CLOSED-LOOP CONTROL OF AN ARTIFICIAL PANCREATIC BETA CELL USING MULTI-PARAMETRIC MODEL PREDICTIVE CONTROL

Matthew W. Percival[†,‡], Eyal Dassau[†,‡], Howard Zisser[†,‡], Lois Jovanovič[†,‡], and Francis J. Doyle III[†,‡,]*
[†] Department of Chemical Engineering, University of California, Santa Barbara
[‡] Sansum Diabetes Research Institute, Santa Barbara
** Corresponding author: frank.doyle@icb.ucsb.edu*

Abstract

A model predictive control algorithm was developed for regulating glycemia in subjects with type 1 diabetes mellitus using exogenous insulin infusion to reject disturbances from oral carbohydrate consumption. A clinically validated nonlinear physiological model of glucose metabolism was linearized and served as the basis for the controller. The control law was developed using multi-parametric programming techniques, so that the online optimization problem was reduced to the evaluation of an affine function. The controller was tuned for optimal response to mixed-meals of up to 40 g carbohydrate. The robustness of the controller was then evaluated for mismatch in the size (25%) and timing (15 minutes) of the meal. Due to the computational ease with which this control law was evaluated online, it would be suitable for implementation on a personal digital assistant with limited computational capacity.

Introduction

Diabetes mellitus is a disease characterized by insufficient endogenous insulin production, leading to poor regulation of plasma glucose concentration. The resulting chronic hyperglycemia can lead to macro- and micro-vascular complications, e.g., stroke, heart attacks, blindness, and kidney disease. An indication of the quality of glycemic regulation, and thus the likelihood of such complications is given by the quantity of glycosylated hemoglobin (A1C). Successful control consists of normalizing glycemia, and thus reducing A1C to below 6.0% [1].

Type 1 diabetes mellitus (T1DM) is an autoimmune disease that results in the destruction of the pancreatic beta cells found in the islets of Langerhans. These cells are responsible for the production of insulin, one of the hormones that regulate plasma glucose concentration. In order to regulate glycemia in T1DM, insulin must be delivered exogenously. Current therapeutics involve manually taking a capillary blood sample (a finger stick measurement) to determine the blood glucose concentration, then injecting an appropriate amount of insulin depending upon the subject and the situation, e.g., at meal time. For this method to be effective, this process must be repeated up to 12 times daily. Such a high level of discipline is burdensome, and thus patient adherence is typically lacking. As a result, glycemic control is inadequate and complications of T1DM occur. As an indication of the worldwide scale of the problem, there are 1.5 million people with T1DM in the United States alone [2].

Recent technological advances have brought continuous real-time subcutaneous glucose sensors and continuous insulin infusion pumps, which in combination with rapid-acting insulin analogs provide a framework for a closed-loop controller suitable for use in ambulatory conditions [3]. Real-time glucose sensors measure the glucose concentration in the

subcutaneous tissue, and typically offer an updated glucose measurement every five minutes; this sample resolution is in line with that required for an algorithm to achieve efficient closed-loop control via subcutaneous insulin delivery [4].

The first algorithms for closed-loop control of glycemia were developed in the 1970s, most notably those used with the glucose controlled insulin infusion devise called the Biostator® [5]. These early algorithms relied on intravenous insulin and dextrose delivery, along with intravenous glucose measurements. Such invasive methods were therefore only suitable for use in an in-patient hospital setting. Currently, robust control algorithms for a closed-loop device suitable for use in ambulatory conditions are not available [6]. The main challenges involve the large lag time in insulin action of subcutaneously delivered insulin on the plasma glucose concentration, and sustained disturbances to glucose levels after the absorption of a meal. For these reasons, model predictive control (MPC) has been identified as a highly promising controller structure, and has been applied in a clinical environment [7]. The MPC framework also allows for the explicit inclusion of constraints [8], which are pertinent to the physical capabilities of the insulin pump.

In order to mitigate the problems involved with solving large optimization problems online, multi-parametric programming techniques can be applied which lead to the solution of a single optimization problem offline; the online problem thus reduces to reading a lookup table, and the evaluation of an affine function [9]. Such techniques are germane to the first generation of closed-loop devices for managing glycemia for two reasons:

(i) The first commercially available artificial beta cells will no doubt have limited computational capacity.

(ii) There should be an easy way to evaluate systematically controller outcomes under a range of expected conditions before application.

