The mathematician’s control toolbox for management of type 1 diabetes

Marie Csete¹ and John Doyle²

¹Huntington Medical Research Institutes, 99 N. El Molino Avenue, Pasadena, CA 91101, USA
²California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA

Blood glucose levels are controlled by well-known physiological feedback loops: high glucose levels promote insulin release from the pancreas, which in turn stimulates cellular glucose uptake. Low blood glucose levels promote pancreatic glucagon release, stimulating glycogen breakdown to glucose in the liver. In healthy people, this control system is remarkably good at maintaining blood glucose in a tight range despite many perturbations to the system imposed by diet and fasting, exercise, medications and other stressors. Type 1 diabetes mellitus (T1DM) results from loss of the insulin-producing cells of the pancreas, the beta cells. These cells serve as both sensor (of glucose levels) and actuator (insulin/glucagon release) in a control physiological feedback loop. Although the idea of rebuilding this feedback loop seems intuitively easy, considerable control mathematics involving multiple types of control schema were necessary to develop an artificial pancreas that still does not function as well as evolved control mechanisms. Here, we highlight some tools from control engineering used to mimic normal glucose control in an artificial pancreas, and the constraints, trade-offs and clinical consequences inherent in various types of control schemes. T1DM can be viewed as a loss of normal physiologic controls, as can many other disease states. For this reason, we introduce basic concepts of control engineering applicable to understanding pathophysiology of disease and development of physiologically based control strategies for treatment.

1. Introduction: the clinical problem

Type 1 diabetes mellitus (T1DM) is a chronic multisystem disease, the most common metabolic disease of childhood, caused by autoimmune destruction of pancreatic beta cells. Long-term consequences of poor insulin control include considerable morbidity from neuropathy, vasculopathy, including coronary and cerebrovascular disease, nephropathy and retinopathy. Type 1 diabetics are also at risk of hypoglycaemia leading to coma and death, when too much insulin for the acute clinical situation is administered. The estimated economic burden of T1DM in the USA is $14.4 billion [1]. It is easy to see that a technology offering reasonable control over glucose levels in T1DM patients could be life-changing. In September 2013, the US Food and Drug Administration (FDA) approved the first centrally monitored clinical automatic glucose controller, a wearable artificial pancreas, following an international research effort involving close collaboration between the control engineering community, computer scientists and expert clinicians. The device hardware employs previously approved glucose sensors and insulin pumps, but as a safety improvement, devices are monitored electronically (via Bluetooth connection to a smartphone) at the University of California Santa Barbara/Sansum Artificial Pancreas facility. The device software is designed to actuate increased insulin delivery when it senses blood glucose levels climbing above a preset range, and to slow down or turn off insulin when it senses glucose heading below a preset range. Wearers of the device must still manually programme boluses for meals and adjust the device for other events such as activity or illness.

So, though not perfect, the artificial pancreas is a major landmark for patients with T1DM, a tour de force of persistence and collaboration between great scientists, engineers and clinicians. Initial clinical trials supported safety of the
device and justified further efficacy testing [2]. The regulatory path for new medical devices involving software in the USA is difficult as the software has to be validated specifically to regulatory standards. (See www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126955.pdf.) In its press release, the Juvenile Diabetes Research Foundation noted that FDA approval of the device ensures a smoother regulatory pathway for future improvements in software and hardware of the artificial pancreas, improvements that will be based on optimizing control algorithms in the device [3].

Here, we describe some of the controls mathematical toolbox behind the artificial pancreas in non-mathematical terms. We focus specifically on the mathematical approaches applied clinically in the development of the artificial pancreas because they illustrate the importance of understanding control systems generally in medicine. Development of the artificial pancreas required some trial and error using different control algorithms, and an international multidisciplinary team, with results from experiments in human subjects fed back to optimize the control algorithms.

