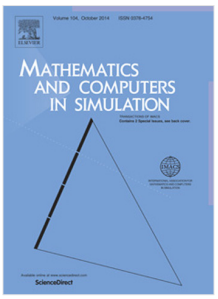
Exploring dynamical properties of a Type 1 diabetes model using sensitivity approaches

Al Ali, H., Daneshkhah, A., Boutayeb, A., Malunguza, N. J. & Mukandavire, Z

DOI 10.1016/j.matcom.2022.05.008
ISSN 0378-4754

Publisher: Elsevier

Journal Pre-proof



Exploring dynamical properties of a Type 1 diabetes model using sensitivity approaches

Hannah Al Ali, Alireza Daneshkhah, Abdesslam Boutayeb, Noble Jahalamajaha Malunguza, Zindoga Mukandavire

PII: S0378-4754(22)00196-3
DOI: https://doi.org/10.1016/j.matcom.2022.05.008
Reference: MATCOM 5716

To appear in: *Mathematics and Computers in Simulation*

Received date : 20 June 2021
Revised date : 2 April 2022
Accepted date : 9 May 2022

Please cite this article as: H. Al Ali, A. Daneshkhah, A. Boutayeb et al., Exploring dynamical properties of a Type 1 diabetes model using sensitivity approaches, *Mathematics and Computers in Simulation* (2022), doi: https://doi.org/10.1016/j.matcom.2022.05.008.

Exploring dynamical properties of a Type 1 diabetes model using sensitivity approaches

Hannah Al Ali[1,2,3]*, Alireza Daneshkhah[1,4], Abdesslam Boutayeb[5], Noble Jahalamajaha Malunguza[6], Zindoga Mukandavire[2,3]

[1] *Centre for Computational Science and Mathematical Modelling, Coventry University, UK*

[2] *Institute of Applied Research and Technology, Emirates Aviation University, Dubai, UAE*

[3] *Centre for Data Science and Artificial Intelligence, Emirates Aviation University, Dubai, UAE*

[4] *School of Computing, Electronics and Mathematics, Coventry University, UK*

[5] *Department of Mathematics, Faculty of Sciences, University Mohamed 1er, Oujda, Morocco*

[6] *Department of Insurance and Actuarial Science, National University of Science and Technology, Bulawayo,Zimbabwe*

Abstract

The high global prevalence of diabetes, and the extortionate costs imposed on healthcare providers necessitate further research to understand different perspectives of the disease. In this paper, a mathematical model for Type 1 diabetes glucose homeostasis system was developed to better understand disease pathways. Type 1 diabetes pathological state is shown to be globally asymptomatically stable when the model threshold $\mathcal{T}_0 < 1$, and exchanges stability with the managed diabetes equilibrium state i.e. globally asymptotically stable when $\mathcal{T}_0 > 1$. Sensitivity analysis was conducted using partial rank correlation coefficient (PRCC) and Sobol′ method to determine influential model parameters. Sensitivity analysis was performed at different significant time points relevant to diabetes dynamics. Our sensitivity analysis was focused on the model parameters for glucose homeostasis system, at 3 to 4 hour time interval, when the system returns to homeostasis after food uptake. PRCC and Sobol′ method showed that insulin clearance and absorption rates were influential parameters in determining the model response variables at all time points at which sensitivity analysis was performed. PRCC method also showed the model subcutaneous bolus injection term to be important, thus identified all parameters in $\mathcal{T}_0$ as influential in determining diabetes model dynamics. Sobol′ method complemented the sensitivity analysis by identifying relationships between parameters. Sensitivity analysis methods concurred in identifying some of the influential parameters and demonstrated that parameters which are influential remain so at every time point. The concurrence of both PRCC and Sobol′ methods in identifying influential parameters (in $\mathcal{T}_0$) and their dynamic relationships highlight the importance of statistical and mathematical analytic approaches in understanding the processes modelled by the parameters in the glucose homeostasis system.

Keywords: Diabetes model, equilibria, stability, Gaussian process, sensitivity analysis.

*Correspondence: Hannah Al Ali, Email: alalih3@uni.coventry.ac.uk,hannah.alali@emirates.com

1 Introduction

As of 2019, a total of US$760 billion had been spent on diabetes, representing 10% of total global health expenditure [1]. This is set to increase as global prevalence of the disease increases [2] and recently due to COVID-19 infection which makes diabetes treatment difficult due to fluctuations in blood glucose levels [1]. Diabetes, has two main forms that are, Type 1 (insulin dependent diabetes) and Type 2 (non-insulin dependent diabetes). Globally, the number of patients with diabetes in 2019 was 463 million, of which 10% were of Type 1 [1].

Type 1 diabetes is classified as an autoimmune disease (a disease where the immune system mistakenly attacks the body [3]). The immune system attacks the β-cells, which are responsible for producing insulin, therefore preventing production of insulin. Thus as the β-cells are destroyed, very few (if any) β-cells remain in the body, resulting in little or no insulin available in the body. Therefore, biologically it is assumed that a Type 1 diabetic has negligible β-cells in their body [4–6]. As blood glucose levels rise due to food uptake, insulin plays a significant role in controlling the blood glucose back to normal levels [6]. Symptoms of the disease are increased thirst, hunger, food intake, urination, weight loss, blurred vision and extreme tiredness [6]. If not treated, diabetes may cause heart disease, kidney failure, nerve damage, coma and eventually death [1, 4]. Chronic elevation of blood glucose levels (hyperglycemia) over long periods of time, due to lack of insulin, results in complications such as cardiovascular disease [7]. Individuals with Type 1 diabetes therefore need daily exogenous insulin dosages in order to control their blood glucose levels. Without the administration of insulin, the individual would die [1, 6]. Insulin injections can be delivered as insulin bolus or continuous insulin injections. Alternatively an insulin pump can also be used [8]. Insulin pumps are open loop devices and are not automated. Recently, an artificial pancreas providing an automated insulin delivery and eliminating the need for human intervention to calculate dosages has gone into trial [9–11].

Mathematical modeling is an important tool to better understand insulin and its analogues in vivo dynamics in order to design future treatment approaches for individuals with Type 1 diabetes [8]. Several types of models have been formulated for Type 1 diabetes, depending on the forms of insulin delivery. Currently there are models for depot injections of insulin analogs, and compartmental and systemic models [12–18]. Most of these models are based on [12] which assumed that insulin absorption is inversely proportional to concentration of insulin in the body [8]. Systems made up of nonlinear differential equation with non-autonomous insulin dosages would be of interest to provide a different perspective on current models [8, 12–21]. Existing models on diabetes do not capture important biological processes. For example, only one mathematical model has so far incorporated the role of the growth hormone [22] and most of the other models represent insulin molecules rather than the system as a whole (i.e glucose, insulin and growth hormone). In addition, existing mathematical models of diabetes have not fully modeled Type 1 diabetes pathway, which describes the zero insulin steady state [8, 17–20]. In this study, we propose a simple Type 1 diabetes model with an insulin bolus injection com-

ponent. Exploring the mathematical properties of such a model is important in understanding the key parameters for insulin management [8]. Glucose homeostasis models, that are sometimes used to model Type 1 diabetes, do not take into account the fact that Type 1 diabetic individuals have no β-cells [22] and this is a major drawback of such models.

In this study rigorous analysis of the model is conducted using classical mathematical analytic approaches and global sensitivity analysis methods. We use the concept of threshold quantities to provide insights on the important model parameters [23–29]. Global sensitivity analysis methods used in this study are partial rank correlation coefficient (sampling-based method) [30] and Sobol′ method (variance-based method) [31].

2 Model formulation

In this study, we developed a diabetes model consisting of the following variables: insulin (I), glucose (G_L) and growth hormone (H). Insulin (I) is secreted by the β-cells and is dependent on the glucose level within the body, therefore if there are no β-cells, no insulin is produced [32]. We thus included a subcutaneous insulin injection term (I_0), which represents a bolus value. The injection is done up to 3 times a day (15-30 minutes before meals depending on blood glucose levels) [11]. The insulin levels in the blood are a product of the amount of insulin externally injected and the absorption rate, ψ. The insulin injection term I_0, is assumed to have an inversely proportional relationship with insulin concentration in the blood (I) [12, 33–35]. We model this relationship using the term $\frac{I_0}{1+I}$. The choice of the function is a new formulation term to clearly capture the state with zero-insulin. The 1 mIU/ml in the term is an assumed shape value to model a zero-insulin state. Overtime, blood insulin level drops as glucose is absorbed by muscle, fat and liver cells and clears at a constant rate δ. Glucose (G_L) level is increased by the growth hormone through suppression of glucose uptake by insulin, at a constant rate c. The parameter a represents average glucose obtained from carbohydrate intake and body production. The growth hormone(H) in model system (1) is increased by the rate of production by the somatotropic cells in the pituitary gland at a constant rate ρ. Growth hormone is decreased by the rate of w due to the absorption by the liver [36]. It has been demonstrated [37], that growth hormone increases glucose production in blood through gluconeogenesis and glycogenosis [38, 39]. Model variables, parameter values and their symbols are provided in Table 1. The model dynamics are governed by the following system of differential equations.

$$\left.\begin{aligned} \frac{dI}{dt} &= \frac{\psi I_0 I}{1+I} - \delta I, \\ \frac{dG_L}{dt} &= a - (b + cI)G_L + cH, \\ \frac{dH}{dt} &= \rho - wH. \end{aligned}\right\} \tag{1}$$

A summary description of model variables and parameter values is given in Table 1.

Parameter/variable definition	Symbol	Baseline value[Range]	Unit	Reference
Biological parameters				
Glucose production rate	a	$864[850-20000]$	$mg/dl\,min$	Assumed
Glucose clearance rate independent of insulin	b	$1.44[1-5]$	min^{-1}	Assumed
Insulin induced glucose uptake rate	c	$0.85[0.1-1]$	$ml/mIU\,min$	Assumed
Growth hormone production rate by somatotropic cells	ρ	$15.06[5-30]$	$mIU/ml\,min$	[36]
Growth hormone clearance rate by the liver	w	$1958.40[2000-4000]$	min^{-1}	[36]
Insulin absorption rate	ψ	$0.2143[0.1-1]$	min^{-1}	Assumed
Insulin clearance rate	δ	$0.0215[0.01-1]$	min^{-1}	[21]
Insulin bolus	I_0	$5[5-30]$	mIU/ml	[21]
Model response variables		**Range**		
β-cells	β	$600-1000$	mg	[4]
Insulin	I	$0-25$	mIU/ml	[40]
Glucose	G_L	$70-200$	mg/dl	[41]
Growth hormone	H	$10-40$	mIU/ml	[42]

Table 1: Model parameters, variables and their definition. *Note that baseline values are from given references and associated ranges are assumed values for sensitivity analysis.