Such prior evaluation would be a requirement of the Food and Drug Administration (FDA) for a device to be approved for commercial use [10]. This algorithmic approach has previously been used to develop a multi-parametric MPC (mpMPC) algorithm in the context of diabetes, albeit for intravenous insulin delivery using a simplified model of insulin action [11].

In this paper, a simulation study was carried out to show the utility of mpMPC when applied to regulation of glycemia. An established physiological model of glucose metabolism [7, 12] characterizing orally ingested carbohydrate and subcutaneously injected insulin was used represent a person with diabetes. A control law was then developed using mpMPC. Simulations were carried out to determine optimal controller tuning in the case of meal (disturbance) rejection. The resulting controller was then tested for robustness by introducing elements of plant/model mismatch typical of a person with diabetes.

Experimental Methods

This section describes the design of the simulation study and the implementation of mpMPC. The simulation study was designed as an *in silico* closed-loop clinical trial. A physiological model represents the subject with T1DM, which is subject to meal ingestion and insulin delivery determined by the control algorithm.

Simulation model

The physiological models of glucose-insulin kinetics and subcutaneous insulin absorption presented by Hovorka et al. [7] and Wilinska et al. [12] formed the virtual T1DM subject. Orally ingested glucose is absorbed by the gut into plasma using a two-compartment fundamental model. Subcutaneously injected insulin is absorbed into the plasma from a fourth

order model incorporating saturation effects. The glucose effects are described by fifth order model including three insulin sensitivity effects, endogenous glucose disposal and production, and insulin dependent and independent effects. The nonlinear effects of the model are assumed representative of those observed *in vivo*.

Controller development

The formulation of the control algorithm is a two part process:

i. Formulation of the objective function as a constrained optimization problem

ii. Reformulation of the objective function as a multi-parametric program and its solution

The objective function for the MPC problem includes quadratic penalties for the vector of predicted deviations from setpoint of glucose measurements (y) and the vector of future insulin delivery rates (u). The cost, J, is written as

$$\min_u J = y^T Q_y y + u^T R u. \tag{1}$$

The model predictions are given by a linearization of the plant around nominal operating conditions, which correspond to fasting and basal insulin delivery; the state-space representation of the model at time k is

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned}, \tag{2}$$

where x represents the state vector, u represents the inputs, and y the measurement; all are deviation variables. Physical restrictions on insulin delivery corresponding to the mechanical limitations of the insulin pump are included as constraints on the manipulated variable

$$u_{min} \leq u \leq u_{max}. \tag{3}$$

Constraints are also given on the region of state-space in which a solution is required:

$$x_{min} \leq x \leq x_{max}, \tag{4}$$

where x_{min} and x_{max} represent the minimum and maximum values permissible for each state in the state vector.

In order to apply multi-parametric programming techniques the objective function must be reformulated as a standard mpQP,

$$\begin{aligned} &\min_z \mu = \frac{1}{2} z^T H z \\ &s.t.\ Gz \leq W + GH^{-1}F^T x_0 \end{aligned} \tag{5}$$

where x_0 is the state vector at the current time, and z is a linear transformation of the manipulated variable, defined as

$$z = u + H^{-1}F^T x_0. \tag{6}$$

Matrices F, G, H and W are obtained from A, B, Q_y, and R. An explicit solution to this optimization problem exists [9]. The solution consists of multiple regions within state-space; within each region the optimal values of the manipulated variables, u, are an affine function of the current state vector x_0. The algorithm used to determine the regions in the solution and the corresponding control law is described by Pistikopoulos et al. [13]. Transformation of the parametric solution back into state-space gives an explicit control law for all possible initial conditions of the state vector. Full mathematical development of this method has been carried out several times in the literature [9, 11, 13].

Controller object

The solution to the multi-parametric programming problem was found using the Multi-Parametric Toolbox [14], and interfacing software, YALMIP [15]. The solution consists of a

polytope defining the control law. The control law consists of regions defined in state-space, essentially a lookup table. Each region, or entry in the lookup table contains an affine function of the current state; the evaluation of this function gives the optimal control moves to be implemented.

In order to evaluate the control law, the state vector is evaluated using the latest measurements, thus determining the appropriate region, or lookup table entry. The affine function associated with that region is then evaluated with the current state vector, thus determining the control move. A graphical example of control law evaluation is given in Appendix A.

Simulated Results

Using the controller and simulated subject described in the previous section, controller tuning was performed and robustness due to expected levels of uncertainty in key variables were investigated. The goal was to investigate the rejection of a single meal and to ensure long-term controller stability over a period of 12 hours.

