Modelling of glucose-insulin dynamics from low sampled data

Aradóttir, Tinna Björk; Boiroux, Dimitri; Bengtsson, Henrik; Poulsen, Niels Kjølstad

Published in:
Proceedings of 18th IFAC Symposium on System Identification

Link to article, DOI:
10.1016/j.ifacol.2018.09.213

Publication date:
2018

Document Version
Peer reviewed version

Preprints,18th IFAC Symposium on System Identification
July 9-11, 2018. Stockholm, Sweden

Modelling of glucose-insulin dynamics from low sampled data ⋆

Tinna B. Aradóttir [*,†] **Dimitri Boiroux** [*] **Henrik Bengtsson** [†]
Niels K. Poulsen [*]

[*] *Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark (e-mail: tiar@dtu.dk, dibo@dtu.dk, nkpo@dtu.dk)*
[†] *Novo Nordisk A/S, DK-2880 Bagsværd, Denmark (e-mail: hbss@novonordisk.com)*

Abstract: In this paper we focus on modelling the glucose-insulin dynamics in the human body for the purpose of controlling the glucose level. Due to the fast dynamics in the glucose-insulin system compared to the natural sampling period (24 h) in a clinical situation, the model structure has to be adapted adequately. This results in a reduced order model with a non-linear output relation. The development of the estimation methodology is based on a simulation study with a continuous time model. The resulting model structure is used for estimating the parameters of the non-linear system, representing the slow dynamics observed from the slow and sparse sampled clinical data.

Keywords: System identification, reduced order models, glucose-insulin dynamics, diabetes control.

1. INTRODUCTION

Diabetes is a chronic condition characterized by raised levels of glucose in the blood, hyperglycemia. The condition affects more than 425 million people today and the numbers are expected to rise to 693 million by 2045 [International Diabetes Federation 2017]. Poorly managed diabetes can lead to cardiovascular diseases, lower limb amputation, blindness and kidney failure. The American Diabetes Association estimates that care for people with diabetes accounts for more than 20% of all health care expenditure in the U.S. [Petersen 2016]. Type 2 diabetes accounts for 90% of all diabetes cases. In type 2 diabetes, the elevated glucose levels are caused by inadequate production and response to the hormone insulin. As opposed to type 1 diabetes, a congenital disease with quick onset most commonly in children, type 2 diabetes is most commonly diagnosed in older adults. But with increased prevalence of obesity, less physical activity and poor diet, type 2 diabetes is becoming more common in young adults, children and adolescents.

When treating type 2 diabetes, a first attempt is through lifestyle changes. This is followed by oral medication if increased physical activity and change in diet is not adequate. When these treatments fail, insulin injections may be needed. There are two main types of insulins, long acting insulin to lower fasting glucose levels, and fast acting insulin to lower glucose levels following food intake. Standards of Medical Care in Diabetes recommend starting with long acting insulin treatment, and adding fast acting insulin if glucose values are still too high [American Diabetes Association 2017]. In this paper, we focus on treatment of type 2 diabetes with long acting insulin, specifically initiation of the long acting insulin treatment.

Initiating long acting insulin treatment is an iterative process since response to insulin is individual. Too large doses of insulin can cause low blood glucose, hypoglycemia, which in severe cases can lead to coma and even death. Health care professionals prescribe small doses and increase dose sizes over time, based on self-monitored blood glucose (SMBG), until a target glucose level is reached. SMBG measurements are glucose measurements performed by the patients by finger-pricking. Doses of long acting insulin are adjusted based on SMBG measurements performed in a fasting state, typically before breakfast. The rules by which health care professionals change dose sizes are typically represented by simple tables that do not account for the great inter-patient variability. Therefore, insulin initiation can take months to years, and in the U.S., more than 60% of type 2 diabetes patients on insulin treatment do not reach recommended treatment goals [Wong et al. 2012].

