One-Day Bayesian Cloning of Type 1 Diabetes Subjects: Toward a Single-Day UVA/Padova Type 1 Diabetes Simulator

Roberto Visentin, Chiara Dalla Man, and Claudio Cobelli*, *Fellow, IEEE*

***Abstract*—*Objective*: The UVA/Padova Type 1 Diabetes (T1DM) Simulator has been shown to be representative of a T1DM population observed in a clinical trial, but has not yet been identified on T1DM data. Moreover, the current version of the simulator is "single meal" while making it "single-day centric," i.e., by describing intraday variability, would be a step forward to create more realistic *in silico* scenarios. Here, we propose a Bayesian method for the identification of the model from plasma glucose and insulin concentrations only, by exploiting the prior model parameter distribution. *Methods*: The database consists of 47 T1DM subjects, who received dinner, breakfast, and lunch (respectively, 80, 50, and 60 CHO grams) in three 23-h occasions (one open- and one closed-loop). The model is identified using the Bayesian Maximum *a Posteriori* technique, where the prior parameter distribution is that of the simulator. Diurnal variability of glucose absorption and insulin sensitivity is allowed. *Results*: The model well describes glucose traces (coefficient of determination $R^2 = 0.962 \pm 0.027$) and the posterior parameter distribution is similar to that included in the simulator. Absorption parameters at breakfast are significantly different from those at lunch and dinner, reflecting more rapid dynamics of glucose absorption. Insulin sensitivity varies in each individual but without a specific pattern. *Conclusion*: The incorporation of glucose absorption and insulin sensitivity diurnal variability into the simulator makes it more realistic. *Significance*: The proposed method, applied to the increasing number of long-term artificial pancreas studies, will allow to describe week/month variability, thus further refining the simulator.**

***Index Terms*—Artificial pancreas, circadian variability, closed-loop control, compartmental modeling, *in silico*.**

I. Introduction

SIMULATION models of the glucose–insulin system are extremely useful for studying type 1 diabetes (T1DM) treatments and in particular for the design, testing, and validation of closed-loop control algorithms for artificial pancreas. In fact, it is widely accepted that preclinical testing using computer simulation accelerates the development of control algorithms, since a number of different experimental scenarios can be easily investigated, allowing relevant time- and cost saving, and avoiding the need of animal testing [1].

A T1DM computer simulator is made up of a mathematical model describing glucose dynamics, a set of *in silico* subjects, i.e., model parameter vectors, summarizing the intersubject variability of the glucose response to a given external perturbation (usually a meal), and an interface that allows the user to set up simulation scenarios, running the simulations, displaying and saving the results. Several simulation tools have been developed (see [1] and [2]–[5] for a review), each one based on a comprehensive mathematical model and equipped with an *in silico* population. In particular, the University of Virginia (UVA) and Padova T1DM Simulator has been accepted by FDA as a substitute for preclinical trials of certain insulin treatments, including closed-loop algorithms [2], [5].

A clinical validation of the UVA/Padova T1DM simulator has been reported [6], where it has been shown that the virtual adult population of the simulator is representative of a T1DM population observed in a clinical trial by comparing 96 measured postprandial glucose profiles with those simulated in 100 *in silico* subjects using well-accepted outcome metrics of diabetes control [6]. Recently, two richer datasets in T1DM have become available, one where the diurnal pattern of insulin sensitivity in 19 T1DM subjects was investigated with the triple tracer meal method with frequent plasma glucose and insulin concentration measurements [7], and one where 47 T1DM subjects in six clinical centers underwent three-randomized 23-h admissions (one open-loop and two closed-loop) with frequent plasma glucose and insulin concentration measurements [8]. With these informative glucose and insulin datasets, we can now attempt to fit, for the first time, the simulator to a T1DM subject. Since in [7], three identical meals (50 g of carbohydrate) were given in order to eliminate confounding effects on the diurnal pattern of insulin sensitivity in the 19 T1DM subjects, we prefer to focus here on the dataset of [8] where the three meals resemble real life breakfast, lunch, and dinner and a much larger number, i.e., 141, of daily traces of glucose and insulin concentrations is available (see Fig. 1).

To do that, Bayesian parameter estimation techniques are very helpful. Different strategies can be adopted. For instance, recently, Haidar *et al.* proposed a Bayesian method, called stochastic e-cloning [9], based on Markov Chain Monte Carlo and regularization, and used it for the *in silico* cloning of 12 young T1DM subjects from plasma glucose and insulin concentration data. Model structure was that incorporated in the T1DM simulator proposed by the same group [3]. However,

Manuscript received July 28, 2015; revised December 3, 2015; accepted February 18, 2016. Date of publication February 26, 2016; date of current version October 18, 2016. This work was supported in part by the Italian Ministero dell'Università e della Ricerca Scientifica (Progetto FIRB 2008) and in part by the FP7-EU Project AP@Home. *Asterisk indicates corresponding author.*

R. Visentin and C. Dalla Man are with the Department of Information Engineering, University of Padova.

*C. Cobelli is with the Department of Information Engineering, University of Padova, Padova 35131, Italy (e-mail: cobelli@dei.unipd.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2016.2535241

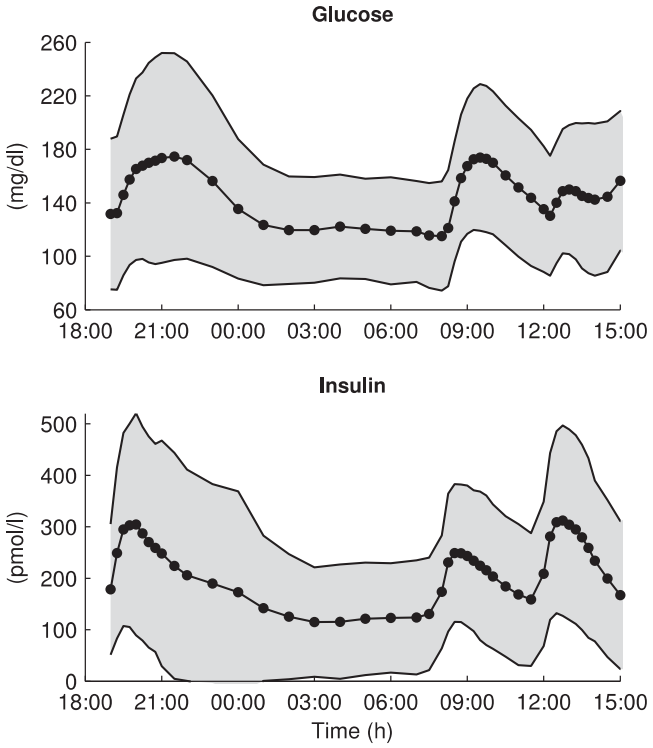

Fig. 1. Mean ± standard deviation of plasma glucose (*upper panel*) and insulin (*lower panel*) concentrations in T1DM subjects ($N = 141$).

new multiplicative and additive fluxes, not included in the original model, had to be introduced to obtain a good fit of glucose and insulin data. The method was then tested on a T1DM population of 12 subjects, and the computational time required to get a virtual clone is quite long (about 5 days with a standard PC) likely due to the use of Markov Chain Monte Carlo, made necessary by the lack of *a priori* information available on model parameter distribution.

