Challenges and Progress in the Development of a Closed-loop Artificial Pancreas

B. Wayne Bequette

Abstract—Pursuit of a closed-loop artificial pancreas that automatically controls the blood glucose of individuals with type 1 diabetes has intensified during the past six years. Here we discuss the progress and challenges in the major steps towards a closed-loop system. Continuous insulin infusion pumps have been widely available for over two decades, but “smart pump” technology has made the devices easier to use and more powerful. Continuous glucose monitoring (CGM) technology has improved and the devices are more widely available. A number of approaches are currently under study for fully closed-loop systems; most manipulate only insulin, while others manipulate insulin and glucagon. Algorithms include on-off (for prevention of overnight hypoglycemia), proportional-integral-derivative (PID), model predictive control (MPC) and fuzzy logic based learning control. Meals cause the major “disturbance” to blood glucose, and we focus on approaches that our group has developed to predict when a meal is likely to be consumed and its effect.

I. BACKGROUND

The alpha and beta cells of the pancreas of a healthy individual regulate the blood glucose concentration to around 80 mg/dL. When the concentration is high, insulin is secreted by the beta cells and when the concentration is low, glucagon is secreted by the alpha cells. Individuals with Type 1 diabetes mellitus (T1DM) no longer produce insulin and, therefore, these individuals must inject insulin to regulate their blood glucose concentration. The importance of tight blood glucose control was not fully appreciated until the results of the Diabetes Control and Complications Trial (DCCT) were published in 1993. The DCCT involved a comparison of conventional therapy (one or two daily insulin injections, and daily monitoring of blood glucose or urine) with intensive insulin therapy (multiple daily injections or an insulin pump, and blood glucose measured at least 4 times per day, with daily adjustments to the insulin), and concluded that intensive therapy resulted in lower mean blood glucose values and significantly reduced complications (retinopathy, nephropathy, and macrovascular disease). The risk of complications is directly related to glycated hemoglobin, known as A1c; further, the A1c is related to the mean blood glucose values during the previous 2-3 months. O’Grady et al. (2011) find that tighter blood glucose levels achievable with a closed-loop artificial pancreas can result in Medicare savings of $1.9 billion over 25 years.

In this paper the emphasis is on subcutaneous delivery of rapid-acting insulin using external continuous insulin infusion pumps, and continuous glucose monitors (sensors) that output a signal that is related to the interstitial glucose (just beneath the skin) and therefore an indicator of the capillary blood glucose concentration. It should be noted that research continues on the use of implantable sensors and pumps and alternative delivery routes such as the intraperitoneal cavity (Renard et al., 2010).

The focus of this review is on articles and approaches that we expect to be applied in clinical studies in a relatively short term. In particular, we concentrate on projects related to the Juvenile Diabetes Research Foundation (JDRF) Artificial Pancreas Program. While many of the papers are based on simulations, we emphasize publications that involve a medical collaborator, again with a realistic plan of clinical implementation. Other recent reviews include Bequette (2005), Doyle et al. (2007), Hovorka (2008), Kumareswaran et al. (2009), El-Youssef et al. (2009), Cobelli et al. (2009), Harvey et al. (2010), and Hovorka (2011). Doyle (2011) presents the artificial pancreas as one of the grand challenges for control.

II. OVERVIEW OF THE TECHNOLOGY AND CHALLENGES

A closed-loop artificial pancreas is shown in Fig. 1; the actuators and sensors are available commercially. There are four continuous glucose monitors (CGM), numerous pumps, and very many self-monitoring (fingertip) blood glucose ( SMBG) meters. Diabetes Forecast (see Jan. 2011, for example) provides a comprehensive list of diabetes care products in its annual Consumer Guide.

