBR [6].

The devices used for gathering the glucose levels of patients were Medtronic iPro2 Recorders[3]. The CGM took readings every 5 minutes, 24 hours a day over 7-10 consecutive days, and the units used to measure the glucose were *mmol/l*. The CGM data was supplemented with additional information provided by the patients, such as carbohydrates ingested (g), the insulin shots (U), alcohol consumption (*True* or *False*), or whether the patient performed some physical activities or not before the meal.

Initially, there were 2404 logbook entries (all the patients together). After a data cleaning process (due to missing values caused by the errors of the glucose

[3] `http://www.professional.medtronicdiabetes.com/ipro2-professional-cgm`

recorder, superfluous entries concerning sensor events, or due to small time interval meals), 1158 entries remained. Since the logbook entries are associated with CGM recordings, we also excluded all the sequences having too many missing data in the CGM, leaving a total of 891 entries.

2.2 Simulated dataset

The simulated dataset has been generated using the *UVA/PADOVA Type 1 Diabetes Simulator* [7, 8].

It consists of 10 virtual patients, for whom 500 days of data have been generated for each. The data consists of CGM measurements every 5 minutes with the associated quantity of carbohydrates ingested and insulin units taken. This represents 1500 entries per virtual patients for a total of 15 000. The necessary pre-processing have been done in order to convert units to the ones of the real dataset. Most where straightforward conversions, except for the insulin dosage. In the simulated dataset, the insulin dosage are reported for insulin pumps, with the following definition:

$$\mathrm{IIRt} = \mathrm{IU} \times \frac{6000}{\mathrm{BW}} + \frac{\mathrm{Basal}}{60} \times 5 \times \frac{6000}{\mathrm{BW}} \tag{1}$$

Where *IIRt* is the value provided by the simulator, *IU* is the insulin units we want to know, *Basal* in the basal insulin, and BW is the patients body weights. In order to use the same type of data as in the real-patients dataset to be able to compare results, the formulae has ben transformed as follow:

$$\mathrm{IU} = \frac{\mathrm{IIRt} \times \mathrm{BW} - \mathrm{Basal} \times 500}{6000} \tag{2}$$

The *IU* value has then been used as the insulin intake feature in the experiments.

3 Methodology and Results

Each log entry was labelled according to Zecchin [9]: the glucose below 70 mg/dL (3.889 mmol/l) was considered hypoglycemia, while glucose above 180 mg/dL (10 mmol/l) was considered hyperglycemia; other glucose levels correspond to the normoglycemia state. In this experiment we regrouped normoglycemia and hyperglycemia in the same class because we only want to predict hypoglycemias.

Since most nadirs occur around 2 hours after ingestion of a carbohydrate meal [10, 9], we looked for hypoglycemias between 1.5 hours and 2.5 hours after a meal for defining the insulin-induced hypoglycemia class labels. This results, as expected, in quite imbalanced datasets: 827 non-hypos for 64 hypos in the real-patients datasets, 14 923 non-hypos for 67 hypos in the simulated one.

Several families of machine learning algorithms have then been tried with the Python *scikit-learn* library[4]: linear classifier, nearest neighbors, random forest,

[4] `http://scikit-learn.org/`

extra trees and SVM. We used typical machine learning good practices: evaluations have been done with 10-folds cross-validation, within each cross-validation loop the models parameters have been fine-tuned with a cross-validated grid-search on the training set, and the classes weights have been set to compensate the class imbalance.

The selection of the scoring method is a more subjective task. The accuracy alone is not useful on imbalanced datasets because answering always the majority class gives high scores without being useful at all. Precision and Recall both have their utility. Precision relates to the number of false alarms, which is important to keep patients adherence to the system. Recall relates to the percentage of detected hypoglycemias, and we are trying to avoid hypoglycemias so it's important to detect the maximum number of them. The F1 score has been selected as it gives the harmonic means between precision and recall, but this choice over other scoring method is arbitrary.

Table 1. F1 scores

Algorithm	**Patients dataset** (Prior=13.4)	**Simulated dataset** (Prior=0.89)
Linear classifier	10.93	**1.15**
Nearest neighbors	5.76	0.55
Random forest	2.41	0.50
Extra trees	6.72	0.58
SVM	**13.70**	0.81

The prior F1-score of the real-patients dataset is 13.4% and the best result is obtained with the SVM classifier, which achieves 13.7%. The majority of non-hypos are classified correctly, as well as the majority of hypos that are also classified correctly. The number of false-alarm and the number of missed hypoglycemias prevent any application in real-word though. On the simulated dataset the results are not really better: the prior F1-score is 0.89% and the linear classifier achieves a score of 1.15%. This does not allow either any use in real-world scenarios.

4 Discussions

The best results obtained on the patients dataset are slightly better than the dataset priors, but are not good enough to be used in any real-world application. One point that have been noted while working on this dataset are unexplained glucose peaks, and the most likely cause of this should be missing carbohydrates information. Another related weak point of such real-life dataset is the fact that accurate carbohydrates estimations are difficult. In order to evaluate if these flaws may be explaining the difficulties of prediction on the real-patients dataset, we reproduced the same experiment on the simulated dataset.

The experiment on the simulated dataset, does not seem to offer significant improvements. Being generated by a simulator, the data should however behave

more predictively than a human body because the human body is much more complex and is sensible to the external environment. The simulator is also giving the exact and complete set of carbohydrates and insulin intakes, in contrast to human annotations.