Simulations were performed with measured disturbances of 10-40 g CHO mixed-meals. For these announced disturbances, the prediction horizon and weighting matrices were tuned. In order to test for robustness, mismatches of 25% were introduced to the CHO measurement and of 15 minutes to the disturbance timing.

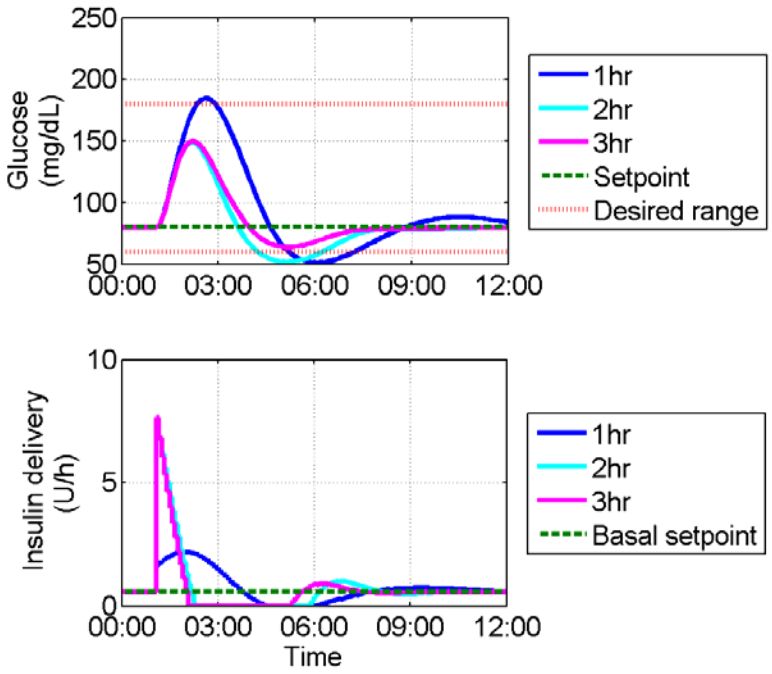

Figure 1: Closed-loop responses to an announced meal of 40 g CHO. TOP: glucose trajectory. Increasing the prediction reduces hyperglycemia and hypoglycemia. BOTTOM: insulin delivery. Increased prediction horizon allows more aggressive control action at meal time.

Prediction horizon

An appropriate choice of prediction horizon is one in which the full effects of current inputs are observed, i.e., the open-loop speed of response. Based on this principle, a prediction horizon of three hours is appropriate, as is corroborated by Figure 1.

Weighting matrices

The weighting matrices on glucose error, Q_y, and insulin delivery about basal, R, determine how aggressive the controller is; if large penalties are applied to glucose error, an aggressive controller will result; if large penalties are applied to insulin delivery about basal, a more conservative controller will result. In the case where Q_y and R are constant over time, the ratio Q_y/R is the single tuning parameter. As a general rule, a controller should be more conservative if there is known to be a significant plant/model mismatch, or if the consequences of large deviations from setpoint are severe.

In this nominal case with a model based on the Jacobian of the plant, the plant model mismatch should be small; hence, the deciding factor in tuning should be avoidance of hypoglycemia. Figure 2 shows the effects of various Q_y/R with a three hour prediction on controller performance. The best Q_y/R is equal to unity, since this setting avoids hypoglycemia and hyperglycemia.

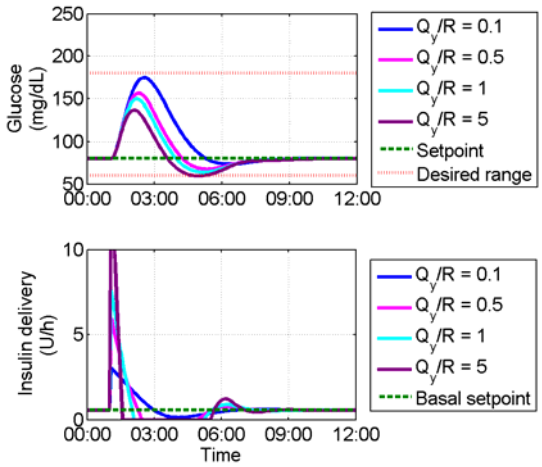

Figure 2: Closed-loop responses to an announced meal of 40 g CHO and the effect of varying the input/output weighting matrices. TOP: a less aggressive controller minimizes the approach to hypoglycemia while remaining below the hyperglycemic threshold. BOTTOM: the more aggressive controllers deliver insulin as a pseudo-bolus in order to minimize the glycemic excursion.

