Journal of Diabetes Science and Technology
Volume 3, Issue 2, March 2009

ORIGINAL ARTICLES

In Silico Preclinical Trials: Methodology and Engineering Guide to Closed-Loop Control in Type 1 Diabetes Mellitus

Stephen D. Patek, Ph.D.,[1] B. Wayne Bequette, Ph.D.,[2] Marc Breton, Ph.D.,[3] Bruce A. Buckingham, M.D.,[4] Eyal Dassau, Ph.D.,[5] Francis J. Doyle III, Ph.D.,[5] John Lum,[6] Lalo Magni, Ph.D.,[7] and Howard Zisser M.D.[8]

Abstract

This article sets forth guidelines for *in silico* (simulation-based) proof-of-concept testing of artificial pancreas control algorithms. The goal was to design a test procedure that can facilitate regulatory approval [e.g., Food and Drug Administration Investigational Device Exemption] for General Clinical Research Center experiments without preliminary testing on animals. The methodology is designed around a software package, based on a recent meal simulation model of the glucose–insulin system. Putting a premium on generality, this document starts by specifying a generic, rather abstract, meta-algorithm for control. The meta-algorithm has two main components: (1) patient assessment and tuning of control parameters, i.e., algorithmic processes for collection and processing patient data prior to closed-loop operation, and (2) controller warm-up and run-time operation, i.e., algorithmic processes for initializing controller states and managing blood glucose. The simulation-based testing methodology is designed to reveal the conceptual/mathematical operation of both main components, as applied to a large population of *in silico* patients with type 1 diabetes mellitus.

J Diabetes Sci Technol 2009;3(2):269-282

Introduction

In 2006 the Juvenile Diabetes Research Foundation (JDRF) created a consortium of research sites that now includes Boston University, Cambridge University (England), Sansum Diabetes Research Institute, Stanford University, University of Colorado, University of Virginia (UVA), and Yale University. The consortium is focused on the development of an artificial pancreas for the control of type 1 diabetes mellitus (T1DM) to be built from off-the-shelf interstitial continuous glucose monitoring (CGM) devices and continuous subcutaneous insulin infusion

Author Affiliation: [1]Department of Systems and Information Engineering, University of Virginia, Charlottesville, Virginia; [2]Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, New York; [3]Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, Virginia; [4]Department of Pediatrics, Pediatric Endocrinology and Diabetes Division, Stanford University, Palo Alto, California; [5]Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, California; [6]Jaeb Center for Health Research Foundation, Inc., Tampa, Florida; [7]Dipartimento di Informatica e Sistemistica, Universitá di Pavia, Pavia, Italy; and [8]Sansum Diabetes Research Institute & University of California Santa Barbara Department of Chemical Engineering, Santa Barbara, California

Abbreviations: (ARX) autoregressive, external input, (BG) blood glucose [concentration], (CGM) continuous glucose monitoring/monitor, (CHO) carbohydrate, (CVGA) control variability grid analysis, (FDA) Food and Drug Administration, (GCRC) General Clinical Research Center, (ICU) intensive care unit, (IDE) Investigational Device Exemption, (JCHR) Jaeb Center for Health Research, (JDRF) Juvenile Diabetes Research Foundation, (LBGI) low blood glucose index, (MPC) Model Predictive Control, (SD) standard deviation, (T1DM) type 1 diabetes mellitus, (UVA) University of Virginia

Keywords: artificial pancreas, computer simulation, diabetes control, *in silico* testing

Corresponding Author: Stephen D. Patek, Ph.D., Department of Systems and Information Engineering, P.O. Box 400747, University of Virginia, Charlottesville, VA 22904; email address patek@virginia.edu

pumps with a feedback control law that computes appropriate insulin infusion doses at periodic intervals in response to glucose measurements. Consequently, the development of control algorithms for such an artificial pancreas has been a very active area of research since 2007,[1–6] adding to an already large body of literature on the automatic control of T1DM via intravenous blood glucose (BG) measurement and intravenous or intraperitoneal actuation (see Parker *et al.*,[7] Bequette,[8] and Renard[9] and references contained therein).

Class III medical devices that involve automated drug delivery are considered high risk by the Food and Drug Administration (FDA), and Investigational Device Exemption (IDE) requests must be approved for any clinical trial protocol that involves automated download of CGM data to a feedback control algorithm, use of a computational platform for computing appropriate insulin dosing, and/or automated upload of insulin actuation commands to an insulin pump. The traditional route to IDE approval involves proof-of-concept testing in animal trials, which are costly, time-consuming, and not compatible with the time horizon for research and development set forth by the JDRF. In an attempt to minimize the time to General Clinical Research Center (GCRC) clinical trials and utilize contemporary computing technology, the JDRF consortium has advanced the notion of replacing animal testing with simulated (*in silico*) tests of artificial pancreas control algorithms. In the fall of 2007, researchers from the University of Padua, Italy, and the University of Virginia developed a closed-loop simulation software package[10] for *in silico* testing of closed-loop artificial pancreas algorithms and, in November 2007, proposed to the FDA the use of this package for proof-of-concept testing in place of animal trials. The FDA accepted this notion in January 2008, accepting the software for this specific purpose.[11]

It should be noted that *in silico* studies have proven useful for the design and analysis of control algorithms in intensive care unit (ICU) blood glucose control. Lin and colleagues[12] studied a virtual (*in silico*) cohort of 200 subjects to analyze the performance of the SPRINT protocol for insulin and glucose infusion (Lonergan and associates[13]). The statistical outcome measures of the virtual trials compared very favorably with clinical data from actual clinical SPRINT studies conducted on 165 critical care patients. Further, Chase and colleagues[14] provided clinical data sets for 20 critical care patients to enable groups without access to clinical data to develop models and design algorithms for glucose control in the ICU. Hovorka and associates[15] proposed a methodology for validating models of critically ill patients for *in silico* testing of ICU glucose controllers.

This article proposes a methodology for *in silico* testing of artificial pancreas algorithms, one that can take a mathematical description of a wide variety of control algorithms and reveal the performance characteristics of that algorithm as applied to a population of *in silico* patients with T1DM. More specifically, the methodology presented here is designed to show that proposed clinical artificial pancreas control algorithms, when implemented in a reliable fashion, perform well in a nominal setting and also in the face of uncertainties due to sensor noise, limited knowledge of patient state at initialization, imprecise knowledge about meals, and so on. The methodology proposed here is not designed to validate the hardware/software *implementation* of an artificial pancreas control algorithm. While simulation could be useful in this regard in general, the software package of Kovatchev and co-workers[10] is not equipped (and is not accepted) to simulate fault modes associated with the physical hardware and software implementation of control algorithms. Consequently, the proposed methodology does not test the response of the controller to software or system faults, such as communication problems, CGM or pump failure modes, or power-out/reboot scenarios that would require a real-time simulation that would include the clinical software and devices in what is referred to as hardware-in-the-loop studies,[16,17] which we leave as the subject of future work. *In silico* testing, as envisioned here, should be taken as just part of the overall strategy for securing FDA IDE approval, and a separate process of hardware and software "system validation" would inevitably be required, including validation of hardware and software features designed to accommodate human error and lack of compliance.