![Fig. 1. Block diagram of a closed-loop artificial pancreas.](image)

II.1 Current State of Care

Note that in the current standard of care, an individual serves as the feedback controller, with measurements and control decisions that are made relatively infrequently. Roughly 20-25% of individuals in the US with T1DM use continuous insulin pumps that continuously deliver microboluses of rapid-acting insulin. A much smaller number of individuals use CGM. Individuals on intensive insulin therapy use SMBG meters to measure their blood glucose levels. Roughly 1-2% of individuals in the US with T1DM use a closed-loop system, which is not applied in the current standard of care.
concentration. Similarly, an insulin to carb (I:C) ratio is used to estimate meal insulin bolus needs. Rapid-acting insulin is used for correction or meal boluses, regardless of whether the individual uses a pump or syringes.

II.2 Input Challenges
There is a significant lag on the insulin effect on glucose uptake, even when rapid-acting insulin is delivered subcutaneously. Since insulin delivered in the recent past continues to have an effect, it is important to consider the remaining “insulin on board” (IOB) when deciding on a current insulin delivery rate. While meals cause a faster response in glucose concentration than insulin delivered, the time scale can be significant and highly variable; in addition, while it would be preferable to use knowledge about meal size to provide feedforward control (“meal announcement”) by injecting insulin (a “meal bolus”), it can be difficult to estimate the amount of carbohydrates in a meal.

II.3 Output Challenges
Currently available continuous glucose sensors suffer from a time lag between the capillary blood and the interstitial fluid, where the sensor is placed; also, there are often periods when the sensor results are biased due to the calibration procedure. Continuous glucose sensors do not eliminate the need for capillary blood glucose measurements (fingersticks); most CGMs require a calibration after a 2-hour “warm-up” period, and calibrating blood glucose measurements every 12 hours. These sensors are currently approved for “adjunctive” use only, that is, any decision to change treatment (such as changing the insulin delivery rate) must be based on a confirmation fingerstick measurement.

II.4 Related Behavior
An individual’s insulin sensitivity and insulin delivery needs vary throughout the day, depending on meals, exercise, stress and normal diurnal variations. An individual using a pump can set different basal rates for different times during the day. The “dawn phenomena” at roughly 4:00 am, results in reduced insulin sensitivity, causing the blood glucose concentration to rise; an individual can compensate for this by programming the pump to provide a higher insulin delivery rate during that time period. Similarly, during periods of exercise, an individual may need to reduce their basal insulin delivery to near 0. During long periods of intense training, say for a marathon, an individual’s daily insulin demand may decrease by 50%.

III. Dynamics of manipulated & disturbance inputs

III.1 Dynamics of Subcutaneously Delivered Insulin
A major challenge in regulating blood glucose levels by manipulating the delivery of subcutaneous insulin, whether manually or automatically, is the long time-scale pharmacodynamic action of even rapid-acting insulin. After a bolus of insulin is delivered, the time before the maximum rate of change in blood glucose uptake (peak action) is roughly 90 minutes, and the insulin continues to have an effect on glucose for 6-8 hours; see Fig. 2 for typical pharmacodynamics profiles for rapid-acting insulin. Bequette (2009) reviews the glucose clamp procedures that are implemented clinically to estimate the profiles. Nearly all commercially available insulin pumps provide an estimate of the “insulin on board” (IOB), which is an indicator of previously delivered insulin that will continue to have an effect on the glucose concentrations in the future (see Fig. 3). Zisser et al. (2008) review the approaches used by four different insulin pumps to estimate the IOB.

III.2 Meal Dynamics
The dynamic effect of a meal on blood glucose can vary depending on a number of factors, including the fat content of the meal. The rate of glucose appearance in the blood is shown in Fig. 4 for several different studies; note that a meal...
effect can continue for 3-8 hours, but usually with a faster “peak” than the insulin pharmacodynamic behavior.

IV. CONTINUOUS GLUCOSE MONITORING (CGM)
It is intuitive that CGM would enable an individual to better regulate their blood glucose values. The JDRF CGM Study Group (NEJM, 2008) showed that the use of CGM improves glycated hemoglobin(A1c) levels in individuals over the age of 25, whether they use continuous insulin infusion (pump) therapy, or multiple daily injections. Bergenstal et al. (2010) further showed that CGM combined with pump therapy resulted in better performance (greater reduction in A1c) than CGM combined with multiple daily injections. A review of this “sensor-augmented” insulin pump therapy is provided by Cengiz et al. (2011).