Robustness

Uncertainty is not only attributed to plant/model mismatch, but also to inputs. These input uncertainties arise whenever a meal is announced, specifically the quantity of carbohydrate in the meal and the exact time of the meal. It is not that meal size and time are impossible to determine exactly, but that the announcement itself is burdensome and a typical subject with T1DM does not record these figures precisely. Mismatches of 25% were introduced to the CHO measurement and of 15 minutes to the disturbance timing. Figures 3 and 4 show how time and size mismatches affect the controller tuned with $Q_y/R=1$. Combining these effects lead to hypoglycemic events with the $Q_y/R=1$, so the controller was further detuned. Figure 5 shows the combined effects of meal size and time mismatch with $Q_y/R=0.5$. Clearly detuning is necessary to avoid hypoglycemia in the case of uncertainty.

Based on the limits of uncertainty previously described, it is important to ascertain the greatest amount of CHO that can be consumed without inducing hypo- or hyperglycemia. Figure 6 shows the upper and lower bounds of the glycemic trajectory for meal sizes of 10, 25, and 40 g CHO. For this simulation model and these conditions of uncertainty, the maximum meal size is 35 g CHO. This meal size is very low in CHO, which serves to illustrate further the difficulties associated with rejection of meal-based disturbances.

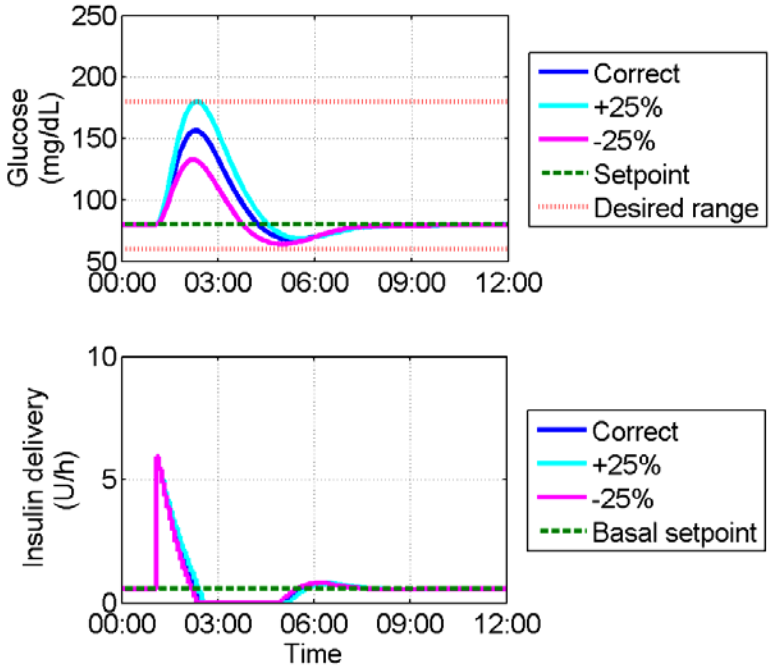

Figure 3: Closed-loop responses to an announced meal of 40 g CHO with the actual meal size of 40 g CHO ± 25%. TOP: glycemic excursions remain within desired range, thus avoiding hyperglycemia and hypoglycemia. BOTTOM: insulin delivery is initially the same in each case; feedback from glucose measurements adjusts the following delivery slightly to compensate for the differences in expected and actual glycemic excursions.