The Closed-Loop Simulation Software Package

The closed-loop simulation software package is based on (1) the oral glucose "meal" model of Dalla Man and colleagues[18] and (2) modifications to the meal model[4,19] to reflect glucose–insulin kinetics in T1DM. Details of the mathematical model, including component models for insulin transport from subcutaneous infusion to blood circulation and CGM sensor characteristics, can be found elsewhere.[10] The simulation software comes with 300 *in silico* patients with T1DM (100 adults, 100 adolescents, and 100 children). As described in Kovatchev and colleagues,[10] the decision of the FDA to accept the software for the purpose of proof-of-concept testing was based on comparisons with clinical data collected from a large number of human subjects with T1DM. The key

observation in this comparative study was that for each human subject with T1DM it was possible to identify at least one *in silico* subject in the same population group whose glucose–insulin response is comparable to that observed *in vivo* under comparable conditions. Thus, it is possible to conclude that an artificial pancreas control algorithm will be safe for human subject trials if the algorithm performs well at least for the subset of *in silico* subjects representative of the study population. In this vein, we point out that it is possible that among the 300 *in silico* subjects some of them show glucose–insulin responses that have not been observed *in vivo*. Any such "outlying" subjects still serve a useful purpose in revealing the stability and performance characteristics of control algorithms, helping establish robustness and limits of performance.

The closed-loop simulation software is implemented in Simulink/MATLAB and provides a well-defined set of interfaces that allow for testing of artificial pancreas control algorithms in user-defined scenarios with prescribed meal profiles. CGM noise is reflected in the simulator, as modeled elsewhere.[20] The software reports a wide variety of control-relevant plots and outcome measures, both per patient and for the various populations under study, including time within the target range of 70–180 mg/dl, the low blood glucose index (LBGI),[21] and the recently developed method of "control variability grid analysis (CVGA)."[22] For each *in silico* subject, there is an associated set of screening questionnaire parameters about the subject that can be used for customizing the controller, including age, weight, fasting BG, basal insulin rate, daily insulin requirement, "optimal" insulin bolus [units of insulin needed to cover 1 gram carbohydrate (CHO) without dropping below 95% of fasting BG], and maximum drop in BG due to 1 unit insulin bolus.

For researchers affiliated with the JDRF Artificial Pancreas Consortium, the Jaeb Center for Health Research (JCHR) plays a key role acting as an objective third party applying this methodology to assess CGM-based control algorithms. Consequently, the software is designed to be user-friendly and allows JCHR technical staff to "plug in" (1) a MATLAB code (an .m file) that implements patient assessment and tuning procedures for component 1, (2) a Simulink model of the warm-up and run-time procedures for component 2, and (3) an ASCII "scenario" file that describes the sequence of meals to be taken during closed-loop control and also allows the user to specify the *in silico* patient's blood glucose state at the onset of component 2. While knowledge of the meal profile of a clinical trial may be used in the design of a control algorithm, the software provides a "meal signal" that gives up to a 30-minute prior warning of an imminent meal. Since the software is not intended for medical decision making, it cannot be used directly in tuning feedback control algorithms for individual patients.

It is important to note that good *in silico* performance of a control algorithm does not necessarily translate into good *in vivo* performance. The claim of the computer simulation system is that it can out rule ineffective algorithms and test the stability and the robustness of control with challenges beyond the realm of physiology. However, simulation cannot guarantee *in vivo* performance. This is because the simulated "subjects" are still model based, which means that the complexity of a living organism is only approximated with a certain degree of accuracy provided by available data. In other words, while we were able to find a simulated "subject" that was reasonably close to each real person during testing of the simulated population against real glucose traces, we do not claim that all simulated subjects correspond to real people. With the accumulation of future data, the simulated population would be refined in terms of both excluding unrealistic subjects and expanding with new properties, such as hypoglycemia counterregulation and variable insulin sensitivity.

At the time of this writing, only the JCHR has access to the full version of the simulator, with all 300 *in silico* subjects, and the JCHR is equipped and authorized to run *in silico* preclinical trials on behalf of JDRF Artifical Pancreas consortium members. A limited version of the simulator, with 30 *in silico* subjects, is available to consortium members to allow for code development in preparation for preclinical trials.

Outline

The following section defines a generic artificial pancreas "meta-algorithm." The next section, "Methods: *In Silico* Testing of Control Algorithms" proposes a methodology for *in silico* testing of artificial pancreas control algorithms. "Results 1: Proof-of-Concept Testing for the UVA/Pavia MPC Algorithm" presents illustrative results for the methodology applied to the model predictive control (MPC) algorithm, developed at the University of Pavia, currently being tested in separate clinical trials at the Universities of Virginia, Padova (Italy), and Montpellier (France). "Results 2: Proof-of-Concept Testing for an Autoregressive, External Input (ARX)-Based MPC Algorithm" briefly presents another case study of the methodology in action.

Finally, the article is summarized briefly and conclusions, as well as directions for future research, are discussed.

Artificial Pancreas Control Algorithm: Meta-Algorithm

To simplify the specification of a general methodology for *in silico* proof of concept testing, it is convenient to first identify the core elements of an artificial pancreas control algorithm and their interactions in abstract, but suitably general, terms. In general, we use the term "control algorithm" to refer to the real-time feedback control law ("controller") that replaces conventional treatment in the GCRC *along with* all of the computational processes that relate to controller tuning based on patient-specific data and controller warm-up based on CGM and other data prior to closed-loop operation.

In general, the feedback control law, which, for simplicity, is denoted by $K(\psi)$, can be described as a dynamic system whose evolution in time may depend on a set of patient-specific control parameters, ψ. The controller's output, $u(t)$, at time t is an insulin infusion rate, which, in practice, will be updated at regular intervals depending on time of day or patient state. Inputs to the controller may include the following.

1. $y(t)$, the history of CGM samples received up to time t from the beginning of the closed-loop portion of the clinical trial

2. $\nu(t)$, the history of insulin and glucose interventions (meals and glucose tablets) up to time t (including conventional open-loop treatment prior to closed-loop operation)

3. $\delta(t)$, "meal signal," a signal from clinical staff that provides a 30-minute countdown to the next meal after time t, along with the size of the meal in grams CHO.

In most instances, the control law will be implemented as a sample data process; for example, CGM samples arrive and insulin pump rates are adjusted (or boluses are applied) at periodic intervals.

While the output of the controller at any time, $u(t)$, must be unambiguous for a given set of inputs $[y(t), \nu(t), \delta(t)]$, for implementation it may be convenient to describe operation of the controller in terms of its explicit dependence on a set of controller state variables, $\xi(t)$.

The "meta-algorithm" described in the following subsections has two main components, as illustrated in **Figure 1**. The first component, described next, relates to all computational and algorithmic processes run prior to execution of the control law. The second component, described under "Component 2: Controller Warm-up and Run-Time Operation," relates to controller warm-up and run-time operation. Decomposition into two components allows for researchers to define and test algorithms that are personalized for each subject involved in an artificial pancreas clinical study. Component 1 allows for the specification of a detailed preclinical assessment protocol that is designed to safely excite all relevant modes of the patient's physiology. (In practice, the preclinical protocol may take place in an outpatient ambulatory setting.) Component 1 also allows for the specification of a "tuning" procedure, which maps preclinical data about the patient into a set of patient-specific parameters for the run-time control algorithm.