In this paper, we wish to understand the behaviour of fasting glucose in response to long acting insulin, through a physiological model of the glucose-insulin regulatory system. Physiological models of healthy humans and people with type 1 and type 2 diabetes exist. Most models are based on the Bergman minimal model [Bergman et al. 1979, Toffolo et al. 1980], but they vary in purpose and level of complexity. A number of models have been developed to simulate the regulatory system in type 1 diabetes, due to high interest in closed-loop control for this

⋆ This project is funded by Innovation Fund Denmark through the Industrial PhD project 5189-00033B, and the Danish Diabetes Academy supported by the Novo Nordisk Foundation

patient group [Kanderian et al. 2009, Hovorka et al. 2004, Dalla Man et al. 2007]. Jauslin et al. [2011] published a model for 24-hour modelling of insulin and glucose profiles in type 2 diabetes which Røge et al. [2014] used to further build a model to simulate effect of a specific type of insulin. These models are similar in that their purpose is 24-hour modelling where fast dynamics, such as increase in glucose due to meals, are well captured. In the current work we are interested in slower dynamics, as we work with fasting glucose and long acting insulin. Aradóttir et al. [2017] published a model for simulating fasting glucose during long acting insulin treatment. This model is based on a type 1 diabetes model by Kanderian et al. [2009] with an endogenous insulin production model by Ruan et al. [2015] to simulate type 2 diabetes. This model used parameters from literature, and here we wish to identify parameters from clinical data.

In Section 2 we outline the problem statement and describe important features of relevant clinical data. In Section 3 we adjust a detailed physiological model such that important parameters are identifiable from available clinical data with low sampling frequency. In Section 4 we discuss identifiability and list results from model identification based on simulated and clinical data.

2. STATEMENT OF THE PROBLEM

In this work we want to create a physiological model of insulin-glucose dynamics in type 2 diabetes. The purpose of this model is for control design for dose guidance in long acting insulin treatment. Such a model should have a physical interpretation, and parameter estimates based on data from a group of patients to capture variability in the patient population.

Clinical trials investigating new drugs are classified into different phases, and vary in purpose and number of participants. The first phases include only a few participants in the clinic, typically healthy volunteers, primarily aimed at testing for safety. Phase II and III studies test for efficacy and effectiveness, include hundreds or a few thousand patients and data is logged in the clinic or at home.

In Phase II and III studies on long acting insulin, glucose measuring frequency is typically one per day and sparse. The low data frequency is insufficient to identify detailed state-of-the-art models, such as the ones by Hovorka et al. [2004] or Kanderian et al. [2009]. For illustration, Figure 1 shows a four day simulation using the Kanderian et al. [2009] model, and glucose measured pre-breakfast (fasting).

Furthermore, excitation of the system in clinical trials is limited due to safety issues. Large doses of insulin can lead to hypoglycemia (too low blood glucose), a dangerous state which can lead to death. Therefore data illustrating insulin-glucose dynamics for glucose levels below 3.9 mmol/L (clinical hypoglycemia, International Hypoglycaemia Study Group [2016]) are rare.

3. MODEL STRUCTURE

We use the model by Kanderian et al. [2009], augmented in Aradóttir et al. [2017], as a starting point. We investigate which parameters are identifiable in long acting insulin treatment in an *in silico* setting, and finally consider whether this is applicable *in vivo*.

3.1 Base model

The following model, from Kanderian et al. [2009], describes insulin-glucose dynamics in type 1 diabetes in four re-named compartments, where the two-compartment meal model has been excluded,

$$\frac{dx_1}{dt} = \frac{1}{p_1}\frac{u}{p_2} - \frac{1}{p_1}x_1 \tag{1a}$$

$$\frac{dx_2}{dt} = \frac{1}{p_1}x_1 - \frac{1}{p_1}x_2 \tag{1b}$$

$$\frac{dx_3}{dt} = p_3 p_4 x_2 - p_3 x_3 \tag{1c}$$

$$\frac{dx_4}{dt} = -(p_5 + x_3)x_4 + p_6 \tag{1d}$$

After changing units to L, U, days and mmol (respectively from mL and dL, μU, min and mg), u is exogenous insulin [U/day], x_1 and x_2 denote subcutaneous and plasma insulin concentrations [U/L], respectively, x_3 is insulin effect [1/day] and x_4 is glucose concentration in plasma [mmol/L]. p_1 is a time constant describing transfer of insulin from the insulin delivery site, subcutaneous compartment, to plasma [day], p_2 is a gain describing insulin clearance [L/day], p_3 is an inverse time constant describing delay in insulin action following increased insulin concentration in plasma [1/day], p_4 is a gain describing insulin sensitivity [L/U·day], p_5 is an inverse time constant describing the effect of glucose at zero insulin to eliminate glucose from plasma [1/day] and p_6 is a contant input describing rate of endogenous glucose production [mmol/L·day].