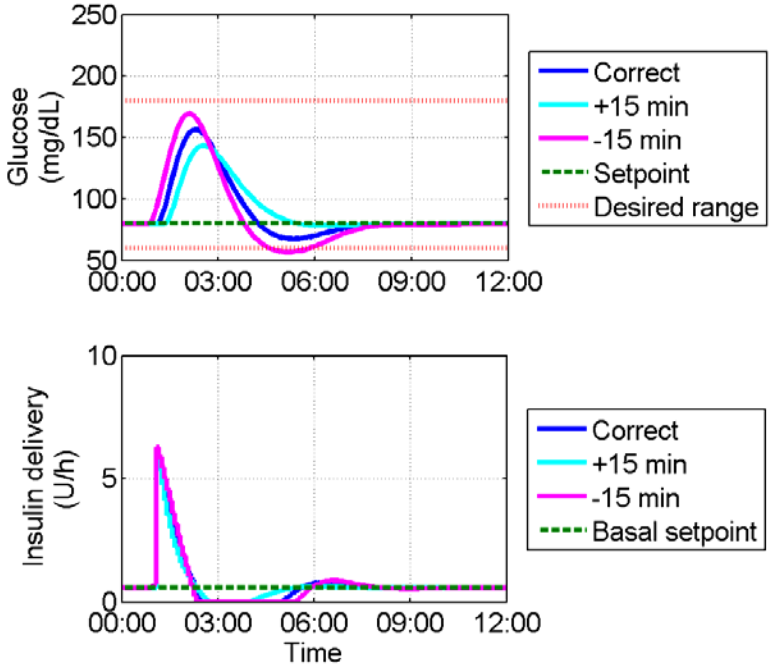

Figure 4: Closed-loop responses to an announced meal of 40 g CHO with the actual meal timing varied by ± 15 minutes. TOP: early consumption of the meal causes a minor hypoglycemic event four hours after meal time. BOTTOM: insulin delivery is initially the same in each case as a direct consequence of meal announcement; feedback from glucose measurements adjusts the following delivery slightly to compensate for the differences in expected and actual glycemic excursions.

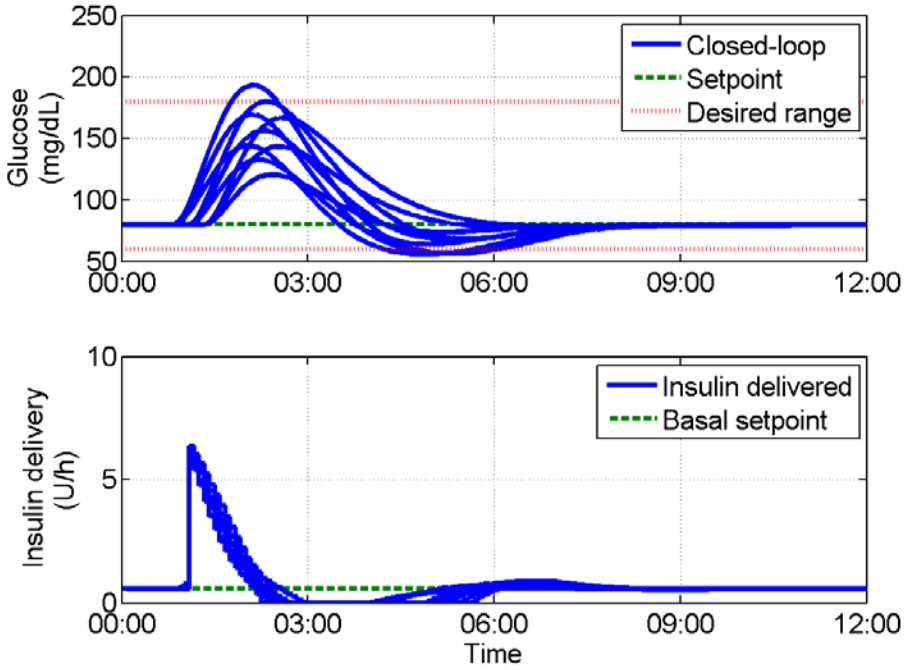

Figure 5: Closed-loop responses for when the mismatches of meal size and meal time are combined. TOP: in some cases minor hyperglycemic excursions and hypoglycemic events occur. BOTTOM: insulin delivery remains similar in all cases since the controller tuning is conservative, thus the major insulin delivery peak is due to the meal announcement.