This experiment supports first that the weakness in patients annotations alone is not enough to explain the difficulties of hypoglycemias prediction, otherwise good results would have been reached on the simulated dataset. Second, the absence of better results on the simulated dataset, despite having been tested with different families of machine learning algorithms, shows that the features used (glucose levels, insulin intakes and carbohydrates estimations) can not predict hypoglycemias as-is. This does not mean however that the problem may not be solved through further features engineering or by using more complex models. A more detailed study of the features and models would help to identify the factors involved in the difficulties of such predictions, and would permit to propose guidelines for improving the acquisition of new Type I diabetes datasets.

References

1. Palerm, C.C., and Bequette, B.W.: Hypoglycemia detection and prediction using continuous glucose monitoring—a study on hypoglycemic clamp data. (2007)
2. Eren-Oruklu, M., Cinar, A., and Quinn, L.: Hypoglycemia prediction with subject-specific recursive time-series models. (2010)
3. Hughes, C.S., Patek, S.D., Breton, M.D., and Kovatchev, B.P.: Hypoglycemia prevention via pump attenuation and red-yellow-green "traffic" lights using continuous glucose monitoring and insulin pump data. Journal of diabetes science and technology 4(5), 1146–1155 (2010)
4. Buckingham, B., Cobry, E., Clinton, P., Gage, V., Caswell, K., Kunselman, E., Cameron, F., and Chase, H.P.: Preventing hypoglycemia using predictive alarm algorithms and insulin pump suspension. Diabetes technology & therapeutics 11(2), 93–97 (2009)
5. Reddy, M., Pesl, P., Xenou, M., Toumazou, C., Johnston, D., Georgiou, P., Herrero, P., and Oliver, N.: Clinical Safety and Feasibility of the Advanced Bolus Calculator for Type 1 Diabetes Based on Case-Based Reasoning: A 6-Week Nonrandomized Single-Arm Pilot Study. Diabetes Technol Ther 18, 487–493 (2016)
6. Aamodt, A., and Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications 7(1), 39–59 (1994)
7. Man, C.D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., and Cobelli, C.: The UVA/PADOVA type 1 diabetes simulator: new features. Journal of diabetes science and technology 8(1), 26–34 (2014)
8. Visentin, R., Dalla Man, C., Kovatchev, B., and Cobelli, C.: The University of Virginia/Padova type 1 diabetes simulator matches the glucose traces of a clinical trial. Diabetes technology & therapeutics 16(7), 428–434 (2014)
9. Zecchin, C.: Online Glucose Prediction in Type 1 Diabetes by Neural Network Models. Università degli Studi di Padova (2014).
10. Brun, J.F., Fedou, C., and Mercier, J.: Postprandial reactive hypoglycemia. Diabetes & Metabolism 26(5), 337–351 (2000)

A clinical decision support system for the diagnosis of Diabetic Retinopathy

Aida Valls[1], Antonio Moreno[1], Domenec Puig[1], Emran Saleh[1], Jose Escorcia-Gutierrez[12] and Jordina Torrents-Barrena[1] Pedro Romero-Aroca[3], Ramon Sagarra-Alamo[3], and Sofia de la Riva-Fernandez[3]

[1] Dep. d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Spain
[2] Universidad Autónoma del Caribe, Colombia
[3] Ophthalmic Service, University Hospital Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira i Virgili, Spain

Abstract. Diabetes has several secondary complications. In this paper we study Diabetic Retinopathy, which consists of the split of some blood vessels on the eye, causing a blur vision or even blindness. Early detection may improve significantly the treatment of this disease and minimize its consequences on the vision loss. For this reason, we are working with a Catalan hospital to build a Clinical Decision Support System to help family physicians to estimate the risk of developing diabetic retinopathy using the information they obtain in the regular visits done to a diabetic patient. Such system may improve the diagnosis in Primary Health Care centers, which enables a better scheduling of other screening tests like the inspection of the eye's fundus with cameras, which is done in specialized ophthalmologist services.

1 Introduction

Diabetic retinopathy (DR) is one of the main complications of diabetes, being a common cause of blindness for this kind of patients. A *Clinical Decision Support System* (CDSS) is being developed in order to help clinicians to estimate the risk of developing diabetic retinopathy as early as possible. It is called RETIPROGRAM. Such system may improve the diagnosis in Primary Health Care centers, reducing the workload of ophthalmologist services. Nowadays, all diabetic patients are screened every two years by taking images of the eyes' fundus. Lesions in the eye can be detected in those images, being the main source for DR diagnosis. However, using a fixed time for screening is not an appropriate model because there are cases in which retinopathy will be detected too late, while there are many patients that do not require this screening because the disease will not appear. In a study made with the patients of the area of South Catalonia (from 2007 to 2014) [2], it was observed that incidence was stable between 2007 and 2011 (around 8.1%) but since 2011 it has continuously increased until almost 9%. It has also been observed that it is still increasing and it will probably be around 10% in 2020. If the screening focuses on this small subset of

the population, resources would be used in a more cost-effectively way. In fact, in [3] a study of the cost of screening shows that it can be reduced by a personalized screening timing based on each patient's risk factors. This information is systematically gathered in the Electronic Health Record (EHR). Thus, the goal is that the CDSS can use data from the EHR to decide when to make the eye analysis, focusing the use of resources on the patients that really need them. In that way, we can avoid unnecessary screenings on patients, saving the time of doctors and patients.