Component 1: Patient Assessment and Tuning of Control Parameters

It is widely believed that, to achieve safe and effective closed-loop control of T1DM, the control algorithm for a given patient must be tuned to patient-specific characteristics affecting both the insulin responses computed by the algorithm and, for model-based control, the underlying glucose–insulin model of the patient. Consequently, many of the proposed artificial pancreas control algorithms involve a preliminary phase of data collection, perhaps during a separate "screening visit" for each patient enrolled in the trial. This phase of operation provides the opportunity to collect relevant biometric data (body weight, height, etc.) and to subject the patient to metabolic tests that allow for ex ante tuning of feedback control parameters and controller model development. Thinking of the algorithm itself as a "device," we seek to avoid ad hoc, human-user tuning of control parameters. Moreover, we consider the algorithmic processes that transform screening data into a "tuned" set of controller parameters for a patient to be a key part of the "control algorithm." Component 1 of the meta-algorithm comprises the specific metabolic tests, data collection protocol, and any computational procedures for processing patient assessment data and tuning controller parameters, ψ.

Assessment Protocol. Upon admission, essential biometric data are collected from the patient in a screening questionnaire. Data received from the questionnaire are denoted in the form of "screening data," η, including both "objective" parameters about the patient that can be assessed perfectly upon admission to the GCRC (such as height and weight, if used by the control algorithm) and other parameters that are derived from long-term

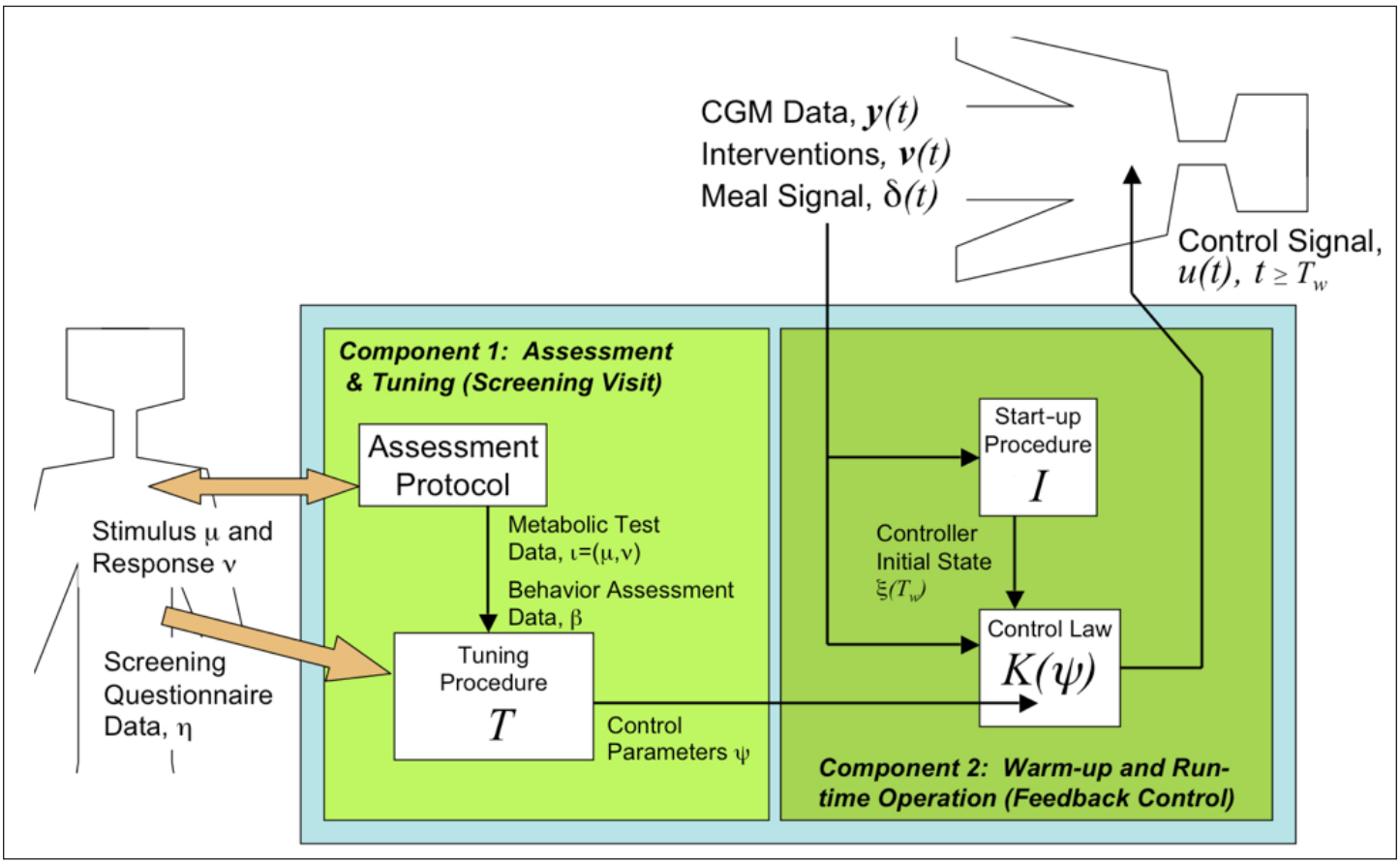

Figure 1. Artificial pancreas meta-algorithm.

observation of the patient under conventional treatment, such as average total daily insulin dose, if used by the control algorithm.

If the algorithmic procedure for tuning the feedback control algorithm requires more information about the patient, then the screening visit may also involve one or more metabolic tests (as would be specified in the clinical protocol). In this case, the patient would be admitted to the GCRC, instrumented with all specified measurement devices, and subjected to a metabolic stimulus, μ, that is designed to excite control-relevant modes of the patient's glucose–insulin system. The stimulus could, for example, take the form of an oral glucose tolerance test. The patient's response, ν, to the stimulus would be recorded over a specified period of time. The response comprises all blood glucose measurements (interstitial or intravenous) and any other clinically relevant transient responses to the stimulus (e.g., plasma insulin concentration, if used). For notational simplicity, we aggregate "metabolic test data" as $\iota = (\mu, \nu)$. Alternatively, metabolic test data, ι, can be collected during an outpatient phase of the protocol or through a combination of a screening visit and outpatient testing.

The assessment protocol may also involve the characterization of patient behaviors, including the patient's daily schedule of meals. We use the notation β to denote all such behavioral assessment data.

Tuning of Control Parameters. If required by the control methodology, patient assessment data are processed to allow for tuning of a patient-specific feedback controller. As an intermediate step, clinically significant characteristics of the patient may be calculated, including insulin sensitivity, clearance, or perhaps other specific parameters of a quantitative model of the patient's metabolic system. Finally, the computational procedure for tuning the control law is denoted by the operator T so that $\psi = T(\eta, \mu, \nu, \beta)$.

Remarks:

1. We assume that the structure of the feedback control law, $K(\psi)$, is fixed and that the process of tuning controller

parameters, ψ, based on η, μ, ν, and β is deterministic. In particular, the proposed meta-algorithm does not allow for "user-tunable" parameters. As an illustration, suppose that the feedback controller is a linear MPC based on an autoregressive model of glucose–insulin dynamics. In this case, the parameters of the model and the feedback gains must be uniquely determined from η, μ, ν, and β. The order of the autoregressive model (and controller) must either be fixed or be determined algorithmically from assessment data.

