UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN BIOINGEGNERIA

Modeling the inter-day variability of insulin sensitivity in patients with type 1 diabetes through Markov models

Relatore:
Dott. Michele Schiavon

Correlatrice:
Prof. ssa Chiara Dalla Man

Laureanda:
Diletta Babato

Matricola:
1242426

04/04/2022

Anno Accademico 2021-2022

Contents

Abstract

Diabetes Mellitus is one of the most common diseases worldwide, and it is estimated to increase to 700 million by 2045. Moreover, this disease represents a big challenge for global healthcare, since a real treatment for its remission does not exists nowadays, and studies are performed to find the best way to reduce disease-related short- and long-term complications. Diabetes Mellitus is divided into two main categories: Type 1 Diabetes (T1D), characterized by a lack of insulin secretion, caused by the destruction of β-cells, and Type 2 Diabetes (T2D), characterized by an inadequate responsitivity of cells to insulin, both leading to chronic hyperglycemia.
An important parameter involved diabetes therapy is insulin sensitivity, representing the ability of insulin to promote glucose utilization and inhibit glucose production. Estimation of such parameter is essential, especially in every day life of T1D patients who need to administer exogenous insulin to keep glucose within the safe range, as well as to study its intra-day and inter-day variability to optimize insulin therapy day by day.
Gold standard methodology to assess insulin sensitivity requires inpatient experiments, hence cannot be used to assess insulin sensitivity in daily life conditions and thus its intra- and inter-day variability. In 2014 a method to assess insulin sensitivity from minimally invasive technologies used by T1D patients has been developed to overcome this issue allowing its estimation in daily life conditions.
In this thesis an update of this formula has been proposed and tested to improve insulin sensitivity estimation in challenging scenarios which can occur in daily life. Then a method to assess intra-day pattern of insulin sensitivity based on statistical methods as well as a stochastic model for estimating inter-day variability of intra-day pattern has been developed.
To do so we used the UVA/Padova T1D simulator, a tool accepted by FDA as a substitute for preclinical trials of certain insulin treatments, by running different experimental scenarios using the virtual adult population of 100 type 1 diabetic subjects. First of all the proposed new formula for the estimation of insulin sensitivity from minimally-invasive technologies is validated against the reference methodology of the oral glucose minimal model. Then, intra-day pattern of insulin sensitivity is estimated for each subject by means of statistical methods and it is compared

against the reference one estimated from the oral glucose minimal model. In the end, subject-specific inter-day variability of intra-day pattern of insulin sensitivity is modeled by using Markov Chains models and compared against the ones randomly generated and included into the simulator.
This work represents the first step toward the estimation of the variability of insulin sensitivity in the long term, which is a key point for the development of algorithms for the automatic adaptation of insulin therapy able to follow patient's insulin sensitivity variation in daily life conditions of T1D.

Sommario

Il diabete mellito è una delle malattie più diffuse a livello mondiale e si stima che arriverà a colpire 700 milioni di persone nel 2045. Inoltre, questa malattia rappresenta una grande sfida per la sanità globale, dal momento che ad oggi non esiste una vera e propria cura per la sua remissione, e gli studi e le ricerche che vengono compiuti si concentrano sul trovare il modo migliore per ridurre le complicazioni a breve e lungo termine collegate alla malattia. Il diabete mellito è diviso in due categorie principali: il diabete di tipo 1, caratterizzato da una mancanza di secrezione di insulina, dovuta alla distruzione delle cellule beta del pancreas, e il diabete di tipo 2, caratterizzato da una inadeguata risposta delle cellule all'insulina. Entrambe le forme di diabete portano ad una condizione di iperglicemia cronica.
Un parametro importante nelle terapie del diabete di tipo 1 è la sensibilità insulinica, ossia la capacità dell'insulina nel promuovere l'utilizzo del glucosio da parte dei tessuti e inibire la sua produzione da parte del fegato. È essenziale stimare questo parametro e la sua variabilità intra-day e inter-day, specialmente nei pazienti affetti da diabete di tipo 1, che necessitano di una amministrazione esogena di insulina.
Il gold standard dei metodi per stimare la sensibilità insulinica richiede esperimenti con pazienti ospedalizzati, e dunque non è utilizzabile per ricavare la sensibilità insulinica e la sua variabilità in condizioni di vita quotidiana. Nel 2014 un nuovo metodo, in cui la sensibilità insulinica viene stimata tramite tecnologie minimamente invasive, è stato introdotto per permettere così la stima della sensibilità insulinica in condizioni di vita quotidiana. In questa tesi viene proposto e testato un aggiornamento di questa formula, per migliorare la stima della sensibilità insulinica in condizioni sperimentali complesse che possono accadere nella vita quotidiana. In seguito, viene definito un metodo per stimare il pattern intra-day della sensibilità insulinica basato su metodi statistici, e infine la variabilità inter-day del pattern intra-day viene stimata tramite un modello stocastico.
Per fare ciò, è stato utilizzato il simulatore di UVA/Padova di diabete di tipo 1, uno strumento approvato FDA come sostituto dei trial preclinici di alcuni trattamenti con insulina, eseguendo diversi scenari sperimentali con una popolazione di 100 adulti diabetici di tipo 1. Innanzitutto, la formula proposta per la stima della sensibilità insulinica con tecniche minimamente invasive è stata validata rispetto al

modello minimo orale del glucosio. In seguito, il pattern intra-day della sensibilità insulinica è stato stimato con metodi statistici e validato rispetto al riferimento stimato con il modello minimo di glucosio. Infine, è stata modellata la variabilità inter-day del pattern intra-day della sensibilità insulinica specifica di ogni soggetto utilizzando i modelli di Markov Chain, e confrontata con quella generata in modo casuale ed inclusa inclusa nel simulatore.
Questo lavoro rappresenta un punto di partenza per la stima della sensibilità insulinica a lungo termine, che è un punto chiave per lo sviluppo di algoritmi per l'adattamento automatico della terapia insulinica, in grado di monitorare le variazioni della sensibilità insulinica del paziente nella sua vita di tutti i giorni.