Continuous glucose monitors currently require frequent calibration by the use of a reference glucose based on a fingerstick. If there is an error in the reference glucose value, due to meter uncertainty, a user mistake in taking the blood sample, or sampling during transient conditions (particularly because of the lag between blood and interstitial fluid glucose), there is often a bias in the CGM signal until the next reference glucose sample is taken. An overview of CGM calibration and sensor signal filtering algorithms is provided by Bequette (2010).

IV.1 Hypoglycemia Detection/Prediction/Prevention
A major concern of any parent of a child with diabetes, is that the child may go hypoglycemic overnight; during the day the symptoms would be more likely to be noticed within a relatively brief period of time, but without a continuous glucose sensor there would be no way to detect overnight hypoglycemia. Palerm and Bequette (2007) perform a retrospective analysis of clinical data to show the effect of different tuning parameters in a Kalman Filter based hypoglycemic alarm system. They show how tuning parameters could be adjusted by an individual based on their own tolerance to false alarms. Dassau et al. (2010) report clinical results for a voting-based strategy that involves several different algorithms, including a Kalman filter and a statistical prediction method (Cameron et al., 2008) to predict hypoglycemia.

IV.2 Pump Shut-off (Low Glucose Suspend)
Buckingham et al. (2005), in a study of the GlucoWatch G2 Biographer CGM (no longer commercially available), note that children only awoke to 29% of alarms, and their parents, when sleeping in the same room, only responded to 37% of alarms. It is desirable to take people out of the loop and simply shut-off the pump rather than sound an alarm; this approach is often called “low glucose suspend.” Buckingham et al. (2009) report clinical pump shut-off results; a linear prediction algorithm with a 45-min horizon prevented hypoglycemia 80% of the time. Buckingham et al. (2010) present a voting-based strategy that prevents 84% of possible hypoglycemic events. To-date, performance metrics have involved blood glucose samples, which are available because all studies have been performed in a clinic. Beck et al. (2011) argue that the outcome measures for outpatient studies of must be based solely on the CGM data.

Choudhary et al. (2011) report results based on the Medtronic Paradigm Veo, which can suspend the basal insulin delivery for up to 2 hours when hypoglycemia is detected by a CGM; this low glucose suspend system is currently available in Europe but not the US. Insulin pump suspensions may occur naturally, as part of a fully closed-loop system, particularly during periods of a rapid decrease in blood glucose. Cengiz et al. (2009) report on pump suspensions that occurred during clinical trials of a PID algorithm, while Elleri et al. (2010) show pump suspension results during trials of a MPC algorithm.

V. OTHER CONTROL-RELATED CONSIDERATIONS

V.1 Risk Measures
Blood glucose control in diabetes is a balancing act between the long-term complications of hyperglycemia and the short-term danger of hypoglycemia. In addition, the hypoglycemic and hyperglycemic risks are asymmetric. A patient at 100 mg/dL faces a much greater risk if the glucose decreases by 50 mg/dL than if glucose increases by 50 mg/dL. An asymmetric risk measure was developed by Kovatchev et al. (1997), adapted by Cameron et al. (2011), and used in a closed-loop MPC strategy detailed in section VI.5. Desborough et al. (2011) have noted that only the log-square and the Cameron et al. measures effectively balance the risks of both hypoglycemia and hyperglycemia.