2 Methods

The system is now under development but the plan is to integrate it in the software tools managed in the Health Care Centers of Catalonia (Spain). The system will get information directly from the databases and will use the patient's data to make an estimation of the risk of developing DR.

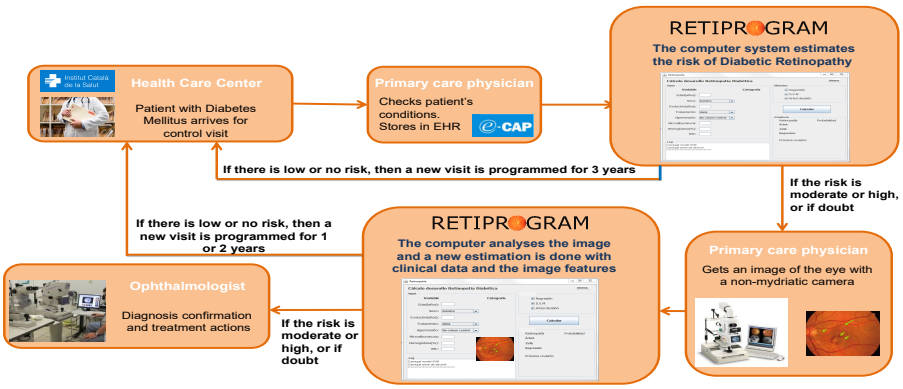

Fig. 1: CDSS architecture for the diagnosis of the risk of DR on the diabetic patients

Figure 1 shows the architecture of the CDSS. It consists of six steps. First, the patient is visited in the Health Care center by a family doctor. If the patient is diabetic, the doctor will use the RETIPROGRAM to get the values of some relevant attributes from the patients EHR. Third, a set of fuzzy decision rules are applied to calculate the risk of DR. If the risk is low, the next visit of the patient is scheduled for two or three years. If risk is moderate or high, a photo of the eyes' fundus is taken using a non-mydriatic camera. This test is neither invasive nor painful, but it needs to be done by a specialist because the light and calibration conditions are very important to get a good image. This image is then fed into the RETIPROGRAM. The system then applies computer vision techniques to automatically detect any kind of lesion in the eye, especially those

near the macula. This information is displayed to the doctor and the next visit is scheduled according to the DR risk level. In fact, if lesions are found, the risk will be set to a high level and an urgent appointment with an ophthalmologist will be programmed. Up to now, there is no CDSS with these characteristics. Although there exist some works on the assistance of RD detection using computer vision techniques, there are no tools that also integrate the analysis of the information in the Electronic Health Record due to the lack of a systematic collection of all data from patients. Since 2007 several analytical, metabolic and demographic data have been systematically collected and stored in the EHR of patients treated at Sant Joan de Reus University Hospital (Catalonia, Spain). This hospital serves an area of Catalonia with a population of 247,174 inhabitants, having 17,792 patients with Diabetes Mellitus. The goal of the current research project is to develop this CDSS. Developing such kind of system needs the use of different Artificial Intelligence techniques. RETIPROGRAM includes two main components:

PART 1. Analysis of the EHR

Due to the uncertainty present in the data and the suitability of linguistic models, a system based on fuzzy linguistic rules is being constructed. The identification of 9 relevant attributes for the diagnosis of DR was done in [2]. Using a dataset with around 2,000 patients we have trained and tested different models. The construction of a single fuzzy decision tree was proposed in [5]. To improve the classification accuracy, an ensemble of classifiers has been obtained using a Fuzzy Random Forest [4]. The rules generated indicate that the most relevant attributes are hypertension, medication and age. As the number of records is not large, we have cases in which the lack of examples makes the diagnosis of DR unfeasible. This case is communicated to the user.

PART 2. Analysis of the eye fundus images

Diabetic Retinopathy causes blindness due to the dilatation of small blood vessels (microaneurysms) which produces intra-retinal haemorrhages and fluid leaking composed by lipoproteins and lipids (exudates). In order to detect these lesions in a 2D image, it is necessary to first remove the normal structures of the eye (blood vessels, optic disc). Until now, we have explored two algorithms (Convexity Shape Prior and GrabCut) for the segmentation [1]. This cleaning of the image should allow us to leave in the image only the pathological elements and the image background color. In an ongoing study, the Convexity Shape Prior method is also being used to locate the macula, because lesions near the macula are the ones that cause blindness. Then, some techniques for identification and counting of the types of lesions will be needed.

Finally, the features extracted with the image processing should be later analyzed together with the EHR data, to improve the accuracy of the system.

3 Conclusions and research challenges

The results obtained up to now give us confidence in the possibility of constructing a powerful decision support system for estimating the risk of developing diabetic retinopathy. Building this CDSS presents some research challenges that we should address: (1) the integration of different methods of rule construction seems necessary as the current results only achieve a sensitivity and specificity of 80%; (2) the dataset is highly imbalanced (with less cases of patients with DR) and this fact requires of special training techniques, both for rule induction as well as for image processing; (3) the combination of the features extracted from the images with the attributes of the EHR must be studied; (4) the applicability of RETIPROGRAM in different populations should also be carefully considered to validate the system.