Chapter 1

Introduction

The regulation of blood glucose is controlled by a complex neuro-hormonal system, and it is important to understand which mechanisms are involved in this chain and the fact that a failure in one of these steps can cause several problems and in the worst case diabetes [5].
The concentration of glucose in blood is called glycemia, and it is usually expressed in mg/dL or mmol/l. In a healthy subject the normal fasting glucose level is in the range 70-110 mg/dL [20].
Hyperglycemia, a condition in which fasting level of glycemia is above 140 mg/dL, is dangerous because it can damage tissues and sometimes bring to diabetes mellitus; on the other hand hypoglycemia, that occurs when fasting level of glycemia is below 60 mg/dL, has negative effects on the functionality of brain and the nervous system, that use glucose as a source of energy.
In human beings the pancreas is a gland which is located in the abdomen, behind the stomach, and has both an endocrine and an exocrine function (Fig. 1.1).
The exocrine pancreas, which is about the 99% of the organ, is composed by acini and ducts that have a digestive function and produce a fluid that contains digestive enzymes. The endocrine pancreas is composed by several groups of cells called pancreatic islets (also called islet of Langerhans), each one producing a different hormone [20]. The most important types are alpha cells and beta cells: when the concentration of glucose blood is low, alpha cells produce glucagon, a hormone that increases blood glucose level, by promoting the breakdown of glycogen to glucose in the liver. On the other hand, when blood glucose is high, beta cells produce insulin, a hormone that promotes the absorption of glucose by tissues. The other two types are delta and F cells, whose function is partially unknown.

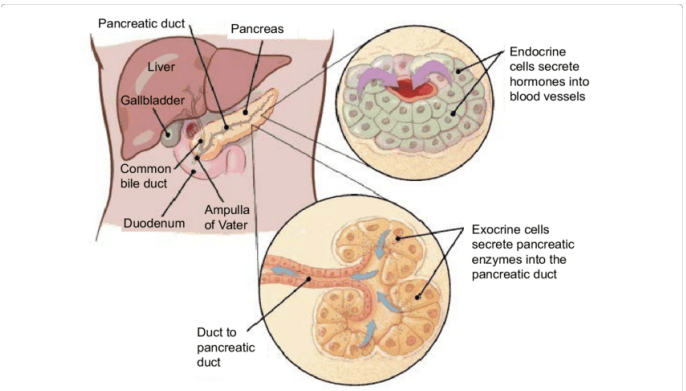


Figure 1.1: Representation of pancreatic cells [13].

The glycemia is controlled by insulin and glucagon, which act with a mechanism of negative feedback. When blood glucose increases, it encourages the production of insulin and inhibits the secretion of glucagon. These two actions cause a decrease of glucose level. Vice versa, the decrease of glycemia causes a reduction of the production of insulin and an increase of the release of glucagon. These two actions enhance the increase of blood glucose.
In our body there are insulin-dependent (muscle and adipose tissues) and insulin-independent (the central nervous system and red blood cells) tissues [5]. Regarding the former insulin enables the transport of nutrients through the membranes thanks to glucose transporters in the cells that are sensitive to insulin.
Regarding the latter it is important to note that when the level of insulin in the blood is low, cells of the nervous system can absorb glucose anyway.
The glucose-insulin control system is represented in the figure:

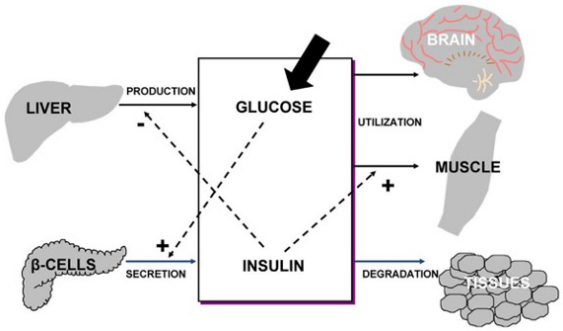


Figure 1.2: Glucose-Insulin Regulation system [5].

Glucose is mainly produced by the liver, which can produce glucose thanks to the glycogenolysis and the gluconeogenesis. As said before, glucose is uptake by different organs, such as brain and muscle and adipose tissue as a source of energy.

1.1 Diabetes Mellitus

Diabetes is a metabolic disorder, characterized by chronic hyperglycemia due to the absence of insulin secretion, Type 1 Diabetes, or impaired insulin action, Type 2 Diabetes [20].
Type 1 Diabetes (T1D) is a disease that affects about 5-10% of the diabetic population and it usually arises in young people (but it can occur in all ages). It is an autoimmune disease of the beta cells of the pancreas that causes their destruction and therefore a reduced production of insulin in blood circulation. Because of this, in patients with this type of diabetes there is the presence of antibodies against proteins of beta-cells. In genetically disposed people, T1D can be triggered by environmental factors or viral infections. People with this type of diabetes are insulin-dependent: they must assume an insulin therapy for all their life.
Type 2 Diabetes (T2D) is a disorder that affects about the 90-95% of diabetic people and it usually occurs in people older than 45 years, although in the last years it is diagnosed in even younger people. This type of diabetes is characterized by an insulin resistance of body tissues against insulin action and a relatively reduced production of insulin, that cannot overcome this problem, thus bringing to hyperglycemia.
T2D has a genetical basis (more than 50% of people with T1D have a member of their family with this disease), but other factors are important for the development of the disorder: lifestyle, obesity, physical activity, diet, and stress.

Diabetes may lead to different acute effects, for example [20]:

- Ketoacidosis: it is a metabolic acidosis due to an uncontrolled production of ketone bodies from muscles. It happens because tissues use fats instead of glucose as a source of energy and it is a consequence of hyperglycemia, that leads glucagon to promote lipolysis.

- Nonketotic hyperosmolar coma: high blood glucose levels bring the water to come out from cells and glucose is thrown out with urine (glycosuria). It causes dehydration, the osmolarity of blood increases thus leading to severe consequences such as coma.

- Hypoglycemia: causing confusion and loss of consciousness. Sometimes if the patient assumes a high amount of insulin he can come into hypoglycemic coma.

Diabetes has also long-term chronic consequences, most of them due to hyperglycemia, and damages concern principally microcirculation. Because of diabetes, blood vessels become weaker and different diseases can arise:

- Diabetes retinopathy: leading to damages on eye blood vessels and that brings to blindness.

- Diabetes nephropathy: leading to a damage of kidney tissue with an end-stage renal disease. This fact brings to a big loss of proteins in the urine and a resulting generalized body swelling (edema).