The aim of this study is thus threefold: 1) propose a method to identify the model included into the latest version of the UVA/Padova T1DM simulator [5] from the 141 daily profiles of T1DM plasma glucose and insulin concentrations measured in [8]; 2) move from a single meal to a breakfast/lunch/dinner meal data scenario, thus accommodating intrasubject variability of glucose absorption and insulin sensitivity; and 3) compare the intersubject parameter variability of the *in silico* population with that estimated from the data, thus possibly improving the *in silico* population included into the simulator.

To achieve the first objective, at variance with [9], we resort to a Maximum *a Posteriori* Bayesian approach, which exploits both the information provided by the experimental data and the *a priori* knowledge on model parameters represented by the joint parameter distribution incorporated in the T1DM simulator. Plasma insulin concentration will be used as model forcing function (i.e., assumed to be known without error) and the identification of the UVA/Padova simulator on a specific T1DM subject will provide an *in silico* clone, compatible with the simulator. Therefore, the possibility to clone a large number of T1DM individuals will allow us to improve the *in silico* population included into the simulator in two respects: we will move from a single meal to a breakfast/lunch/dinner meal scenario accommodating intrasubject variability (aim 2), and will better describe intersubject variability (aim 3).

II. Database and Protocol

Forty-seven T1DM subjects (Age = 42.0 ± 10.1 years, BW = 77.5 ± 13.4 kg, BMI = 24.4 ± 0.1 kg/m^2) were recruited in six clinical centers [Academic Medical Center Amsterdam, NL ($N = 7$); CHRU Montpellier, FR ($N = 8$); Medical University Graz, AT ($N = 8$); Profil Institute for Metabolic Research GmbH, GER ($N = 8$); University of Cambridge, U.K. ($N = 8$); University of Padova, IT ($N = 8$)], within the AP@home FP7-EU project, in which subjects underwent three randomized 23-h admissions: one open-loop and two closed- loop sessions. During the open-loop visit, subjects had their usual insulin therapy through an insulin pump, while the insulin infusions were managed by a control algorithm during the closed-loop admissions. For each admission, subjects received dinner D (19:00, day1), breakfast B (08:00, day2), and lunch L (12:00, day2), respectively, containing 80, 50, and 60 g of carbohydrates (CHO), and did a moderate physical activity session (15:00, day2). Throughout the admissions, venous blood samples were collected for measurements of plasma glucose and insulin concentrations every 15 min in the first 2 h after each meal, every 1 h at night and every 30 min elsewhere (average measures of plasma glucose and insulin are reported in Fig. 1). Plasma glucose was measured using YSI 2300 STAT Plus Analyzer (YSI incorporated, Yellow Springs, OH, USA) and plasma insulin was measured using an insulin chemiluminescence assay (Invitron Ltd., Monmouth, U.K.) (The Institute of Life Sciences, Swansea University, S. Luzio). For a detailed description of the clinical protocol, we refer to [8].

III. Model

The model included into the T1DM simulator is shown in Fig. 2. Briefly, this model consists of glucose, insulin, and glucagon subsystems, and puts in relation the measured plasma concentrations, i.e., glucose *G*, insulin *I*, and glucagon *H*, with the glucose fluxes (meal rate of appearance Ra_{meal}, endogenous production *EGP*, utilization *U*, renal extraction *E*), the insulin fluxes (rate of appearance of subcutaneous insulin Ra_I, degradation D_I), and the glucagon fluxes (secretion SR_H, degradation D_H). The model consists of 18 differential equations and 39 parameters. For sake of clarity, we report model equations in the Appendix section, but we refer to [5] for a more detailed description.

IV. Bayesian Model Identification

A. Fundamentals

Given the complexity of the model, the sole availability of plasma glucose and insulin data makes impossible to reliably identify the model by using nonlinear least squares or maximum likelihood identification techniques. In fact, one can obtain a good description of plasma glucose and insulin data with many

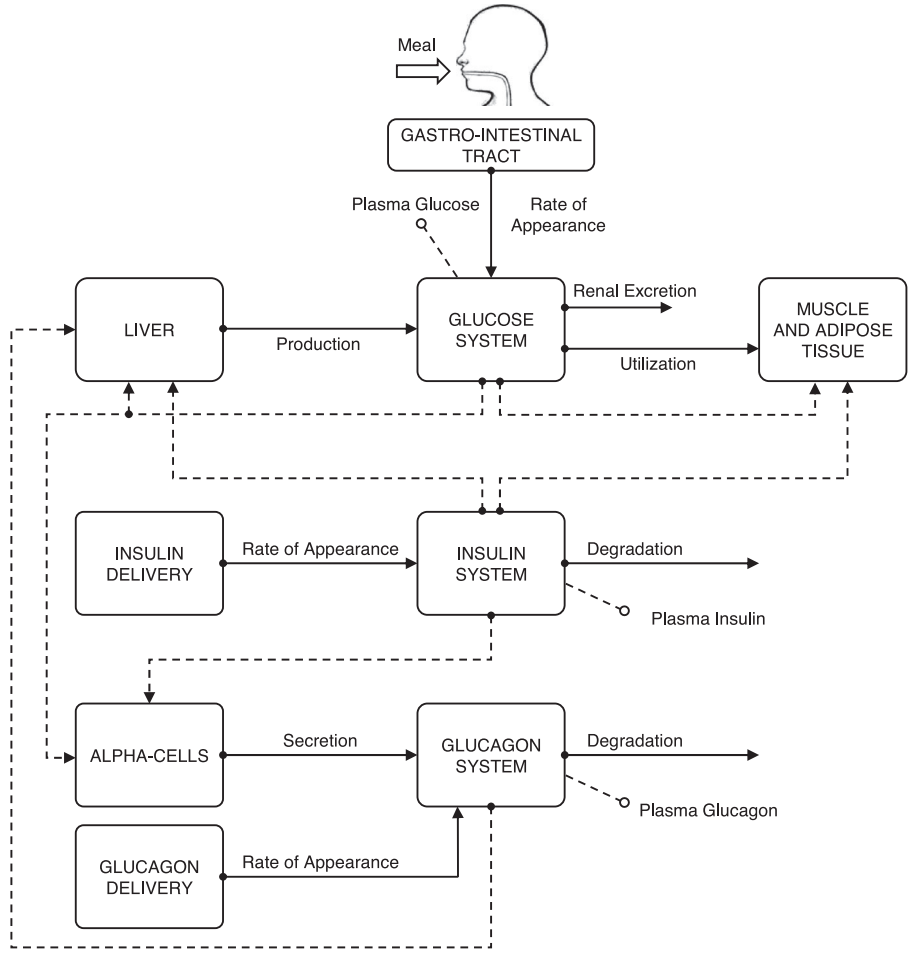

Fig. 2. Scheme of the T1DM simulator.