3 Mathematical analysis and results

3.1 Model basic properties

The model system (1) has an initial condition given by $I(0) \geq 0, G_L(0) \geq 0$, and $H(0) \geq 0$. Since the model represents fluid concentrations in the human body, all variables should be non-negative for biological feasibility in the following region,

$$\mathcal{D} = \{(I, G_L, H) \in \mathbb{R}^3_+\}.$$

Hence we establish the following result and proof in Theorem 1.

Theorem 1. *The region $\mathcal{D} \subset \mathbb{R}^3$ is positively invariant with respect to the system of equations and non-negative solutions exist $\forall\ 0 < t < \infty$. Let the initial data be $I(0) > 0, G_L(0) > 0$, and $H(0) > 0$, then solutions $(I(t), G_L(t), H(t))$ of model system (1) with positive initial data will remain positive $\forall\ t > 0$.*

Proof. Suppose that $t_1 = \sup\{t > 0 : \beta > 0, I > 0, G_L > 0, H > 0, \in [0, t]\}$. Under the given initial conditions it can be shown that solutions of model system (1) are positive for $t > 0$. We show that this is true $\forall\ t > 0$ by proceeding as follows. The first equation in model system (1) is given by

$$I'(t) = \frac{\psi I_0 I}{1 + I} - \delta I,$$

which gives

$$\frac{d}{dt} \ln I(t) \geq -\delta \implies I(0) \exp\left\{ -\int_0^{t_1} \delta\ dt \right\} > 0.$$

It follows that the solution to the equation is positive $\forall\ t > 0$. In a similar fashion, we provide the proof for each equation in model system (1) as follows. For G_L we have

$$\frac{d}{dt}\ln G_L(t) \geq -(b+cI) \implies G_L(0)\exp\left\{-\int_0^{t_1}\left(b+cI\right)\,dt\right\} > 0.$$

Similarly H gives

$$\frac{d}{dt}\ln H(t) \geq -w \implies H(0)\exp\left\{-\int_0^{t_1} w\,dt\right\} > 0.$$

Thus, solutions for model system (1) are positive $\forall\ t > 0$ hence the model is biologically well-posed. □

3.2 Model equilibria

Model system (1) has two steady states which are as follows:

The **diabetes (pathological) equilibrium state** is given by

$$P_0(I^*, {G_L}^*, H^*) = \left\{\ 0, \frac{aw+c\rho}{bw}, \frac{\rho}{w}\ \right\}. \tag{2}$$

The **managed diabetes equilibrium state** is given by

$$P_1(I^{**}, {G_L}^{**}, H^{**}) = \left\{\ \frac{\psi I_0}{\delta} - 1, \frac{c\rho\delta + aw\delta}{\psi c I_0 w + bw\delta - cw\delta}, \frac{\rho}{w}\ \right\}. \tag{3}$$

From equation (3), the managed diabetes state exists if $\frac{I_0\psi}{\delta} - 1 > 0$ implying that $\frac{I_0\psi}{\delta}$ is the threshold parameter, $\mathcal{T}_0$. $\mathcal{T}_0 = 1$ becomes a bifurcation point above which a diabetic individual has control of diabetes and below which, the individual is diabetic and failing to manage the disease as they will be in a state of hyperglycemia. There are two solutions for I,

$$I^* = 0 \quad \text{and} \quad I^{**} = \frac{I_0\psi}{\delta} - 1 = \mathcal{T}_0 - 1.$$

When $I^* = 0$, we obtain a Type 1 pathological state (no insulin) and when $I^{**} = \mathcal{T}_0 - 1$ we get the managed diabetes equilibrium state.

3.3 Stability of equilibria

We use the threshold parameter $\mathcal{T}_0$ to investigate the stability of both P_0 and P_1.

3.3.1 Local stability of P_0

Linearising model system (1) gives the following Jacobian matrix.

$$\mathcal{J} = \begin{pmatrix} \dfrac{\psi I_0}{I^*+1} - \dfrac{\psi I^* I_0}{(I^*+1)^2} - \delta & 0 & 0 \\ -cG_L^* & -b - cI^* & c \\ 0 & 0 & -w \end{pmatrix} \tag{4}$$

We use the $\mathcal{J}$ to determine local stability of the steady states in the following sections. The equilibrium state P_0 is a pathological steady state as the individual has Type 1 diabetes ($I^* = 0$). Solving $\mathcal{J}$ at the pathological equilibrium P_0 gives the following eigenvalues, $\lambda_1 = -b, \lambda_2 = -w$ and $\lambda_3 = \frac{\psi I_0}{(1)^2} - \delta$. Equilibrium P_0 is defined as stable if $\lambda_3 < 0$ which occurs when $\mathcal{T}_0 < 1$.

Lemma 1. *The pathological state P_0 is locally stable for $\mathcal{T}_0 < 1$.*

Theorem 2. *The managed diabetes steady state P_1 of system (1) is locally asymptotically stable whenever it exists.*

Proof. Linearising the system at P_1 we obtain the following eigenvalues at P_1, $\lambda_1 = -b - c\mathcal{T}_0 + c < 0$, $\lambda_2 = -w < 0$ and $\lambda_3 = \delta\left(\dfrac{\delta}{\psi I_0} - 1\right)$. Therefore for P_1 to be stable, $\lambda_3 < 0$. We can rewrite λ_3 as the following, $\lambda_3 = \delta\left(\dfrac{1}{\mathcal{T}_0} - 1\right)$. On solving $\lambda_3 < 0$ we obtain $\mathcal{T}_0 > 1$. □

This means that when diabetic individuals are in the managed diabetic state, they will remain in that state for as long as the threshold quantity $\mathcal{T}_0 > 1$.

3.3.2 Global stability of P_0 and P_1

Theorem 3. *If $\mathcal{T}_0 < 1$, the pathological state P_0 is globally asymptotically stable.*

Proof. Define a Lyapunov function

$$\mathcal{L}(I(t), G_L(t), H(t)) = I + k_1 G_L + k_2 H$$

with constants k_1 and k_2 to be defined such that the derivative of $\mathcal{L}(t)$ is negative definite. Let

$$k_1 = \frac{bw(\psi I_0 - \delta)}{aw + c\rho} \text{ and } k_2 = \frac{w(\psi I_0 - \delta)}{\rho}.$$

$$\begin{aligned} \mathcal{L}(I^*, G_L^*, H^*) &= 0 + \frac{bw(\psi I_0 - \delta)}{aw + c\rho}\left(\frac{aw + c\rho}{bw}\right) + \frac{w(\psi I_0 - \delta)}{\rho}\left(\frac{\rho}{w}\right) \\ &= 2(\psi I_0 - \delta) \end{aligned}$$

Thus $\mathcal{L}(I, G_L, H) = \mathcal{L}(I^*, G_L^*, H^*) = 0$ if and only if $\psi I_0 = \delta$ i.e the insulin absorption bolus will be equal to the clearance rate and $\mathcal{L}(I, G_L, H) > 0$ hence $\mathcal{L}(I, G_L, H) \geq 0$ in P_0
Then

$$\begin{aligned}
\frac{d}{dt}\mathcal{L}(t) &= \frac{d}{dt}I(t) + k_1\frac{d}{dt}G_L(t) + k_2\frac{d}{dt}H(t),\\
&= \left(\frac{\psi I_0}{1+I} - \delta\right) I + \frac{bw(\psi I_0 - \delta)}{aw + c\rho}\frac{d}{dt}G_L(t) + \frac{w(\psi I_0 - \delta)}{\rho}\frac{d}{dt}H(t),\\
&\leq \left(\frac{\psi I_0}{1+I} - \delta\right) I + (\psi I_0 - \delta)\left[\left(\frac{bw}{aw + c\rho}\right) G_L(t) + \left(\frac{w}{\rho}\right) H(t)\right],\\
&\leq (\psi I_0 - \delta)\, I + (\psi I_0 - \delta)\left[\left(\frac{bw}{aw + c\rho}\right) G_L(t) + \left(\frac{w}{\rho}\right) H(t)\right],\\
&= (\psi I_0 - \delta)\left[I + \left(\frac{bw}{aw + c\rho}\right) G_L(t) + \left(\frac{w}{\rho}\right) H(t)\right],\\
&= \delta\left(\mathcal{T}_0 - 1\right)\left[I + \left(\frac{bw}{aw + c\rho}\right) G_L(t) + \left(\frac{w}{\rho}\right) H(t)\right],\\
&\leq 0.
\end{aligned}$$

Using the Lyapunov stability theorem $\frac{d\mathcal{L}(t)}{dt}$ is negative definite. The ω-limit set of each solution is the largest invariant set for which $I = I^*$, $G_L = G_L^*$ and $H = H^*$ for which P_0 is a singleton. By LaSalle's invariance principle [43], the pathological state P_0 is globally asymptotically stable in $\mathcal{D}$. □

This shows us that individuals with managed diabetes will remain in this state whenever the threshold quantity $\mathcal{T}_0 < 1$. This confirms that Type 1 diabetes is a non-reversible condition when it exists.

Theorem 4. *The managed diabetes state P_1 is globally asymptotically stable for $\mathcal{T}_0 > 1$.*

Proof. Let $I = x_1$,$G_L = x_2$ and $H = x_3$ and consider a possible Lyapunov function

$$\mathcal{V}(x) = \left(x_1 - x_1^{**} - x_1^{**}\ln\left[\frac{x_1}{x_1^{**}}\right]\right) + \left(x_2 - x_2^{**} - x_2^{**}\ln\left[\frac{x_2}{x_2^{**}}\right]\right) + \left(x_3 - x_3^{**} - x_3^{**}\ln\left[\frac{x_3}{x_3^{**}}\right]\right).$$

At steady state $x_3^{**} = \frac{\rho}{w} \Rightarrow \rho = wx_3^{**}$. Thus

$$\begin{aligned}
\dot{\mathcal{V}} &= \left(\frac{\psi I_0}{1+x_1} - \delta\right)(x_1 - x_1^{**}) + [a + cx_3 - (b + cx_1)x_2]\left(1 - \frac{x_2^{**}}{x_2}\right) + \rho - \rho\frac{x_3^{**}}{x_3} - wx_3 + wx_3^*,\\
&= \left(\frac{\psi I_0}{1+x_1} - \delta\right)(x_1 - x_1^{**}) + [a + cx_3 - (b + cx_1)x_2]\left(1 - \frac{x_2^{**}}{x_2}\right) - \frac{w}{x_3}\left(x_3^{**} - x_3\right)^2,\\
&\leq 0
\end{aligned}$$

since the expressions $\left(\frac{\psi I_0}{1+x_1} - \delta\right)$ and $[a - (b + cx_1)x_2 + cx_3]$ are positive by definition of model system (1) and $x_i \leq x_i{}^{**}$ everywhere in $\mathcal{D}$. We used the Lyapunov stability theorem to show that $\dot{\mathcal{V}} < 0$ for all $(I^*, G_L{}^{**}, H^{**}) > 0 \in \mathcal{D}$ and the strict equality $\dot{\mathcal{V}} = 0$ holds only for

$I = I^{**}, G_L = {G_L}^{**}$ and $H = H^{**}$. The equilibrium state P_1 is the only positively invariant set of the solution for model system (1) contained entirely in $\mathcal{D}$. By the asymptotic stability theorem in [43], the managed diabetes state P_1 is globally asymptotically stable. □

P_1 equilibrium state is also shown to be globally stable when $\mathcal{T}_0 > 1$. This is a case where an individual has well managed diabetes. This state is in line with biological findings that individuals with well managed diabetes will have a balanced glucose homoeostasis system.