Ruan et al. [2015] suggested a number of models for endogenous insulin production in type 2 diabetes, as a

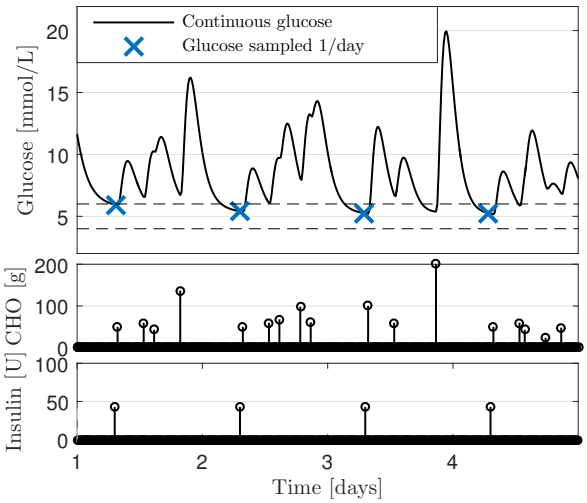

Fig. 1. Four days of simulated glucose data where carbohydrates (CHO) are ingested during the day and long acting insulin is injected bre-breakfast. The blue glucose curve indicates a full evolution of glucose concentration during the period, while the red markers indicate pre-breakfast measured glucose. The dashed lines indicate target range for fasting glucose.

Table 1. Parameter values for the base model (2). Mean (sd) for $p_2 - p_6$ as presented in Kanderian et al. [2009], and chosen or derived parameter values for p_1 and p_7.

Parameter	Unit	Mean (Standard deviation)
p_1 *	[day]	0.5
p_2	[L/day]	1800 (760)
p_3	[1/day]	15.8 (6.2)
p_4	[L/U·day]	792 (560)
p_5	[1/day]	3.31 (3.17)
p_6	[mmol/L·day]	96.7 (63.1)
p_7 †	[U/mmol]	1.4×10^{-3}

function of glucose concentration. The simplest model consists of two parts, a basal rate and a linear increase in production with elevated glucose levels. In this work, we assume a linear relationship between glucose and insulin production, so (1) becomes

$$\frac{dx_1}{dt} = \frac{1}{p_1}\frac{u}{p_2} - \frac{1}{p_1}x_1 \tag{2a}$$
$$\frac{dx_2}{dt} = \frac{1}{p_1}x_1 - \frac{1}{p_1}x_2 \tag{2b}$$
$$\frac{dx_3}{dt} = p_3p_4(x_2 + p_7x_4) - p_3x_3 \tag{2c}$$
$$\frac{dx_4}{dt} = -(p_5 + x_3)x_4 + p_6 \tag{2d}$$

where p_7 is a parameter describing glucose sensitivity of the insulin producing cells in the pancreas [U/mmol] and insulin production is assumed to increase linearly with fasting glucose. We refer to this model as the Base model.

To determine which parameters are identifiable, and for in silico data generation, we use parameter values published by Kanderian et al. [2009]. The time constant p_1 in this paper describes fast acting insulin, and can therefore not be used here. We choose a value for p_1 to roughly describe long acting insulin-glucose dynamics. The parameter describing glucose sensitivity of insulin production is a result of the other parameters, chosen such that fasting glucose has a specific value, $x_4(t=0) = x_{4,0}$. Setting the last two equations of (2) equal to zero, $x_1 = x_2 = 0$ and rewriting, we get

$$p_7 = \frac{1}{p_4x_{4,0}}\left(\frac{p_6}{x_{4,0}} - p_5\right) \tag{3}$$

which is used to determine the value for p_7 based on the mean values of the other parameters.

3.2 Modification for identifyability

The area of greatest interest is around the target glucose values, ranging from approximately 4 to 6 mmol/L. Values for x_3 in this area are of order 10. Typical values for p_5 range from an order of 10^{-5} to 1. We therefore neglect this inverse time constant as it is small relative to the dynamic inverse time constant x_3. The Base model (2) has two gains, insulin clearance rate p_2 and insulin sensitivity p_4. Considering the assumptions in Section 2, data from clinical trials will only contain insulin input, u, and glucose measurements, x_4. Therefore only one gain is identifiable.