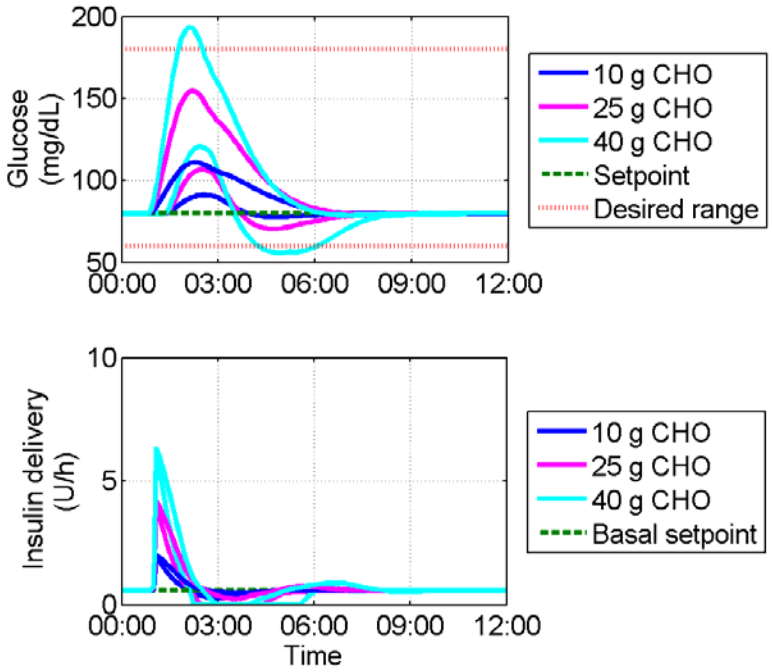

Figure 6: Maxima and minima of closed-loop responses under uncertainty for meals ranging from 10, 25, and 40 g CHO. TOP: the limits of the glucose trajectories show that the maximum meal size permissible in order to avoid hyperglycemia and hypoglycemia under these conditions of uncertainty is less than 40 g CHO.

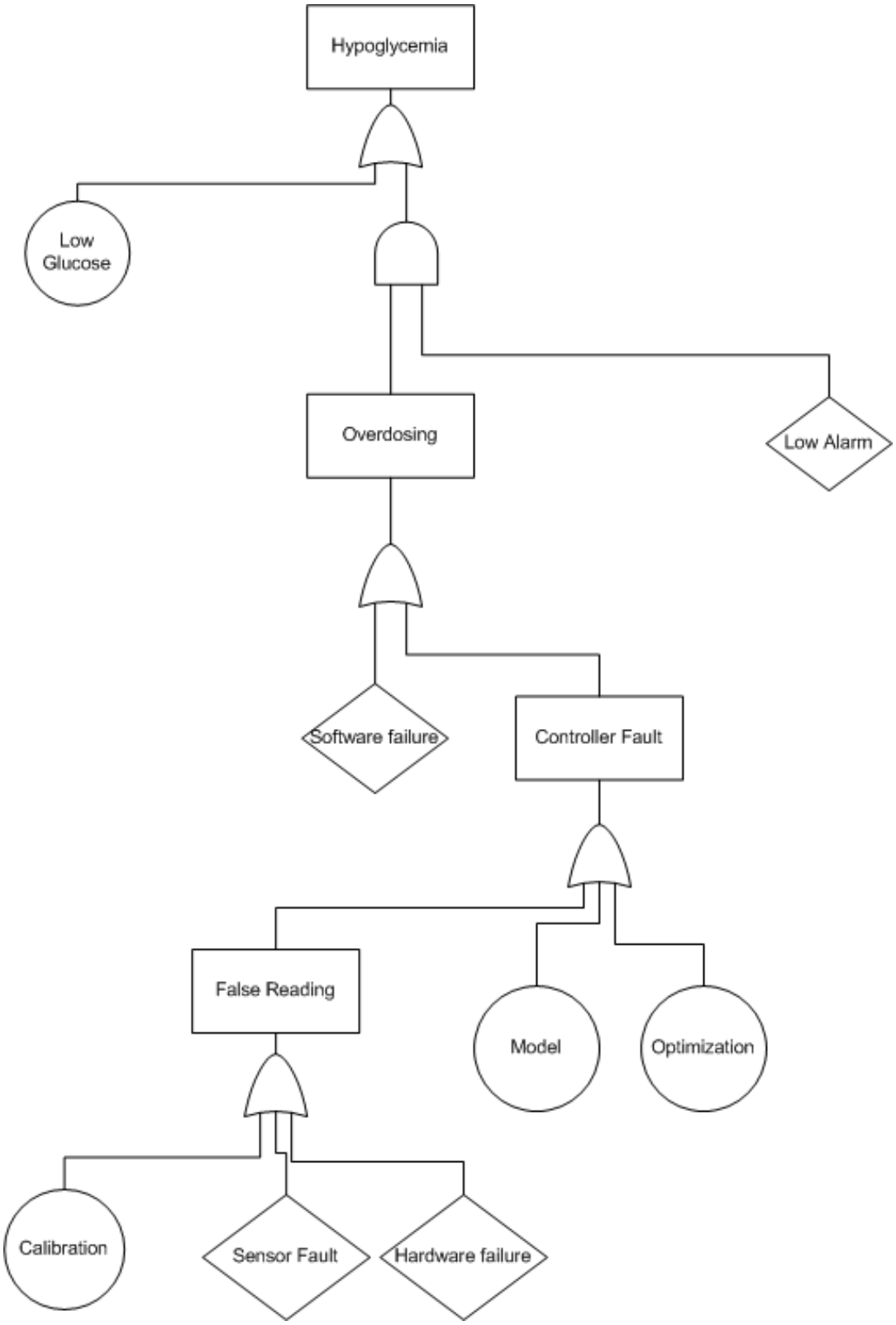

Figure 7: A simplified fault tree analysis for an artificial pancreatic beta cell with a top event defined as hypoglycemia. This analysis shows the causes of hypoglycemia by different components of the artificial pancreas.