V.2 Meal Detection/Prediction
Since meal dynamics (disturbance input) can have a significant time-scale and certainly the insulin (manipulated input) pharmacodynamics time scale is long, it is desirable that feedforward control be used to compensate for a meal; this is often called meal announcement. It is known, however, that people often forget to provide a meal-related insulin bolus (Burdick et al., 2004) resulting in higher A1c values and long-term complications. It is desirable, therefore to have a scheme that detects or predicts a meal and at least partially provides feedforward action. Dassau et al. (2008) present clinical results on a voting algorithm based procedure that detects a meal, on average, within 30 minutes after the onset of the meal. Lee and Bequette (2009) and Lee et al. (2009) present alternative procedures that also estimate the meal size. Cameron et al. (2009) present a probabilistic approach to meal detection and estimation of the rate of glucose appearance.

The meal detection methods discussed above are not “anticipatory” in nature, that is, they are based on changes in blood glucose due to meals that have already occurred (although “unannounced”). Better closed-loop performance can be obtained if meals are anticipated, perhaps through knowledge of common mealtimes. Particularly with model predictive control, if there is a high probability of a meal occurring during the “prediction horizon,” then this can be considered by the control action. Hughes et al. (2011) present a stochastic MPC algorithm that is based on a probabilistic description of the individual’s daily meal habits. Cameron et al. (2011) take a slightly different
approach, by assigning probabilities of future meals based on the time that has lapsed since the last meal.

V.3 Modeling for Control
While it would be ideal to use existing glucose sensor and insulin infusion data, with meal information, to develop models, this is difficult in practice. For one, the recording of meal times and carbohydrate amounts is notoriously bad and the sensor and insulin pump times are not necessarily synchronized with any written recordings. Also, since an insulin bolus is typically given at the same time as a meal, it is difficult for identification algorithms to distinguish the difference in the two effects. Finan et al. (2007), and Lee and Bequette (2009) show that even the sign of the gain between insulin and glucose concentration can be wrong; it should be negative, but identification techniques may yield a positive gain, due to the simultaneous effect of the meal.

Lee and Bequette (2009) propose a “human-friendly” identification-based approach to improve the development of control-relevant models; the methodology is analogous to “plant-friendly” techniques that have been developed for the process industries. Finan et al. (2009) analyzed data from nine subjects with type 1 diabetes in ambulatory conditions and found that identified ARX models yielded marginal improvements in glucose predictions, compared to simply assuming that the current glucose concentration remains constant into the future; this is another indicator of the difficulty of using normal “free living” data, rather than data generated specifically for control-relevant model development. An alternative approach is to develop model parameters based on clinically relevant information, such as total daily insulin, I:C ratio and CF, as proposed by Percival et al. (2010), who apply this approach to a first-order+deadtime model.

V.4 Simulation-based Testing
A realistic simulation environment with a wide variety of simulated subjects enables the development of control strategies that are robust and reliable. Indeed, the FDA approved the use of the UVa-Padova simulator (Kovatchev et al., 2009) for use in simulated clinical trials, enabling participants to skip the animal trial stage; this simulator is based on a model presented by Dalla Man et al. (2007). Patek et al. (2009) discuss this approach for simulated closed-loop clinical trials. The UVa-Padova simulator contains 300 subjects (100 each of children, adolescents and adults), and includes sensor errors representative of two CGMs and the discrete resolution from two insulin pumps. Wilinska et al. (2009, 2010) discuss the use of simulation studies for evaluating model predictive control strategies in simulated clinical trials; their simulation studies involve 18 different subject parameter sets.

V.5 Control System Platforms for Clinical Studies
While there are many commercially available subcutaneous insulin infusion pumps, and four currently available continuous glucose monitors that are approved by the US FDA, there are currently no standards to connect them with a device that contains a control algorithm. Many of the initial clinical studies have used manual entry of the CGM reading into a laptop computer, followed by manual implementation of the calculated control action into the insulin pump; this is perhaps one reason that many of the initial studies are based on a sample time of 15 minutes.

For the fully automated clinical studies that have been performed, a great deal of effort has gone into the development of hardware and software to form the closed-loop. For example, the artificial pancreas system (APS) platform developed at the University of California at Santa Barbara (UCSB) has been used by a number of groups involved in the JDRF AP consortium (Dassau et al., 2008, 2009); currently two CGMs and three insulin pumps are supported by the APS.