2. The closed-loop simulation software of Kovatchev and colleagues[11] is accepted by the FDA only for proof-of-concept testing. In particular, the software is not approved for use in designing/tuning control algorithms for specific human subjects.

Component 2: Controller Warm-up and Run-Time Operation

Component 2 of the meta-algorithm comprises the protocol for collecting and processing data for initialization of the feedback controller just prior to closing the feedback loop and run-time implementation of the feedback control law.

Run-Time Protocol. Upon admission for the closed-loop part of the clinical protocol, essential metabolic state data are collected from the patient as part of initializing the controller. First, the patient is instrumented with one or more glucose measurement devices and then open-loop treatment is administered in accordance with the clinical protocol. In the time prior to activation of the controller, prerun patient data are collected: (1) prerun glucose readings $\{\boldsymbol{y}(t)\}_{0\le t\le T_w}$ (i.e., open-loop CGM readings), where T_w is the length of time over which prerun data are collected, and (2) glucose and insulin interventions $\{\boldsymbol{v}(t)\}_{0\le t\le T_w}$ including conventional open-loop treatment applied prior to the controller being turned on. Prerun data are collected in the interval $[0,T_w]$ and closed-loop operation commences at time $t = T_w$.

Controller Initialization. Prerun data are used to initialize the internal states of the controller so that closed-loop operation may begin at time $t = T_w$. The algorithmic process by which controller states are initialized from these data is a critical part of the control algorithm overall. Controller initial states are computed as a function of prerun data: $\xi(T_w) = I(\{\boldsymbol{y}(t)\}_{0\le t\le T_w}, \{\boldsymbol{v}(t)\}_{0\le t\le T_w})$, where the operator I denotes the computational procedure for initializing the controller.

Run-Time Operation. Once the initial state of the controller is set, the control law, $K(\psi)$, can be set into motion.

Methods: *In Silico* Testing of Control Algorithms

This section proposes a simulation-based methodology for proof-of-concept testing of control algorithms that conform to the meta-algorithm described earlier. The goal is to define a test procedure that uses closed-loop simulation software to generate *in silico* trial data to assert that GCRC human subject clinical trials may proceed without a preliminary phase of animal testing. The test procedure serves to provide evidence that a specific control algorithm will perform well within a specific clinical protocol. The procedure itself involves three software inputs: (1) a set of MATLAB "controller setup" m-files that implement the tuning procedure T of component 1 of the control algorithm, in which the software package can be used to simulate the metabolic tests (if any) needed to tune the control law for individual patients; (2) MATLAB/Simulink code that encapsulates an implementation of component 2, specifically start-up procedure I and control law $K(\psi)$; and (3) a set of scenarios to be run that collectively represent the battery of tests to be performed.

This section, which is divided into four subsections, begins by identifying some key elements of the clinical protocol that should be specified for the test procedure to proceed. It then discusses sources of variability associated with the preparation of test subjects for GCRC clinical trials and proposes a corresponding set of "test instances" for each *in silico* subject involved in the study. Further, the third subsection, entitled "Generating Test Scenarios," outlines the procedure by which test cases should be generated, based on primary sources of uncertainty and on attributes of the proposed clinical protocol. The resulting set of closed-loop experiments generates a sufficient number of *in silico* trials to evaluate the performance characteristics of control algorithms and reveals the sensitivity to both patient initial state and sensor noise. Finally, the subsection entitled "Outcome Measures" lists outcome measures from the *in silico* study that should be reported for establishing proof of concept.

Two Key Features of Clinical Protocols

The "meta-algorithm" described earlier referred to two aspects of GCRC clinical protocols for artificial pancreas control algorithms: (1) the "assessment protocol," which describes clinical processes relating to patient-specific tuning of the control law, and (2) the "run-time protocol," which describes clinical processes relating to open-loop control of the patient prior to switching on the controller, CGM and blood glucose measurements taken

before and after controller initialization and during run-time operation, and meals. This subsection defines two features of the run-time protocol that should be specified as part of the *in silico* test procedure.

Meal Profile. Clinical studies relating to control algorithms for blood glucose regulation will typically specify that the patient will take one or more meals while subject to closed-loop control. The exact sequence of meals, including times at which they start and end, along with their carbohydrate content (grams CHO), is referred to as the "meal profile" of the study. From a control perspective, the meal profile serves as a disturbance that must be rejected by the control law implemented within the artificial pancreas algorithm.

Commutation versus Regulation Periods of Closed-Loop Control. Often the transient period after a controller is "switched on" is not representative of the long-term performance of the controller. Indeed, for the artificial pancreas, it will take some time for the controller to settle into a regular pattern of disturbance rejection, which could be a problem if the metrics used to assess the safety and performance of the algorithm are dominated by controller initial conditions (e.g., unusually low or high initial BG). We refer to the transient period after the controller is switched on as the "commutation period," and we refer to the subsequent period of time, which is representative of the long-term performance of the systems, as the "regulation period." In making use of this methodology, it is important to clearly specify commutation and regulation periods that are appropriate for their protocol and algorithm. The closed-loop simulation software is designed to assess outcome measures for any user-specified regulation period.

Sources of Variability in GCRC Trials

Even though general clinical research centers provide a stable environment for the implementation of clinical trials, control algorithm performance can still be impacted by a number of factors that should be treated as sources of uncertainty.

Variability in Patient Initial State in Component 2 (Warm-up and Run Time). Controller initialization is a key aspect of the artificial pancreas algorithm overall, as represented by start-up procedure I in component 2 of the meta-algorithm. The patient's metabolic state upon admission for the run-time phase of the protocol is perhaps the single most important source of variability in controller initialization. In practice, the patient's metabolic state when prerun data collection commences cannot be stipulated or known in advance as part of the protocol and must be treated as variable. We define three "admit states" in **Table 1**. The nominal case, "admit state 1," is defined by assuming that (1) the patient is initially steady at 100 mg/dl for any metabolic tests run during the screening visit and (2) the patient is initially steady at 100 mg/dl at the onset of collecting prerun data for controller initialization (just prior to switching on the controller). For "admit state 2" and "admit state 3," we set the state variables and inputs of the model so that the patient is admitted in steady state at 80 and 180 mg/dl, respectively, at the onset of prerun data collection.

Sensor Noise in Collecting Prerun Data and in Run-Time Operation. CGM sensor noise is an important factor in both controller initialization and run-time operation. The CGM noise model implemented within the closed-loop simulation software package serves to emulate the impact of sensor noise in *in silico* testing of closed-loop algorithms. Recall that if closed-loop operation begins at simulation $t = T_w$, then open-loop prerun data collection begins at simulation time $t = 0$, where T_w is the duration of time over which prerun data are collected. (Simulation time for component 2 begins at $t = 0$.)

Meal Profile Variability. To test robustness to meal profile variability, we subject all *in silico* patients to meals that deviate from the prescribed meal profile in terms of both timing and meal size. Since meal profiles will vary significantly for different studies, it is impossible to specify exactly what variations should be introduced. However, in developing an *in silico* test plan, it is important to explore the impact of both meal size variability (e.g., light and heavy CHO amounts) and, for algorithms that attempt to take into account information about meal timing, meal time variability (e.g., early and late meals). We suggest five scenarios per meal: (1) nominal meal, (2) early meal/nominal size, (3) late meal/nominal size, (4) light meal/nominal time, and (5) heavy meal/nominal

Table 1.
Patient Admit States for Controller Initialization

Admit state	Description
1	Patient steady at 100 mg/dl upon admission for component 2 in collecting prerun data for controller initialization
2	Patient steady at 80 mg/dl upon admission for component 2 in collecting prerun data for controller initialization
3	Patient steady at 180 mg/dl upon admission for component 2 in collecting prerun data for controller initialization

time, where "early," "late," "light," and "heavy" should be defined appropriately with respect to the objectives of the clinical protocol, but meal amounts should constitute at least 25% deviation from the nominal protocol values (**Table 2**).