Regulatory body approval of medical devices

An artificial pancreas (AP) for the regulation of glycemia is considered a high-risk device (class III) by regulatory bodies such as the FDA. As such, a thorough risk analysis and/or fault tree analysis (FTA) is required that demonstrates the safety of both the system and its subcomponents. In particular, the control algorithm and software that run the system need to demonstrate both safety and efficacy to be robust to both system disturbances other abnormalities [16, 17, 18].

The use of MPC as a controller of such a device opens the question of controller stability, for example, if constrained MPC is implemented or if an output constraint has been implemented there can be a situation of unsolvable optimization [19, 20]. As demonstrated on a simplified FTA (Figure 7), a top event defined as hypoglycemia can be developed due to overdosing of insulin that is the outcome of either software failure or a controller fault. MPC with real time optimization can therefore produce an error that can be a trigger for a catastrophic event.

The FDA has asked researchers conduct clinical trials to submit an Investigative Device Exemption (IDE) before clinical trials are carried out; this document should provide the necessary information for them to approve the device for clinical studies. As part of the requirements the control algorithm needs to be described in a way that will demonstrate its stability. This requirement is challenging if an optimization problem is solved in real time. One way to fulfill this requirement is to use unconstrained MPC or a PID algorithm, where controller output is then clipped in order to produce feasible instructions to the hardware; the control action is then no longer optimal. The use of mpMPC thus permits evaluation of controller stability and performance prior to any *in silico* evaluation of the algorithm based on an approve test simulator [21] and the IDE submission.

Conclusions

Development of a linear model from the Jacobian of the plant produced an efficacious control law. The use of a disturbance model is of paramount importance in this application, due to the large lag time for insulin action. Without such a model, controller performance is inadequate under conditions of model uncertainty similar to those likely to be experienced with intra-subject variation.

Evaluation of controller performance considering hypo- and hyperglycemia showed that these conditions could be avoided with meal sizes of up to 35 g CHO, even under the conditions of uncertainty expected. In the United States, a 70 kilogram man typically consumes over 70 grams of carbohydrate per meal plan. Thus this upper limit on meal size indicates that optimal therapy requires some form of patient adherence to the maximum amount of CHO consumed at one meal.

Offline optimization, although computationally expensive for high-order systems, has been shown to be a feasible technique for applying mpMPC to glycemic regulation. Indeed, there is an increase in the number of controller regions as the control law becomes more aggressive, and as more controller features, such as a reference trajectory, are included. For the controllers demonstrated here, the offline computation of the control law took a time of the order minutes.

Since the glucose-insulin kinetics in subjects with T1DM is known to be time-varying, regular updates of the offline optimization would be required to ensure a sufficiently accurate model. This would be feasible on a daily basis due to the time required for the offline

optimization. Future work will include output constraints implemented through mixed integer programming, to satisfy hierarchical objectives on avoidance of hypoglycemic events and extended periods of hyperglycemia.

A distinctive feature of mpMPC is that it provides the necessary ability to ensure controller stability to different inputs as required by the FDA and demonstrated on several applications in [19] without the need to compromise the controller design to a less sophisticated one such as PID or unconstrained MPC. This implementation also provides means to analyze the algorithm before implementation, which is critical to satisfying FDA requirements concerning risk analysis.

Acknowledgements

This work was supported by the Juvenile Diabetes Research Foundation (JDRF) grant 22-2006-1115, the Institute for Collaborative Biotechnologies ICB, and the Otis Williams Fund at the Santa Barbara Foundation.