- Diabetes neuropathy: affecting peripheral nerves such as sensory neurons, motor neurons and the autonomic nervous system. Moreover, there is a loss of proprioception, and if the patients have an injury on the feet or legs there is the risk of developing ulcers and infections which can bring to amputation in the worst cases.

In the end, diabetes can hit also big vessels and people become more subjected to strokes, heart attacks and cardiomyopathy.
Nowadays about 537 million people worldwide have diabetes, compared to the 108 million in 1980, the majority living in low-and middle-income countries, and 1.5 million deaths are directly attributed to diabetes each year [27],[28] . Both the number of cases and the prevalence of diabetes have been steadily increasing over the past few decades, and it is estimated that they will increase to 700 million by 2045 [15]. According to International Diabetes Federation at least USD 966 billion dollars are spent for diabetes.
According to ISTAT in Italy in 2016 more than 3,200 million people have diabetes, twice as 30 years ago. In Italy 8.26 billion euros are invested by the Sistema Sanitario Nazionale to treat diabetic people (Fig. 1.3) [26].

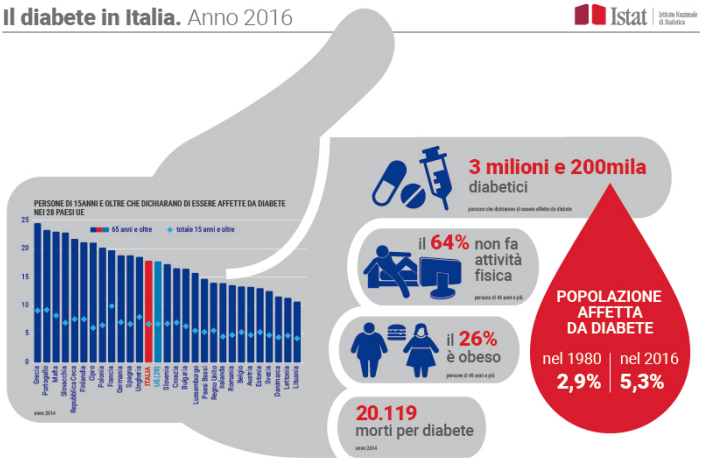


Figure 1.3: Diabetes in Italy in 2016 [26].

1.2 Diabetes Mellitus therapies

There are different therapies available according to the type of diabetes but all focus on mantaining the blood glucose level in the normal range [20].
In the case of T2D, patients can mantain a standard level of glycemia with a diet poor of carbohydrates, fats and other foods that raise blood glucose level. Most people with T2D are obese and they have to control how much food they assume and their weight, because obesity itself affects insulin resistance. Also physical activity and an active lifestyle are important to manage insulin resistance and weight. However, it can occur that patients can require a pharmacological treatment. Metformin is the most used drug for T2D patients, because it helps to mantain the glycemia in a normal range by lowering the amount of glucose produced in the liver and promoting the amount of insulin used in the body. In the recent years also other drugs have been developed, but in the worst cases patients can require insulin therapy [29].
However, in this thesis we will focus on the management and the therapy of T1D: this disease requires insulin treatment, that is mainly provided by insulin injections, whose purpose is controlling and keeping constant the system output (blood glucose) despite disturbances, that consists mostly in meals and physical exercise, but also other factors such as illness, emotions, stress and so on. In this regulation system control input is represented by insulin administration that can be carried out mostly in two ways: by insulin pens or insulin pumps [6].
In the first case 1 or 2 subcutaneous injections of long-acting insulin (at bed and/or wake-up time) are performed to guarantee a costant presence of insulin in the circulation. In addition, before and during meal time patients have to inject rapid acting insulin, that is called "meal bolus", because a fast increase of insulin concentration is needed to prevent postprandial hyperglycemia [18].
In the second case, i.e. with insulin pumps only rapid-acting insulin is used, where the basal level of insulin is mantained by programming the pump to administer microboluses every few minutes, while the meal bolus is delivered before each meal to face the rise of blood glucose concentration due to meal injection.
While in the first case injections are performed manually through Multiple Daily Injections (MDI), tuned by the physician according to patient's parameters, in the second one they are made with Continuous Subcutaneous Insulin Injections (CSII) thorugh pumps, that consist on wearable devices with a refillable reservoir of insulin connected through an intradermal catheter where the insulin flows into the subcutaneous tissue [21]. In some cases if blood glucose concentration is still high after meal, patients can assume an additional quantity of insulin, that is called "correction bolus". The dose can be calculated by the patients themselves, thanks to some

parameters established by physicians and blood glucose level.
In traditional T1D therapy glucose sensing is performed through devices for Self Monitoring Blood Glucose (SMBG), which take a small sample of blood from capillaries. On the other hand, in the last decades new devices like Continuous Glucose Monitoring (CGM) have been developed: they provide frequently interstitial glucose concentration measures about every 5-10 min.
CGM sensors are minimally invasive systems that consist of a needle electrochemical sensor. This is a great advantage compared to SMBG devices, that require a finger-prick every time the patient wants to know his blood glucose level.

Figure 1.4: Model of CGM sensor produced by Dexcom [25].

CGM devices are able to show glucose trends, they can collect patients' data and give a support to medical decisions. There are also algorithms for automatic optimization of insulin dosing parameters and for automatic insulin delivery. Thanks to predictive algorithms CGM sensors are able to predict 30-45 minutes before an event of hypo- hyperglycemia is going to happen, and generate alarms so that people can intervene before the incident occurs.
CGM sensors can also be used toghether with insulin pumps, i.e. Sensor Augmented Pump Therapy (Fig. 1.5), which represents the best open loop therapy or, in presence of a control algorithm, their combination is called artificial pancreas [12].

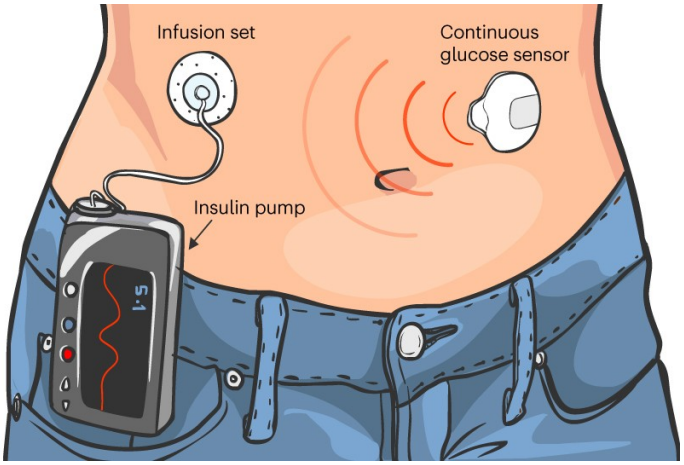


Figure 1.5: Scheme of Sensor Augmented Pump Therapy [17].