Protocols with Multiple Meals. Clearly, the number of scenarios will grow quite rapidly for protocols that involve more than one meal under closed-loop control. In such cases it will be necessary to identify the most relevant combinations of admit state and meal variation for robustness testing. These choices will, in general, be protocol specific.

Generating Test Scenarios

Based on the sources of variability listed under "Sources of Variability in GCRC Trials," it is possible to generate a set of test scenarios that explore the impact of patient admit states, CGM noise, and meal profile variability for the population(s) being investigated, adults with T1DM, adolescents, and/or children.

Illustration Protocol. The appropriate test scenarios for *in silico* proof-of-concept testing are specific to both the control algorithm and the proposed clinical protocol. As an illustration, consider the following clinical protocol, whose run-time phase is designed to test the ability of a controller to regulate blood glucose for a 24-hour period. Suppose that the protocol excludes both adolescents and children. Suppose further that the protocol stipulates the following:

1. At approximately 18:00 Day 1, while undergoing conventional open-loop treatment, the subject takes a dinner meal containing 85 grams CHO, with insulin injections and CGM measurements fed into start-up procedure I of component 2 of the control algorithm
2. At approximately 21:00 Day 1, the basal rate function of the subject's pump is set to zero
3. From the time that the basal rate function is set to zero (21:00 Day 1) until 21:00 Day 2, the subject receives a 1-minute bolus every 15 minutes according to calculations made by the control law, $K(\psi)$
4. At approximately 7:30 Day 2, the subject takes a breakfast meal containing 50 grams CHO
5. At approximately noon Day 2, the subject takes a lunch meal containing 65 grams CHO
6. At approximately 18:00 Day 2, the subject takes a dinner meal containing 85 grams CHO
7. At 21:00 Day 2, the control algorithm is turned off, and the subject's normal basal pump rate is resumed.

Table 2.
Meal Variability

Meal variant	Description
1	Nominal meal (exactly as specified in the clinical protocol)
2	Early meal/nominal size [meal arrives earlier than expected but size of the meal (in grams CHO) is as specified in the protocol]
3	Late meal/nominal size
4	Light meal/nominal time
5	Heavy meal/nominal time

We suggest that an appropriate set of *in silico* test scenarios for this hypothetical protocol would be as shown in **Table 3**. Note that this test suite accounts for (1) variability in patient initial conditions at the beginning of closed-loop operation and (2) meal profile variability for the dinner meal of Day 2 (the largest meal experienced under closed-loop control).

In setting up the simulation, we would initialize simulation time $t = 0$ (minutes) to correspond to 18:00 Day 1 so that closed-loop operation would begin at simulation time $t = T_w = 180$ (minutes), corresponding to 21:00 Day 1. In assessing outcome metrics, we would specify that the commutation period begins at 21:00 Day 1 (just after closed-loop operation begins) and end at 00:00 Day 2 at which time we expect controller initialization transients to have decayed. The regulation period, where we are interested primarily in evaluating controller performance, would begin for this illustration protocol at 00:00 Day 2 and end at completion of the closed-loop portion of the study at 21:00 Day 2.

Outcome Measures

Various metrics are generated from the simulation-based procedure described earlier. First, to facilitate regulatory approval, we recommend that nominal-case BG traces be presented for each *in silico* test subject. (For the illustration protocol given earlier, BG plots for test scenario 1 would be presented for all 100 adults.) In addition, CVGA plots should be presented for each test scenario, reflecting the aggregate performance of the control algorithm for the target population. (For the illustration protocol, we would present the CVGA plot for test scenario 1 illustrating the nominal performance of the control algorithm for all 100 adult CVGA plots separately for the remaining test scenarios: 2–15.) Finally, summary statistics should be reported for each *in silico* population group. **Table 4** suggests some relevant metrics based on Kovatchev and

Table 3.
Test Scenarios for Illustration Protocol

Test scenario	Description
1	Admit state 1 (BG = 100 mg/dl at 18:00 Day 1), breakfast meal variant 1 (50 g CHO at 7:30 Day 2), lunch meal variant 1 (65 g CHO at noon Day 2), dinner meal variant 1 (85 g CHO at 18:00 Day 2)
2	Admit state 1, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 2 (85 g CHO at 17:30 Day 2)
3	Admit state 1, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 3 (85 g CHO at 18:30 Day 2)
4	Admit state 1, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 4 (63.75 g CHO at 18:00 Day 2)
5	Admit state 1, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 5 (106.25 g CHO at 18:00 Day 2)
6	Admit State 2 (BG = 80 mg/dl at 18:00 Day 1), breakfast meal variant 1, lunch meal variant 1, dinner meal variant 1
7	Admit state 2, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 2
8	Admit state 2, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 3
9	Admit state 2, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 4
10	Admit state 2, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 5
11	Admit State 3 (BG = 180 mg/dl at 18:00 Day 1), breakfast meal variant 1, lunch meal variant 1, dinner meal variant 1
12	Admit state 3, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 2
13	Admit state 3, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 3
14	Admit state 3, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 4
15	Admit state 3, breakfast meal variant 1, lunch meal variant 1, dinner meal variant 5

associates.[23] Report forms should be presented separately for the nominal case and for all of the deviation cases combined. [For the illustration protocol, for nominal case data, mean, standard deviation (SD), and median of each outcome measure for adults would be computed from all data collected in test scenario 1. For the deviation cases, mean, SD, and median of each outcome measure would be computed from all data collected in test scenarios 2–15.] When normative case data are available, it may be reasonable to compare summary statistics for each population group based on combined nominal and deviation case data.

Establishing Proof-of-Concept from in Silico Test Data

While there are no specific thresholds to establish control proof of concept, it is necessary to argue on the basis of *in silico* test data generated earlier that an algorithm would perform comparably to conventional treatment. Roughly speaking, an algorithm will "pass" if the primary outcome measures (% time within target range of 70–180 mg/dl and the LBGI) are within reasonable target limits relative to normative data. The recently published JDRF randomized clinical trial of CGM in children, adolescents, and adults provides normative CGM data on glucose control in well-motivated patients who were doing an average of more than six capillary blood glucose tests each day.[24] These data are presented in **Table 5** and

Table 4.
Sample Outcome Measures Report Form (Primary Outcome Measures Shown in Bold)

	Population group		
	(e.g., adults, adolescents, children)		
Measure	Mean	SD	Median
Mean BG			
Mean premeal BG			
Mean postmeal BG			
% time <50			
% time <70			
% time in [70-180]			
% time >180			
% time >300			
% time in [70–145](fasting)			
Postprandial area under curve/g CHO			
LBGI			
High blood glucose index			
BG rate of increase			
SD of BG rate of change			

could serve as one measure of contemporary normative data with conventional treatment based on glucose values derived from continuous glucose monitoring.

