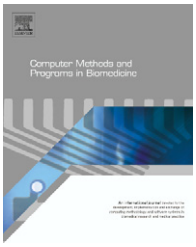
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 109 (2013) 116–125

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



An Actor–Critic based controller for glucose regulation in type 1 diabetes

Elena Daskalaki[a], *Peter Diem*[b], *Stavroula G. Mougiakakou*[a,*]

[a] *ARTORG Center for Biomedical Engineering Research, Diabetes Technology Research Group, University of Bern, Murtenstrasse 50, 3010 Bern, Switzerland*

[b] *Division of Endocrinology, Diabetes and Clinical Nutrition, University Hospital, Inselspital, University of Bern, 3010 Bern, Switzerland*

ARTICLE INFO

Article history:
Received 16 April 2011
Received in revised form
29 February 2012
Accepted 8 March 2012

Keywords:
Closed-loop control
Glucose control
Actor–Critic
Reinforcement learning

ABSTRACT

A novel adaptive approach for glucose control in individuals with type 1 diabetes under sensor-augmented pump therapy is proposed. The controller, is based on Actor–Critic (AC) learning and is inspired by the principles of reinforcement learning and optimal control theory. The main characteristics of the proposed controller are (i) simultaneous adjustment of both the insulin basal rate and the bolus dose, (ii) initialization based on clinical procedures, and (iii) real-time personalization. The effectiveness of the proposed algorithm in terms of glycemic control has been investigated *in silico* in adults, adolescents and children under open-loop and closed-loop approaches, using announced meals with uncertainties in the order of ±25% in the estimation of carbohydrates.

The results show that glucose regulation is efficient in all three groups of patients, even with uncertainties in the level of carbohydrates in the meal. The percentages in the A + B zones of the Control Variability Grid Analysis (CVGA) were 100% for adults, and 93% for both adolescents and children.

The AC based controller seems to be a promising approach for the automatic adjustment of insulin infusion in order to improve glycemic control. After optimization of the algorithm, the controller will be tested in a clinical trial.

1. Introduction

The external artificial pancreas (AP) is intended to act as a substitute of the pancreatic inability to produce insulin and thus to support glucose regulation in patients with type 1 diabetes (T1D). AP is a closed loop control system, consisting of a continuous glucose monitor (CGM), a continuous insulin infusion pump and a control algorithm for the estimation of the optimal insulin infusion. AP has been at the forefront of diabetes research since the 1970s [1]. There has been extensive work on possible control algorithms [2–4], combined with significant advances in CGMs and short acting continuous subcutaneous (sc) insulin infusion pumps. The most often discussed control algorithms include proportional–integral–derivative (PID) control, model predictive control (MPC), the run-to-run approach, optimal control (OC), sliding mode (SM) and fuzzy logic techniques. These algorithms have been evaluated either *in silico* or in clinical environments.

The PID controller won increased acceptance after association of its response with the biphasic nature of the pancreatic β-cells [5,6]. This was developed to provide an algorithm for external physiologic insulin delivery (ePID) [7,8]. In addition, an extended version with an insulin feedback module (ePID-IFK) has also been proposed [9,10]. Several versions of the PID controller have been developed and evaluated *in silico*, usually

* *Corresponding author at*: Murtenstrasse 50, Postfach 44, CH-3010 Bern, Switzerland. Tel.: +41 31 632 7592; fax: +41 31 632 7576.
E-mail address: stavroula.mougiakakou@artorg.unibe.ch (S.G. Mougiakakou).
0169-2607/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2012.03.002

combined with feed-forward or open-loop techniques, such as a PID controller which incorporates a switching strategy for initialization after a meal and insulin bolus, together with time-dependent trajectories and CGM denoising filters [11]. Finally, some researchers have proposed the use of hybrid PID controllers in order to overcome the problems related to delays in insulin action *e.g.* fuzzy-PID [12,13].