3.4 Numerical simulations and results

In order to illustrate some of the mathematical analysis, numerical simulations of model system (1) are conducted using a code in R programming environment [44] and parameter values in Table 1. Figures 1 and 2 illustrate the time series plots based on simulating the model with different initial conditions. Figure 1 shows the solution profiles for the concentration of I, G_L and H for $\mathcal{T}_0 < 1$. Simulation results in Figure 1 show that solutions will converge to the Type 1 diabetic steady state (as in Lemma 1). The glucose levels are approximately 500 $\frac{mg}{dl}$ and insulin levels are at zero, a hyperglycemic state. Figure 2 shows the solution profiles for the concentration of I, G_L and H for $\mathcal{T}_0 > 1$ and this also confirms the non-diabetic steady state is also stable (as in Theorem 2). The glucose and insulin levels are within the normal range, and no hyperglycemic state is occurring. Both results in Figures 1 and 2 agree with the biology of the disease that Type 1 diabetes is a non-reversible stable state as illustrated by a forward bifurcation in Figure 3.

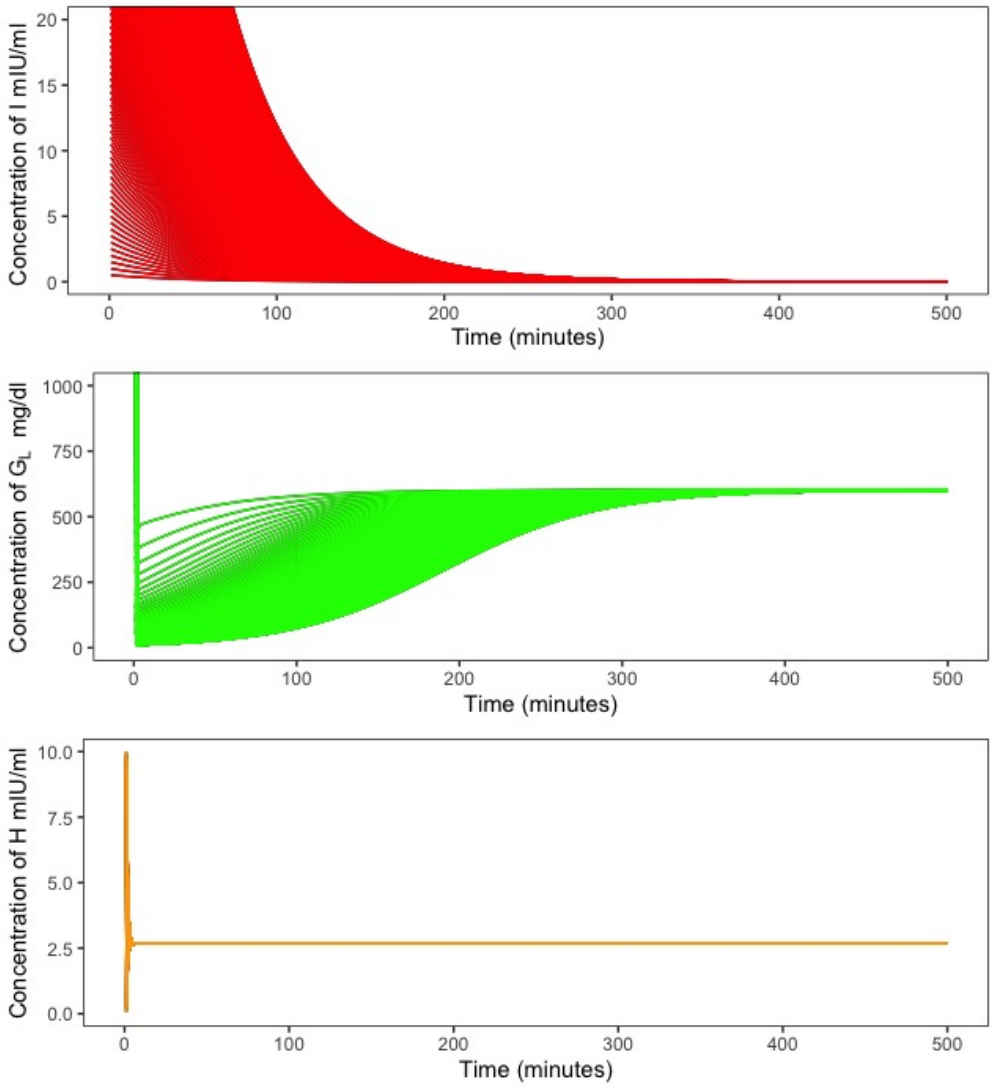


Figure 1: Simulations of model system (1) with different initial conditions for $\mathcal{T}_0 = 0.049837 < 1$. Parameter values used are as in Table 1 with $I_0 = \frac{j}{2}$ and $\psi = 2.143 \times 10^{-4}$ where j is the step value which is varied in the range $1 - 200$. Note that the y-axis scale for the figures is different in order to make figures clearer.

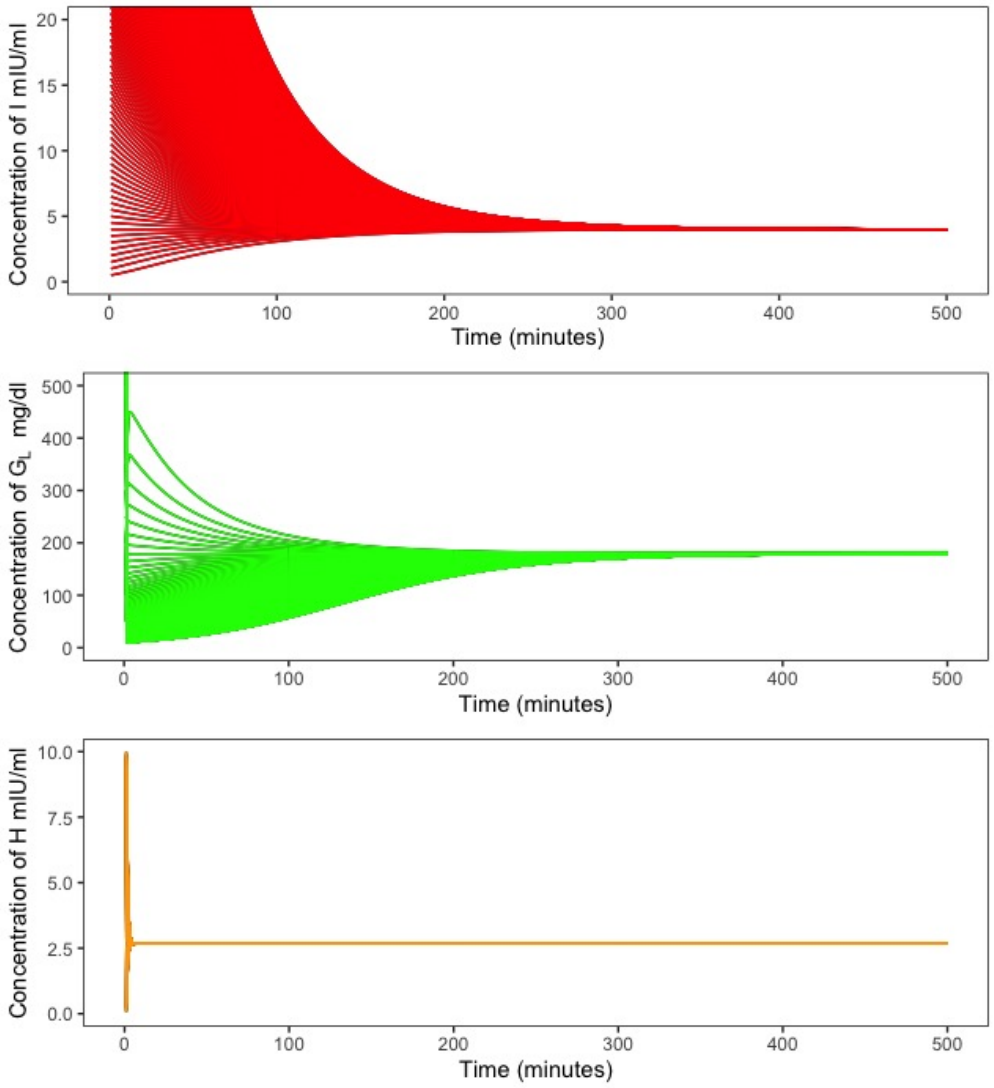


Figure 2: Simulations of model system (1) with different initial conditions for $\mathcal{T}_0 = 4.983721 > 1$. Parameter values used are as in Table 1 with $I_0 = \frac{j}{2}$ and $\psi = 2.143 \times 10^{-2}$ where j is the step value which is varied in the range $1 - 200$. Note that the y-axis scale for the figures is different in order to make figures clearer.

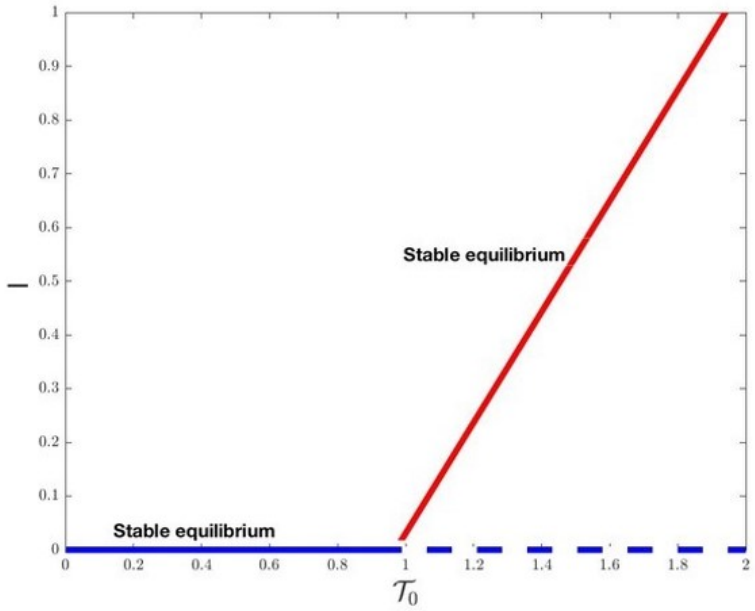


Figure 3: Bifurcation diagram showing a forward transcritical bifurcation occurring. The bold blue line represents a Type 1 diabetic, the dashed blue line represents a unstable diabetic equilibrium and the red represents a stable non-diabetic equilibrium.

The solution for $I^* = 0$ is given by the bold blue/dashed line and a Type 1 diabetic. The second solution for $I^{**} = \mathcal{T}_0 - 1$ is given by the red line and is a non-diabetic state.