We rewrite the model by setting $\tilde{x}_1 = x_1p_2$ [U/day], $\tilde{x}_2 = x_2p_2$ [U/day], $\tilde{x}_3 = x_3p_2/p_4$ [U/day], $\tilde{p}_7 = p_7p_2$ [U·L/mmol·day] and the modified model becomes

$$\frac{d\tilde{x}_1}{dt} = \frac{1}{p_1}u - \frac{1}{p_1}\tilde{x}_1 \tag{4a}$$
$$\frac{d\tilde{x}_2}{dt} = \frac{1}{p_1}\tilde{x}_1 - \frac{1}{p_1}\tilde{x}_2 \tag{4b}$$
$$\frac{d\tilde{x}_3}{dt} = p_3(\tilde{x}_2 + \tilde{p}_7x_4) - p_3\tilde{x}_3 \tag{4c}$$
$$\frac{dx_4}{dt} = -\tilde{p}_4\tilde{x}_3x_4 + p_6 \tag{4d}$$

where we have reduced to one gain, a ratio between the two original gains, $\tilde{p}_4 = p_4/p_2$ [1/U].

3.3 Discretization and model reduction

In order to determine whether the model can be reduced, we linearize the system and investigate properties of a discretization of the system. We linearize the modified model (4) such that

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + Bu(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0$$
$$y(t) = C\mathbf{x}(t)$$

where

$$A = \begin{bmatrix} -\frac{1}{p_1} & 0 & 0 & 0 \\ \frac{1}{p_1} & -\frac{1}{p_1} & 0 & 0 \\ 0 & p_3 & -p_3 & p_3\tilde{p}_7 \\ 0 & 0 & -\tilde{p}_4x_{4,ss} & -\tilde{p}_4x_{3,ss} \end{bmatrix}$$
$$B = \begin{bmatrix} \frac{1}{p_1} & 0 & 0 & 0 \end{bmatrix}^T, \quad C = [0\ 0\ 0\ 1]$$

and $x_{4,ss}$ and $x_{3,ss}$ are steady state values. In this work we linearize around $x_{4,ss} = 5$ mmol/L. All eigenvalues of the matrix A have negative real parts, and the system is asymptotically stable.

Discretizing this system yields

$$\mathbf{x}_{k+1} = A_d\mathbf{x}_k + Bu_k$$
$$y_k = C_d\mathbf{x}_k$$

where

$$\begin{bmatrix} A_d & B_d \\ 0 & I \end{bmatrix} = \exp\left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} T_s\right), \quad C_d = C$$

and T_s is the sampling frequency. Since in Section 2 we assume that the sampling frequency is 1/day, we investigate eigenvalues of A_d for $T_s = 1$ day for the mean parameter values in Table 1. We observe that the real part of two eigenvalues are close to zero. This means that some time constants are not identifiable, and we may want to assume that two compartments reach steady state immediately. In the next section we residualize the modified model (4).

3.4 Residualized model

From Table 1 we observe that $1/p_3 = 0.06$ [day] which is of an order lower than p_1 and small compared to $T_s = 1$ day. We perform residualization for model reduction as

* p_1 is roughly assessed based on knowledge about long acting insulin, e.g. Heise et al. [2017] describe a half-life of approximately 24 hours.

† p_7 is an estimated parameter from mean parameter values. In this work we set $x_4(0) = 8$ mmol/L. This may however be chosen to fit the actual fasting glucose level at zero insulin input.

presented by Skogestad and Postlethwaite [1996], and start by assuming that the compartment $\tilde{x}_3$ reaches steady state immediately following change in $\tilde{x}_1$, $\tilde{x}_2$ and x_4. Setting (4c) to zero yields

$$\tilde{x}_{3,ss} = \tilde{x}_2 + \tilde{p}_7 x_4 \tag{5}$$

Inserting (5) into (4d) gives

$$\begin{aligned} \frac{dx_4}{dt} &= -\tilde{p}_4\tilde{p}_7 x_4^2 - \tilde{p}_4\tilde{x}_2 x_4 + p_6 \\ &= -(\tilde{p}_4\tilde{p}_7 x_4 + \tilde{p}_4\tilde{x}_2)x_4 + p_6 \end{aligned} \tag{6}$$

The value for $x_{2,ss}$ around the area of interest, $x_4 \in [4, 6]$ mmol/L, is $x_2 \in [21.5, 44.8]$ U/day.