MPC is probably the most commonly applied algorithm, due to its intrinsic combination of future glucose estimation and insulin dose optimization in a specified prediction horizon. An MPC controller using a non-linear model for glucose prediction and adaptive techniques for individualized tuning of its parameters has been developed and evaluated in two clinical trials for overnight glucose regulation [14,15]. Reduced overnight time spent in hypoglycemia and increased time spent in normoglycemia was found. An MPC with linear and non-linear prediction models has been developed and assessed for its robustness when the intake of carbohydrates (CHO) and the times of meals are uncertain. The algorithm was first evaluated *in silico* [16] and two clinical evaluations have followed using the unconstrained version of the algorithm; these showed reduction in nocturnal hypoglycemic events during closed-loop control [17,18]. An MPC combined with a model for meal estimation, insulin-on-board constraints and pump shut-off modules has been proposed, which gave lower mean glucose concentrations than when no meal estimation model was used [19]. Furthermore, a combination of an MPC with iterative learning techniques has been presented, in order to achieve adaptation to inter- and intra-patient variability and robustness to random variations in meal times and CHO content estimation [20]. In order to improve glycemic control, a zone-MPC has been developed, in which the control variable objective was a zone and not a fixed point or trajectory. Additionally, the controller has been enhanced with models for sc insulin and CHO absorption, to be used as indicators of the system's history and to provide constraints in insulin dosing [21]. An extended version of MPC based on the minimization of a combined risk index of hypo- and hyperglycemia has been developed which gave lower blood glucose risk indices and higher times spent in normoglycemia than with PID or standard MPC, with the closest possible adherence to the lowest risk bound, as defined by an optimally tuned basal/bolus controller [22]. Finally, a multi-parametric MPC has been developed and assessed for robustness against insulin sensitivity changes and meal size uncertainties, which gave low values of the low blood glucose index (LBGI) [23].

Run-to-run algorithms have the ability to learn from previous experience and adjust their parameters to improve performance. Run-to-run algorithms use patient data collected over a predefined period in order to update the control parameters over the following control period. Such a technique has been applied for the first time in T1D in a limited clinical trial, which demonstrated the applicability of run-to-run control in managing insulin administration for meal management [24]. The same group has proceeded to develop a run-to-run strategy for the adjustment of basal insulin infusion rates. The algorithm has been evaluated in an *in silico* trial in which the meal size was the same every day and the insulin:CHO (IC) ratios were fixed. The results showed improvement in the glycemic profile and convergence of the method, depending on the simulation model used [25]. One year later, the same group presented a run-to-run strategy for the daily adaptation of IC ratios based on sparse blood glucose measurements. The ability of the algorithm to improve post-prandial blood glucose levels has been assessed in a 2-week clinical trial with nine diabetic patients [26]. The results showed improvement in post-prandial blood glucose levels, although it was uncertain whether the algorithm could handle large and variable meals. Finally, an automatic bolus and adaptive basal approach has been proposed involving a meal detection algorithm for prandial insulin infusion and adaptive basal insulin estimation [27].

Furthermore, OC techniques are based on the optimization of the system's performance through the minimization of an appropriate cost function. OC has been used in diabetes management, in an effort to handle the uncertainties of the glucoregulatory system and to establish robust control. In [28], the time-dependent uncertainties of the system were modeled as stochastic processes and the deterministic and stochastic OC approaches were compared. Finally, a linear quadratic OC was developed in [29] and evaluated *in silico* using linear and non-linear models of the glucoregulatory system.

In SL techniques, a sliding surface is defined along which the process can slide to its desired value. SL control has been combined with internal model control in an *in silico* study aiming at the development of robust glucose regulation [30].

Finally, an MD-Logic controller, based on fuzzy-logic theory, has been developed combined with a learning algorithm for the periodical adjustment of insulin treatment [31]. The system has been evaluated *in silico* under conditions of meal uncertainty and insulin sensitivity changes.

Glucose regulation faces multiple challenges, due to the intrinsic complexity of the glucoregulatory system and its interaction with the environment. The high inter- and intra-variability of the diabetic population expressed in terms of age, insulin sensitivity, life-style *etc.* raises the demand for personalized approaches in insulin delivery, while the variety of external and internal disturbances which are usually unexpected or poorly estimated, such as meal intake, physical activity and stress, necessitates fast reactions and robust behavior. However, the current technological disadvantages of the CGMs introduce major inaccuracies, while the sc–sc route is associated with delays – both in insulin absorption and in glucose measurement. One major concern involves the delay of *ca.* 30 min in the action of subcutaneously infused insulin after delivery. This significantly restricts the reaction time and necessitates some kind of predictive ability. Furthermore, when glucose is regulated by insulin infusion, the exact glucose values at each measurement are not necessarily important, while features of the glucose profile in longer periods (such as trends, minima or maxima) may incorporate more useful information on subsequent changes. PID controller, in order to handle non-linear systems such as the gluco-regulatory system requires enhancement with additional non-linear or artificial intelligence components [8,11]. Furthermore, MPC, despite its intrinsic ability to handle non-linearities, solves an open-loop optimization problem at each time step, an approach that performs suboptimally in the presence of non-deterministic components [32]. On the other