4 Sensitivity analysis and results

In this section, various sensitivity analysis (SA) methods were used to assess the relative importance of the input parameters when varied over wide ranges (as given in Table 1) to the model outputs (I, G_L, H) which are derived by solving model system (1). Here we mainly focus on the SA of G_L and I which are the most important components in the glucose homeostasis system and in managing diabetes. We will begin by conducting SA using the partial rank correlation coefficient (PRCC) and then proceed to use probabilistic SA methods. The partial rank correlation coefficient, as one of the widely used global SA approaches will be briefly introduced in Section 4.1. The PRCC values for each input parameter and their corresponding p-values are computed in Matlab Statistics and Machine Learning toolbox (R2019b) [45]. We introduce several probabilistic SA methods, including main and interaction effects and Sobol′ method in Section 4.2, including the Gaussian process. We also develop a computational algorithm using the Gaussian process emulator to efficiently evaluate these SA measures. The SA measures proposed for the Sobol′ method are computed using *tgp* package in R [46].

4.1 Partial Rank Correlation Coefficient

PRCC and their corresponding p-values are used to evaluate parameter importance on the model outputs. The method is combined with Latin hypercube sampling and explores the

entire parameter space [30]. The PRCC values illustrate the correlation between the model outputs (I, G_L, H) and the input parameters. PRCC will give the singular effect of each input parameter on the model output of interest. The corresponding p-values highlight the level of uncertainty of each input parameter on the model output. The input parameters with larger PRCC values are those which have more impact on the model output, and the ones with relatively insignificant values could be removed from the model as they are regarded of being less important (see [30, 47] for similar analysis). The input parameters with $p < 0.05$ are regarded to have significant impact on the model output. Scatter plots were also generated to visually illustrate the relationship between input parameters and model outputs at time $t = 210$ minutes. Scatter plots showing sensitivity analysis results of input parameters $(a, b, c, \delta, I_0, \psi, \rho, w)$ against I are in *Supplementary Figures S1 and S2*. The PRCC results for the entire time period and corresponding p-values for all the parameters against I are shown in Table 2 and illustrated in Figure 4 (a). The results suggest that the parameters that are most influential on I were δ, I_0 and ψ. In exploring most influential parameters on I, we calculate the PRCC and p-values at different time points. Initial time point ($t = 5$ minutes), is called the "*fasting*" level in an individual and usually observed in the morning. However, it can also represent 3 hours post food as the system should reach homeostasis within 3-4 hours. The second time period of interest is immediately after food, when glucose is high due to the ingested source of glucose entering the blood stream. This is at $t = 10$ minutes, where we assume the meal is taken within 5-10 minutes after waking up. The third time is $t = 60$ minutes, an hour postpandrial. This is when glucose level should be reducing towards homeostasis. Time points $t = 90$ and 180 minutes, corresponds to 2 and 2.30 hours postpandrial meaning if an individual was not diabetic or had good management of their diabetes the glucose, insulin and growth hormone level should be nearly at homeostasis levels. Finally $t = 210$ minutes when glucose level should be normal. The remaining PRCC tables for each time point are in *Supplementary Tables S1-S31*. Results demonstrate that, regardless of time point, the parameters which are significant remain significant. Parameters identified as influential are parameters that make up the Type 1 diabetes threshold quantity.

Scatter plots showing sensitivity analysis results of input parameters $(a, b, c, \delta, I_0, \psi, \rho, w)$ against G_L are in *Supplementary Figures S3 and S4*. The PRCC results and corresponding p-values for all the parameters against G_L are shown in Table 3 and illustrated in Figure 4 (b). The results suggest that the parameters that are most influential on G_L were $\delta, \psi, I_0, \rho, w$. Parameters δ and ρ have positive PRCC values suggesting that these parameters have positive effect on glucose concentration thus are important in maintaining glucose homeostasis. These results also show the importance of growth hormone in the glucose homeostasis system as parameter w has shown to influence G_L. Model parameters which have shown to be significant remain significant at different time points and after $t > 90$ minutes, two extra parameters are highlighted as significant and these are ρ and w. The remaining PRCC tables are shown in *Supplementary Tables S32-S61*.

The scatter plots showing sensitivity analysis results against H are shown in the *Supplementary*

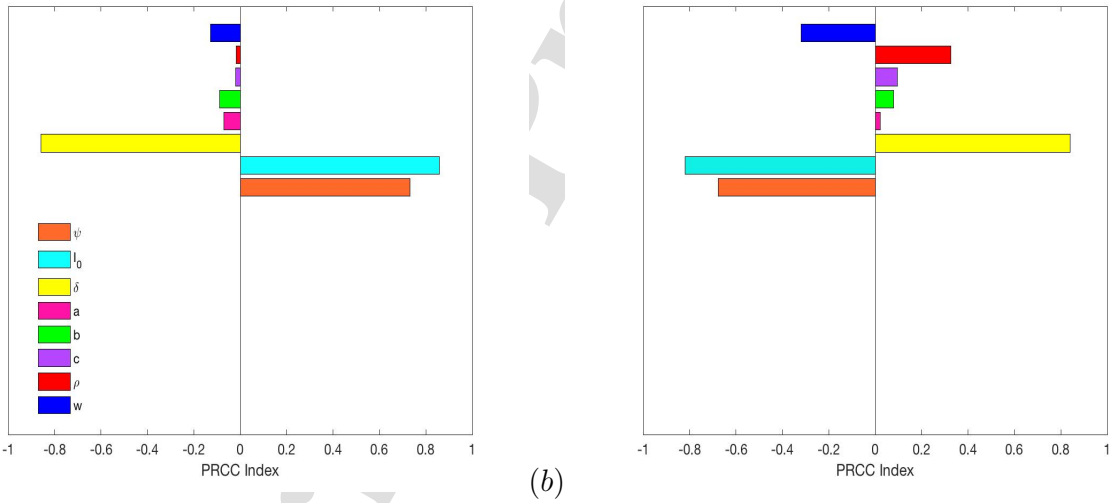
Figures S3 and S4. The PRCC results and corresponding p-values for all the parameters against H are shown in *Supplementary Tables S62-S91* and illustrated in *Supplementary Figures S5 and S6.* The results show that the parameters that are most influential against H are ρ and w.

Parameter	p-values	PRCC
ψ	$p < 0.0001$	0.7758
I_0	$p < 0.0001$	0.8550
δ	$p < 0.0001$	-0.8009
w	$p = 0.3003$	0.0629
c	$p = 0.4578$	0.0422
ρ	$p = 0.5555$	-0.0654
a	$p = 0.8186$	-0.0634
b	$p = 0.8397$	0.0278

Table 2: PRCC sensitivity analysis of parameters ranked in terms of importance to the model variable I for entire time period.

Parameter	p-values	PRCC
ψ	$p < 0.0001$	-0.8307
I_0	$p < 0.0001$	-0.8795
δ	$p < 0.0001$	0.8816
ρ	$p < 0.0001$	0.4534
w	$p < 0.0001$	-0.3081
c	$p = 0.0104$	-0.0238
b	$p = 0.9627$	0.2222
a	$p = 0.9856$	0.2869

Table 3: PRCC sensitivity analysis of parameters ranked in terms of importance to the model variable G_L for entire time period.



(a) (b)

Figure 4: Plot (a) shows a tornado plot of the parameters with their PRCC values showing the effect of input parameters on I and (b) is a tornado plot of the parameters with their PRCC values showing the effect of input parameters on G_L.

4.2 Probabilistic sensitivity analysis

In addition to the PRCC method, we employ the variance-based SA methods as more efficient global SA methods to evaluate the relative importance of input parameters when they are altered extensively. This would allow us to take into account inputs uncertainty as they vary over a wide range. One of the motivations to use these efficient probabilistic SA methods is that the system is complex in regard to the relationships between the inputs and output that are highly non-linear. In addition, PRCC as the common regression analysis-based global SA

method, assumes that there must be a monotonic relationship between the output and each input parameter of interest, which is often violated by the underlying input-output relationship exhibited by the system of interest in this paper [48, 49]. Furthermore, the PRCC-based approach is not capable to evaluate the uncertainty levels of each input parameter affecting the model outputs. Finally, the variance-based SA methods are able to allocate the variance of the output and quantify the effect of high-order interactions between input parameters, but PRCC method is not able to evaluate the impact of the interactions between inputs.

The probabilistic global SA methods of interest in this study is based on the analysis of variance of the model response variable [31]. The approach can capture the fraction of the model response variable variance explained by model input on its own or by a group of model inputs. In addition, it can also provide the total contribution to the output variance of a given input (i.e. its marginal and cooperative contribution). The main challenge of this approach, for the costly system under study, is in computing the Sobol′ indices, and other variance-based SA measures, including main effects, the variance contributions of each input parameter to the model output, and corresponding uncertainty levels. There are different computational techniques to perform Sobol′ method SA [31,50,51]. This study reports the final results of sensitivity indices computed using the emulator-based method [52, 53], which will be briefly discussed in Section 4.2.2.

To perform the variance-based SA methods, we will examine how a function $f(\mathbf{x})$ depends on its input variables. For the case of this study, $f(.)$ will typically be the function that computes I, G_L and H as a function of a vector of biological input parameters illustrated in Table 1. Important notations that will appear in the next sections are introduced in the following. We denote a d-dimensional random vector as $\mathbf{X} = (X_1, \ldots, X_d)$, where X_i is the i_{th} element of $\mathbf{X}$, the subvector (X_i, X_j) is shown by $\mathbf{X}_{i,j}$. In general, if p is a set of indices, then $\mathbf{X}_p$ can be written for the subvector of $\mathbf{X}$ whose elements have those indices. $\mathbf{X}_{-i}$ is defined as the subvector of $\mathbf{X}$ containing all elements except X_i. Similarly, $\mathbf{x} = (x_1, \ldots, x_d)$ denotes the corresponding observed random vector $\mathbf{X}$. Here, $\mathbf{X}$ is considered as an input vector consists of all biological input parameters discussed in Table 1. The output, denoted by Y, represents either I, G_L or H variables.

4.2.1 Variance-based sensitivity analysis methods

In this section we briefly introduce the variance-based SA methods of interest. These methods generally measure the sensitivity of model output, Y (i.e., I, G_L or H), to the variation of an individual input X_i. In other words, they measure the sensitivity of model output, when the model inputs are varied over a wide range, in terms of reduction in the variance of Y.