different descriptions of Ra_{meal}, *EGP*, and *U*, i.e., a good model fit can be achieved with several combinations of model parameters. To overcome this limitation, here, we adopt a Bayesian approach, i.e., the estimation of the parameter vector *p* takes into account both the information provided by the data vector *z*, i.e., the *a posteriori* information, and the knowledge on the *a priori* joint distribution of *p,* assumed independent form *z*. In particular, the Maximum *a Posteriori* (MAP) Bayesian estimator provides a point estimate *p* so that, once fixed *z*, the *a posteriori* probability density of *p* is maximum.

$$\hat{p}_{\mathrm{MAP}} = \underset{p}{\operatorname{argmax}} f_{(p|z)}\left(p|z\right) \tag{1}$$

where

$$f_{(p|z)}\left(p|z\right) = \frac{f_{(z|p)}\left(z|p\right) f_p\left(p\right)}{f_z\left(z\right)} \tag{2}$$

with $f_p(p)$ denoting the *a priori* probability density of *p*, considered random, $f_z(z)$ the *a priori* probability density of *z*, and $f_{(z|p)}(z|p)$ the *a posteriori* probability density of *z*. Assuming that *z* is affected by measurement error *v,* Gaussian, with zero mean and covariance $\sum_v$, and *p* is extracted from a Gaussian distribution with mean μ_p and covariance $\sum_p$, (1) can be rewritten as

$$\hat{p}_{\mathrm{MAP}} = \underset{p}{\operatorname{argmin}} \Big\{ \left[z - G\left(p\right)\right]^T \Sigma_v^{-1} \left[z - G\left(p\right)\right] + \left[p - \mu_p\right]^T \Sigma_p^{-1} \left[p - \mu_p\right] \Big\} \tag{3}$$

with *G(p)* denoting the model prediction. In other words, the first term in (3) is the model fit, while the second term represents the distance of the estimated parameters from their joint distribution.

In addition, to guarantee the nonnegativity of model parameters, parameter distributions are assumed to be lognormal

$$s = \log\left(p\right). \tag{4}$$

Therefore, the estimated parameter vector can be expressed as

$$\hat{p} = \exp\left(\hat{s}\right) \tag{5}$$

where

$$\hat{s} = \underset{s}{\operatorname{argmin}} \Big\{ \left[z - G\left(\exp\left(p\right)\right)\right]^T \Sigma_v^{-1} \left[z - G\left(\exp\left(p\right)\right)\right] + \left[s - \mu_s\right]^T \Sigma_s^{-1} \left[s - \mu_s\right] \Big\} \tag{6}$$

with μ_s and $\sum_s$ denoting the mean and the covariance matrix in logarithmic form. For more details on MAP estimation, we refer to [10].

B. Model Identification

The T1DM simulator does not yet account for the effect of physical activity on glucose dynamics. Thus, the model identification was performed excluding data on physical exercise sessions, which however occurred at the end of the experiment. Hence, the model was identified on 20-h plasma glucose data using the MAP estimator implemented in MATLAB R2013b [11]. In order to avoid local minima, the minimization of the objective function was performed using a cascade of a direct search method followed by a gradient-based algorithm. Measurement error was assumed to be additive, Gaussian, with zero mean and constant coefficient of variation (CV) of 2%. Plasma insulin concentration was the model forcing function and was assumed to be known without error. As it will be discussed in the Conclusion section, the choice of using plasma insulin as model input rather than the insulin pump delivery rate is motivated by necessity to set up the methodology in a simpler context, i.e., in this case only the glucose subsystem model parameters are estimated. Similarly, being glucagon measurements not available, the average glucagon model parameters were used.

In addition, model identification provides an estimation of the glucose fluxes, i.e., Ra_{meal} calculated from the model (A3), and the net rate of disappearance, defined as $Rd_{\mathrm{net}} = U + E - EGP$. In particular, Rd_{net} is considered due to the data information, i.e., no tracers are employed during the experiment, therefore, it is difficult to distinguish *U*, *E*, and *EGP* contributes.

The *a priori* information in (6) is the joint parameter distribution used to generate the adult *in silico* population included into the T1DM simulator [5]. To account for the physiological intra-subject parameter variations, diurnal variability of parameters describing glucose absorption [k_{abs}, $k_{\max}$, and $k_{\min}$ in (A3) and (A4)] and insulin sensitivity [k_{p3} and V_{mx} in (A5) and (A10)] was permitted. In particular, gastric absorption parameters were allowed to assume different values at B, L, and D, while insulin sensitivity was allowed to be different at B with respect to L and D, while L and D were similar. These assumptions are based on the fact that different meal compositions have a different impact on glucose absorption dynamics, and that T1DM subjects exhibit an insulin sensitivity trend lower at B compared to L and D [7]. The intraday variability was implemented for each time-varying parameter as an almost step-wise-line signal that varies three times a day (i.e., before each meal).

C. Model Assessment

In order to assess the goodness of the model fit, the distribution of some performance metrics were analyzed, such as the coefficient of determination (R^2) and the performance index (FIT) defined as

$$\mathrm{FIT} = 1 - \sqrt{\frac{\sum_{k=1}^{N}\left(G^{\mathrm{meas}}(t_k) - G^{\mathrm{pred}}(t_k)\right)^2}{\sum_{k=1}^{N}\left(G^{\mathrm{meas}}(t_k) - G^{\mathrm{mean}}\right)^2}} \quad (7)$$

where G^{meas} is the measured and G^{pred} is the simulated glucose concentration, G^{mean} is the average measured glucose, and *N* is the number of samples. The precision of parameter estimates is expressed by the coefficient of variation (CV, defined as the ratio between the standard deviation of the estimated parameter and the parameter value), which is related to how much a variation of a specific parameter influences the model prediction (the lower the CV, the higher the sensitivity of model prediction to the parameter).