2. Normal control of glucose homeostasis

Blood glucose levels are normally controlled by feedback loops between the blood, pancreas, liver and peripheral tissue (especially skeletal muscle). High blood glucose levels result in insulin release from the pancreas, facilitating glucose uptake into tissues and conversion of glucose to glycogen in the liver. Together, these processes contribute to a reduction of blood glucose levels. Low blood glucose levels promote glucagon release from the pancreas; glucagon taken up in the liver promotes glucose production from glycogen, and then release of glucose into circulation. On the surface, then, regimens for keeping glucose in normal physiological range seem ideally suited to control engineering, just as temperature in our homes is controlled by thermostats working with similar feedback loops (figure 1). In these similarly wired control systems, the thermostat (pancreas) senses a disturbance (temperature high or low) and signals a heating or cooling element to ‘release’ warm air (insulin) or cold air (glucagon) and return the target element (temperature/glucose) to a preset range. An obvious and clinically important difference between the artificial pancreas and the thermostat is the active control the thermostat can initiate by actuating air conditioning or heating. The artificial pancreas does not have the option of actuating glucagon release and instead, when low glucose is sensed, can only hold insulin delivery. Acute active correction of low blood glucose requires intervention by the patient (sensing the low blood glucose) or bystanders who observe abnormal behaviour of the patient.

Although these two control systems for blood glucose and temperature control are superficially similar, the artificial pancreas was much more difficult to engineer than a thermostat because of the complexities of quantifying input (amount of carbohydrate in a meal, the effects of stress, exercise and illness), heterogeneity in insulin responsiveness from person to person, and changes in insulin formulations and delivery methods. Furthermore, the control diagram for blood glucose does not capture all the elements of glucose control—some known and some still to be discovered. For example, details at the level of dopaminergic signalling in the pancreas [4] are not specifically modelled in the control algorithms.

3. Modelling of normal glucose control versus engineered control of glucose levels

3.1. Early translationally focused models of glucose homeostasis

Mathematical modelling of glucose homeostasis in health and disease has a long history. From the 1960s through 1990s, most models were based on ordinary differential equations and partial differential equations, with a variety of network approaches introduced in the past 25 years. A recent review of the diverse approaches to non-clinical and clinical models of diabetes highlights the diversity of mathematical approaches, models focused on a wide range of physiologic scales, covering various subsystems of the glucose regulatory machinery, and various levels of complexity [5].

The artificial pancreas consists of a subcutaneous glucose monitor (sensor), an algorithm that controls the pattern of insulin delivery and a pump (actuator) for insulin administration subcutaneously. The ‘brains’ of the artificial pancreas are the control algorithms that govern the timing and rate of insulin delivery. In silico testing of insulin delivery systems has been critical in speeding up and refining development of the artificial pancreas. The first clinical models generally involved comparison of glucose control achieved by an algorithm in silico versus actual results in experimental subjects. An in silico patient simulation model of a ‘healthy’ diabetic requires more components than those implied by figure 1—because of interindividual heterogeneity in each component of glucose regulation. For example, individuals absorb carbohydrate with different kinetics and each component of glucose regulation is affected differently by perturbations to the system (such as stress). For that reason subcomponent modelling of glucose kinetics,
subcutaneous glucose kinetics, gut absorption, insulin absorption and kinetics is more realistic than a global control model [6], and represents a long-term ideal for glucose controllers beyond simple glucose sensing.

Early algorithms meant to mimic the best clinician's intuition were hampered by sparse glucose information (before continuous glucose sensing became available) and no availability of rapid-acting insulin formulations. Heterogeneity in individual glucose–insulin responses (the insulin to carbohydrate ratio in particular) required use of control algorithms that were optimized by 'learning' an individual patient's responses; for example to adjust glucose levels after a meal, information was used on the glucose responses to a meal the day before, each meal adjusted independently. These first algorithms did not aim for tight control, rather for safe insulin delivery regimens with some improvement in glucose homeostasis, but they fell short even in idealized situations where input and stressors were well controlled, situations that are difficult to control in real life.

3.2. Technology advances to overcome sparse information for modelling

Continuous glucose sensing provided more information to modellers, so that the patterns of glucose over days could be used to determine safe baseline insulin infusion rates. Subcutaneous glucose readings lag behind blood glucose readings, the lag time differs between sensors and this lag has to be incorporated into control models. Similarly, insulin onset is delayed after delivery depending on delivery device, site of delivery and other variables, and these variables must be accounted for in control algorithms.