1.3 Insulin Sensitivity: State of the art

To better understand the control chain of the glucose-insulin system and the amount of insulin that must be delivered through insulin pumps or pens, it is important to define the concept of insulin sensitivity, that is the ability of insulin to stimulate glucose utilization and inhibit glucose production [19]. This value gives a quantitative estimate of insulin's ability to lower blood glucose concentration by both improving its utilization by muscles and adipose tissues and reducing its production by the liver [14].
Nowadays there are a lot of different methods to quantify the role of insulin sensitivity on glucose metabolism in different pathological states.

1.3.1 Insulin Sensitivity from plasma glucose and insulin data

The gold standard method is represented by the euglycemic clamp technique [9]: the plasma insulin concentration is raised thanks to an insulin infusion and, at the same time, glucose level is mantained constant at basal level by a variable glucose infusion. When the steady state of euglycemia is reached the glucose infusion rate is the same as the quantity uptake by tissues.
With the clamp technique insulin sensitivity is given by the formula:

$$S_I^{clamp} = \frac{M}{\Delta I \cdot G_b} \qquad with\, \Delta I = I_{ss} - I_b \tag{1.1}$$

Where M is the glucose infusion rate, ΔI is the difference between insulin in the

steady state and basal insulin and G_b blood glucose.
An additional model-based technique developed in 1979 is the "glucose minimal model", based on an IVGTT experiment [1]. This method, though it is invasive and non-physiological as the clamp techinque, is based on a mathematical model where the system is partitioned into a glucose and insulin subsystem, and insulin sensitivity is estimated by fitting the model on glucose data as output, while insulin is used as input.
Another method that is more physiologic and easier to apply than the previous one, is the "oral glucose minimal model method" [4],[7]. The oral glucose minimal method allows to estimate insulin sensitivity from a mixed-meal tolerance test (MTT) and the oral glucose tolerance test (OGTT) [4]. Both of them are less invasive and easier to administer than the IVGTT and the clamp technique, and they have been validated against them [8].

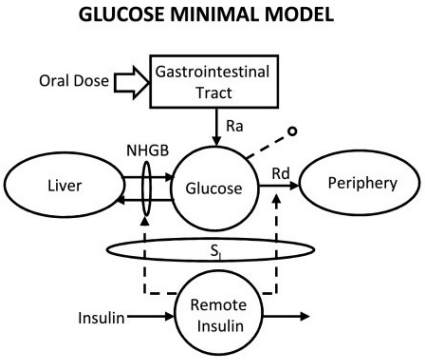


Figure 1.6: Oral Glucose Minimal Model [4].

The oral glucose minimal model is described by the following equations:

$$\begin{cases} \dot{Q}(t) = -\big[S_G + X(t)\big] \cdot Q(t) + S_G \cdot Q_b + Ra(\alpha, t) & Q(0) = Q_b \\ \dot{X}(t) = -p_2 \cdot X(t) + p_3 \cdot \big[I(t) - I_b\big] & X(0) = 0 \\ G(t) = \frac{Q(t)}{V_G} \end{cases} \tag{1.2}$$

Where Q is the plasma glucose mass, V_G is the glucose distribution volume, X is insulin action on glucose disposal and production, I is plasma insulin concentration and I_b its basal value. S_G is fractional glucose effectiveness measuring glucose ability per se to promote glucose disposal and inhibit glucose production, p_2 and p_3 rate constants describing insulin dynamics and magnitude. Ra is described as a piecewise linear function with known break point t_i and unknown amplitude α_i :

$$Ra(\alpha,t)=\begin{cases}\alpha_{i-1}+\frac{\alpha_i-\alpha_{i-1}}{t_i-t_{i-1}}\cdot(t_i-t_{i-1}) & t_{i-1}\le t\le t_i,\ i=1,...,8\\ 0\end{cases} \tag{1.3}$$

In the oral glucose minimal method insulin sensitivity is given by the formula:

$$S_I^{OMM}=\frac{p_3}{p_2}\cdot V_G \qquad (dl/kg/min \quad per\,\mu U/mL) \tag{1.4}$$

Insulin sensitivity has been calculated both with MTT and OGTT and the resutls of these two methods have been validated against the clamp technique, with a correlation of 0.86 and 0.81, respectively. MTT results have been validated also against IVGTT, with a correlation of 0.74 [4].

1.3.2 Insulin Sensitivity from Sensor-Augmented Pump Data in Type 1 Diabetes

A more recent method for the estimation of insulin sensitivity in T1D has been introduced by Schiavon et al. in 2014 [19]. Thanks to this method it is possible to evaluate insulin sensitivity in T1D patients from CGM sensor and insulin pump data in their daily life. This technique takes origin from the oral minimal model equations, that are adopted to work with SAP data, i.e. CGM and CSII data, instead of plasma glucose and insulin concentrations. The equations obtained by substituting $\mathrm{X(t)}=\frac{\mathrm{p_3}}{\mathrm{p_2}}\cdot\mathrm{X'(t)}$ and $S_G=\frac{GEZI}{V_G}+\frac{S_I}{V_G}\cdot I_b$ are:

$$\begin{cases}\dot{G}(t)=-\left[\frac{GEZI}{V_G}+\frac{S_I}{V_G}\cdot I_b+\frac{S_I}{V_G}\cdot X'(t)\right]\cdot G(t)+\left(\frac{GEZI}{V_G}+\frac{S_I}{V_G}\cdot I_b\right)\cdot G_B+\frac{Ra_G(t)}{V_G}\\ \qquad\qquad G(0)=G_b\\ \dot{X}'(t)=-p_2\cdot X'(t)+p_3\cdot\left[I(t)-I_b\right]\\ \qquad\qquad X(0)=0\end{cases} \tag{1.5}$$