We start by introducing the main and interaction effects. Follow Sobol′ [31], it can be shown that any quadratically integrable function $f(\cdot)$ can be decomposed in terms of its main effects and interactions as follows:

$$y = f(\mathbf{x}) = z_0 + \Sigma_{i=1}^{d} z_i(x_i) + \Sigma_{i<j} z_{i,j}(\mathbf{x}_{i,j}) + \ldots + z_{1,2,\ldots,d}(\mathbf{x}) \tag{5}$$

where $f(.)$ is a function of uncertain quantities $\mathbf{x}$, and its expected value is denoted by $z_0 = E[f(\mathbf{X})]$. The function $z_i(x_i)$ presented in equation(5) is so-called the *main effect* of the i^{th} variable, x_i. The main effect, $z_i(x_i)$ is the function of x_i that best approximates $f(.)$ in terms of minimizing the variance (calculated over the other variables) [55, 56]. It is defined as:

$$z_i(x_i) = E[f(\mathbf{X}) \mid x_i] - E[f(\mathbf{X})] \tag{6}$$

The *first order interaction* between x_i and x_j, which is denoted by $z_{i,j}(\mathbf{x}_{i,j})$ in equation (5), and is given in equation (7).

$$z_{i,j}(\mathbf{x}_{i,j}) = E[f(\mathbf{X}) \mid \mathbf{x}_{i,j}] - z_i(x_i) - z_j(x_j) - E[f(\mathbf{X})]. \tag{7}$$

Similarly the *second order interaction* between x_i and x_j is denoted by $z_{i,j,k}(\mathbf{x}_{i,j,k})$, and so on. The details of higher order interactions given in equation (5) can be found in [52, 53].

The main effects, the first-order interaction and their plots can be considered as a powerful visual tool to investigate how the model output responds to each individual input, and how those inputs interact in their influence on the model output. The variance of main effect can be interpreted as the amount by which the overall variance of $f(.)$ would be reduced if we knew X_i. A useful SA measure which is given in equation (8), can be considered as the expected amount by which the uncertainty in Y will be reduced if we learn the true value of X_i.

$$V_i = var\{E(Y \mid X_i)\}. \tag{8}$$

It should be also noted that V_i given in equation (8) can be written as $V_i = var(z_i(X_i))$ which is a function of the main effect of X_i.

The second measure, proposed in [54], can be written as:

$$V_{T_i} = var(Y) - var\{E(Y \mid \mathbf{X}_{-i})\} \tag{9}$$

which is the remaining uncertainty in Y that is unexplained after everything has been learnt except X_i.

These two measures, given in equations (8) and (9), can be converted into scale invariant measures by dividing by $var(Y)$ as follows:

$$S_i = \frac{V_i}{var(Y)}, \qquad S_{T_i} = \frac{V_{T_i}}{var(Y)} = 1 - S_{-i} \tag{10}$$

where S_i can be considered as the *main effect index* of X_i, and S_{T_i} is the *total effect index* of X_i.

4.2.2 Emulators-based sensitivity analysis

To compute the variance-based methods in previous sections, we use an emulator to reduce computation costs. The reason we do this is that the function $f(\mathbf{x})$ (the Type 1 diabetes

model) is a complex case as the outputs must be computed by solving the nonlinear model hence computation is costly if done without an emulator.

If $f(\mathbf{x})$ is not complex (computationally cheap), the standard Monte Carlo (MC) methods would be sufficient to estimate $var(Y)$ and other SA measures described in Section 4.2. The computation techniques proposed in [31, 50] require many function evaluations meaning they are not suitable with complex, costly functions. We use a further developed methodology based on the Bayesian paradigm that was proposed in [52, 55, 57] in order to overcome the computational complexity. By using Bayesian method we are able to estimate all the quantities of interest required to examine the SA in modelling diabetes.

The functional relationship, $f(.)$, is unknown for any particular input configuration $\mathbf{x}$ until the model is run for those inputs, therefore we specify a prior distribution for the values taken by $f(\mathbf{x})$ at different values of $\mathbf{x}$ within the Bayesian setting. This prior is then updated according to the usual Bayesian paradigm, using the generated data, $\mathcal{D} = \{(\mathbf{x}_i, y_i) : y_i = f(\mathbf{x}_i), i = 1, \ldots, n\}$, from a set of runs of the model. The result will be then a posterior distribution for $f(.)$, which is used to make formal Bayesian inferences about the SA measures. Although we are still uncertain about the function $f(\cdot)$ at parameter values where it was not evaluated, the uncertainty can be further reduced by taking into account the correlation of function values from one point to another. The expected value of the posterior distribution is used as a point estimate for $f(\cdot)$. There are two different distributions being used in the SA computation. The first is the distribution of input parameters which represents the uncertainty in the model parameters $\mathbf{x}$, and which is propagated to the output values through the function $f(\cdot)$. The second is the posterior distribution on $f(\cdot)$ which plays a pure computational role, and can be reduced as much as required by computing the function $f(\cdot)$ by increasing the training points $\mathbf{x}$, and does not have any operational interpretation.

4.2.3 Gaussian process emulators

Gaussian processes are a class of supervised machine learning algorithms, that describe a functional relation as a multivariate Gaussian distribution and can thus be used for non-linear regression and classification problems. The key requirement to use the Gaussian process is that $f(\cdot)$ should be a smooth function, so if we know the value of $f(\mathbf{x})$ we should then have some idea about the value of $f(\mathbf{x}')$ for $\mathbf{x}$ close to $\mathbf{x}'$. The advantages of the Gaussian process assuming a smooth, continuous function is that it is computationally much quicker and cheaper than using the standard MC methods. This approach usually ignores the expected proximity of the function values evaluated at close by points.

The mean of $f(\mathbf{x})$ conditional on the hyper-parameters β, is modelled as

$$E[f(\mathbf{x})|\boldsymbol{\beta}] = \mathbf{h}(\mathbf{x})^T\boldsymbol{\beta} \tag{11}$$

where $\mathbf{h}(\cdot)$ is a vector of q known functions of $\mathbf{x}$, and β is a vector of coefficients. The choice of $h(\cdot)$ is arbitrary, but it should be chosen to incorporate any beliefs that we might have about

the form of $f(\cdot)$. The covariance between $f(\mathbf{x})$ and $f(\mathbf{x}')$ is given by,

$$cov(f(\mathbf{x}), f(\mathbf{x}')|\sigma^2) = \sigma^2 c(\mathbf{x}, \mathbf{x}') \tag{12}$$

where $c(\cdot,\cdot)$ is a monotone correlation function on $\mathbb{R}^+$ with $c(\mathbf{x},\mathbf{x}) = 1$, and decreases as $|\mathbf{x}-\mathbf{x}'|$ increases. Furthermore, the function $c(\cdot,\cdot)$ must ensure that the covariance matrix of any set of outputs $\{y_1 = f(\mathbf{x}_1), \ldots, y_n = f(\mathbf{x}_n)\}$ is positive semi-definite. Throughout this paper, we use the following correlation function which satisfies all the conditions mentioned above and is widely used for its computational convenience,

$$c(\mathbf{x}, \mathbf{x}') = \exp\{-(\mathbf{x}-\mathbf{x}')^T \mathbf{B}(\mathbf{x}-\mathbf{x}')\}, \tag{13}$$

where $\mathbf{B}$ is a diagonal matrix of positive smoothness parameters, $\{(\sqrt{2}b_i)^{-2}\}_{i=1}^d$, and d is the dimension of $\mathbf{x}$. The matrix $\mathbf{B}$ has the effect of re-scaling the distance between $\mathbf{x}$ and $\mathbf{x}$ and $\mathbf{x}'$. Thus $\mathbf{B}$ determines how close two inputs $\mathbf{x}$ and $\mathbf{x}'$ need to be such that the correlation between $f(\mathbf{x})$ and $f(\mathbf{x}')$ takes a particular value. Oakley and O'Hagan [52] suggest, for fixed hyper-parameters $\mathbf{z}, V, a$ and d, the following conjugate prior, the normal inverse gamma distribution, for $(\boldsymbol{\beta}, \sigma^2)$

$$p(\boldsymbol{\beta}, \sigma^2) \propto (\sigma^2)^{-\frac{1}{2}(d+q+2)} \exp\{-\{(\boldsymbol{\beta}-\mathbf{z})^T V^{-1}(\boldsymbol{\beta}-\mathbf{z}) + a\}/(2\sigma^2)\}$$

The output of $f(\cdot)$ is observed at n design points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ to obtain $\mathbf{y} = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$ considered as data. It should be noticed that these points, in contrast with MC methods, are not chosen randomly but are selected to give good information about $f(.)$. The design points will usually be spread to cover $\mathcal{X}$, the input space of $\mathbf{X}$. Since $\mathbf{X}$ is unknown, the beliefs about $\mathbf{X}$ is represented by the probability distribution $G(\mathbf{X})$. Therefore, the choice of the design points will also depend on $G(.)$ (the choice of design points is discussed in [60]). The standardised posterior distribution of $f(\cdot)$ given $\mathbf{y} = \{f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)\}$ is

$$\frac{f(\mathbf{x}) - m^*(\mathbf{x})}{\hat{\sigma}\sqrt{c^*(\mathbf{x}, \mathbf{x}')}} \mid \mathbf{y} \sim t_{d+n} \tag{14}$$

where t_{d+n} is a student t random variable with $n+d$ degrees of freedom and d is the dimension of $\mathbf{x}$, the posterior mean is given by,

$$m^*(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \hat{\boldsymbol{\beta}} + t(\mathbf{x})^T A^{-1}(\mathbf{y} - H\hat{\boldsymbol{\beta}}), \tag{15}$$

the updated correlation function described in equation (13) given the observed data can be written as,

$$\begin{aligned} c^*(\mathbf{x}, \mathbf{x}') = c(\mathbf{x}, \mathbf{x}') - \mathbf{t}(\mathbf{x})^T A^{-1} \mathbf{t}(\mathbf{x}') + (\mathbf{h}(\mathbf{x})^T \\ - \mathbf{t}(\mathbf{x})^T A^{-1} H)(H^T A^{-1} H)^{-1}(h(x')^T - \mathbf{t}(\mathbf{x}')^T A^{-1} H)^T \end{aligned} \tag{16}$$

and

$$\begin{aligned} \mathbf{t}(\mathbf{x})^T &= (c(\mathbf{x}, \mathbf{x}_1), \ldots, c(\mathbf{x}, \mathbf{x}_n)), \\ H^T &= (\mathbf{h}^T(\mathbf{x}_1)^T, \ldots, \mathbf{h}^T(\mathbf{x}_n)^T), \end{aligned} \tag{17}$$

$$A = \begin{pmatrix} 1 & c(\mathbf{x}_1, \mathbf{x}_2) & \dots & c(\mathbf{x}_1, \mathbf{x}_n) \\ c(\mathbf{x}_2, \mathbf{x}_1) & 1 & & \vdots \\ \vdots & & \ddots & \\ c(\mathbf{x}_n, \mathbf{x}_1) & \dots & & 1 \end{pmatrix} \tag{18}$$

$$\boldsymbol{\beta} = V^*(V^{-1}\mathbf{z} + H^T A^{-1}\mathbf{y}),$$

$$\hat{\sigma}^2 = \frac{\{a + \mathbf{z}^T V^{-1}\mathbf{z} + \mathbf{y}^T A^{-1}\mathbf{y} - \hat{\boldsymbol{\beta}}^T (V^*)^{-1}\hat{\boldsymbol{\beta}}\}}{(n + d - 2)}$$

$$V^* = (V^{-1} + H^T A^{-1} H)^{-1}.$$

The outputs corresponding to any set of inputs will now have a multivariate t-distribution, with covariance between any two outputs given by equation (14). Note that the t-distribution arises as a marginal distribution for $f(.)$ after integrating out the hyper-parameters $\boldsymbol{\beta}$ and σ^2. In practice, further hyper-parameters, the smoothness parameters $\mathbf{B}$, will be associated with the modelling of the correlation function, $c(\cdot, \cdot)$. It is not practical to give $\mathbf{B}$ a fully analytical Bayesian treatment, as it is nearly always impossible to integrate the posterior distribution analytically with respect to these further parameters. We can keep $\mathbf{B}$ fixed as the simplest option. An alternative approach is to use a numerical method to integrate the posterior distribution. It is possible to integrate numerically, in particular, by using Markov chain Monte Carlo (MCMC) sampling however it is a highly intensive computational task. We can estimate the hyper-parameters of $c(\cdot, \cdot)$ from the posterior distribution and then to substitute these estimates into $c(\cdot, \cdot)$ wherever they appear in the above formulae, this is a more robust approach proposed in [52]. These estimates can be obtained by using the posterior mode in combination with a cross validation approach [58]. The GEM-SA [59] is capable of estimating the smoothness parameters using both methods.