Therefore, there are still some important open questions to work on before deploying the system into real use.

Acknowledgements. This work is supported by Spanish projects from Instituto de Salud Carlos III (PI15/01150), FEDER funds and URV grants (2015PFR-URV-B2-60, 2016PFR-URV-B2-60). Mr. Saleh has a predoctoral FI grant by Generalitat de Catalunya. Mr. Escorcia-Gutiérrez has a predoctoral grant by Fundación Carolina.

References

1. Escorcia-Gutierrez, J., Torrents-Barrena, J., Romero-Aroca, P., Valls, A., Puig, D.: Interactive Optic Disk Segmentation via Discrete Convexity Shape Knowledge Using High-Order Functionals, Artificial Intelligence Research and Development, vol. 288, pp. 39–44. IOS Press (2016)
2. Romero-Aroca, P., de la Riva-Fernandez, S., Valls-Mateu, A., Sagarra-Alamo, R., Moreno-Ribas, A., Soler, N.: Changes observed in diabetic retinopathy: eight-year follow-up of a spanish population. British Journal of Ophthalmology 100(10), 1366–1371 (2016)
3. Romero-Aroca, P., de la Riva-Fernandez, S., Valls-Mateu, A., Sagarra-Alamo, R., Moreno-Ribas, A., Soler, N., Puig, D.: Cost of diabetic retinopathy and macular oedema in a population, an eight year follow up. BMC Ophthalmology 16, 136 (2016)
4. Saleh, E., Moreno, A., Valls, A., Romero-Aroca, P., de la Riva-Fernandez, S.: A Fuzzy Random Forest Approach for the Detection of Diabetic Retinopathy on Electronic Health Record Data, Artificial Intelligence Research and Development, vol. 288, pp. 169–172. IOS Press (2016)
5. Saleh, E., Valls, A., Moreno, A., Romero-Aroca, P., de la Riva-Fernandez, S., Sagarra-Alamo, R.: Diabetic Retinopathy Risk Estimation Using Fuzzy Rules on Electronic Health Record Data, Modeling Decisions for Artificial Intelligence, vol. 9880, pp. 263–274. Springer (2016)

Towards a Formal Model of Type 1 Diabetes for Artificial Intelligence

D. Brown, C. Martin, D. Duce, A. Aldea, R. Harrison

Department of Computing and Communication Technologies,
Oxford Brookes University, Wheatley, Oxford, OX33 1HX, England
dbrown@brookes.ac.uk

Abstract. Artificial Intelligence (AI) is potentially useful for cost effective diabetes self-management. One research priority for the development of robust and beneficial AI concerns the use of formal verification techniques to model such self-modifying systems. In the context of diabetes, formal methods may also have a role in fostering trust in the technology as well as facilitating dialogue between a multidisciplinary team to determine system requirements in a precise way. In this paper we show how the formal modelling language Event-B can be used to capture safety-critical constraints associated with AI systems for diabetes management.

1 Introduction

Most people with Type 1 Diabetes (T1D) have to perform complex insulin dose calculations several times a day. The computations must consider multiple factors and can occur in a variety of contexts that might affect cognitive load. Adaptive solutions that use artificial intelligence (AI) to replace this human decision-making are emerging [5, 6], but the application of such technology to safety-critical healthcare problems is often met with high levels of resistance [4]. Individuals are concerned about loss of control and question whether the perceived risk is outweighed by the potential benefit.

Such systems must therefore behave robustly and deliver the intended benefits consistently in order to foster trust. Some computer science research priorities that have been identified towards this goal include formal *verification* and *validation* [7]. Although features of AI systems, such as nondeterminism, can make them difficult to verify, it should be possible to build them from individual components that have been proved correct. Additionally, they might be used within the confines of a deterministic safety systems that have been formally verified.

Another dimension is introduced by the user interface of AI systems. Tools exist to support the formal design and analysis of human-machine interfaces, including the behaviour of objects such as buttons and keyboards [3]. Such technologies can be used to give a lightweight, formal analysis of the safety requirements of the interface. They may therefore have a role in facilitating dialogue between a multidisciplinary team, which could include clinicians and computer scientists, in evolving a common understanding of these requirements in a precise

way. In this paper we present a simplified formal specification of part of a T1D bolus recommender system in an attempt to show how some properties might be expressed. This might form part of a model that a trustworthy AI system must guarantee not to violate.

2 Formal Specification

Event-B is a formal modelling language developed by Jean-Raymond Abrial as an evolution of the B-Method [1]. The purpose of modelling with Event-B is to prove that a model will work prior to implementation and for early identification of problems with the system requirements. The Rodin Platform is an Eclipse-based integrated development environment (IDE) for Event-B modelling. It has an associated repository of resources to aid existing and new users of the platform, including a plug-in to integrate ProB into the IDE, which provides a method to systematically check the model for errors through animation. Other plug-ins include automated code generation from the Event-B specification and external provers to assist the automatic proofs. For example, the Atelier B prover plug-in was used in the work described below.

An Event-B model is validated through the successful discharge of *proof obligations*, which imply that the invariants and type declarations of the model are not violated by any aspect of the model. For example, if a variable x has type $\mathbb{N}1$, the assignment of 0 to x would fail to discharge the proof obligation created by $x \in \mathbb{N}1$ since x must be greater than or equal to 1. A key feature of the Event-B language is its incremental modelling process through *refinement*. Refinement allows gradual development of the specification starting at an abstract level, which can then be refined to include new functionality and gradually move the model towards the concrete implementation.