Now we define a time constant in (6) such that

$$\frac{dx_4}{dt} = -\frac{x_4}{\tau} + p_6, \quad \tau = \frac{1}{\tilde{p}_4\tilde{p}_7 x_4 + \tilde{p}_4\tilde{x}_2}$$

We find that $\tau \in [0.05, 0.1]$ day which is small compared to p_1, the time constant in the first two compartments. Therefore we assume that x_4 reaches steady state immediately following change in $\tilde{x}_1$ and $\tilde{x}_2$. Setting (6) to zero gives a steady state expression for x_4 (after eliminating the conjugate root due to x_4 is a concentration and therefore $x_4 > 0$),

$$x_{4,ss} = -\frac{\tilde{p}_4\tilde{x}_2 - \sqrt{4\tilde{p}_4\tilde{p}_7 p_6 + \tilde{p}_4^2\tilde{x}_2^2}}{2\tilde{p}_4\tilde{p}_7} \tag{7}$$

and the residualized model becomes

$$\frac{d\tilde{x}_1}{dt} = \frac{1}{p_1}u - \frac{1}{p_1}\tilde{x}_1 \tag{8a}$$

$$\frac{d\tilde{x}_2}{dt} = \frac{1}{p_1}\tilde{x}_1 - \frac{1}{p_1}\tilde{x}_2 \tag{8b}$$

$$y = h(\tilde{x}_2) = -\frac{\tilde{p}_4\tilde{x}_2 - \sqrt{4\tilde{p}_4\tilde{p}_7 p_6 + \tilde{p}_4^2\tilde{x}_2^2}}{2\tilde{p}_4\tilde{p}_7} \tag{8c}$$

where y is the measured glucose concentration in plasma. We refer to this model as the Residualized model. Figure 2 illustrates the output function in the area of interest (blue markers).

4. MODEL IDENTIFICATION

We investigate the identifiability of $\tilde{p}_4, p_6$ and $\tilde{p}_7$ in (8c) through the Fisher Information matrix. For values of $\tilde{x}_2$ in the area of interest, we observe that the Fisher information matrix is nearly singular, so we can not estimate all three parameters simultaneously.

The goal is to make a model of fasting glucose in response to long acting insulin in an area of interest, where parameters and states have a physiological meaning. Inspired by the expression in (8c), we suggest estimating $y = h(\tilde{x}_2)$ with a model of the form

$$\hat{y} = \alpha + \beta\tilde{x}_2 + \gamma\sqrt{1 + \tilde{x}_2} \tag{9}$$

where we might interpret α as fasting glucose at baseline [mmol/L], β as a form of insulin sensitivity [mmol·day/U·L], and γ contributes to describing glucose sensitivity of the insulin producing beta cells [day$^{1/2}$·mmol/L·U$^{1/2}$]. Figure 2 illustrates the output function $h(\tilde{x}_2)$ for the mean parameter values in Table 1 where $h(0) = 8$ mmol/L (blue markers). Fitting the model (9) to these points with a least squares method gives

$$\alpha = 8.89, \quad \beta = -0.003, \quad \gamma = -0.71 \tag{10}$$

The fit is illustrated with a red line in Figure 2.

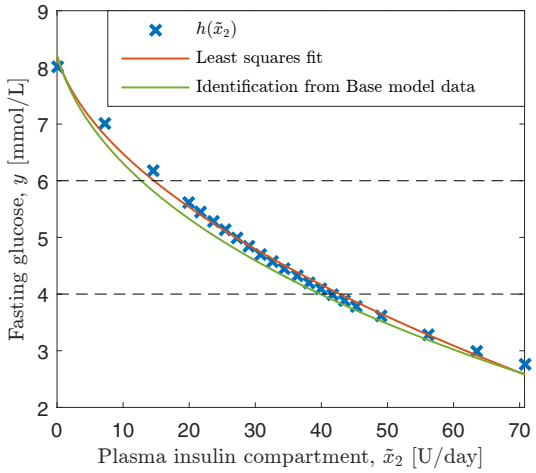

Fig. 2. Output function of the Residualized model in blue, a least squares fit of the model (9) and a fit to the data generated by Base Model in (2).

4.1 Parameter estimation in CTSM-R

For model identification we use CTSM-R, Continuous Time Stochastic Modelling for R, an open source platform for identifying parameters of linear and non-linear grey-box models. Given discrete time series data, CTSM-R can identify parameters of stochastic differential equations. The general model structure is a state space model on the form

$$d\mathbf{x}_t = f(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})d\mathbf{w}_t \tag{11a}$$

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}) + \mathbf{e}_k \tag{11b}$$

where $\boldsymbol{\theta}$ is an l-dimensional set of parameters to estimate, θ is the set of parameters to identify, $\mathbf{u}_t$ is the input at time t, $\boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})^2$ is process noise covariance matrix and $\mathbf{w}_t$ is a Brownian motion path. $\mathbf{y}_k$ is discrete observations and $\mathbf{e}_k$ is the measurement error (assumed to be white Gaussian noise). The model identification is based on maximum likelihood, where the likelihood function is the joint probability density