Results 1: Proof-of-Concept Testing for the UVA/Pavia MPC Algorithm

The methodology described previously was used to generate *in silico* test data for the linear MPC algorithm currently being tested in separate human subject clinical trials at the Universities of Virginia, Padova, and Montpellier. FDA IDE approval for the trial at the UVA was granted in April 2008. The MPC algorithm, developed at the University of Pavia, is based on analytical derivations[6] and involves only one tuning parameter, namely, q, the weight placed on quadratic state deviation away from the nominal operating point of the patient relative to insulin utilization. Only screening questionnaire data, η, are used to compute the appropriate q value for the patient via a simple nonlinear regression. Patient behavioral data, β, specifically meal size and meal timing data, are used by the controller to obtain an anticipatory insulin effect prior to the meal.

The clinical protocol is similar to the illustrative protocol discussed earlier in the section entitled, "Methods: *In Silico* Testing of Control Algorithms" but is only designed to test the ability of a controller to regulate blood glucose overnight and to compensate for breakfast in the morning. In particular, the protocol stipulates the following.

1. At approximately 18:00 Day 1, while undergoing conventional open-loop treatment, the subject takes a dinner meal containing 85 grams CHO, with insulin infusion rates and CGM measurements fed into start-up procedure I of component 2 of the control algorithm
2. At approximately 21:00 Day 1, the basal rate function of the subject's pump is set to zero
3. From the time that the basal rate function is set to zero (21:00 Day 1) until 11:00 Day 2, the subject receives a 1-minute bolus every 15 minutes according to calculations made by the control law, $K(\psi)$
4. At approximately 7:30 Day 2, the subject takes a preplanned breakfast meal containing 50 grams CHO
5. At 11:00 Day 2, the control algorithm is turned off, and the subject's normal basal pump rate is resumed.

The protocol excludes both adolescents and children.

Based on the outline of the protocol just given, it is possible to define a suite of 15 test scenarios that are analogous to test scenarios 1–15 for the illustrative protocol discussed under "Generating Test Scenarios." Whereas test scenarios for the illustrative protocol focused on variability associated with the dinner meal of Day 2, test scenarios for the UVA protocol focused instead on variability associated with the breakfast meal of Day 2, i.e., the only meal that is meant to be covered under closed-loop control. A subset of *in silico* test results is presented.

Table 5.
Benchmark Performance for Conventional Treatment[24]

Mean	Adults	Adolescents	Children
% time <50	2	3	1
% time <70	6	7	3
% time in 70–180	59	48	45
% time >180	35	45	52
% time >250	10	19	24
Mean mg/dl/min	0.73	0.85	0.84

Table 6.
Results from *in Silico* Testing: Nominal Scenario (Primary Outcome Measures Shown in Bold)

	Mean	SD	Median
Mean BG	121.87	9.73	121.45
Mean premeal BG	106.04	10.33	107.39
Mean postmeal BG	154.59	24.79	155.38
% time <50	0	0	0
% time <70	0.06	0.58	0
% time in [70–180]	**97.3**	**5.55**	**100**
% time >180	2.65	5.55	0
% time >300	0	0	0
Postprandial area under curve/g CHO	0.51	0.07	0.512
LBGI	**0.32**	**0.52**	**0.12**
High blood glucose index	1.08	0.77	1.03
BG rate of increase	1.4	0.94	1.21
SD of BG rate of change	0.53	0.21	0.52

Testing the Nominal Scenario

Table 6 presents summary statistics from the testing of $N = 100$ simulated adults.

Because of the tightly controlled conditions, including prior knowledge of a breakfast meal, the *in silico*

performance of the control algorithm suggests that the control algorithm is effective and safe for the nominal scenario. The primary effectiveness parameter, time within the target range of 70–180 mg/dl, exceeds 97%, which compares favorably to the benchmark figure in **Table 5**. The primary safety parameter, the low BG index indicating the exposure of subjects to hypoglycemia, is nearly zero. In addition, there are no readings above 300 mg/dl and only 0.06% of the readings are below 70 mg/dl.

Figure 2 presents the CVGA plot from testing of the adult *in silico* population. Consistent with the numerical results of **Table 4**, 99% of all glucose excursions (in terms of their 95% confidence intervals) are in the accurate A zone or are in benign error B zones. None of the *in silico* patients land within the extreme error D or E zones. Not shown here are the glucose traces for all of the *in silico* patients. Review of the traces indicates strict overnight control, tight control of the breakfast meal, and no hypoglycemic episodes.

Testing against Meal Size Variability

In the nominal scenario the subject eats a 50-gram CHO breakfast meal at 7:30 Day 2, and the linear MPC algorithm makes use of this information as behavioral data, β, in producing an anticipatory insulin effect. To investigate the sensitivity of the control algorithm to the size of the meal, we tested meals with a carbohydrate content of 25% above and 25% below the expected meal amount for all 100 *in silico* adult subjects. **Table 7** presents summary statistics from testing $N = 100$ *in silico* adult subjects, with the breakfast meal of Day 2 having 25% more and 25% fewer grams CHO than specified in the nominal scenario.

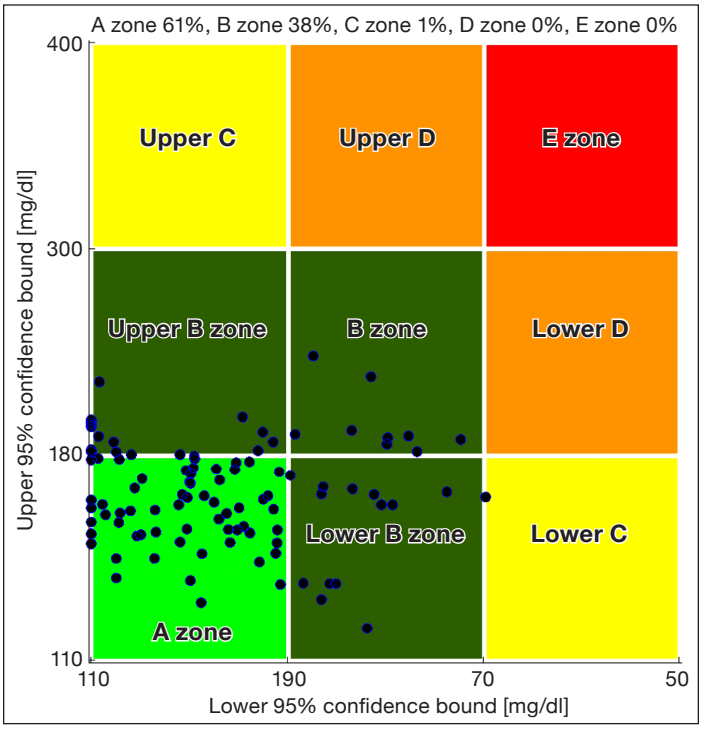

Figure 2. Nominal-case CVGA of *in silico* adults.

All results given earlier are computed for the duration of an *in silico* protocol emulating the conditions of the clinical trial. Even with variations in meal size, the *in silico* performance of the linear MPC algorithm is still quite good. For the case of 25% more CHO, the primary effectiveness parameter, the time within the target range of 70–180 mg/dl is greater than 94% and only 0.06% of the readings are below 70 mg/dl. For the case of 25% less CHO, the time within the target range of 70–180 mg/dl exceeds 99% and again only 0.06% of the readings are below 70 mg/dl. In both cases there are no readings above 300 mg/dl.