D. Assessment of a Priori Parameter Distribution

The Bayesian MAP estimation technique is a very powerful tool to identify models whenever the available data are scarce in comparison to model complexity. However, the choice of the prior distribution is critical and influence the final parameter estimates. If the prior is too informative (small variance) the risk is that parameter estimates collapse into the prior and the model does not fit the data well. On the other hand, if the prior is not enough informative (large variance) the parameter estimates may be imprecise. Ideally, the *a priori* information and the *a posteriori* information should be well balanced. This is reasonably achieved if the fit of the data is good and the parameter estimates are precise. However, this performance does not guarantee that the *a priori* parameter distribution is representative of our type 1 diabetic population. To investigate the agreement of the prior with the posterior and the stability of the solution, the distributions are compared by a statistical test (see Statistical Analysis) and the iterative two-stage method. Briefly, for sake of convenience, let us call: *Stage1* the model identification performed using the prior included into the simulator, *Stage2* the model identification performed using the posterior obtained from *Stage1* as new prior, and so on. We compared the posteriors with the priors after each iteration, i.e., *Stage2* versus *Stage1*, *Stage3* versus *Stage2*, etc. In particular, we calculated, for each subject, the absolute relative difference $\delta\mathbf{p}_{\mathrm{j}}$:

$$\delta\mathbf{p}_j = [\delta p_{1,j}, \ldots, \delta p_{25,j}], \quad j \geq 2 \quad (8)$$

with

$$\delta p_{i,j} = \left|\frac{p_{i,j} - p_{i,j-1}}{p_{i,j-1}}\right|, \quad i \in [1, \ldots, 25] \quad (9)$$

where $p_{i,j}$ is the value of *i*th parameter estimated at the *j*th stage. The lower the $\delta p_{i,j}$ the better the agreement between prior and posterior.

E. Statistical Analysis

Results are presented as median and interquartile range (IQR), if not differently stated. Two sample comparisons were done by Wilcoxon Signed Rank test, with significance level set at $P = 0.05$.

V. Results

A. Fit, Parameters, and Fluxes

The model was identified in all the available 141 glucose traces, and well fitted the glucose data, as proved by the average weighted residuals time course, shown in Fig. 3. The analysis

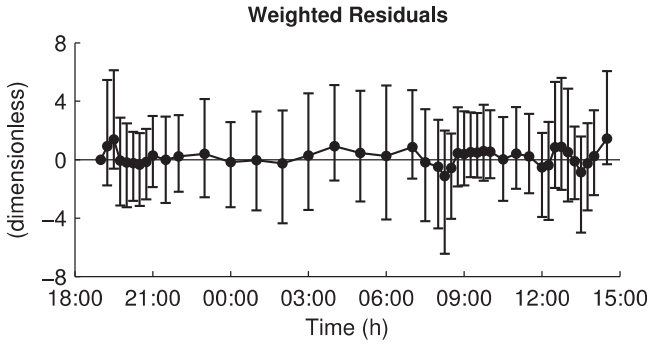

Fig. 3. Average weighted residuals [vertical bars represent standard deviation (SD)] of model fit on plasma glucose data ($N = 141$).

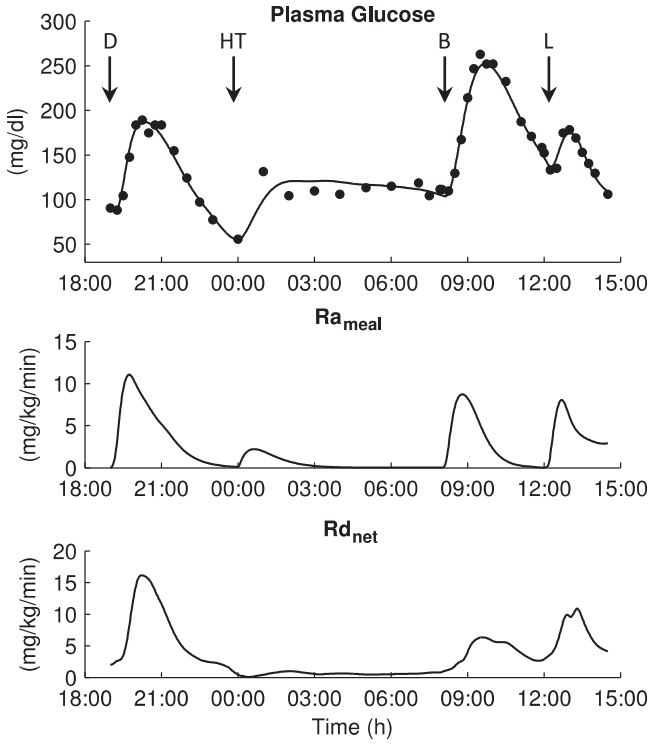

Fig. 4. Model prediction versus plasma glucose concentration (*upper panel*) and the corresponding glucose fluxes, i.e., $Ra_{\rm meal}$ (*middle panel*) and $Rd_{\rm net}$ (*lower panel*), provided by the model in an illustrative subject. Data are indicated with dots, while model prediction and fluxes are shown with continuous lines. B, L, D, and HT, respectively, indicate breakfast (50 CHO grams), lunch (60 CHO grams), dinner (80 CHO grams), and hypoglycemia treatment (10 CHO grams).

of performance metrics supported the quality of model fit: mean $\pm$ SD (min – max) of R^2 and FIT were 0.962 ± 0.027 (0.854 – 0.996) and 0.812 ± 0.066 (0.615 – 0.934), respectively.

An example of the glucose model fit in a representative subject is shown in Fig. 4, *upper panel*. In addition, we provide the model-derived time courses of $Ra_{\rm meal}$ and net rate of disappearance $Rd_{\rm net}$ (see Fig. 4 *middle* and *bottom panels*).

As expected, having resorted to a Bayesian approach, model parameters are estimated with good precision ($CV = 1.3\% \pm 0.2\%$), despite the complexity of the model with respect to available data.

B. Intra- and Intersubject Variability

The *a posteriori* distribution of model parameters is generally in agreement with that included in the UVA/Padova simulator. However, some differences occur. In particular, glucose gastric emptying parameters at B are significantly different from those at L and D (see Fig. 5 *left panel*), reflecting a more rapid glucose dynamics at B, possibly due to the different meal compositions of the B versus L and D. On the other hand, insulin sensitivity varies during the day in a subject-specific fashion, without showing a consistent pattern in the population (see Fig. 5 *right panel*), in agreement with what already reported in [7]. The complete lists of glucose absorption and insulin sensitivity parameter estimates are reported in Table I.