VI. CONTROL ALGORITHMS

VI.1 On-Off
On-off control is used by the low glucose suspend (or pump shut-off) types of systems. The decision to shut-off the pump can be based on a threshold (basically a hypoglycemic detection), or a projected violation of a threshold (a prediction that a low glucose will occur within a prediction horizon). In on-going in-patient studies, and in planned outpatient studies, we are using a Kalman filter based prediction for the pump shut-off algorithm; this provides more flexibility for handling sensors with different sample times, and naturally handles brief sensor drop-outs.

VI.2 Control-to-Range
A low glucose suspend controller seeks to maintain blood glucose above some minimum value. Control-to-range represents the next level of control, where the objective is to regulate blood glucose between upper and lower bounds. Kovatchev et al. (2009) present a basic structure for this approach, while Grosman et al. (2010) show simulation results for a zone model predictive control strategy.

VI.3 Proportional-Integral-Derivative (PID)
The Medtronic external physiological insulin delivery (ePID) system includes a PID controller that has been used in animal and human studies. Weinzimer et al. (2008) applied ePID to 17 adolescents and found that a small insulin priming bolus (feedforward or “meal announcement”), provided 15 minutes before meals, reduced the postprandial glucose peaks. The recent approach used by Medtronic involves model-based feedback of insulin concentration, creating a cascade type of strategy (Palem, 2011; Steil et al., 2011), called ePID-IFB. van Bon et al. (2010) use a PD controller in a clinical study of six subjects.

Gopakumaran et al. (2005) developed a fading memory proportional derivative (FMPD) controller that is roughly equivalent to PID. Ward et al. (2008) manipulate both insulin and glucagon in studies conducted on rats.

VI.4 Fuzzy-Logic
A fuzzy logic-based approach that uses a combination of control-to-range and control to setpoint strategies is incorporated into the MD-Logic Artificial Pancreas System (Atlas et al., 2010), with a 5-min sample time; this has been
tested in a trial on seven adults, without the use of meal announcement. Miller et al. (2011) describe the learning algorithm that extends the MD-Logic strategy to better handle interpatient variability; the strategy is tested in simulation studies.

Mauseth et al. (2010) describe a fuzzy logic based controller, with a 15-min sample time, that uses BG, its rate-of-change and its acceleration as inputs, and is tuned based on a personalization factor. A preliminary version was tested (without a personalization factor) on four subjects, before enhancements were made and performance demonstrated in simulation studies.

VI.5 Model Predictive Control (MPC)

MPC is a basic framework or strategy that can involve many different types of models and objective functions. Hovorka et al. (2004a) presented an approach based on a nonlinear model and Bayesian techniques to estimate parameters in simulation studies. Clinical studies were performed under fasting conditions by Hovorka et al. (2004b), based on i.v. measurements that were delayed by 30 min to mimic the time lag associated with a s.c. sensor. Hovorka et al. (2010) performed overnight studies using an MPC strategy and manually entering CGM data in the algorithm and transferring results to a pump at 15-min intervals; the major outcome was a reduction in nocturnal hypoglycemia compared to standard pump treatment. Hovorka et al. (2011) and Elleri et al. (2011) presented overnight studies based on a fully-automated MPC strategy that was initiated immediately after either dinner or a late night snack; again with a 15-min sample time.

Magni et al. (2007) present an unconstrained MPC strategy, where the model is a linearization of a nonlinear model, obtained at an average value of the population parameters. In simulation studies with a sample time of 30 minutes, they show that a single parameter, the weighting on the output predictions in the objective function, can be tuned for each individual for better performance. This approach is used in Bruttomesso et al. (2009) in a trial with six subjects; parameters included a sample time of 15 minutes and a prediction horizon of 240 minutes.