4.3 Sobol′ method results

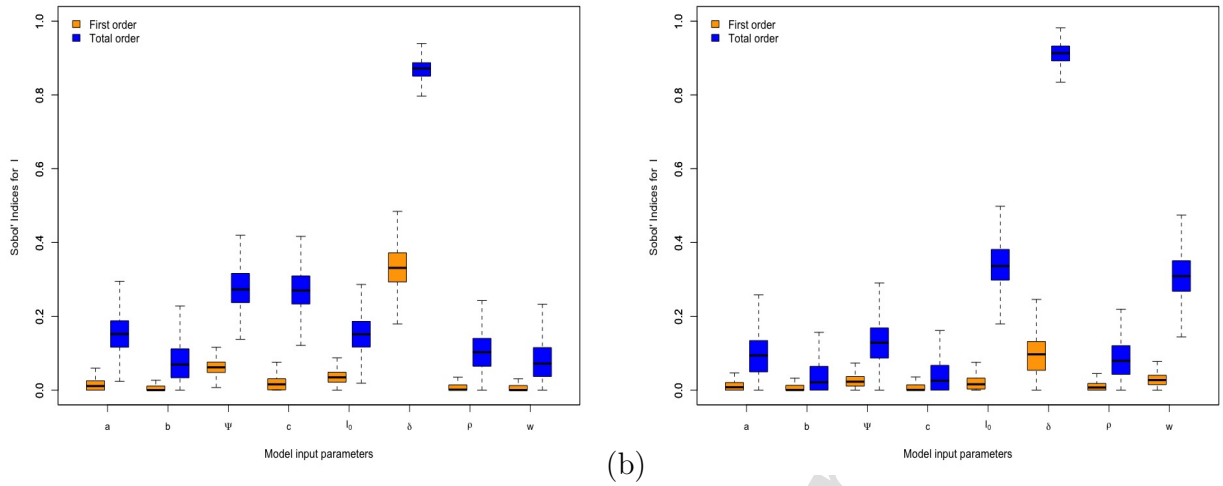
In order to compute the emulator-based SA measures, we first evaluated the outputs of model system (1) for 100 data points selected over a range of input parameters in Table 1 using the Latin hypercube sampling [53] which is a space filling design originally proposed in [60]. We then compute first and total order variance-based sensitivity indices using the Gaussian process emulator at significant time points. Parameters with sensitivity greater than 0.05 were considered to be significant [61]. The Sobol′ indices are analysed for the insulin bolus injection term, I_0, to investigate if model influential parameters are affected by the amount of insulin injected. It should be noted that the insulin bolus injection (I_0) is a dosage level such as $I_0 = 5$ and is a parameter, thus should not be confused with the output variable insulin concentration (I).

Figure 5 and Table 4 illustrates the first order variance-based sensitivity indices for insulin bolus level of $I_0 = 5$ for insulin concentration (I). Results were produced for different insulin bolus levels of $I_0 = 10, 15, 20, 25, 30$ and are presented in *Supplementary Tables S93-S97*. Total effect sensitivity indices for I at insulin bolus levels ($I_0 = 5, 10, 15, 20, 25, 30$) are in *Supplementary Tables S109-S114*. The figure shows that, parameters δ and ψ are critical in influencing I.

Parameter δ is the most influential with a total order index of 0.9127 (at $t = 60$ minutes) and this is followed by parameter ψ, other model parameters are not significant in determining I. Additionally, parameter δ has an extremely high total order index in comparison with its first order, implying there is lots of interaction with other parameters suggesting the significance of using global methods or SA methods that evaluate parameter relationships between each other. The first order and total order indices did show some change at different time points, nevertheless the parameters which are influential remained significant. Results also showed that regardless of bolus amount injected, δ remains significant and its significance increases over time points. Explanation for this increase of clearance term δ is due to the fact that the increased insulin bolus I_0 injected requires higher clearance to maintain the glucose level once homeostasis is reached. Conversely, ψ and I_0 indices, although remain significant, they reduce over time when $t > 60$ minutes. This can be explained by the fact I_0 is highest when first injected and gradually decreases as it is used up and the absorption term ψ is therefore less significant due to clearance of insulin.

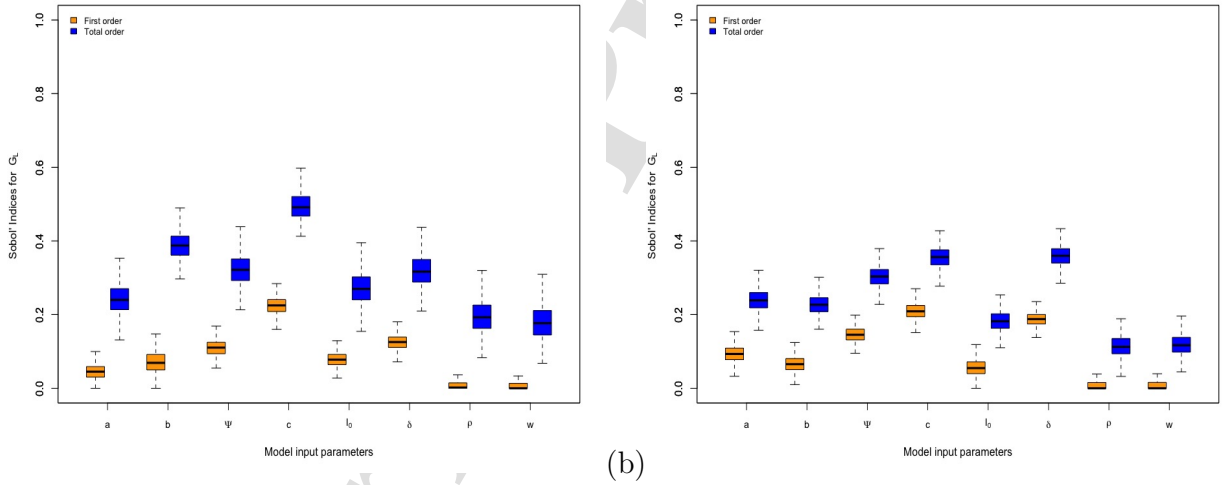
Figure 6 and Table 5 illustrate the first order variance-based sensitivity indices at insulin bolus level $I_0 = 5$ for G_L. The remaining values of I_0 within the range given in Table 1, on first order variance-based sensitivity indices for G_L are in *Supplementary Tables S98-S102)*. Total effect sensitivity indices for G_L are presented in *Supplementary Tables S115-S120*. Sensitivity analysis results show that parameters δ (with total order index of 0.3126 at $t = 60$ minutes), c and ψ are the most influential parameters on G_L. The total order and first order indices for these parameters are similar, suggesting that these parameters have no interaction with other parameters. Sensitivity analysis results based on Sobol′ method for H are shown in *Supplementary Figures S6 and S5, Tables S103-S108, S121-S126*. Results also show the change in variance of each parameter on the variable over the different time points converges. Parameters which have the most influence at a certain time have shown to have the most effect. Results of the Sobol′ method for different insulin bolus injection terms I_0, showed that significant parameters remained significant throughout variations in values for the insulin injection term. The insulin induced glucose uptake rate (c) is not substantially affected by changes in the insulin bolus term I_0. Parameter c clears the insulin therefore as bolus levels increase the parameter performs its role at an increased rate (i.e. up-taking the glucose). At 10 minutes the absorption term (ψ) significance is increased as this is when the bolus is injected. The significance of parameter δ increases over time. However, after 180 minutes, δ importance begins to decrease and this is linked to the need to clear insulin once blood glucose level is maintained. The process is usually achieved in 2 hours and consequently insulin clearance is reduced as homeostasis is maintained [40, 41].

The results showed that all the parameters for H have first order indices close to zero (< 0.05) with exception of ρ and w. Using both PRCC and Sobol′ method sensitivity analysis was also conducted at different times and similar conclusions were reached.



(a) (b)

Figure 5: Plot (a) shows first order and total effects sensitivity indices of the model parameters (a, b, ψ, c, I_0, δ, ρ and w) on insulin for model system (1) using Sobol′ method at time $t = 210$. Plot (b) similarly shows first order and total effects sensitivity indices of the model parameters (a, b, ψ, c, I_0, δ, ρ and w) on insulin for model system (1) using Sobol′ method at time $t = 60$.



(a) (b)

Figure 6: Plot (a) shows first order and total effects sensitivity indices of the model parameters (a, b, ψ, c, I_0, δ, ρ and w) on glucose for model system (1) using Sobol′ method at time $t = 210$. Plot (b) similarly shows first order and total effects sensitivity indices of the model parameters (a, b, ψ, c, I_0, δ, ρ and w) on glucose for model system (1) using Sobol′ method at time $t = 60$.

	Time(in minutes)					
Parameter	$t = 5$	$t = 10$	$t = 60$	$t = 90$	$t = 180$	$t = 210$
a	0.0040	0.0003	0.0017	0.0012	1.92×10^{-5}	0.0051
b	0.0051	0	0.0003	0.0050	0.003	0.0351
ψ	0.2335	0.1485	0.0938	0.0299	0.0196	0.0034
c	5.21×10^{-5}	0	0.0002	0.0007	0.0004	0.0036
I_0	0.2312	0.1619	0.0574	0.0028	0.0132	0.0278
δ	0.3431	0.4957	0.6084	0.2149	0.7357	0.0352
ρ	0	0	0.0006	0.0038	0.0216	0.0359
w	0.0097	0	0.0002	0.1024	0	0.0241

Table 4: First order Sobol′ indices of each model parameter for I at significant time periods with $I_0 = 5$.