Early results applying the first control algorithms to the clinic unmasked the importance of information from all components of the glucose maintenance machinery. For example, knowledge of the carbohydrate content of a meal was important input data, and with that knowledge an in vitro controller worked well. But when actual patients ate meals with a defined carbohydrate content, the model failed to be as expert in glucose control as an experienced clinician. The algorithm had been based on liquid glucose meals that are absorbed differently from carbohydrates in solid food, highlighting the need to explicitly model gut absorption [7]. Once again, these results pointed to the need for control models to explicitly deal with subsystems using separate controllers to manage the data in an organized hierarchy. One example is the compartment model developed by Hovorka that includes a subsystem model of gut carbohydrate absorption and one of subcutaneous insulin absorption. In this model, Bayesian parameter estimation was used [8]. Another approach to the issue of patients taking bolus insulin and meals at the same time was development of a linear model using clinical data to identify model parameters. The results of the clinical trial embedded in the development process were similar to the Hovorka model, but had the advantage that it could be personalized for the common clinical scenarios of a snack without an insulin bolus and a ‘correction’ bolus [9]. Early clinical studies also helped define how the controller gain (change in input versus change in output) could be adjusted to best match expert clinical recommendations for insulin delivery using simple linear rules. High gain mediates a faster loop response but potentially more unstable oscillations.

3.3. Control methodologies for glucose control systems

Clinical studies on the first automated closed-loop insulin delivery system were reported starting in 2006 [10]. These pumps are based on a proportional–integral–derivative (PID) controller with input data from hyperglycaemic clamp tests. PID controllers are by far the most commonly used controls in chemical, oil and paper industries. Hyperglycaemic clamp tests are used to determine beta-cell sensitivity to glucose. In this test, a subject's blood glucose is raised to abnormally high levels and held there using intravenous glucose infusions. With the high glucose levels held constant, the rate of glucose infusion necessary to maintain the target level reflects the rate of glucose metabolism. These studies led to the understanding that the plasma insulin response to sustained hyperglycaemia is biphasic, with a quick burst of insulin release then a slower release phase [11].

A PID controller uses past information, present information and prediction of future control errors to take a control action [12]. For a PID controller, \( P \) refers to proportional action, here delivery of insulin in proportion to the difference between the measured glucose level and the target glucose level. This \( P \) component of the controller does not contribute to the insulin levels required to maintain fasting glucose at a certain target. I refers to integral and D to derivative: here they mediate the second-phase rise (I) and rapid first-phase rise (D) observed during hyperglycaemic clamps. The I component of the controller ensures that when glucose veers off target, there is a change in insulin delivery. The D component stabilizes the system by responding to changes in glucose with changes in insulin delivery. Insulin delivery controlled by the three PID components, calculated separately, is the sum total of insulin delivered [12]. In general terms, PID controllers exhibit good performance in systems that do not require a steady-state change in the controller response to get to the target output. PID and other control methodologies, with their relative strengths and weaknesses are summarized in table 1, with control synthesis methods listed at the bottom of the table. Overall, the commercial PID controller insulin pump, using insulin feedback, was a big advance for controlling glucose levels in T1DM patients, but at the price of relatively high post-prandial glucose levels. Various methods for tuning PID controllers for optimization of control have been developed for use in other systems [13].