2.1 Event-B Specification of T1D System Constraints

The specification begins with an abstract specification of the system, focussing on the constraints of the system variables. Figure 1 provides an extract of the variables and invariants that includes the maximum bolus dose, and upper and lower target blood glucose ranges. A full specification of the variables and invariants is included in [2]. Event-B is limited to natural and integer numerical types, which presents a problem when decimal values are required. To overcome this, all natural numbers or integers in the specification are represented through multiplication by the power of 10, allowing precision to one decimal place (e.g. 5.5 is represented as 55). Comments in the specification will be used help to reinforce this representation.

The types of *maxBolus*, *targetRangeUpper* and *targetRangeLower* are defined in Figure 1 as natural number ($\mathbb{N}$). This invariant alone implies that the value of these variables must be greater than or equal to 0, and that negative values are not permitted. *inv2* states that *maxBolus* is restricted to a maximum bolus value of 50.0 and *inv7* ensures that the target range is non-empty.

MACHINE t1dm_m0
VARIABLES
`maxBolus` Maximum bolus dose limit
`targetRangeUpper` Target blood glucose range upper value
`targetRangeLower` Target blood glucose range lower value
INVARIANTS
`inv1` : $maxBolus \in \mathbb{N}$
`inv2` : $maxBolus \leq 500$
The maximum bolus dose must be in the range [0.0,50.0] IU
`inv3` : $targetRangeUpper \in \mathbb{N}$
`inv4` : $targetRangeUpper \geq 55 \land targetRangeUpper \leq 150$
Target blood glucose range upper value must be in the range [5.5,15.0]
`inv5` : $targetRangeLower \in \mathbb{N}$
`inv6` : $targetRangeLower \geq 30 \land targetRangeLower \leq 80$
Target blood glucose range upper value must be in the range [3.0,8.0]
`inv7` : $targetRangeLower \leq targetRangeUpper$
Target blood glucose range upper value must be in the range [3.0,8.0]

Fig. 1. Extract of machine variables and invariants

Both of the invariants for *maxBolus* state that $0 \leq maxBolus \leq 50$ insulin units. Additionally, the invariants for *targetRangeUpper* and *targetRangeLower* also define the constraints: $5.5 \leq targetRangeUpper \leq 15.0$ mmol/L, and $3.0 \leq targetRangeLower \leq 8.0$ mmol/L. These invariants prevent the variables from breaching the constraints imposed by the Accu-Chek® Aviva Expert blood glucose meter.

MACHINE t1dm_m0
EVENTS
Event $bolusCalc \mathrel{\widehat{=}}$
Bolus calculator event for instances where s is > 0 and $<$ maxBolus
any
s
s is the parameter of bolus suggestion
where
`grd1` : $s \in \mathbb{N}$
`grd2` : $s \leq maxBolus$
The bolus solution suggestion is $\leq$ maxBolus
then
`act1` : $bolusSuggestion := s$
Sets the bolus suggestion value to parameter s
end

Fig. 2. Abstract bolus calculation event

An abstract event for the bolus suggestion *bolusCalc* is described by Figure 2. One parameter is specified for the event, the bolus suggestion s. The event has two guards which check that parameter s is a natural number and that s is less

than or equal to the maximum bolus dose $maxBolus$:

$$bolusSuggestion \in \mathbb{N} \land bolusSuggestion \leq maxBolus$$

The *bolusCalc* event does not satisfy the requirements of the application alone as it is possible that the bolus suggestion may be negative or greater than the maximum bolus dose. In these circumstances, the bolus dose should be set to 0 and the *maxBolus* respectively. To model this, two new events are added to the model to allow for these circumstances *bolusCalcNeg* and *bolucCalcMax*. These events are shown in Fig. 3.

MACHINE t1dm_m0
EVENTS
Event *bolusCalcNeg* $\widehat{=}$
Bolus calculator event for instances where s is < 0
any
s
s is the parameter of bolus suggestion
where
`grd1` : $s \in \mathbb{Z}$
`grd2` : $s < 0$
The bolus solution suggestion is < 0
then
`act1` : $bolusSuggestion := 0$
Sets the bolus suggestion to 0 as negative values are not permitted
end
Event *bolusCalcMax* $\widehat{=}$
Bolus calculator event for instances where s is $>$ maxBolus
any
s
s is the parameter of bolus suggestion
where
`grd1` : $s \in \mathbb{N}$
`grd2` : $s > maxBolus$
The bolus solution suggestion is $>$ the maximum bolus dose
then
`act1` : $bolusSuggestion := maxBolus$
Sets the bolus suggestion to maxBolus as values $>$ maxBolus are not permitted
end

Fig. 3. Additional abstract bolus calculation events

At present, the model does not include the computation to be performed to determine the bolus dose, but instead defines the abstract events required. This abstract model can now be refined into a more concrete one. In [2] it is refined to include a case base and deterministic bolus calculator. All proofs are automatically discharged by the prover and animation with ProB indicates that

the events are only activated by parameters which satisfy the guards. The same technique could be applied to the refinement for a nondeterministic calculator including AI.