	Time(in minutes)					
Parameter	$t = 5$	$t = 10$	$t = 60$	$t = 90$	$t = 180$	$t = 210$
a	0.0951	0.0498	0.0610	0.0599	0.0803	0.0246
b	0.0073	0.0346	0.1043	0.0386	0.0457	0.0049
ψ	0.0389	0.1470	0.1140	0.1328	0.1531	0.0296
c	0.2080	0.3535	0.1863	0.3190	0.1706	0.0429
I_0	0.0990	0.1027	0.0557	0.0699	0.1120	0.1383
δ	0.0676	0.1337	0.1140	0.1807	0.2157	0.0162
ρ	0.0045	0	0.0020	0.0007	0	0.0009
w	0.0150	0.0008	0.0105	0.0036	0	0.0102

Table 5: First order Sobol′ indices of each model parameter for G_L at significant time periods with $I_0 = 5$.

4.4 Comparison with PRCC results

We compare the results obtained from PRCC method with the variance-based methods used in this study. We note that the PRCC highlighted more influential parameters (i.e. more parameters are shown to affect the model outputs; insulin, glucose and growth hormone). Sobol′ method identified a smaller set of influential parameters than the PRCC method, as noted in other studies [49, 67]. The explanation for this could be that the PRCC assumes a monotonic input and output relationship, unlike the Sobol′ method. However, the Sobol′ method is able to quantify the effect of the high-order interactions between input parameters, thus providing us with further insight on understanding the model system [49]. The results from the Sobol′ method showed consistently the parameters that are influential against I and G_L were δ and ψ, these parameters were also identified as influential by the PRCC method. The PRCC method identified I_0 as significant constantly, however Sobol′ method only highlighted its significance at one time point. However, the Sobol′ method identified the high interaction of δ with other

parameters, something PRCC method was unable to show. Both methods concurred completely on the influential parameters against H and the parameters identified were ρ and w.

5 Discussion

Mathematical models of Type 1 diabetes [8, 12–20] have been developed to understand the disease and develop more effective treatment methods in order to provide better lifestyles for diabetes patients. Current treatment methods are invasive, inconvenient and require constant monitoring. Presently, the treatment methods include daily self injections, constant recording of blood glucose levels, carbohydrate counting and even transplant of islets [11]. Several diabetes models [8, 10, 12–20] have managed to describe molecular dynamics in a Type 1 diabetic and lead to the development of open-loop insulin pumps based on mathematical algorithms. However, the current mathematical models of diabetes do not consider the condition with zero insulin concentration in the blood as expected in a Type 1 diabetic [12, 62]. In light of this limitation, we developed a new mathematical model to fully capture Type 1 diabetes dynamics.

The results of the mathematical analysis showed that the model has two stable steady states. The model threshold quantity $\mathcal{T}_0$ was derived and it was shown that the pathological equilibrium was locally asymptotically stable for values of $\mathcal{T}_0 < 1$. The importance of the model threshold $\mathcal{T}_0$ is in determining key parameters governing the dynamics of Type 1 glucose homeostasis system. Further, pathological equilibrium was shown to be globally stable. The managed diabetes equilibrium was shown to be globally stable for values of $\mathcal{T}_0 > 1$. Numerical analysis of the model showed a transcritical bifurcation (Figure 3), confirming results illustrated by the time series plots (Figures 1 and 2).

Sensitivity analysis was conducted to systematically evaluate key parameters influencing the model at different time points using PRCC and Sobol′ methods. PRCC method identified all key parameters which appeared in the Type 1 diabetes threshold quantity as important. The parameters which were significant remained so for different time points. However, for glucose concentration, at $t = 90$ minutes two additional parameters, growth hormone clearance rate and growth hormone production rate, were identified as significant. The growth hormone role in influencing glucose concentration is most important at 2 hours postprandial. The PRCC and Sobol′ methods concurred in many scenarios but showed some differences. Both methods identified the importance of insulin clearance rate (δ) and insulin absorption rate (ψ) as influential parameters in determining insulin concentration (see Figures 4(b), *Supplementary Figures S1, S2*, Table 2, *Supplementary Tables S1-S31* for PRCC, Figure 5 and Table 4, *Supplementary Tables S93-S97, S109-S114* for Sobol′ indices). However, for the Sobol′ method the insulin bolus term (I_0) was only shown to be significant at $t = 60$ minutes, unlike in the PRCC method where it was shown to be significant throughout all time points. Furthermore, the difference of the total order and first order indices for δ against insulin concentration was large, implying that there was interaction with other parameters. Global sensitivity analysis methods which allow us to explore relationships between parameters are important in understanding the overall

influence of each parameter in a model. PRCC method also showed that the most influential parameters on blood glucose concentration are insulin clearance (δ) and insulin absorption rates (ψ) (see Figures 4(b), *Supplementary Figures S3, S4*, Table 3, *Supplementary Tables S32-S61*). Sobol′ method also confirmed these findings and highlighting that, the parameters have little interactions with other parameters (Figure 6 and Table 5, *Supplementary Tables S98-S102, S115-S120).*

Our results showed that, PRCC method managed to identify all key parameters in glucose and insulin concentration dynamics. However, Sobol′ method managed to provide insights on parameter interaction, thus demonstrating the importance of using other global SA methods along with the PRCC method. One advantage of using variance based methods such as Sobol′ method is that they are computationally more efficient and suitable for complex models. It is clear that classical mathematical analysis alone is not sufficient in understanding model dynamics and parameter interactions, thus SA methods should be used to fill this gap. Sensitivity analysis insights on important model parameters varied by method [63–65]. Other studies [63, 66] have proposed selection of SA methods to be used based on model complexity, characteristics and research question.

Findings from this study have some similarities and differences from those in [22]. Both models, although different, highlighted the significance of using both PRCC and Sobol′ methods as these methods can offer different but important model insights. For example, in this study it is clear that the PRCC method identified more influential parameters including all parameters in the threshold, $\mathcal{T}_0$. However, Sobol′ method provided insights on interaction between model parameters. In [22], both methods managed to identify the importance of all the parameters in the model threshold quantity, but similarly revealed that Sobol′ method provides more insights on parameter interaction.

The results of this study are important in informing the building of suitable mathematical algorithms to use in an artificial pancreas. This model provides a potential open-loop algorithm framework which captures a zero-insulin state, a condition which occurs frequently in individuals with Type 1 diabetes and has so far not been considered in previous models. An artificial pancreas is vital to ease the life of Type 1 diabetic individuals by offering better treatment to those suffering with the disease. However, the building of suitable artificial pancreases requires accurate and efficient mathematical algorithms.

This study has several limitations, the model is parameterised using estimates in [68] and other assumed parameters and only considers sensitivity analysis using PRCC and Sobol′ methods. Nonetheless, it would be interesting to explore how these results vary when the model is calibrated using experimental data and if other global sensitivity analysis approaches are used. We also assumed that, insulin injection is inversely proportional to insulin concentration in the blood following [12], however it would be interesting to confirm the validity of the assumption using robust data-driven modeling approaches. Experimental data on the relationship between

insulin injection and insulin in the blood would be key in informing the building accurate and efficient diabetes models. Additionally, the model is based on short term injection, whereas it would be interesting to explore if having a long-term insulin injection term instead would affect the results. In light of the current COVID-19 pandemic and increased risk of developing COVID-19 complications among diabetic individuals, it would be interesting to develop an in-vivo model to understand COVID-19 infection and diabetes dynamics. Despite these limitations, this study presents a new way of modelling Type 1 diabetes and provides an important framework for understanding nonlinear model parameters using a combination of mathematical and sensitivity analysis approaches.

Acknowledgments

HAA acknowledges financial support from Emirates Aviation University. The authors acknowledge Dr Abhinav Vepa, at Milton Keynes University Hospital NHS Foundation Trust for reviewing the paper and validating results.

References

[1] International Diabetes Federation https://idf.org/aboutdiabetes/type-1-diabetes.html (Accessed May 2021).

[2] Saeedi P., Petersohn I., Salpea P. *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, *Diabetes Research and Clinical Practice*, (2019).

[3] Watson S, Autoimmune diseases: Types, symptoms, causes, and more, https://www.healthline.com/health/autoimmune-disorders, (Accessed February 2021).

[4] Boutayeb W., Lamlili M.E.N., Boutayeb A. and Derouich M. Mathematical modelling and simulation of β-cell mass, insulin and glucose dynamics: Effect of genetic predisposition to diabetes, *Journal of Biomedical Science and Engineering*, 330(7), (2014).

[5] Diabetes.co.uk, Beta Cells - What They Do, Role in Insulin, http://www.diabetes.co.uk/body/beta-cells.html (Accessed 11th May 2021).

[6] Cantley J. and Ashcroft F.M. Q&A: insulin secretion and type 2 diabetes: why do β-cells fail?, *BMC Biology*, 33(13), (2015).

[7] Mayo Clinic. 2020. Diabetes - Symptoms And Causes. https://www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444 (Accessed 25th May 2020).

[8] Li J. and Johnson J.D. Mathematical models of subcutaneous injection of insulin analogues: a mini-review. *Discrete and Continuous Dynamical Systems*, Series B, 12(2), 401, (2009).

[9] Haidar, A., Tsoukas, M.A., Bernier-Twardy, S., Yale, J.F., Rutkowski, J., Bossy, A., Pytka, E., El Fathi, A., Strauss, N. and Legault, L. A novel dual-hormone insulin-and-pramlintide artificial pancreas for type 1 diabetes: a randomized controlled crossover trial. *Diabetes Care*, 43(3), 597-606, (2020).

[10] Brown S.A., Kovatchev B.P., Raghinaru D., Lum J.W., Buckingham B.A., Kudva Y.C., Laffel L.M., Levy, C.J., Pinsker J.E., Wadwa R.P. and Dassau E., Six-month randomized, multicenter trial of closed-loop control in type 1 diabetes. *New England Journal of Medicine*, 381(18), 1707-1717, (2019).

[11] Pathak V., Pathak N.M., O'Neill C.L., Guduric-Fuchs J. and Medina R.J., Therapies for type 1 diabetes: current scenario and future perspectives. *Clinical Medicine Insights: Endocrinology and Diabetes*, 12, 1179551419844521, (2019).

[12] Li J. and Kuang Y. Systemically modeling the dynamics of plasma insulin in subcutaneous injection of insulin analogues for type 1 diabetes. *Mathematical Biosciences and Engineering: MBE*, 6(1), 41, (2009).

[13] Kraegen E.W. and Chisholm D.J. Insulin responses to varying profiles of subcutaneous insulin infusion: kinetic modelling studies. *Diabetologia*, 26(3), 208-213, (1984).

[14] Puckett W.R. and Lightfoot E.N. A model for multiple subcutaneous insulin injections developed from individual diabetic patient data. *American Journal of Physiology-Endocrinology and Metabolism*, 269(6), E1115-E1124, (1995).