Proof-of-Concept Assessment of the Control Algorithm

From the statistics and the figures presented in the subsections "Testing the Nominal Scenario" and "Testing against Meal Size Variability," along with *in silico* results for all of the remaining test scenarios (not shown), it appears that the linear MPC algorithm is stable overnight and performs well for the nominal scenario with prescribed meal timing and meal amounts. We point out that while the *in silico* results assure stability overnight, it is difficult to assess overnight controller performance, as the closed loop simulation software of Kovatchev and colleagues[11] does not at this time take into account diurnal variation in insulin sensitivity. The results given in "Testing against Meal Size Variability," along with the full suite of *in silico* results for the remaining test scenarios (not shown), suggest that the controller is robust to variations in meal size, meal timing, and patient initial conditions. Overall the linear MPC algorithm appears to be stable with respect to intersubject variation, even with just one patient-specific tuning parameter, q.

Results 2: Proof-of-Concept Testing for an Autoregressive, External Input (ARX)-Based MPC Algorithm

This section briefly presents another case study of the methodology for the *in silico* testing of artificial pancreas algorithms. Specifically, we illustrate a proof of concept in which two variants of MPC (with and without an insulin on board dynamic constraint) are evaluated on a clinical

protocol using the simulator environment **(Figure 3)**. This controller was designed based on an ARX model that was identified on a 4-day scenario following a preclinical protocol consisting of three meals at 8 a.m., noon, and 6 p.m. with 20, 40, and 70 grams of carbohydrates with matching boluses.[25]

Table 7.
Results from *in Silico* Testing: Breakfast Meal with 25% More and 25% Less CHO Than Expected Meal Size from Nominal Scenario (Primary Outcome Measures Shown in Bold)

	25% More CHO (62.5 g)			25% Less CHO (37.5 g)		
	Mean	SD	Median	Mean	SD	Median
Mean BG	124.77	10.57	124.46	118.93	8.94	118.61
Mean premeal BG	106.04	10.33	107.39	106.04	10.33	107.39
Mean postmeal BG	168.24	29.88	167.68	140.93	19.72	141.58
% time <50	0.00	0.00	0.00	0.00	0.00	0.00
% time <70	0.06	0.58	0.00	0.06	0.58	0.00
% time in [70–180]	**94.34**	**7.96**	**100.00**	**99.29**	**2.52**	**100.00**
% time >180	5.60	7.98	0.00	0.65	2.47	0.00
% time >300	0.00	0.00	0.00	0.00	0.00	0.00
Postprandial area under curve/g CHO	0.44	0.07	0.44	0.64	0.08	0.64
LBGI	**0.31**	**0.52**	**0.12**	**0.32**	**0.53**	**0.12**
High blood glucose index	1.49	1.01	1.40	0.73	0.55	0.64
BG rate of increase	1.80	1.15	1.58	1.06	0.77	0.81
SD of BG rate of change	0.63	0.25	0.62	0.43	0.17	0.41

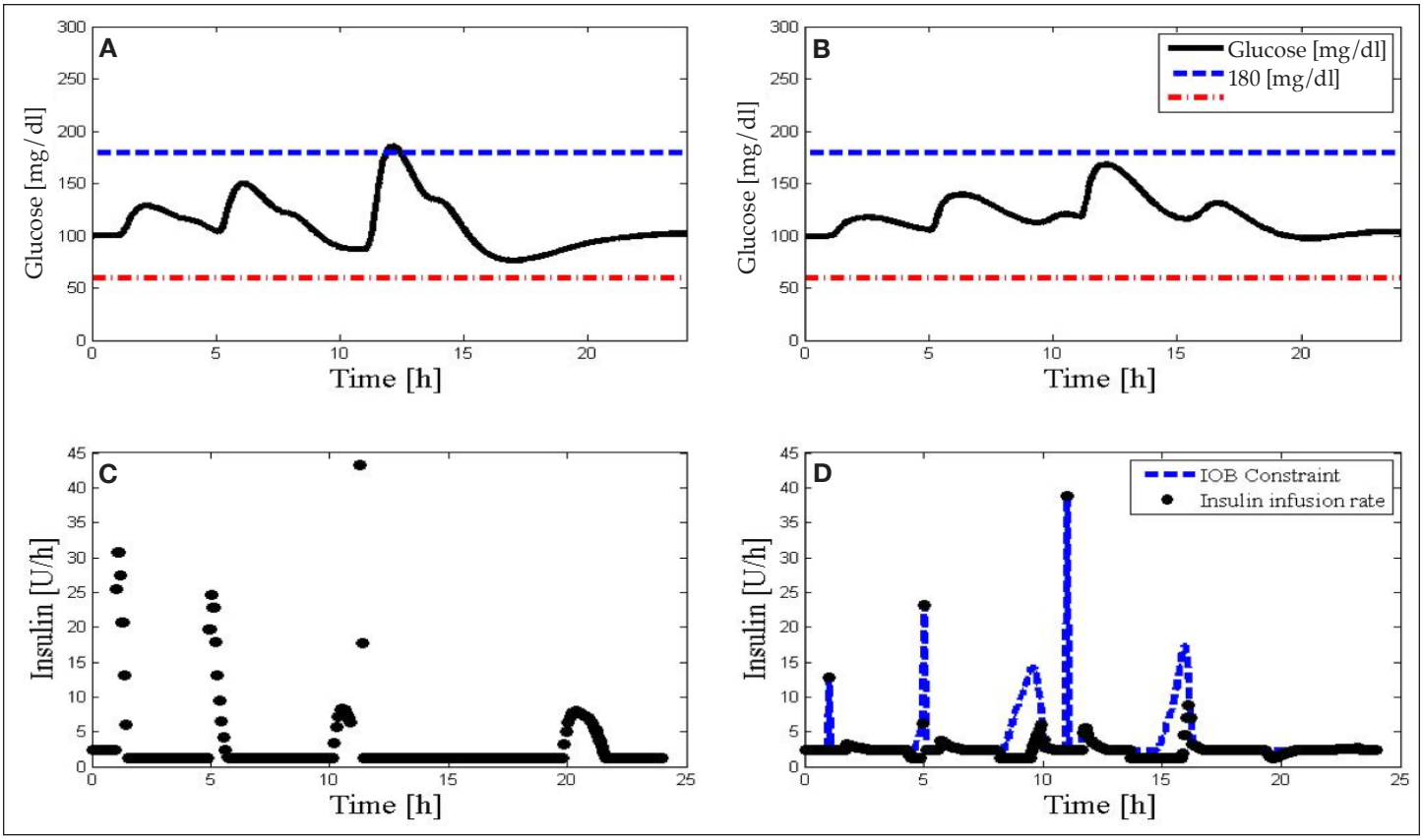

Figure 3. ARX-based MPC on subject 9 without and with the insulin on board (IOB) constraint. (Based on table entry from Ellingsen and colleagues.[25]) The 24-hour scenario starts at 7 a.m. at steady state followed by a protocol of three meals at 8 a.m., noon, and 6 p.m. with 20, 40, and 70 grams of carbohydrates, respectively. Glucose trajectories without (**A**) and with (**B**) the IOB constraint are presented. Controller moves without (**C**) and with (**D**) the IOB constraint are shown. Dashed lines represent values of hyperglycemia and hypoglycemia. The controller that incorporated the IOB constraint shows more conservative behavior (**D**) than the controller that was missing the IOB constraint (**C**).