3 Conclusion

This short example illustrates how critical safety constraints of a T1D bolus calculator involving AI can be captured formally. We believe that the construction of a validated, definitive core task model for the T1D individual and associated recommender system, verified using formal techniques, should be a research priority for artificial intelligence in diabetes. However, there are more difficult challenges related to the ability to verify AI systems that modify themselves, possibly repeatedly, over time. It is not yet known whether straightforward verification tools can be applied to this broader setting [7].

Acknowledgments

This work has received funding from the EU Horizon 2020 research and innovation programme under grant agreement No 689810.

References

1. Abrial, J.R., Abrial, J.R.: The B-book: assigning programs to meanings. Cambridge University Press (2005)
2. Brown, D.: Temporal Case-based Reasoning for Insulin Decision Support. Ph.D. thesis, Department of Computing and Communication Technologies, Oxford Brookes University (2015)
3. Fayollas, C., Martinie, C., Palanque, P., Masci, P., Harrison, M.D., Campos, J.C., e Silva, S.R.: Evaluation of Formal IDEs for Human-machine Interface Design and Analysis: the case of CIRCUS and PVSio-web. In: Dubois, C., Masci, P., Méry, D. (eds.) Proceedings of the Third Workshop on Formal Integrated Development Environment, F-IDE. EPCTS 240 (2016)
4. Hengstler, M., Enkel, E., Duelli, S.: Applied artificial intelligence and trustthe case of autonomous vehicles and medical assistance devices. Technological Forecasting and Social Change 105, 105–120 (2016)
5. Herrero, P., Lopez, B., Martin, C.: PEPPER: Patient Empowerment Through Predictive Personalised Decision Support. . In: Proceedings of the First ECAI Workshop on Artificial Intelligence (2016)
6. Herrero, P., Pesl, P., Reddy, M., Oliver, N., Georgiou, P., Toumazou, C.: Advanced insulin bolus advisor based on run-to-run control and case-based reasoning. Biomedical and Health Informatics, IEEE Journal of 19(3), 1087–1096 (2015)
7. Russell, S., Dewey, D., Tegmark, M.: Research priorities for robust and beneficial artificial intelligence. AI Magazine 36(4), 105–114 (2015)

Developing a Dynamic Bio-inspired Intelligent System to Support Diabetic Patients

Mireya Munoz-Balbontin, Tjeerd V. olde Scheper, and Arantza Aldea

Department of Computing and Communication Technologies - Oxford Brookes University
Wheatley Campus, Oxford, OX33 1HX - United Kingdom

Abstract.
This paper proposes a new biologically-inspired phenomenological model of insulin release that can be used to continuously monitor a person with diabetes and therefore recommend insulin dose and predict events. Given all the parameters and conditions that have to be met in order to achieve stability, the proposed insulin-release system will be considered as highly non-linear and dynamical. The resulting model will study the mechanisms that keep the system from reaching a chaotic state (diabetes), based on the property of self-organised criticality. This property was selected since it has been previously identified and studied in other dynamical biological systems. The proposed model will then be used as part of a model-based reasoning system that aims to support patients with diabetes.

1 Introduction

Regulation of insulin and glucose in the human body is part of a complex dynamical system, whose understanding and study is highly relevant for the management of diabetes. This system is in turn influenced by factors such as meal intake, heart rate, exercise, sleep, and other factors. Research and development on the insulin release system, also known as the artificial pancreas, can be traced back at least 50 years [1]. The models, signals, and control algorithms involved in the design of the insulin-release clinical system are the three most relevant areas of study. This research has facilitated the advancement of continuous glucose monitoring technologies, and the development of medical devices that assist type 1 diabetic patients in self-management of their condition.

Significant improvements have been made in continuous glucose monitoring technologies, and several control techniques have been tested and implemented on medical devices. However, in the context of modelling, it was found that a significant discrepancy exists between the physiological understanding and the low-level mechanisms that permit the physiological state [2]. This finding is the main motivation for the development of the model proposed in this paper. The modelling approach will be improved by studying the phenomenological aspects of insulin release using mathematical properties.

This paper summarizes the preliminary research that was conducted, as well as the modelling approaches that will be considered in order to construct the

research proposal. The model will then be used to identify how individual diabetes patients are influenced by different factors such as carbohydrate intakes, hormones, physical exercise and so on, and therefore support them in their required insulin doses.

2 Preliminary research on Bio-Inspired Models

To better understand the features of the model proposed in this research, physiological and mathematical models were studied. Among the physiological models analysed, minimal and maximal models were identified as the main models considered for CGM technology development. For the purpose of this research project, further studying of the phenomenological aspects of insulin-release using mathematical properties will be conducted. This was decided given a few misrepresentations found in physiological models, which are briefly also detailed in this section.

2.1 Physiological models

A physiological model is defined as the qualitative or quantitative representation of an actual physiological system. Quantitative physiological modelling is preferred, and considered to be more useful, as it provides a mathematical representation characterized by differential equations [5]. The most commonly used approach for the artificial pancreas system is compartmental modelling [4], a special case of physiological modelling where the movement of a substance between compartments is described. Furthermore, there are two classes of physiologically-based models for the artificial pancreas system, both described in this section: maximal models and minimal models.