[15] Rossetti P., Pampanelli S., Fanelli C., Porcellati F., Costa E., Torlone E., Scionti L. and Bolli, G.B. Intensive replacement of basal insulin in patients with type 1 diabetes given rapid-acting insulin analog at mealtime: a 3-month comparison between administration of NPH insulin four times daily and glargine insulin at dinner or bedtime. *Diabetes Care*, 26(5), 1490-1496, (2003).

[16] Shimoda S.E., Nishida K.E., Sakakida M.I., Konno Y.U., Ichinose K.E., Uehara M.A., Nowak T.A., Shichiri M.O. Closed-loop subcutaneous insulin infusion algorithm with a short-acting insulin analog for long-term clinical application of a wearable artificial endocrine pancreas. *Frontiers of medical and biological engineering: International Journal of the Japan Society of Medical Electronics and Biological Engineering*, 8(3), 197-211, (1997).

[17] Mukhopadhyay A., De Gaetano A. and Arino O. Modeling the intra-venous glucose tolerance test: a global study for a single-distributed-delay model. *Discrete and Continuous Dynamical Systems-B*, 4(2), 407, (2004).

[18] Nucci G. and Cobelli C. Models of subcutaneous insulin kinetics. A critical review. *Computer Methods and Programs in Biomedicine*, 62(3), 249-257, (2000).

[19] Tarin C., Teufel E., PicÛ, J., Bondia, J. and Pfleiderer, H.J. Comprehensive pharmacokinetic model of insulin glargine and other insulin formulations. *IEEE Transactions on Biomedical Engineering*, 52(12), 1994-2005, (2005).

[20] Wilinska M.E., Chassin L.J., Schaller H.C., Schaupp, L., Pieber T.R. and Hovorka R. Insulin kinetics in type-1 diabetes: continuous and bolus delivery of rapid acting insulin. *IEEE Transactions on Biomedical Engineering*, 52(1), 3-12, (2004).

[21] Al Ali H., Boutayeb W., Boutayeb A. and Merabet N. A mathematical model for type 1 diabetes, on the effect of growth hormone. *IEEE, 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO) 15-17 April 2019*, `https://ieeexplore.ieee.org/document/8902178`, (2019).

[22] Al Ali H., Boutayeb W., Boutayeb A. and Merabet N. A mathematical model on the effect of growth hormone on glucose homeostasis. *ARIMA Journal*, 30: 31-42, (2019).

[23] Mukandavire Z., Smith D.L. and Morris J.G, Jr. Cholera in Haiti: reproductive numbers and vaccination coverage estimates. *Scientific Reports*, 3:997, (2013).

[24] Mukandavire Z., Liao S., Wang J., Gaff H., Smith D.L. and Morris J.G., Jr. Estimating the reproductive numbers for the 2008-2009 cholera outbreaks in Zimbabwe. *Proceedings*

of the National Academy of Sciences of the United States of America, 108(21):8767-8772, (2011).

[25] Li J., Blakeley D., Smith R.J. The failure of R_0. *Computational and Mathematical Methods in Medicine* 2011:527610; (2011).

[26] Smith D.L., McKenzie F.E., Snow R.W. and Hay S.I. Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLoS Biology* 5: e42., (2007).

[27] Dietz K. The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research*, 2(1), 23-41, (1993).

[28] Anderson R.M. and May R.M. Infectious Diseases of Humans, *Oxford University Press*, London/New York, (1991).

[29] Diekmann O., Heesterbeek J.A.P. and Metz J.A.P. On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in hetergeneous populations. *Journal of Mathematical Biology*, 28, 365ñ382, (1990).

[30] Blower S. M. and Dowlatabadi H. Sensitivity and uncertainty analysis of complex-models of disease transmissionóan HIV model, as an example, *International Statistical Review*, 62, 229-243, (1994).

[31] Sobol I.M. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4), 407-414, (1993).

[32] Medlineplus Medical Encyclopedia, Type 1 Diabetes, *Medlineplus,* `https://medlineplus.gov/ency/article/000305.htm` (Accessed 29 June 2020).

[33] Koutny, T. Modelling of glucose dynamics for diabetes. *In International Conference on Bioinformatics and Biomedical Engineering*, 314-32, Springer, Cham, (2017).

[34] Magdelaine, N., Chaillous, L., Guilhem, I., Poirier, J.Y., Krempf, M., Moog, C.H. and Le Carpentier, E. A long-term model of the glucose–insulin dynamics of type 1 diabetes. *IEEE Transactions on Biomedical Engineering*, 62(6), 1546-1552, (2015).

[35] Sanofi-Aventis, Product monograph apidra insulin glulisine, `http://products.sanofi.ca/en/apidra.pdf`, Accessed (25th October 2021).

[36] Taylor A., Finster J. and Mintz D. Metabolic clearance and production rates of human growth hormone, *The Journal of Clinical Investigation*, 48(12), 2349-2358, (1969).

[37] Schwarz J., Mulligan K., Lee J., Lo J., Wen M., Noor M., Grunfeld C. and Schambelan M. Effects of recombinant human growth hormone on hepatic lipid and carbohydrate metabolism in HIV-infected patients with fat accumulation, *The Journal of Clinical Endocrinology and Metabolism*, 87(2), 942-945, (2002).

[38] Kim S. and Park M. Effects of growth hormone on glucose metabolism and insulin resistance in human, *Annals of Pediatric Endocrinology and Metabolism*, 22(3), 145, (2017).

[39] Holly J.M.P., Amiel S.A., Sandhu R.R., Rees L.H. and Wass J.A.H. The role of growth hormone in diabetes mellitus, *Journal of Endocrinology*, 118(3), 353-364, (1988).

[40] Buppajarntham S., Junpaparp P., Salameh R., Anastasopoulou C. and Eric B.S. Insulin: Reference Range, Interpretation, Collection and Panels, https://emedicine.medscape.com/article/2089224-overview (Accessed 23rd March 2019).

[41] Mayo Clinic, Diabetes; Diagnosis and treatment, *Mayo Foundation for Medical Education and Research,* https://www.mayoclinic.org/diseases-conditions/diabetes/diagnosis-treatment/drc-20371451 (Accessed 11th May 2021).

[42] Wisse B., Zieve D. Growth hormone test: MedlinePlus Medical Encyclopedia, *A.D.A.M,* https://medlineplus.gov/ency/article/003706.htm (Accessed 23rd March 2019).

[43] LaSalle J.P.The Stability of Dynamical Systems, *CBMS-NSF Regional Conference Series in Applied Mathematics 25*, SIAM: Philadelphia, (1976)

[44] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org, (2013).

[45] MATLAB and Statistics Toolbox Release 2019b, The MathWorks, Inc., Natick, Massachusetts, United States, (2019).

[46] Gramacy R.B. *tgp*: An R Package for Bayesian Nonstationary, Semiparametric Nonlinear Regression and Design by Treed Gaussian Process Models. *Journal of Statistical Software, 19(9),* http://www.jstatsoft.org/v19/i09, (2007).

[47] Malunguza, N.J., Hove-Musekwa, S.D. and Mukandavire, Z. Projecting the impact of anal intercourse on HIV transmission among heterosexuals in high HIV prevalence settings, *Journal of Theoretical Biology*, 437, 163-178, (2018).

[48] Zi Z. Sensitivity analysis approaches applied to systems biology models, *IET Systems Biology*, 5(6), 336-346, (2011).

[49] Chen, X., Wang, W., Xie, G., Hontecillas, R., Verma, M., Leber, A., Bassaganya-Riera, J. and Abedi, V. Multi-resolution sensitivity analysis of model of immune response to helicobacter pylori infection via spatio-temporal metamodeling. *Frontiers in Applied Mathematics and Statistics*, 5, 4, (2019).

[50] Saltelli A., Tarantola S. and Chan, K. S. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1), 39-56, (1999).

[51] Sudret B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, 93(7), 964-979, (2008).

[52] Oakley J.E. and O'Hagan A. Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3), 751-769, (2004).

[53] Daneshkhah, A. and Bedford, T. Probabilistic sensitivity analysis of system availability using Gaussian processes. *Reliability Engineering and System Safety*, 112, 82-93, (2013).

[54] Homma, T., and Saltelli, A. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52(1), 1-17, (1996).

[55] Daneshkhah, A., Stocks, N. G., and Jeffrey, P. Probabilistic sensitivity analysis of optimised preventive maintenance strategies for deteriorating infrastructure assets. *Reliability Engineering and System Safety*, 163, 33-45, (2017).

[56] Daneshkhah, A. and Bedford, T. Probabilistic sensitivity analysis of system availability using Gaussian processes. *Reliability Engineering and System Safety*, 112, 82-93, (2013).

[57] Daneshkhah A., Hosseinian-Far A., and Chatrabgoun O. Sustainable maintenance strategy under uncertainty in the lifetime distribution of deteriorating assets. *Strategic Engineering for Cloud Computing and Big Data Analytics,* 29-50. Springer, Cham, (2017).

[58] O'Hagan, A., Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M. Uncertainty analysis and other inference tools for complex computer codes (1998).

[59] Kennedy M. C. and Petropoulos G. P. GEM-SA: the Gaussian emulation machine for sensitivity analysis, In: Sensitivity Analysis in Earth Observation Modelling, ,emphElsevier, 341-361, 2017.

[60] Sacks J., Welch W. J., Mitchell T. J. and Wynn H. P. Design and analysis of computer experiments. *Statistical Science*, 4(4), 409-423, (1989).

[61] Zhang X.Y., Trame M.N., Lesko L.J. and Schmidt S. Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models. *CPT: pharmacometrics & systems pharmacology*, 4(2), 69-79, (2015).

[62] Mosekilde E, Jensen KS, Binder C, Pramming S, Thorsteinsson B. Modeling absorption kinetics of subcutaneous injected soluble insulin. J. *Journal of Pharmacokinetics and Biopharmaceutics*, 17, 67ñ87, (1989).

[63] Wu J., Dhingra R., Gambhir M. and Remais J.V. Sensitivity analysis of infectious disease models: methods, advances and their application, *Journal of the Royal Society Interface*, 10, 20121018, (2013).

[64] Iooss B. and Prieur C. Shapley effects for sensitivity analysis with correlated inputs: comparisons with Sobol'indices, numerical estimation and applications. *International Journal for Uncertainty Quantification*, 9(5), (2019).

[65] Iooss B. and LemaÓtre P. A review on global sensitivity analysis methods. In Uncertainty management in simulation-optimization of complex systems, *Springer*, Boston, MA, 101-122, (2015).

[66] Brevault L., Balesdent M., BÈrend N. and Le Riche R. Comparison of different global sensitivity analysis methods for aerospace vehicle optimal design. *In: 10th World Congress on Structural and Multidisciplinary Optimization, WCSMO-10*, (May 2013).

[67] Marino, S., Hogue, I.B., Ray, C.J. and Kirschner, D.E. A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of theoretical biology, 254(1), 178-196, (2008).

[68] Topp B., Promislow K., Devries G., Miura R.M. and Finegood T.D. A model of β-cell mass, insulin, and glucose kinetics: pathways to diabetes, *Journal of Theoretical Biology*, 206(4), 605-619, (2000).