Discussion

We have proposed a methodology for *in silico* testing of artificial pancreas algorithms based on the closed-loop simulation software package described in Kovatchev and colleagues.[11] The software is designed to take a mathematical description of control algorithms that conform to the constraints of the "meta-algorithm" discussed under "Artificial Pancreas Control Algorithm: Meta-Algorithm" and reveal the performance characteristic of that algorithm as applied to a population of *in silico* patients with T1DM. The test procedure described serves to establish the "proof-of-concept" for control algorithms, a key step toward FDA IDE approval, without requiring animal trials. As discussed in under "Results 1: Proof-of-Concept Testing for the UVA/Pavia MPC Algorithm," researchers at the University of Virginia used this methodology to validate an MPC algorithm for an overnight closed-loop protocol, and FDA IDE approval was granted within 4 months after FDA acceptance of the closed-loop simulation software. With clinical trials underway, we will have the opportunity to further validate and refine the *in silico* population based on controller performance *in vivo*.

It is worth noting that while the software is equipped with a model for CGM noise, it is not equipped to simulate fault modes associated with the physical hardware and software *implementation* of control algorithms, and thus the methodology proposed here is not designed to validate the hardware/software implementation of an artificial pancreas control algorithm. Safety in the face of human errors and lack of compliance has to be assured through a complete systems engineering process: hazard analysis, requirements, specifications, and testing. The methodology of this article addresses just a part of the testing procedure, specifically the performance of the system in the presence of systematic uncertainty (CGM noise, initial states, etc.) given that all hardware, software, and human interactions are reliable. Future research efforts may seek to develop enhanced simulation capabilities that include new features, such as the ability to represent other important factors, such as exercise, infection, and diurnal variation in insulin sensitivity, all of which could affect the performance and safety of an algorithm *in vivo*. New simulation tools would have many uses beyond proof-of-concept testing, including hardware/software system validation (as suggested in Dassau and associates[17]) and training.

Acknowledgements:

The authors gratefully acknowledge financial support from the Juvenile Diabetes Research Foundation Artificial Pancreas Project. We thank Dr. Boris Kovatchev for many helpful suggestions in preparing this manuscript.

References:

1. Hovorka R. Continuous glucose monitoring and closed-loop systems. Diabet Med. 2006;23(1):1-12.
2. Steil GM, Rebrin K, Darwin C, Hariri F, Saad MF. Feasibility of automating insulin delivery for the treatment of type 1 diabetes. Diabetes. 2006;55(12):3344-50.
3. El-Khatib FH, Jiang J, Damiano ER. Adaptive closed-loop control provides blood glucose regulation using dual subcutaneous insulin and glucagon infusion in diabetic swine. J Diabetes Sci Technol. 2007;1(2):181-92.
4. Magni L, Raimondo DM, Bossi L, Dalla Man C, De Nicolao G, Kovatchev BP, Cobelli C. Model predictive control of type 1 diabetes: an *in silico* trial. J Diabetes Sci Technol. 2007;1(6):804-12.
5. Patek SD, Breton MD, Chen Y, Solomon C, Kovatchev B. Linear quadratic gaussian-based closed-loop control of type 1 diabetes. J Diabetes Sci Technol. 2007;1(6):834-41.
6. Magni L, Raimondo DM, Dalla Man C, De Nicolao G, Kovatchev B, Cobelli C. Model predictive control of glucose concentration in subjects with type 1 diabetes: an *in silico* trial. 17th IFAC World Congress; 2008. p. 4246-51.
7. Parker RS, Doyle III FJ, Peppas NA. A model-based algorithm for blood glucose control in type 1 diabetic patients. IEEE Trans Biomed Eng. 1999;46:148-57.
8. 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.
9. Renard E. Implantable closed-loop glucose-sensing and insulin delivery: the future for insulin pump therapy. Curr Opin Pharmacol. 2002;2(6):708-16.
10. Kovatchev BP, Breton M, Dalla Man C, Cobelli C. *In silico* preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. J Diabetes Sci Technol. 2009;3(1):44-55.
11. 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.
12. Lin J, Lee D, Chase JG, Shaw GM, Le Compte A, Lotz T, Wong J, Lonergan T, Hann CE. Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care. Comput Methods Programs Biomed. 2008;89(2):141-52.
13. Lonergan T, LeCompte A, Willacy M, Chase JG, Shaw G, Hann CE, Lotz CT, Lin J, Wong X. A pilot study of the SPRINT protocol for tight glycaemic control in critically ill patients. Diabetes Technol Ther. 2006;84(4):449-62.
14. Chase JG, LeCompte A, Shaw GM, Blakemore A, Wong J, Lin J, Hann CE. A benchmark data set for model-based glycemic control in critical care. J Diabetes Sci Technol. 2008;2(4):584-94.
15. Hovorka R, Chassin LJ, Ellmerer M, Plank J, Wilinska ME. A simulation model of glucose regulation in the critically ill. Physiol Meas. 2008;29(8):959-78.
16. Dassau E, Zisser H, Palerm CC, Buckingham BA, Jovanovic L, Doyle III FJ. Modular artificial β-cell system: a prototype for clinical research. J Diabetes Sci Technol. 2008;2(5):863-72.

17. Dassau E, Palerm CC, Zisser H, Buckingham BA, Jovanovic L, Doyle III FJ. *In silico* evaluation platform for artificial pancreatic β-cell development: a dynamic simulator for closed-loop control with hardware-in-the-loop. Diabetes Technol Ther. 2009 Feb 3.

18. Dalla Man C, Rizza RA, Corbelli C. Meal simulation model of the glucose-insulin system. IEEE Trans Biomed Eng. 2007;54:1740-9.

19. Dalla Man C, Raimondo DM, Rizza RA, Cobelli C. GIM, simulation software of meal glucose: insulin model. J Diabetes Sci Technol. 2007;1(3):323-30.

20. Breton M, Kovatchev BP. Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. J Diabetes Sci Technol. 2008;2(5):1-10.

21. Kovatchev BP, Cox DJ, Gonder-Frederick LA, Clarke W. Symmetrization of the blood glucose measurement scale and its applications. Diabetes Care. 1997;20(11):1655-8.

22. Magni L, Raimondo DM, Dalla Man C, Breton M, Patek S, De Nicolao G, Cobelli C, Kovatchev BP. Evaluating the efficacy of closed-loop glucose regulation via control variability grid analysis. J Diabetes Sci Technol. 2008;2(4):630-5.

23. Kovatchev BP, Clarke WL, Breton M, Brayman K, McCall A. Quantifying temporal glucose variability in diabetes via continuous glucose monitoring: mathematical methods and clinical application. Diabetes Technol Ther. 2005;7(6):849-62.

24. JDRF Continuous Glucose Monitoring Study Group. Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med. 2008;359(14):1464-76.

25. Ellingsen C, Dassau E, Zisser H, Grosman B, Percival MW, Jovanovič L, Doyle III, FJ. Safety constraints in an artificial β cell: an implementation of model predictive control with insulin on board. J Diabetes Sci Technol DTM. A50;2008.