Kovatchev and colleagues use a one-min sensor sample time and a 15-min actuator sample time. Kovatchev et al. (2010) report that clinical studies (each 22 hr, with 14.5 hr in closed-loop) involving 20 adults reduced nocturnal hypoglycemic events from 23 to 5, and increased the amount of time within the target range from 64% to 78% compared to standard open-loop treatment. Simulation-based studies were used to design the controllers before implementation in the clinical studies. Clarke et al. (2009) revise this approach in an overnight study of eight subjects, using individualized models based on weight, total daily insulin dose and a BG CF measured during admission.

Ellingsen et al. (2009) use MPC based on ARX models and a 5-min sample time; IOB constraints based on I:C and CF were implemented in a simulation study.

An adaptive generalized predictive control (GPC) approach (based on recursive identification of ARX models) is taken by El-Khatib et al. (2007) in their studies involved diabetic swine; in addition to insulin, their strategy adjusts glucagon to improve control at lower blood glucose levels. Also, El-Khatib et al. (2007) include a prediction of the insulin concentration and include it in an objective function to avoid problems associated with IOB. In the human clinical studies reported in El-Khatib et al. (2010), a PD controller that is active under certain glucose concentrations is used for glucagon.

Lee et al. (2009) and Lee and Bequette (2009) use subspace identification techniques to develop discrete state space models, and incorporate IOB constraints in MPC; additional features include a pump shut-off algorithm to avoid hypoglycemia, and meal detection and meal size estimation algorithms to handle un-announced meals. More recently, we (Cameron et al., 2011) have developed a multiple model probabilistic predictive control (MMPPC) strategy that minimizes an asymmetric risk function subject to satisfying hypoglycemic constraints; the controller is forced to be more conservative when uncertainties are high. We are using a similar approach in on-going clinical trials.

VII. SAFETY AND FAULT DETECTION

There are many problems that can arise in a closed-loop artificial pancreas. Examples include sensors signals that drift or dropout, or are poorly calibrated, insulin infusion sets can fail, a planned meal that is not consumed, etc.

VII.1 Sensor dropout

There are often brief periods when CGM signals are either lost, or attenuated. Bequette (2010) discusses how a Kalman filter based approach can be used for model predictions without the measurement updates; when the state covariances indicate that there is too much uncertainty, an alarm could be activated. If this occurs overnight, the closed-loop system could be placed in some default basal delivery mode until morning.

Sensor attenuation (drift) is a difficult problem to detect, but similar approaches to the infusion set failure detection problem below could be used. An obvious solution is to use multiple sensors, as proposed by Castle and Ward (2010).

VII.2 Infusion Set Failure

A common problem encountered by diabetic patients on continuous insulin therapy is insulin infusion set failure, IISF, when Teflon catheters or steel needle infusion sets are worn for long periods of time. Common causes of IISF include blocked or dislodged sets, inflammation, or insulin leakage back to the skin surface. Rojas et al. (2011a,b) use bivariate classification, principal component analysis and a combined approach to detect simulated faults in 10 subjects. These approaches will be used in upcoming clinical trials for insulin set failure.

VII.3 Announced Meals That Are Not Consumed

While better closed-loop performance can be achieved when meal announcement is used to provide an insulin bolus, there is some risk that the meal will not be consumed; this provides a strong argument for a using a smaller “priming bolus.” An alternative is to extend the probabilistic
strategies of Cameron et al. (2011) and Hughes et al. (2011). Even when a meal is announced, it is not necessary to assign a prior probability of 100% to the meal algorithm; a lower value would enable a risk-based controller to provide a smaller meal bolus to provide safe behavior over the prediction horizon.

CONCLUSION

Diabetes technology has advanced considerably during the past five years. The path to a fully closed-loop artificial pancreas is proceeding in stages, with hypoglycemic alarms naturally leading to pump shut-off (low glucose suspend) systems, which in turn leads to control-to-range. Simulation studies are providing important results that should enable fewer clinical trials to accomplish given performance goals.

ACKNOWLEDGMENT

My on-going collaboration with Bruce A. Buckingham, Darrell M. Wilson and Fraser Cameron is gratefully acknowledged.

LITERATURE CITED