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) = \left(\prod_{k=1}^{N} p(\mathbf{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta})\right) p(\mathbf{y}_0, \boldsymbol{\theta}) \tag{12}$$

where $\mathcal{Y}_k = [\mathbf{y}_k, \mathbf{y}_{k-1}, \ldots, \mathbf{y}_1, \mathbf{y}_0]$ is a sequence of measurements $\mathbf{y}_k$. CTSM-R considers stochastic differential equations that are driven by Wiener processes and so the conditional densities are approximated by Gaussian densities. A continuous-discrete extended Kalman filter is used for smoothing to determine an estimate for the measurements, $\hat{\mathbf{y}}_{k|k-1} = E[\mathbf{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta}]$, its covariance $\mathbf{R}_{k|k-1} = V[\mathbf{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta}]$, and the innovation where $\epsilon_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$. The density is then

$$p(\mathbf{y}_k|\mathcal{Y}_{k-1}, \boldsymbol{\theta}) = \frac{\exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_k^T \mathbf{R}_{k|k-1}^{-1}\boldsymbol{\epsilon}_k\right)}{\sqrt{\det(\mathbf{R}_{k|k-1})}\left(\sqrt{2\pi}\right)^l} \tag{13}$$

The software allows missing observations. CTSM-R outputs estimates for the parameters, initial conditions and noise, standard deviance of the estimates and the t-statistic, to name a few. The methods and software are described in more detail in [CTSM-R 2015].

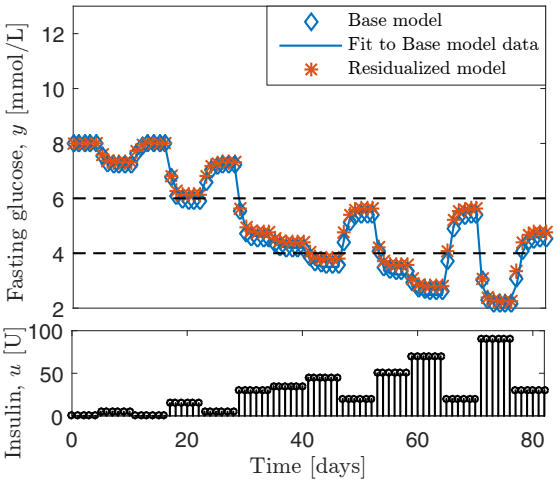

Fig. 3. Simulated glucose data. The blue diamonds indicate a simulation using the Base model, and the red stars indicate data from the Residualized model. The blue line shows the model fit estimated by CTSM-R to the Base model data.

Table 2. Parameter estimation from simulated data by the Base model.

θ	θ^*	$\hat{\theta}$	95% confidence interval *	$p(>\|t\|)$
p_1	0.5	0.55	[0.52, 0.59]	< 0.05
α	8.9	9.09	[8.98, 9.22]	< 0.05
β	-0.003	0.013	[0.008, 0.018]	< 0.05
γ	-0.71	-0.88	[-0.93, -0.83]	< 0.05

5. RESULTS

5.1 Simulated data

We use CTSM-R to identify p_1, α, β and γ from series of simulated data, generated from the Base model. We excite the system with a range of insulin injections such that fasting glucose levels span clinically relevant glucose concentrations; hyperglycemia (> 6 mmol/L), normoglycemia (< 6 mmol/L and > 3.9 mmol/L) and hypoglycemia (< 3.9 mmol/L). This is illustrated in Figure 3.

Table 2 shows the results from identifying the model (8a)-(8b) and (9) in CTSM-R using simulated data generated using the Base model. We observe that all parameter estimates are significant, although θ^* is not inside the confidence interval in all cases. Figure 3 illustrates the identified model in red compared to the least squares fit to the output function (8c).

5.2 Clinical data

In a Phase III clinical trial published by Zinman et al. [2012], 773 adult type 2 diabetes patients were initiated on insulin degludec, once daily injections. During the following months, the treatment was intensified until pre-breakfast (fasting) glucose was at a clinically recommended level. Dose sizes were adjusted once per week, based on home-logged SMPG measurements.

As mentioned in Section 2, clinical data are prone to practical issues regarding data capture and excitation. In

* Confidence intervals calculated as mean $\pm$ 1.96 $\cdot$ SD.