The comparison among parameter distributions included in the simulator and those obtained at B, L, and D in this study are shown in Fig. 6, for one of the glucose absorption parameters ($k_{\rm max}$) and insulin sensitivity ($V_{\rm mx}$). In particular, the distributions of $k_{\rm max}$ (*left panel*) have the same shape, but the mean for L and D are significantly lower than those of B and the simulator ($P < 0.0001$ when comparing L&D against both B and the simulator prior). On the other hand, distributions of insulin sensitivity at B, L, D, and the simulator (*right panel*) are statistically identical ($P > 0.05$ when comparing L&D against both B and the simulator prior).

C. Iterative Two Stage

After the first iteration, the posteriors obtained at *Stage1* and *Stage2* were statistically the same except for $k_{\rm abs}$ at breakfast ($P = 0.03$). As concerns the relative difference, we found that $\delta\mathbf{p}_2$ was 0.059 ± 0.044 (mean $\pm$ SD), $\delta\mathbf{p}_3 = 0.045 \pm 0.035$ and $\delta\mathbf{p}_4 = 0.034 \pm 0.028$.

VI. Discussion

The UVA/Padova T1DM Simulator [5] has recently undergone a clinical validation by proving that the virtual subjects are representative of the T1DM population observed in a clinical trial [6], but, given the complexity of its model, it has never been identified in T1DM individuals so far. The objective of this study was thus threefold: 1) propose a method to identify the model included into the simulator from plasma glucose and insulin concentration data; 2) move from a single meal to a breakfast/lunch/dinner meal data scenario, thus accommodating intrasubject variability of glucose absorption and insulin sensitivity; and 3) compare data versus *in silico* intersubject variability, and, if necessary, improving the *in silico* population included into the simulator.

To address point 1), the method was based on a Bayesian approach, which is able to cope with the identification of a model having a large number of parameters, by using the *a priori* knowledge on model parameter distributions. The natural choice was to use, as *a priori* information, the joint parameter distributions included into the UVA/ Padova simulator.

In order to tackle points 2) and 3), concerning the intra-/ intersubject variability, we employed an appropriate dataset, consisting of 47 T1DM subjects, each studied three times with

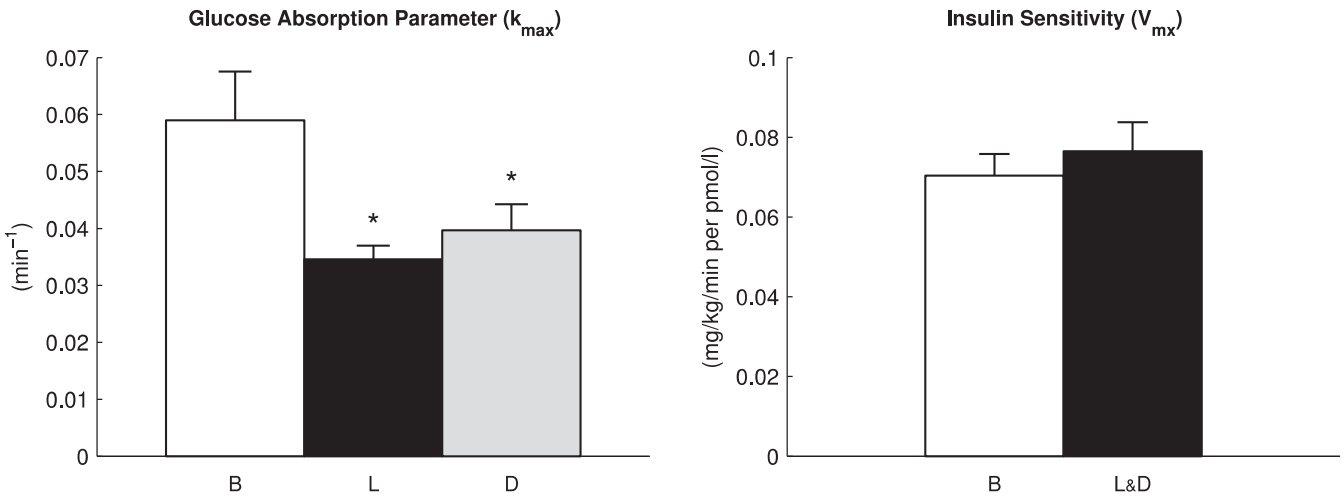

Fig. 5. Absorption parameter, $k_{\max}$, (*left panel*), and insulin sensitivity, V_{mx}, (*right panel*), estimated at breakfast B, lunch L, and dinner D [vertical bars represent standard error (SE)]. $^{*}P<0.05$ with respect to B, from Wilcoxon Signed Rank Test.

TABLE I
A) GLUCOSE ABSORPTION ($k_{\mathrm{abs}}, k_{\max}, k_{\min}$) AND B) INSULIN SENSITIVITY (V_{mx}, k_{p3}) PARAMETER ESTIMATES

A: Glucose Absorption Parameters			
Parameter	B	L	D
k_{abs} (min^{-1})	0.130 [0.092–0.174]	0.130 [0.076–0.216] (NS)	0.147 [0.098–0.209] (NS)
$k_{\max}$ (min^{-1})	0.040 [0.027–0.059]	0.028 [0.021–0.041] (<0.0001)	0.030 [0.021–0.043] (<0.0001)
$k_{\min}$ (min^{-1})	0.015 [0.009–0.019]	0.010 [0.005–0.015] (<0.0001)	0.008 [0.005–0.011] (<0.0001)
B: Insulin Sensitivity Parameters			
Parameter	B	L&D	
V_{mx} (mg/kg/min per pmol/L)	0.051 [0.034–0.090]	0.058 [0.037–0.080] (NS)	
k_{p3} (mg/kg/min per pmol/L)	0.015 [0.006–0.025]	0.014 [0.008–0.033] (NS)	

Values are reported as median [IQR] (*P* value with respect to B, from Wilcoxon Signed Rank Test). NS, not significant ($P > 0.05$).