Another approach to glucose control is the use of fuzzy logic, promoted by an Israeli group, in the MD-Logic artificial pancreas [14]. Fuzzy logic modelling incorporates use of a human ‘expert’ operator’s knowledge in control of a system without need for detailed knowledge of underlying dynamics of a system. The controller is based on fuzzy rules with linguistic variables, using IF (input)–THEN (output) structure. Fuzzy logic controllers were developed for (physical) systems that are difficult to model mathematically, and are not amenable to more analytic control design schemes [15]. The first clinically tested fuzzy-logic-based glucose control system used a combination of control-to-target and a fuzzy control-to-range strategy. Seven adult type 1 diabetics who were historically well controlled were selected for the pilot trial. Input data were obtained from each patient’s history and treatment profile. Performance of the controller in these selected patients was reasonable. This controller was subsequently tested in overnight sessions in seven T1DM patients and showed promise for control of overnight glucose levels [16], because no patients in this small study had overnight hypoglycaemia.
A zone model predictive control (MPC) strategy has advantages over previous open-loop pre-adjusted insulin treatment strategies, and is the basis for the FDA-approved wearable artificial pancreas. MPC is also referred to as moving horizon or receding horizon control. This treat-to-range, closed-loop adaptive control is useful for preventing extremes in glucose levels. It is particularly well suited to the multivariable nature and inherent constraints of the glucose regulatory system. Originally used in the oil industry, MPC is suitable to a system in which the goal is to control output within a specified range (or zone), rather than control to a fixed point. MPC uses iterative optimization to predict future changes in dependent variables caused by changes in independent variables, in the case of T1DM, personalized predcitions based on updated process data. Another advantage of MPC is that it identifies solutions for the current state of the system at every programmed unit time interval (rather than identifying one solution for all time points), in a short time frame using finite horizons. The particular zone MPC underlying the artificial pancreas uses mapped input data (clinical insulin administration and meal history, treated using two different second-order transfer functions) to predict the best next cycle of insulin treatment. This technique overcame a problem arising from the opposite effects on glucose levels of meals and insulin delivered close together. The timing of peak effect of insulin on glucose and timing of peak effect of meals on glucose were used to map each into new states, amenable to analysis with differential equations [17].

3.4. Constraints on control design and synthesis methods to optimize control

Unlike type 2 diabetes, T1DM patients run the risk of deadly hypoglycaemia if insulin is administered too aggressively. Tight control of glucose is critical for prevention of end-organ complications in patients with T1DM, but early algorithms used to control glucose aggressively were associated with a high incidence of hypoglycaemic events. Nocturnal hypoglycaemia is also a major clinical problem in these patients. Because of the risk of seizures and death from hypoglycaemia, considerable effort has gone into developing real-time algorithms that predict hypoglycaemia using continuous glucose monitoring data. These include statistical and linear prediction algorithms [18], Kalman filter [19], hybrid infinite impulse response filter [20], numerical logical algorithm and linear projection, and combinations of these algorithms [21]. Filters are generally applied to implement signal transforms, and filter design is critical in predicting a system’s state. Kalman filter (or linear quadratic estimation) is a mathematical technique to estimate the state of a process, minimizing the mean of squared error. The power of Kalman filtering is that it can predict a state using time series data (here glucose and rate of change of glucose) from a noisy system even if some aspects of the modelled system are poorly modelled. Kalman filters are applied in navigation, computer vision and many other applications. Numerical logical algorithm, using techniques that are robust to differing sample rates, estimates the rate of change of glucose using numerical estimations. These various algorithms are synthesis methods, that can be combined in a controller, to formalize the mathematics of the controller. Table 1 lists the broad categories of synthesis methods in use by control engineers.

3.5. Simulation to support clinical studies

Pre-clinical and clinical experiments are increasingly supplemented and refined using simulation (table 2). A major advance in modelling was FDA approval of a T1DM patient simulator for pre-clinical development of a clinical product [22]. The simulator was a critical factor in refining models of control efficiently (obviating need for pre-clinical animal

Table 1. Controller tuning methods. For a description of these methods, see [12].

<table>
<thead>
<tr>
<th>methods</th>
<th>positives</th>
<th>negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>PID</td>
<td>simple to tune</td>
<td>inflexible</td>
</tr>
<tr>
<td></td>
<td>uses process models or online tuning</td>
<td>restricted class of controllers</td>
</tr>
<tr>
<td></td>
<td>exploits dynamics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>can be analysed with</td>
<td></td>
</tr>
<tr>
<td>standard methods</td>
<td>easy, intuitive</td>
<td>limited</td>
</tr>
<tr>
<td>fuzzy logic</td>
<td>exploits user knowledge</td>
<td>performance especially in dynamic setting</td>
</tr>
<tr>
<td></td>
<td>requires no mathematics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>model predictive</td>
<td>flexible</td>
</tr>
<tr>
<td>control/receding</td>
<td>exploits models</td>
<td>harder to use</td>
</tr>
<tr>
<td>horizon control</td>
<td>can deal with nonlinearities</td>
<td>harder to analyse</td>
</tr>
<tr>
<td>synthesis methods to mathematically formalize tuning methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—classical control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—optimal control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—robust control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—nonlinear control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—adaptive control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>—distributed control</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Controller analysis methods.