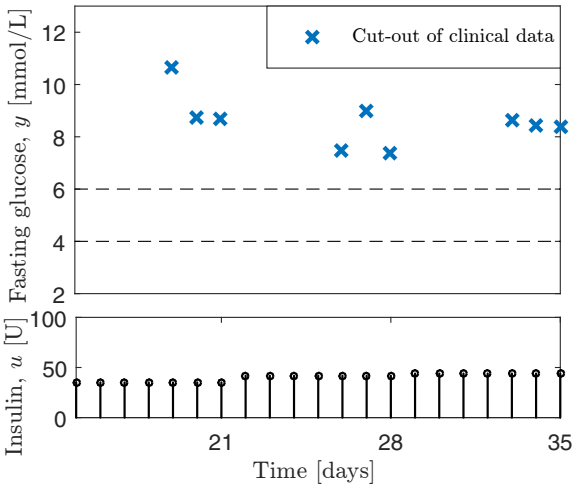

Fig. 4. A cut-out from measured clinical data for one patient in the study by Zinman et al. [2012].

the study by Zinman et al. [2012], pre-breakfast glucose was measured and logged the last three days of the week, along with the corresponding insulin dose. Figure 4 shows a 3 week cut-out from measured data for one patient in the study. This figure illustrates the sparsity of glucose data compared to the insulin data. The full data set for the patient is illustrated in Figure 5. Focus in the clinical trials that we consider relevant is not on system identification but rather on efficacy and safety of the drug. Therefore many data sets do not contain all the appropriate data for model identification. Notice that none of the measured data are lower than 4 mmol/L. The reason is that glucose levels below this value are clinically considered too low, and are therefore not frequently seen in clinical trials. This means that we can expect to only have data for $y \geq 3.9$ mmol/L.

We use CTSM-R to estimate α, β and γ from the measured data in Figure 5. Since p_1 is not identifiable from the data we set $p_1 = 0.5$ days. The resulting parameter estimates, confidence intervals and p-values are listed in Table 3 and Figure 5 illustrates the model fit. We observe that all three parameter estimates are significant.

Table 3. Parameter estimation using measured clinical data from one patient.

θ	$\hat{\theta}$	95% confidence interval *	$p(>\|t\|)$
α	14.6	[12.4, 16.7]	< 0.05
β	0.06	[0.0003, 0.12]	< 0.05
γ	-1.52	[-2.25, -0.79]	< 0.05

6. DISCUSSION

We have identified parameters of a non-linear output function using clinical data. The sampling frequency and range of values in the data are representative for what may be expected from clinical data in relevant clinical studies. This restricted us in estimating the time constant.

The interpretation of the parameters in Section 4 remains open for discussion. We mentioned that β could be interpreted as insulin sensitivity, and that γ could contribute to glucose sensitivity of the insulin producing beta cells.

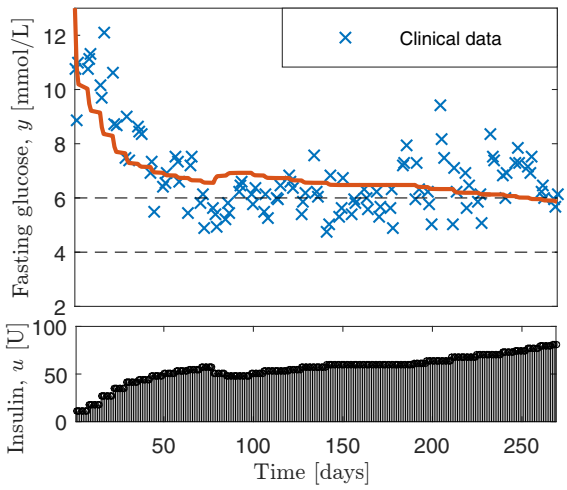

Fig. 5. Measured clinical data from Zinman et al. [2012] and model fit.

We would expect that endogenous insulin production decreases for lower values of glucose. Therefore the total glucose lowering effect of an insulin injection is lower for low values of glucose, and the output function should level off. We might therefore consider β and γ as a combination of insulin sensitivity of the glucose elimination and glucose sensitivity of the insulin production.