Minimal models focus on explaining the key components of the system and measuring essential processes involved in healthy state and diabetes. One of the earliest models documented is the one developed by Bergman and colleagues back in the 70s [1]. The aforementioned model has undergone several modifications to include additional compartments and incorporate the effects of disturbances such as meal or exercise. Other examples of minimal models include the compartmental approaches proposed by Insel, et. al. and Cobelli et. al. in [2] which attempt to measure and understand glucose metabolism and its regulation by insulin.

Maximal models attempt to simulate the glucose insulin system using all available knowledge of functionality, thus allowing the researcher to carry out experiments in particular scenarios [3]. Useful information on the general functions of the system has been collected. One of the most representative examples is that of the GIM (Glucose Insulin Model) simulation software, which was designed based on the model proposed by Chiara Dalla Man, Robert A. Rizza, and Claudio Cobelli [4], consisting of 12 differential equations and 35 parameters. This software has been useful to simulate healthy state subjects, type 2 diabetic

subjects, and impaired glucose-tolerant subjects in two main scenarios: meal and daily life. The model developed by Grodzky, Jonkers and Henquin [9], was able to demonstrate the number of active cells is a sigmoidal function of glucose concentration, and suggest that insulin release is subject to different thresholds .

Minimal models are commonly used to analyse the data obtained from glucose tolerance tests. Although the results obtained prove useful, they tend to misrepresent the actual dynamics of the glucose-insulin system given that they are modelled in steady state. Some misrepresentation is also observed in maximal models, given that several of the parameters used in the system are derived from steady-state conditions.

Several other studies have been conducted in an attempt to understand the insulin release phenomenon in the human pancreas, these include glucose tolerance tests and in vitro studies (see [2] and [9]).

After revising the state-of-the-art in physiological modelling for the artificial pancreas, it was concluded that these models alone would not be sufficient for the development of this research project. The studied models have been able to provide useful information about the phenomenon of insulin release in the pancreas, and they provide solid evidence of progress in the field. However, it is considered that the models do not accurately represent the physiological state as they provide a generalized mathematically reduced approach, rather than a detailed representation on the dynamics of insulin-release. It was thus decided to explore a different set of methods to further develop the research project. These are presented in the subsequent section of this article.

2.2 Computational models and properties

This section introduces the property of self-organized criticality, Turing mechanisms, and two reaction-diffusion systems that will be considered as the foundation for the development of the model. These properties and systems have been chosen since they have been previously studied and identified in biological systems. While physiological models focus on representing the physical interactions of the system, the following systems and properties focus on describing complex mathematical scenarios, such as chemical reactions and/or the interaction between an activator and an inhibitor.

The property of self-organized criticality is particular to dynamical systems that are attracted to a critical point; for the past decade it has been identified and studied in biological systems. It has proven useful in defining and understanding neuronal communication, the functionality of protein families, pattern formation in biological systems, among others [11].

Turing mechanisms have been useful in research related to pattern formation, particularly in biological patterns such as the scales of fish, the formation of a

foetus and, at a more molecular level, pattern formation in cellular slime mould and calcium activity [7]. The model which can assist in studying these mechanisms is the two-variable Lengyel-Epstein model; a model for the photosensitive chlorine dioxide-iodine-malonic acid reaction (CDIMA). It has proven useful for investigating the continuous effect of external influence on Turing pattern formation. (see [7] and [8])

Reaction-diffusion systems and models are also used to describe pattern formation phenomena (including Turing patterns), particularly in systems that involve the interaction of many components. The model of interest for the purpose of this project is the Gray-Scott model, given the wide range of spatio-temporal dynamics it supports [10]. This model describes auto-catalytic chemical reactions of type: $A + 2B \rightarrow 3B$; $B \rightarrow C$, where the first equation represents an auto-catalytic process. It is possible to observe a wide range of dynamics when the input to the reactor is a continuous uniform flow of species A. The system can have up to three homogeneous steady-states: one trivial, and two non-trivial.

Another system of interest is the Brusselator. This system represents the interaction between a reactor and an inhibitor, and is nowadays considered one of the simplest reaction-diffusion systems capable of generating complex spatial patterns [13]. The proposed activator-inhibitor interaction is a well-known principle to explain pattern formation in chemical, ecological, physical, and biological systems.

3 Criticality analysis of the insulin-release system

Three possible approaches to improve the modelling methodology were identified during the preliminary research stage. The approaches were identified based on findings from physiological models that indicate the presence of criticality in the system. The approaches are described in the following subsections and will have to be analysed and simulated in order to fulfil the aim of the research project.

3.1 Finding 1: Non linearity and dynamical requirements

In mathematical terms, the insulin release system has to be considered as highly non-linear and dynamical. Given all of the parameters involved and the conditions needed to reach stability, it qualifies as one of the dynamical systems that is attracted to a critical point, this critical point being a normal glucose level. The low- level mechanisms of insulin release which control the system so that it does not become chaotic (diabetes), have to be studied to gather a wider knowledge on the effect they have over the global state of stability [12]. This will be achieved by studying the property of criticality observed in the insulin-release system.

3.2 Finding 2: Identification of critical patterns

A recent study has found that local feedback has a direct effect on the orientation of the Turing pattern in a system, and can change the global or local dynamics in it [7]. These findings can assist in understanding and defining the communication mechanisms followed by beta cells which lead to the release of insulin. This will be achieved by adding constraints to the insulin-release system, such as thresholds and external stimuli (e.g. meal and exercise). By adding these disturbances to the system, it will be possible to observe the effect on the insulin-release pattern, and whether or not it converges to a critical point.