References

1. Hirsch IB, Brownlee M. Should minimal blood glucose variability become the gold standard of glycemic control? *J Diabetes Complications,* 2005; 19(3):178–181.
2. Eiselein L, Schwartz HJ, Rutledge JC. The challenge of type 1 diabetes mellitus. *ILAR J,* 2004; 45(3):231–236.
3. Bequette BW. A critical assessment of algorithms and challenges in the development of a closed–loop artificial pancreas. *Diabetes Technol Ther,* 2005; 7(1):28–47.
4. Hovorka, R. Continuous glucose monitoring and closed-loop systems. *Diabet Med,* 2005*, 23*, 1-12
5. Clemens AH, Chang PH, Myers RW: The development of Biostator, a glucose controlled insulin infusion system (GCIIS). *Horm Metab Res*, 1977; S7:23–33.
6. Hovorka R, Wilinska ME, Chassin LJ, Dunger DB. Roadmap to the artificial pancreas. *Diabetes Res Clin Prac,* 2006, 74, S178-S182.
7. Hovorka R, Canonico V, Chassin LJ, Haueter U, Massi-Benedetti M, Federici MO, Pieber TR, Schaller HC, Schaupp L, Vering T, Wilinska ME. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol Meas,* 2004; 25(4):905–20.
8. Garcia CE, Prett DM, Morari M. Model predictive control – theory and practice – a survey. *Automatica,* 1989; 25(3):335–348.
9. Bemporad A, Morari M, Dua V, Pistikopoulos EN. The explicit linear quadratic regulator for constrained systems. *Automatica,* 2002; 38(1):3–20.
10. General Principles of Software Validation; Final Guidance for Industry and FDA Staff. [Internet] Jan 11 2002; http://www.fda.gov/cdrh/comp/guidance/938.html.
11. Dua P, Doyle III FJ, Pistikopoulos EN. Model-based blood glucose control for type 1 diabetes via parametric programming. *IEEE Trans Biomed Eng*, 2006; 53(8):1478–1491.
12. Wilinska ME, Chassin LJ, Schaller HC, Schaupp L, Pieber TR, Hovorka R. Insulin kinetics in type 1 diabetes: continuous and bolus delivery of rapid acting insulin. *IEEE Trans Biomed Eng,* 2005; 52(1):3–12.
13. Pistikopoulos E, Georgiadis M, Dua V. Multi-parametric model-based control, *Wiley-VCH,* 2007.

14. Kvasnica M, Grieder P, Baotić M. Multi-Parametric Toolbox (MPT). 2004.
15. Löfberg J. YALMIP: A toolbox for modeling and optimization in MATLAB. *In proceedings of the CACSD Conference*. Taipei, Taiwan, 2004.
16. Fries RC, Reliable Design of Medical Devices: CRC Press, 2005.
17. Food and Drug Administration, Guide To Inspection Of Computerized Systems In Drug Processing, *Food and Drug Administration*, Washington, DC, 1983.
18. Food and Drug Administration, Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices, *Food and Drug Administration*, Washington, DC, 2005.
19 Dua P, Kouramas K, Dua V, Pistikopoulos EN. MPC on a chip—Recent advances on the application of multi-parametric model-based control. *Computers & Chemical Engineering,* vol. 32, no. 4-5, pp. 754-765, April, 2008.
20. A. Bemporad, F. Borrelli, and M. Morari, On the optimal control law for linear discrete time hybrid systems, *Hybrid Systems: Computation and Control*, Lecture Notes in Computer Science, pp. 105-119, 2002.
21. Kovatchev BP, Breton MD, Dalla Man C, Cobelli C. In silico model and computer simulation environment approximating the human glucose/insulin utilization. *Food and Drug Administration Master File* MAF 1521, 2008.

Appendices

Appendix A: Graphical example of control law implementation

Consider an arbitrary second order linear state space model, with state vector given by

$$x = [x_1 \quad x_2]^T, \tag{A1}$$

where x_1 and x_2 represent the model states. The lookup table in this case is a two-dimensional polytope, with regions described graphically in Figure A1. The lookup table is searched until the region containing the state vector, x, is found. The next control move is calculated using the affine function associated with the current region using

$$I_D = Fx + g. \tag{A2}$$

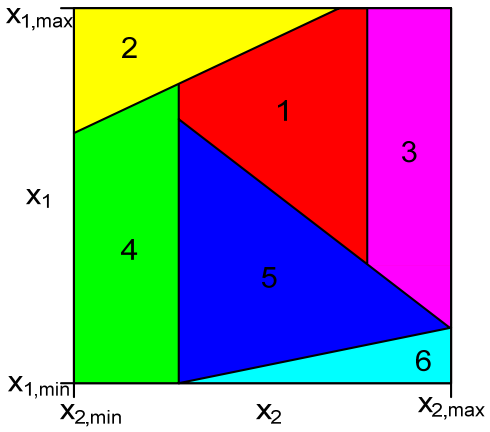

Figure A1: Graphical representation of the controller regions for a hypothetical controller across two states. Associated with each region is an affine function, representing the control law, which is evaluated at each time step using the state vector.