7. CONCLUSION

This paper suggests a model of glucose-insulin dynamics in type 2 diabetes using data of low frequency. The purpose of such a model is to simulate fasting glucose levels in long acting insulin treatment, to enable control algorithm development for dose guidance. As a starting point we used the four-compartmental physiological model for 24-hour simulations of glucose concentration following meal intake and injection of fast and long acting insulin. We investigated identifiability of parameters when the sampling period is 1 day, and reduced the model to a two-compartmental model with a non-linear output function. We use the open source software CTSM-R for model identification. All parameter estimates are significant when fitting to data simulated by the six-compartmental physiological model, as well as for the chosen set of clinical data.

Design of experiment in clinical trials remains a challenge for model identification. This work is one step of many in an iterative process to making a dynamical system of insulin-glucose dynamics in long acting insulin treatment of type 2 diabetes.

REFERENCES

American Diabetes Association (2017). *Standards of medical Care in Diabetes.*

Aradóttir, T.B., Boiroux, D., Bengtsson, H., Kildegaard, J., Orden, B.V., and Jørgensen, J.B. (2017). Model for Simulating Fasting Glucose in Type 2 Diabetes and the Effect of Adherence to Treatment. *IFAC-PapersOnLine*, 50(1), 15086–15091.

Bergman, R.N., Ider, Y.Z., Bowden, C.R., and Cobelli, C. (1979). Quantitative estimation of insulin sensitivity. *Am J Physiol*, 23, E667–E677.

CTSM-R (2015). Continuous Time Stochastic Modelling in R, User's Guide and Reference Manual. URL `http://ctsm.info/`.

Dalla Man, C., Rizza, R.A., and Cobelli, C. (2007). Meal simulation model of the glucose- insulin system. *IEEE Trans Biom Engin*, 54(10), 1740–1749.

Heise, T., Nørskov, M., Nosek, L., Kaplan, K., Famulla, S., and Haahr, H.L. (2017). Insulin degludec: Lower day-to-day and within-day variability in pharmacodynamic response compared with insulin glargine 300 U/mL in type 1 diabetes. *Diabetes, Obesity and Metabolism*, 19(7), 1032–1039.

Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., and Wilinska, M.E. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, 25(4), 905–920.

International Diabetes Federation (2017). *IDF diabetes atlas, Eighth edition.*

International Hypoglycaemia Study Group (2016). Glucose Concentrations of Less Than 3.0 mmol/L (54 mg/dL) Should Be Reported in Clinical Trials: A Joint Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes: Table 1. *Diabetes Care*, 40(1), 155–157.

Jauslin, P.M., Frey, N., and Karlsson, M.O. (2011). Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes. *Journal of clinical pharmacology*, 51(2), 153–164.

Kanderian, S., Weinzimer, S., Voskanyan, G., and Steil, G.M. (2009). Identification of intraday metabolic profiles during closed-loop glucose control in individuals with type 1 diabetes. *Journal of diabetes science and technology*, 3(5), 1047–57.

Petersen, M. (2016). Economic costs of diabetes in the U.S. in 2012. *Diabetes Care*, 39(7), 1033–1046.

Røge, R.M., Klim, S., Kristensen, N.R., Ingwersen, S.H., and Kjellson, M.C. (2014). Modeling of 24-hour glucose and insulin profiles in patients with type 2 diabetes mellitus treated with biphasic insulin aspart. *Journal of Clinical Pharmacology*, 54(7), 809–817.

Ruan, Y., Thabit, H., Wilinska, M.E., and Hovorka, R. (2015). Modelling endogenous insulin concentration in type 2 diabetes during closed-loop insulin delivery. *Biomedical engineering online*, 14(1), 19.

Skogestad, S. and Postlethwaite, I. (1996). *Multivariable Feedback Control, Analysis and Design.*

Toffolo, G., Bergman, R.N., Finegood, D.T., Bowden, C.R., and Cobelli, C. (1980). Quantitative estimation of beta cell sensitivity to glucose in the intact organism: aminimal model of insulin kinetics in the dog. *Diabetes*, 29, 979–990.

Wong, K., Glovaci, D., Malik, S., Franklin, S.S., Wygant, G., Iloeje, U., Kan, H., and Wong, N.D. (2012). Comparison of demographic factors and cardiovascular risk factor control among U.S. adults with type 2 diabetes by insulin treatment classification. 26, 169–174.

Zinman, B., Philis-Tsimikas, A., Cariou, B., Handelsman, Y., Rodbard, H.W., Johansen, T., Endahl, L., and Mathieu, C. (2012). Insulin Degludec Versus Insulin Glargine in Insulin-Naive Patients With Type 2 Diabetes. *Diabetes Care*, 35(12), 2464 LP – 2471.