Where G is plasma glucose concentration (mg/dL) and G_b its basal value. X, I, I_b, Ra_G, V_G, p_2 and p_3 are defined as in the minimal model method.
By integrating the differential equation from the time in which the meal is ingested t_{meal} to the end of the experiment t_{end} and by rearranging them, the result is:

$$S_I=\frac{\int_{t_{meal}}^{t_{end}}Ra_G(t)dt-GEZI\int_{t_{meal}}^{t_{end}}\Delta G(t)dt-V_G\int_{t_{meal}}^{t_{end}}\dot{G}(t)dt}{\int_{t_{meal}}^{t_{end}}X'(t)\cdot G(t)dt+I_b\int_{t_{meal}}^{t_{end}}\Delta G(t)dt} \tag{1.6}$$

The first integral element in the numerator can be written as:

$$\int_{t_{meal}}^{t_{end}} Ra_G(t)dt = \frac{D \cdot f(t_{end})}{BW} \tag{1.7}$$

Where D (mg) is the quantitiy of glucose absorbed during the meal and $f(t_{end})$ is the fraction of ingested dose which at $t = t_{end}$ has reached plasma and can be measured with CGM. The insulin action X' on glucose metabolism is usually unknown because model parameters are unknown, so that the denominator in the equation is substituted as follows:

$$\frac{1}{(t_{end} - t_{meal})} \cdot \int_{t_{meal}}^{t_{end}} I(t)dt \cdot \int_{t_{meal}}^{t_{end}} |\Delta G(t)|dt \tag{1.8}$$

Where the value $|\Delta G|$ represents plasma glucose concentration above basal level. Finally, S_I is given by:

$$S_I^{SP} = \frac{\frac{D \cdot f(t_{end})}{BW} - GEZI \cdot AUC(\Delta G) - V_G \cdot \left[G(t_{end}) - G(t_{meal})\right]}{AUC(I) \cdot \left[\frac{AUC(|\Delta G|)}{t_{end} - t_{meal}}\right]} \tag{1.9}$$

Where AUC is the area under the curve calculated from the beginning to the end of the experiment. The formula is patient-specific, infact body weight (BW, kg) height (m) and age (years) are used to estimate plasma insulin clearance (CL). Moreover, BW influences glucose rate of appearance, as shown above. In addition to patient-specific factors, parameters fixed to the population are set. They are $GEZI$, that is glucose effectiveness at zero insulin, that is fixed to 0.01 dL/kg/min for diabetic people; and V_G, volume of glucose distribution, that is fixed to 1.45 dL/kg [19].
G, ΔG and I are not directly available, but they're estimated from CGM and CSII data. Moreover, the value given by the product of $D(t)$ and $f(t)$ assumed during the meal can be generalized as $AoC(meal)$, that is the amount of carbohydrates ingested during the meal.
In the end, $AUC(I)$ is obtained by adding basal insulin rate during the period $[t_{meal}\, t_{end}]$, the bolus injected before the meal and insluin that can be still present

in plasma since before meal, insulin on board (IOB). Because all these measures are taken from plasma, they must be divided by plasma insulin clearance (CL). By adding these informations, the formula (1.9) can be rewritten as:

$$S_I^{SP} = \frac{\frac{AoC(meal)}{BW} - GEZI \cdot AUC(\Delta CGM) - V_G \cdot \left[CGM(t_{end}) - CGM(t_{meal})\right]}{\left[\frac{AUC(Basal)+AUC(Bolus)+IOB}{CL}\right] \cdot \left[\frac{AUC(|\Delta CGM|)}{t_{end}-t_{meal}}\right]} \tag{1.10}$$

In this formula it is important to define its domain of validity. In fact in certain cases S_I can assume negative values, so it is important to pay attention that recalibrated ΔCGM is lower than a certain threshold (150 mg/dL) after 6h from meal ingestion, otherwise this method gets out from its domain of validity. This method has been validated against the oral minimal method with a correlation of 0.825 [19].

1.4 Aim of the thesis

This thesis aims to develop a method for modeling the intra-day and inter-day variability of insulin sensitivity in patients with Type 1 Diabetes. In particular, the work has been done in a simulation enviroment using the UVA/Padova T1D simulator. Three experiments have been performed on the 100 virtual adult population by simulating 1, 30 and 90 days with different levels of variability of insulin sensitivity. CGM and insulin pump data were simulated and used to develop and test a new formula for the estimation of insulin sensitivity from the one introduced by Schiavon et al. [19], and results have been validated against the oral glucose minimal model insulin sensitivity estimated from glucose and insulin concentration data using the first dataset of one day of simulation. Then this methodology has been used to assess intra-day pattern of insulin sensitivity using the dataset of 30 days of simulation and validated against the pattern estimated from the oral glucose minimal model. In the end, the 90 days of simuation have been used for the estimation of the inter-day variability of insulin sensitivity, by means of Markov Chain models which was then compared against the one randomly generated and included into the simulator for each subject.
The thesis is structured as follows:
In Chapter 2 there is the description of the databases used in the study.
In Chapter 3 methods used to estimate insulin sensitivity and to model its intra- and inter-day variability are exposed.
In Chapter 4 results obtained by applying the methods presented are showed.
Finally, in Chapter 5 results obtained in this thesis are discussed and future devel-

opments are presented.

Chapter 2

Database

To estimate intra- and inter- subject variability of insulin sensitivity it is important to assess it from data with an extended duration as well as from a large cohort mimicking the variability of an entire population. With this purpose, mathematical models have been created to reproduce a heterogeneous population of individuals mimicking the variability of a real population. The use of modeling and simulation has accelerated results and achievements in pharmacological studies.

2.1 The UVA/Padova T1D Simulator

The UVA/Padova Type 1 Diabetes simulator is a tool that allows to design and test certain insulin treatmets for type 1 diabetic patients. It has been approved in 2008 by FDA as a substitute for preclinical trials of certain insulin treatments [11].
The introduction of computer simulation represents an effective alternative to animal experiments, and a great saving of time and money, thanks to the possibility to perform in silico tests. Also the evolution of artificial pancreas control and prediction algorithms has been sped up thanks to the UVA/Padova T1D simulator. The software consists of a population of 300 subjects: 100 adults, 100 adolescents and 100 children, characterized by intra and interpersonal differences and real-life variability of T1D population.