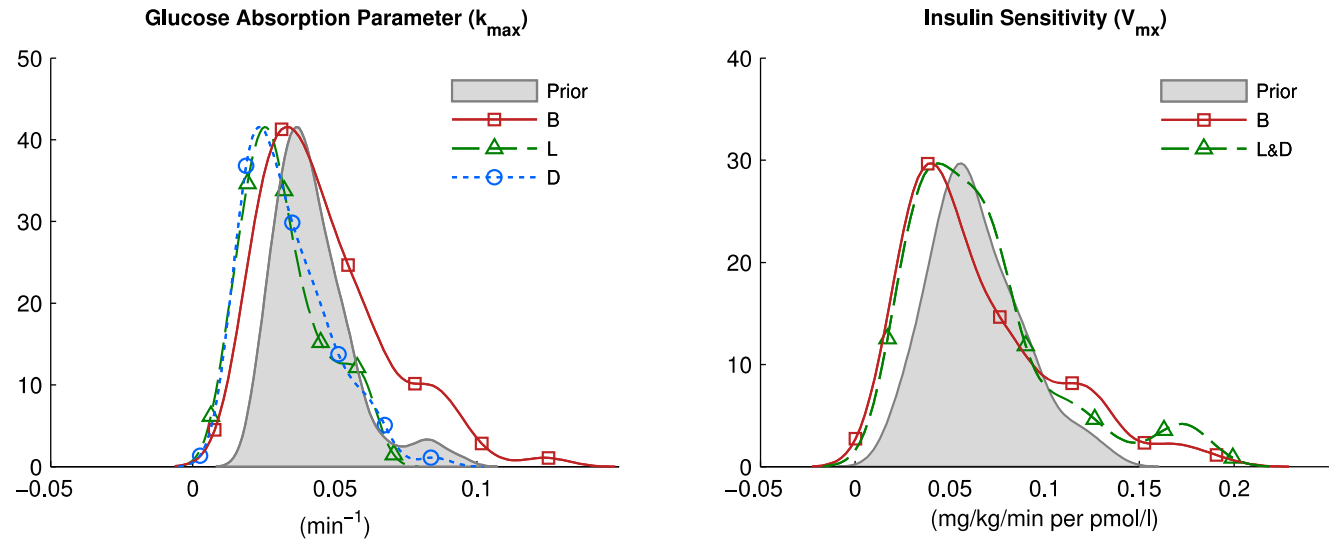

Fig. 6. Distribution of $k_{\max}$, (*left panel*), and V_{mx}, (*right panel*), at breakfast B (red squares), lunch L (green triangles), and dinner D (blue circles) compared to the prior (gray region).

a 23-h admissions, in which meals times were scheduled and CHO contents were determined with precision. In particular, the length of the experiment allows us to evaluate the variability of some relevant parameters during the day, i.e., meal glucose absorption and insulin sensitivity, while the large number of data analyzed ($N = 141$) permits to reliably infer about the intersubject variability.

The glucose time courses were well predicted and the estimated postprandial glucose fluxes, e.g., meal rate of appearance and net rate of disappearance, were also calculated (see Fig. 4). Model parameters were estimated with precision and their posterior distribution was in agreement with that included in the current simulator. In order to evaluate the adequacy of the prior distribution, we tested the method using an iterative two-stage approach. Results show that the relative differences in model parameters were modest, getting smaller, and smaller with the iterations. Paralleling, the posterior distributions were statistically identical except for some of the absorption parameters. This result is reassuring, since these parameters were expected to vary more than the others, since the meal composition in this dataset (solid meal) is different from that used for generating the prior (liquid meal [12]).

Taken together, these results confirm that the choice of the prior is key. Here, we used the parameter distribution included into the UVA/Padova T1DM Simulator, which was derived from multiple tracer data [12], [13]. In principle, other distributions can be used, and it is likely that different *a priori* parameter distributions would lead to different parameter estimates, and thus, to different *a posteriori* distributions. We believe that the power of the proposed method resides in the availability of a good prior (the one of the simulator), which provides repeatable parameter solutions: in fact, we demonstrated that the distribution of constant parameters remains the same among the iterations, and, at the same time, the fitting procedure is sensitive enough to capture the variation of the important parameters, e.g., those related to the meal composition and insulin sensitivity. In this regard, it is important to point out that the significant differences observed in distributions of glucose absorption parameter (see Fig. 6 *left panel*) were expected: in fact, the parameter distributions of the simulator were obtained from subjects who received a mixed meal containing rapidly absorbed carbohydrates. Hence, it is reasonable that the posterior distribution estimated at B is virtually superimposable to the prior, since, usually breakfast contains fast carbs, while the posterior distributions estimated at L and D are statistically different from the original ones, since usually those meals are characterized by a slower absorption. This, again, supports the notion that the information content in the data is rich enough to observe this important variation in model parameters and that the prior is not too constraining, so that model parameters are allowed to move away from it, if data say so.

This is an important step forward with respect to [14], where a model of intraday variability of insulin sensitivity only was introduced into the T1DM simulator: the novelty here is the additional incorporation into the simulator of the concomitant diurnal variation of meal glucose absorption parameters by distinguishing the breakfast CHO absorption versus lunch and dinner.

Our findings confirmed what has been found by Haidar and coworkers [9], i.e., that the use of time-varying model parameters helps to better describe the intraday variability. However, at variance with [9], the introduction of intraday variation of some model parameters (three describing meal glucose absorption and two representing insulin sensitivity) allowed us to improve the model fit, but still retaining the physiological plausibility of the model structure, parameters, and model-derived variables. In particular, we found that insulin sensitivity was lower, on average, at breakfast (B) compared to lunch (L) and dinner (D), but not significantly, i.e., it varied during the day in a subject-specific fashion, without showing a consistent pattern in the population (see Fig. 5 *right panel*), as already reported in Hinshaw *et al*. [7] and Visentin *et al*. [14]. These variations in insulin sensitivity could implicitly include the effect of dawn phenomenon, in agreement with what observed in T1DM subjects [15], i.e., an increase of insulin requirement from 2–3AM to 9–10AM, likely due to a decrease in insulin sensitivity.

Finally, our method proved to be also very efficient from a computational point of view: model identification of a single subject required about 3 h on average, using a 2.6-GHz computer.

This contribution is a first step in developing a method to provide a virtual clone of a specific T1DM subject. As with all modeling studies, there are limitations that will hopefully be tackled in future work. To set up the methodology, we have started from plasma glucose and insulin data and used plasma insulin as a forcing function, thus this method currently provides only a partial clone of a T1DM subject, since the model identification is conducted on the sole glucose subsystem, while the insulin subsystem is bypassed by using plasma insulin data as forcing function. However, plasma insulin is only measured in inpatient clinical studies and this limits the wider applicability of the method. This also precludes the use of the method for prediction and control personalization, which require the capability of "cloning" the subject from glucose sensor and insulin pump data. Therefore, the next step will definitely be the inclusion of the insulin subsystem and using the insulin pump delivery rate as a forcing input. This will also open up much larger databases on T1DM subjects using a sensor-augmented pump insulin therapy. Another refinement will consists in moving form plasma glucose measurements to subcutaneous glucose sensor data. This will permit to exploit the important information provided by the large and increasing number of long-term outpatient studies. This will clearly require some appropriate signal preprocessing, e.g., retrospective fitting [16], which proved to be a valid tool to obtain a reliable estimate of the original blood glucose signal, thus improving the accuracy of sensor measurements. In moving to weekly/monthly scenarios (longer than a single day), some refinements of the method will be considered, for example, the model identification could focus on a subset of model parameters, i.e., the "highly sensitive" parameters that are expected to play a key role in mid-/long-term variability, such as insulin sensitivity and glucose absorption parameters.