3.3 Finding 3: Insulin delivery as a wave function

In the study by Jonkers and Henquin [9] it was demonstrated how beta cells are recruited in accordance with the sensed concentration of glucose, given that they have different threshold sensitivities to it. The study also shows how the behaviour of a single cell can be extrapolated to the behaviour of the islet (cluster) it is associated with. This property can indicate that the propagation of insulin secretion behaves as a wave function. The control of such a system has been studied by Y. N. Kyrychko et. al [10]. Where an activator/inhibitor control was applied on a Gray-Scott model based system, convergence to stability was achieved regardless of the chaotic nature of the system.

In the following stage of development, the low level mechanisms of insulin release will be simulated to observe their influence on stability at a global level. This will be achieved using coupled oscillators to represent the interaction of pancreatic β-cells, which represents the foundation for the development of the proposed model.

4 Bio-Inspired Model-Based Reasoning

In model based reasoning, an explicit model of the behaviour of the system is used to predict its reaction given a set of initial conditions and constraints. The predicted behaviour is then compared with the actual system and if discrepancies are found and assuming that the model is accurate then the reasoning mechanism can adapt the model to detect and identify the discrepancies. According to Herrero et al, [6] model-based reasoning has been proposed in different areas related to diabetes: Safety of monitoring devices, artificial pancreas, and prediction of hypoglycaemia.

In this research we propose to use a dynamic bio-inspired mathematical model to better model patient assimilation of glucose. The model will then be part of a MBR system that will be used to monitor a patient, detect when insulin is needed and identify the doses required.

The simulation and validation of the model will be performed using real patient data, which will be facilitated given the collaboration with Pepper project EU (http://www.pepper.eu.com/)

5 Conclusion

This paper proposes the development of a bio-inspired mathematical model of insulin release. Three approaches to assist in studying the property of criticality in the insulin-release system have been identified, and will be considered to develop the model. There is no evidence suggesting that the mathematical properties and models listed in this document have been used to model the phenomenological aspects of insulin release in the pancreas. A major advantage of using them is their capability of explaining complex dynamical systems using a simpler mathematical approach. Implementing this novel methodology could help fill in some gaps in physiological modelling, such as the inclusion of disturbances and non-linearities present in the system.

6 Acknowledgements

M.M.B. acknowledges financial support from Oxford Brookes University, and CONACyT. The authors also acknowledge the close collaboration with H2020 Pepper Project (http://www.pepper.eu.com/), and its partners. This work has received funding from the EU Horizon 2020 research and innovation programme under grant agreement No 689810.

References

[1] Cobelli C, Renard E and Kovatchev B. Artificial pancreas: Past, present, future. *Diabetes* 60(11):2672-2682, 2011.

[2] Cobelli C, Man CD, Sparacino G, Magni L, Nicolao G De and Kovatchev BP, Diabetes: Models , Signals , and Control, 2:54-96, 2009.

[3] Dalla Man C, Raimondo DM, Rizza RA and Cobelli C, GIM, Simulation Software of Meal Glucose-Insulin Model. *Journal of diabetes science and technology (Online)* 1(3):323-30, 2007.

[4] Dalla Man C, R. A. Rizza, and C. Cobelli, Meal simulation model of the glucose-insulin system, *IEEE Trans. on Biomedical Eng.*, vol. 54, no. 10, pp. 1740-1749, 2007.

[5] Enderle J, and Bronzino J. *Introduction To Biomedical Engineering* . 3rd ed. London: Academic Press, 2011. pp. 127-131.

[6] Herrero P, Lopez B, and Martin C, PEPPER: Patient Empowerment Through Predictive Personalised Decision Support. Proceedings of the *1st ECAI Workshop on Artificial intelligence for Diabetes*, *22nd European Conference on Artificial Intelligence (ECAI 2016)*, pages 8-9.

[7] Ji L and Li QS, Effect of local feedback on Turing pattern formation, *Chemical Physics Letters* 391(1-3): 176-180, 2004.

[8] Ji L and Li QS, Turing pattern formation in coupled reaction-diffusion systems: Effects of sub-environment and external influence. *Chemical Physics Letters* 424(4-6):432-436, 2006.

[9] Jonkers FC and Henquin JC, Measurements of cytoplasmic $Ca2^{+}$ in islet cell clusters show that glucose rapidly recruits β-cells and gradually increases the individual cell response, *Diabetes* 50(3): 540-550, 2001.

[10] Kyrychko YN, Blyuss KB, Hogan SJ and Schöll E, Control of spatiotemporal patterns in the Gray-Scott model. *Chaos* 19(4):1-7, 2009.

[11] Mora T and Bialek W, Are Biological Systems Poised at Criticality? *Journal of Statistical Physics* 144(2):268-302, 2011.

[12] olde Scheper T, Why metabolic systems are rarely chaotic. *BioSystems* 94(1-2):145-152, 2008.

[13] Wang K, Steyn-Ross M, Steyn-Ross D, Wilson M, Sleigh J, and Shiraishi Y, "Simulations of pattern dynamics for reaction-diffusion systems via SIMULINK." *BMC Systems Biology*, vol. 8, no. 1, p. 45, 2014.

[END OF PROCEEDINGS]