	Adults			Adolescents			Children		
Parameter	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max
Body weight (kg)	75.2 (12.1)	52.6	108.4	50.8 (8.8)	30.4	80.3	30.7 (5.3)	18.3	49.8
Insulin (U/day/kg)	0.56 (0.16)	0.32	1.15	0.68 (0.22)	0.30	1.32	0.56 (0.16)	0.34	1.18
CR breakfast (g/U)	16.9 (7.9)	6.0	40.0	14.5 (7.1)	4.0	32.0	21.8 (8.7)	9.0	45.0
CR lunch (g/U)	19.7 (8.2)	4.0	38.0	17.0 (7.0)	6.0	33.0	25.1 (8.5)	7.0	45.0
CR dinner (g/U)	20.6 (9.1)	5.0	40.0	17.8 (7.4)	5.0	35.0	26.4 (10.2)	6.0	51.0
Fasting BG (mg/dL)	119.1 (7.1)	102.1	134.5	120.8 (6.6)	106.4	140.5	119.6 (6.0)	106.0	133.6

Table 2.1: Key Demographic and Metabolic parameters of the In Silico Subjects Available in the Simulation Enviroment [22].

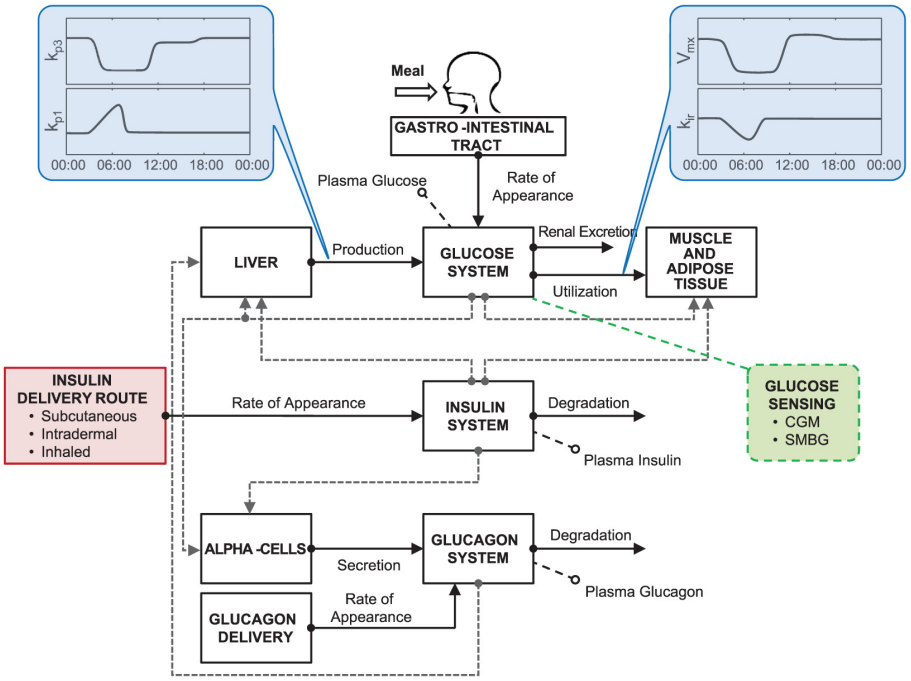


Figure 2.1: The UVA/Padova T1D simulator in his last version (S2017) [22].

The last version of the T1D simulator (S2017) [22] is able to describe a series of factors that could affect glucose metabolism during the day, like intra-day variability of insulin sensitivity, which was assessed by using a triple-tracer method with three meals (breakfast, lunch and dinner) with the same amount and composition, so that the estimation of the intra-day pattern of insulin sensitivity was not influenced by differences of quantity and dimension of meals, but only by the insulin sensitivity itself. Because of the large inter-subject variability, each in silico subject belongs to one among seven possible variability classes, that represents seven different patterns of insulin sensitivity (Fig. 2.2).

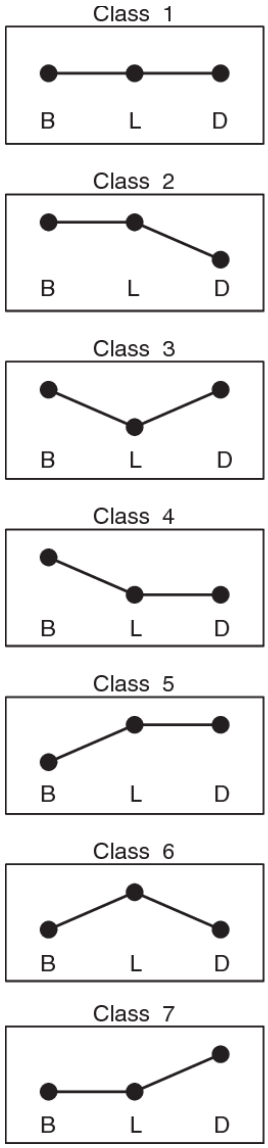


Figure 2.2: Table of the seven possible combinations of the values 'high' and 'low' for Breakfast (B), Lunch (L) and Dinner(D). Each pattern is associated to a class from 1 to 7 [23].

Another important concept introduced in the most recent development of the software is dawn phenomenon, which is defined as the rise in blood glucose concentration that occurs in T1D subjects during early morning, due to an increasing in endogenous glucose production and insulin requirement.
Finally, each in silico individual is characterized by a daily pattern of basal insulin rate that regulates the effects of insulin sensitivity and dawn phenomenon on glucose oscillations. Also carbohydrate-to-insulin ratio (CR) and the correction factor (CF) are considred as time-varying parameters and calculated in the model.

2.2 In silico experiments

In this thesis three datasets using the 100 adult population have been generated:

1. The first dataset simulates data from one day of observations for every subject and each one gets three meals: breakfast at 7:00, lunch at 13:00 and dinner at 19:00, with an amount of 40, 80 and 60 g of carbohydrates rispectively. It is used to test a new method to estimate insulin sensitivity from CGM and pump and validated against the oral minimal model method at breakfast lunch and dinner.