VII. Conclusion

In conclusion, here, we presented a Bayesian method to identify the model of the recent version of the UVA/Padova T1DM simulator [5]. This method exploits the *a priori* joint parameter distribution used for the generation of the virtual subjects included into the simulator. The method proved to be robust, having been successfully applied on a large population of T1DM subjects. Moreover, the *a posteriori* parameter distribution agreed with that included into the simulator, thus proving the validity of the prior itself. The incorporation of glucose absorption and insulin sensitivity diurnal variability into the simulator allowed us to improve the real life resembling of the simulator, by extending the simulation environment from a "single-meal" to a more realistic "single-day" scenario. The method, applied to the increasing number of long-term artificial pancreas studies, will allow to describe the week/month variability, thus further refining the simulator.

Appendix

A. Model Equations

1) Glucose Subsystem:

$$\begin{cases} \dot{G}_p(t) = \mathrm{EGP}(t) + Ra_{\mathrm{meal}}(t) - U_{\mathrm{ii}}(t) - E(t) \\ \qquad\qquad -k_1 \cdot G_p(t) + k_2 \cdot G_t(t) \\ G_p(0) = G_{\mathrm{pb}} \\ \dot{G}_t(t) = -U_{\mathrm{id}}(t) + k_1 \cdot G_p(t) - k_2 \cdot G_t(t) \\ G_t(0) = G_{\mathrm{tb}} \\ G(t) = G_p(t)/V_G \\ G(0) = G_b. \end{cases} \tag{A1}$$

Insulin subsystem (not considered in this study due to the use of plasma insulin as forcing function):

$$\begin{cases} \dot{I}_p(t) = -(m_2 + m_4) \cdot I_p(t) + m_1 \cdot I_l(t) + Ra_I(t) \\ I_p(0) = I_{\mathrm{pb}} \\ \dot{I}_l(t) = -(m_1 + m_3) \cdot I_l(t) + m_2 \cdot I_p(t) \\ I_l(0) = I_{\mathrm{lb}} \\ I(t) = I_p(t)/V_I \\ I(0) = I_b. \end{cases} \tag{A2}$$

Glucose rate of appearance:

$$\begin{cases} Q_{\mathrm{sto}}(0) = Q_{\mathrm{sto1}}(t) + Q_{\mathrm{sto2}}(t) \\ Q_{\mathrm{sto}}(0) = 0 \\ \dot{Q}_{\mathrm{sto1}}(t) = -k_{\max} \cdot Q_{\mathrm{sto1}}(t) + \mathrm{Dose} \cdot \delta(t) \\ Q_{\mathrm{sto1}}(0) = 0 \\ \dot{Q}_{\mathrm{sto2}}(t) = -k_{\mathrm{empt}}(Q_{\mathrm{sto}}) \cdot Q_{\mathrm{sto2}}(t) + k_{\max} \cdot Q_{\mathrm{sto1}}(t) \\ Q_{\mathrm{sto2}}(0) = 0 \\ \dot{Q}_{\mathrm{gut}}(t) = -k_{\mathrm{abs}} \cdot Q_{\mathrm{gut}}(t) + k_{\mathrm{empt}}(Q_{\mathrm{sto}}) \cdot Q_{\mathrm{sto2}}(t) \\ Q_{\mathrm{gut}}(0) = 0 \\ Ra_{\mathrm{meal}}(t) = \frac{f \cdot k_{\mathrm{abs}} \cdot Q_{\mathrm{gut}}(t)}{BW} \\ Ra_{\mathrm{meal}}(0) = 0 \end{cases} \tag{A3}$$

with

$$k_{\mathrm{empt}}(Q_{\mathrm{sto}}) = k_{\min} + \frac{k_{\max} - k_{\min}}{2} \cdot \{\tanh[\alpha(Q_{\mathrm{sto}} - \beta \cdot \mathrm{Dose})] - \tanh[\beta(Q_{\mathrm{sto}} - c \cdot \mathrm{Dose})] + 2\}. \tag{A4}$$

Endogenous glucose production:

$$\mathrm{EGP}(t) = k_{p1} - k_{p2} \cdot G_p(t) - k_{p3} \cdot X^L(t) + \xi \cdot X^H(t) \tag{A5}$$

$$\dot{X}^L(t) = -k_i \cdot \left[X^L(t) - I'(t)\right] \quad X^L(0) = I_b \tag{A6}$$

$$\dot{I}'(t) = -k_i \cdot [I'(t) - I(t)] \quad I'(0) = I_b \tag{A7}$$

$$\dot{X}^H(t) = -k_H \cdot X^H(t) + k_H \cdot \max[(H(t) - H_b), 0]$$

$$X^H(0) = 0 \tag{A8}$$

Glucose utilization:

$$U_{\mathrm{ii}}(t) = F_{\mathrm{cns}} \tag{A9}$$

$$U_{\mathrm{id}}(t) = \frac{[V_{m0} + V_{\mathrm{mx}} \cdot X(t) \cdot (1 + r_1 \cdot \mathrm{risk})] \cdot G(t)}{K_{m0} + G_t(t)} \tag{A10}$$

with

$$\dot{X}(t) = -p_{2U} \cdot X(t) + p_{2U} \cdot [I(t) - I_b] \quad X(0) = 0 \tag{A11}$$

$$\mathrm{risk} = \begin{cases} 0, & \text{if } G \geq G_b \\ 10 \cdot [f(G)]^2, s & \text{if } G_{\mathrm{th}} \leq G < G_b \\ 10 \cdot [f(G_{\mathrm{th}})]^2, & \text{if } G < G_{\mathrm{th}} \end{cases} \tag{A12}$$

$$f(G) = \log\left(\frac{G}{G_b}\right)^{r_2}. \tag{A13}$$

Renal excretion:

$$E(t) = \begin{cases} k_{e1} \cdot [G_p(t) - k_{e2}], & \text{if } G_p(t) > k_{e2} \\ 0, & \text{if } G_p(t) \leq k_{e2}. \end{cases} \tag{A14}$$