<table>
<thead>
<tr>
<th>method</th>
<th>positives</th>
<th>negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>field testing</td>
<td>best validation</td>
<td>inflexible</td>
</tr>
<tr>
<td>experiments</td>
<td>can closely match field testing</td>
<td>inflexible</td>
</tr>
<tr>
<td></td>
<td>useful feedback for design</td>
<td>expensive</td>
</tr>
<tr>
<td></td>
<td>good validation</td>
<td></td>
</tr>
<tr>
<td>simulation</td>
<td>fairly extensive</td>
<td>limited certainty</td>
</tr>
<tr>
<td></td>
<td>flexible</td>
<td>requires good</td>
</tr>
<tr>
<td></td>
<td>relatively inexpensive</td>
<td>mathematical models</td>
</tr>
<tr>
<td></td>
<td>can try one scenario at a time</td>
<td></td>
</tr>
</tbody>
</table>

rsfs.royalsocietypublishing.org/
3.6. Beyond insulin

A bihormonal (fast-acting insulin and its counter-regulatory hormone, glucagon) artificial pancreas has been tested in the clinic. The first patients treated with the controller experienced hypoglycaemia, requiring adjustment of the insulin pharmacokinetic models embedded in their control system, so that the new model accounted for accumulation of insulin in plasma as well as its appearance at the subcutaneous injection site. Subsequent patients treated with the modified controller did not experience hypoglycaemia [27]. The initial trial was done using blood glucose measurements and a laptop computer (rather than smartphone), but control system and product development are underway [28], as the addition of glucagon to the system should ultimately yield a more physiologically relevant control strategy. These studies with two hormones delivered suggest that very small amounts of glucagon are effective at blunting post-prandial hyperglycaemia experienced using the first generation wearable artificial pancreas.

4. Discussion

Control of glucose regulation is a surprisingly complex engineering problem, and the mathematical and software infrastructure necessary for development of the first wearable artificial pancreas was itself complex. Here, we highlighted some control engineering tools used to achieve a safe, patient-friendly, closed-loop artificial pancreas. The device is not perfect, but incorporation of new discoveries in basic and clinical science will certainly drive better controllers. Clearly, parts of the glucose homeostatic regulatory mechanisms are not explicitly modelled because they are poorly understood. For example, a greater understanding of why gastric bypass so dramatically improves glycaemic control in type 2 diabetes may lead to more complete models of the role of the gut and gut hormones in glycaemic control. In vitro systems will certainly yield information about the heterogeneity of beta-cell function that can also inform future models and simulations.

Glucose control was an obvious target for control engineering, because glucose homeostasis has well-defined feedback loops at its core, with identifiable sensors and actuators. Furthermore, diabetes is a major clinical problem, driving research and technology development. Normal physiology is replete with control feedback loops amenable to modelling and control, and the lessons from development of the artificial pancreas should be useful in developing treatment strategies for other diseases. For example, satiety is controlled by a series of mechanical, and chemical and hormonal feedback loops between the gut, viscera, adipocytes and brain. Modelling these systems [29] may lead to insights into treatment of obesity. Circadian rhythms are based on well-known feedback loops. But surprisingly components of both these systems (satiety and circadian oscillations) are still being discovered, and missing components make modelling more challenging. In addition, lessons from the development of the artificial pancreas show that progress in complex modelling and simulation necessary for pre-clinical studies will require a multidisciplinary and probably international effort between mathematicians, engineers and clinicians with deep expertise in several clinical subspecialties, as well as the support of patient advocates and funding agencies. This paper was inspired by the scientists and clinicians who adopted control engineering to an important clinical problem, and is meant as an introduction to control engineering concepts applicable to medicine generally. Fortunately, for readers who want more detail on the particular algorithms used in development of the wearable artificial pancreas, a review by the control engineering group involved was recently published [3].

References

7. Palem CC, Zisser H, Bevier WC, Ivanovic L, Doyle III JF. 2007 Prandial insulin dosing using run-to-run control. Application of clinical data and medical expertise to
define a suitable performance metric. Diabetes Care 30, 1131 – 1137. (doi:10.2337/dbc08-2115)