2. The second dataset has the same protocol, but observations last 30 days. This dataset was used to test the ability of the method to estimate S_I and its intra-day variability as simulated from the UVA/Padova T1D simulator [22]. Of note, each subject is characterized by a specific intra-day pattern of insulin sensitivity to which a random gaussian noise was added allowing to vary the daily values of insulin sensitivity by $\pm 20\%$ of its nominal values thus possibly resulting in random variations of intra-day pattern from their nominal one.

3. In the third database, in addition to the patient-specific intra-day pattern, a Markov Chain model was used to simulate inter-day variability of intra-day pattern of insulin sensitivity and 90 days of simulation were performed. In particular, a Markov Chain model was developed and incorporated into the simulator to randomly generate variations of the intra-day pattern of insulin sensitivity and used to validate the ability of the proposed methods to reconstruct the variations of intra-day pattern of insulin sensitivity.

For all the datasets, CGM and insulin pump data, together with plasma glucose and insulin concentrations, meal amounts and subject demographics, were generated to allow the estimation of insulin sensitivity with the proposed methods. The use of three datasets is necessary because, as it was anticipated in the Section 2.1, insulin sensitivity shows both circadian (intra-day) and inter-day variations. In fact it has been demonstrated that also in T1D patients insulin sensitivity is subjected to changes in its trend due to meals, physical activity and modifications in glucose metabolism.

For every patient the following informations are available:

- Subject's specific model parameters;
- CGM and insulin pump dose;

- Glucose and insulin concentration dose;
- Amount of carbohydrates ingested.

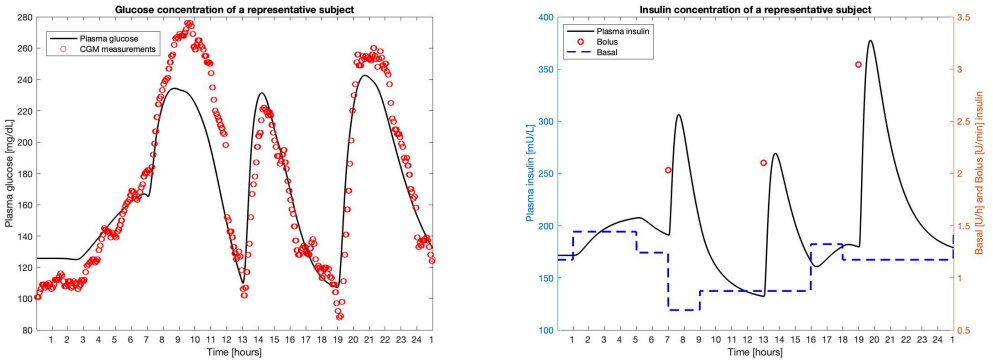


Figure 2.3: Glucose and Insulin of a representative subject in one day of simulation.

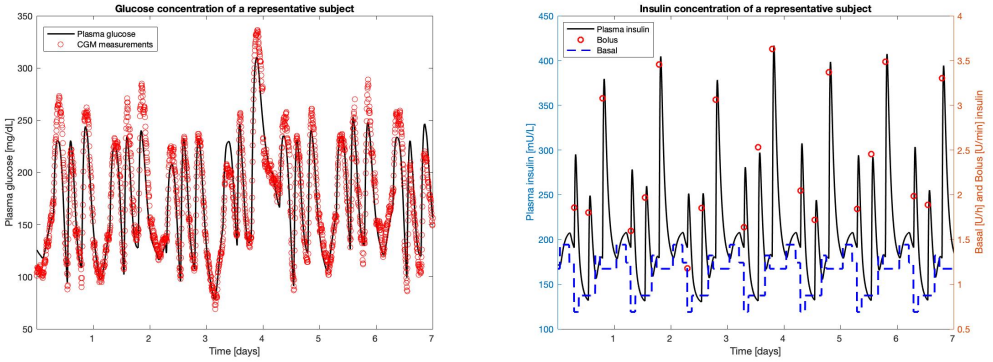


Figure 2.4: Glucose and Insulin of a representative subject in the first seven day of 30 days of simulation.

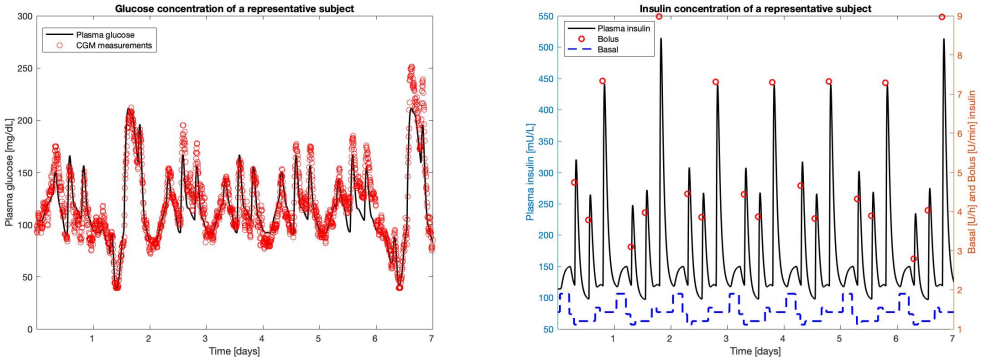


Figure 2.5: Glucose and Insulin of a representative subject in the first seven day of 90 days of simulation.

Chapter 3

Methods for the Estimation of the Variability of Insulin Sensitivity

In this Chapter all the methods used to assess insulin sensitivity and its variability are explained in detail.
As introduced in Subsection 1.3.2, in 2014 Schiavon et al. introduced a new method to calculate insulin sensitivity from minimally invasive technologies (i.e. SAP therapy) [19], thus allowing to calculate it in everyday conditions.
In this Chapter a new version of the formula is developed, allowing to extend its use also in uncontrolled experiments conditions possibly occuring in daily life, and validated against the oral glucose minimal model method.
The following step is to use this methodology to estimate the intra-day variability of insulin sensitivity of each subject of the UVA/Padova T1D simulator, by assessing the diurnal pattern of insulin sensitivity defined by a specific class of variability.
Lastly, Markov models are used to simulate the inter-day variability of diurnal pattern of insulin sensitivity in the UVA/Padova T1D simulator. By using the proposed methodology, we try to reconstruct the inter-day variability of intra-day pattern of insulin sensitivity here ad hoc incorporated into the simulator and compare it against the original one for each subject.