Glucagon kinetics and secretion (average model considered in this study):

$$\dot{H}(t) = -n \cdot H(t) + SR_H(t) + Ra_H(t) \quad H(0) = H_b \tag{A15}$$

with

$$SR_H(t) = SR^s_H(t) + SR^d_H(t) \tag{A16}$$

$$\dot{S}R^s_H(t) = \begin{cases} -\rho \cdot [SR^s_H(t) - \max(\sigma_2 \cdot [G_{\mathrm{th}} - G(t)] + SR^b_H, 0)] & \text{if } G(t) \geq G_b \\ -\rho \cdot \left[SR^s_H(t) - \max\left(\frac{\sigma \cdot [G_{\mathrm{th}} - G(t)]}{I(t)+1} + SR^b_H, 0\right)\right] & \text{if } G(t) < G_b \end{cases} \tag{A17}$$

$$SR^d_H(t) = \delta \cdot \max\left(-\frac{dG(t)}{dt}, 0\right). \tag{A18}$$

TABLE II
LIST OF MODEL PARAMETERS

Glucose Kinetics	
V_g	distribution volume of glucose (dL/kg)
k_1, k_2	rate parameters (min^{-1})
Rate of Appearance	
k_{abs}	rate constant of intestinal absorption (min^{-1})
k_{max}, k_{min}	maximum and minimum levels of gastric emptying rate (min^{-1})
b, d	fraction of dose corresponding to the flexes of gastric emptying curve (dimensionless)
f	fraction of intestinal absorption appearing in plasma (dimensionless)
Endogenous Production	
k_i	rate parameter accounting for delay between insulin Signal and Insulin Action (min^{-1})
k_{p1}	extrapolated EGP at zero glucose and insulin (mg/kg/min)
k_{p2}	hepatic glucose effectiveness (min^{-1})
k_{p3}	hepatic insulin sensitivity (min^{-1})
ξ	hepatic responsivity to glucagon (mg/kg/min per ng/L)
Utilization	
F_{cns}	insulin-independent glucose utilization (mg/kg/min)
V_{mx}	insulin sensitivity on glucose utilization (mg/kg/min per pmol/L)
K_{m0}	glucose mass appearing in Michaelis–Menten relation (mg/kg)
p_{2u}	rate constant of insulin action on the peripheral glucose utilization (min^{-1})
r_1, r_2	parameters of risk function (dimensionless)
Renal Excretion	
k_{e1}	glomerular filtration rate (min^{-1})
k_{e2}	renal threshold of glucose (mg/dL)
Insulin Kinetics	
V_I	distribution volume of insulin (L/kg)
m_1, m_2	rate parameters (min^{-1})
m_3	liver degradation rate (min^{-1})
m_4	peripheral degradation rate (min·kg/pmol)
Ra_I	appearance rate of external insulin (pmol/kg/min)
Glucagon Kinetics	
n	glucagon clearance rate (min^{-1})
Glucagon Secretion	
ρ	rate parameter accounting for the delay between static glucagon secretion and plasma glucose (min^{-1})
σ, σ_2	alpha-cell responsivities to glucose level (ng/L/min per mg/dL·L/pmol)
δ	alpha-cell responsivity to glucose rate of change (ng/L·mg/dL)

REFERENCES

[1] C. Cobelli *et al.*, "Diabetes: Models, signals, and control," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 54–96, Jan. 2009.

[2] B. P. Kovatchev *et al.*, "In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes," *J. Diabetes Sci. Technol.*, vol. 3, pp. 44–55, Jan. 2009.

[3] M. E. Wilinska *et al.*, "Simulation environment to evaluate closed-loop insulin delivery systems in type 1 diabetes," *J. Diabetes Sci. Technol.*, vol. 4, pp. 132–144, Jan. 2010.

[4] S. S. Kanderian *et al.*, "Identification of intraday metabolic profiles during closed-loop glucose control in individuals with type 1 diabetes," *J. Diabetes Sci. Technol.*, vol. 3, pp. 1047–1057, Sep. 2009.

[5] C. Dalla Man *et al.*, "The UVA/PADOVA type 1 diabetes simulator: New features," *J. Diabetes Sci. Technol.*, vol. 8, pp. 26–34, Jan. 2014.

[6] R. Visentin *et al.*, "The university of Virginia/Padova type 1 diabetes simulator matches the glucose traces of a clinical trial," *Diabetes Technol. Ther.*, vol. 16, no. 7, pp. 428–434, Jul. 2014.

[7] L. Hinshaw *et al.*, "Diurnal pattern of insulin action in type 1 diabetes: implications for a closed-loop system," *Diabetes*, vol. 62, pp. 2223–2229, Jul. 2013.

[8] Y. M. Luijf *et al.*, "Day and night closed-loop control in adults with type 1 diabetes: A comparison of two closed-loop algorithms driving continuous subcutaneous insulin infusion versus patient self-management," *Diabetes Care*, vol. 36, pp. 3882–3887, Dec. 2013.

[9] A. Haidar *et al.*, "Stochastic virtual population of subjects with type 1 diabetes for the assessment of closed loop glucose controllers," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3524–3533, Dec. 2013.

[10] C. Cobelli and E. Carson, *Introduction to Modeling in Physiology and Medicine*. New York, NY, USA: Academic, 2008.

[11] MATLAB, version 8.2.0.29 (R2013b), The Math-Works Inc., Natick, MA, USA, 2013.

[12] R. Basu *et al.*, "Effects of age and sex on postprandial glucose metabolism: Differences in glucose turnover, insulin secretion, insulin action, and hepatic insulin extraction," *Diabetes*, vol. 55, no. 7, pp. 2001–2014, Jul. 2006.

[13] C. Dalla Man *et al.*, "Meal simulation model of the glucose-insulin system," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 10, pp. 1740–1749, Oct. 2007.

[14] R. Visentin *et al.*, "Circadian variability of insulin sensitivity: physiological input for an in silico artificial pancreas," *Diabetes Technol. Ther.*, vol. 17, no. 1, pp. 1–7, Jan. 2015.

[15] G. Scheiner and B. A. Boyer, "Characteristics of basal insulin requirements by age and gender in Type-1 diabetes patients using insulin pump therapy," *Diabetes Res. Clin. Pract.*, vol. 69, no. 1, pp. 14–21, Jul. 2005.

[16] S. Del Favero *et al.*, "Improving accuracy and precision of glucose sensor profiles: Retrospective fitting by constrained deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 4, pp. 1044–1053, Apr. 2014.

Authors' photographs and biographies not available at the time of publication.