3.1 Insulin Sensitivity from CGM and insulin pump (S_I^{SP}): a new formulation

As stated above, the estimation of insulin sensitivity with SAP has the advantage to calculate S_I in daily life conditions of T1D patients. The other important aspect is that it allows to get patient-specific estimates, in fact it is calculated for each person according to his/her own parameters.
However the original S_I^{SP} index (2014) [19] was developed and validated on data

collected in strictly-controlled experimental conditions. Here the formula has been updated by adding some components allowing to be used in more real-life conditions, like not well-spaced consecutive meals and/or in presence of hypoglycemic events. To do this, two modifications have been introduced: the first one is the ability to account for the peculiar dynamics of insulin action in the hypoglycemic range by adding:

$$\begin{cases} 0 & if\ G(t) \geq G_b \\ 10 \cdot \left[log\left(\frac{G(t)}{G_b}\right)^{r_1} \right]^2 & if\ G_{th} \leq G(t) \leq G_b \\ 10 \cdot \left[log\left(\frac{G_{th}}{G_b}\right)^{r_2} \right]^2 & if\ G_{th} < G(t) \end{cases} \tag{3.1}$$

Where G_{th} is the hypoglycemic threshold, set at 60 mg/dL. The second modification is the introduction of r_1 and r_2, which are model parameters that account separately for basal and bolus insulin determining insulin action.
Thus the new formulation of insulin sensitivity (S_I^{SP}), calculated for each meal of the simulation, is the following:

$$\mathrm{S_I^{SP}} = \frac{\frac{\mathrm{AoC(meal)}}{\mathrm{BW}} - \mathrm{GEZI} \cdot \mathrm{AUC}(\Delta\mathrm{CGM}) - \mathrm{V_G}\left[\mathrm{CGM}(\mathrm{t_{end}}) - \mathrm{CGM}(\mathrm{t_{meal}})\right]}{\frac{1}{\Delta \mathrm{t}} \cdot \left\{ \frac{\mathrm{AUC(Basal)}}{\mathrm{CL}} \cdot \mathrm{AUC}(\Delta\mathrm{CGM}) + \frac{\Delta\mathrm{IOB}}{\mathrm{CL}} \cdot \mathrm{AUC}(\Delta\mathrm{CGM}) + \frac{\mathrm{AUC(Bolus)}}{\mathrm{CL}} \cdot \mathrm{AUC}\left[(1 + \mathrm{r_1} \cdot \mathrm{risk}) \cdot \mathrm{CGM}\right] \right\}} \tag{3.2}$$

Where: $AoC(meal)$ [mg] is the amount of carbohydrates ingested during the meal, BW is the weight of the patient [kg], $GEZI$ is glucose effectiveness in absence of insulin, fixed to 0.01 [dL/kg/min], V_G is the volume of glucose distribution, fixed to 1.45 [dL/kg], $AUC(...)$ is the area under the curve of the parameter, calculated in the interval between t_{meal} and t_{end}, CGM is the value of plasma glucose measured by the sensor, CL is the clearance of plasma insulin [L/min], $Basal$ and $Bolus$ represent the amount of prandial insulin and insulin injected [mU/min] respectively, ΔIOB is insulin still active in plasma in the period between 6 hours before t_{meal} and t_{end} and $risk$ the function described in (3.1).
However as done in [19], since the method may be not robust enough if some conditions are not met, it is important to define the conditions where the formula falls outside its domain of validity. Specifically, S_I^{SP} can not be calculated if on or more of the following conditions occur:

- AUC(ΔG(pos))/AUC(ΔG(neg)) < 0.6: if the ratio between glucose excursions above or below basal level is below 60%, it possibly lead to compensatory mechanisms not accounted in the model;

- |der(G)| > 2 mg/dL at meal time: the system is not stable at meal time due to glucose oscillations;
- $G_b < 60$ mg/dL or $G_b > 200$ mg/dL: basal level are far from the target values;
- $S_I \leq 0$: not meaningful value possibly due to noise on data;
- meal length < 180 min: not sufficient information to calculate S_I;
- no basal infusion: missing basal information;
- no meal bolus: missing bolus information;
- distance between consecutive CGM samples > 30 min: missing glucose information.

3.1.1 Validation against the Oral Glucose Minimal Model S_I (S_I^{OMM})

To assess the reliability of results obtained with the formula (3.2), a dataset of 100 subjects from the UVA/Padova T1D simulator was used, and results have been validated against S_I^{OMM}, obtained with the oral glucose minimal model.
In this step CGM and insulin pump, together with glucose and insulin concentrations were simulated in one-day scenario, so that for each subject three values of S_I^{SP} and S_I^{OMM} were obtained, one for breakfast, one for lunch and one for dinner respectively.
In particular, S_I^{OMM} values were obtained by fitting the oral glucose minimal model on plasma glucose data with plasma insulin as forcing function, using a Bayesian Maximum a Posteriori Estimator with prior information derived from [7]. Then a regression analysis was done comparing S_I^{SP} versus S_I^{OMM} and the parameters α and β of the linear regression were estimated:

$$S_I^{SP} = \alpha * S_I^{OMM} + \beta \tag{3.3}$$

In Matlab coefficients have been calculated with the command `fitlm`.
Later, the Pearson correlation coefficient R between S_I^{SP} and S_I^{OMM} has been obtained by using the formula:

$$R(S_I^{SP}, S_I^{OMM}) = \frac{cov(S_I^{SP}, S_I^{OMM})}{\sigma_{S_I^{SP}} \sigma_{S_I^{OMM}}} \tag{3.4}$$

Where $cov(S_I^{SP}, S_I^{OMM})$ is the covariance between the vectors S_I^{SP} and S_I^{OMM} , $\sigma_{S_I^{SP}}$

is the standard deviation of S_I^{SP} and $\sigma_{S_I^{OMM}}$ the standard deviation of S_I^{OMM}.
In Matlab it was simply obtained with the command `corrcoef`.
It is important to note that in the correlation analysis we included only S_I^{OMM} values estimated with an acceptable precision, as estimated with a coefficient of variation less or equal to 50%. In fact, the oral glucose minimal model must have reliable results, since it is considered as the point of reference for the validation of S_I^{SP}.

3.2 Classification of intra-day pattern of Insulin Sensitivity

After the correlation between S_I^{SP} and S_I^{OMM} has been assessed, the estimation of intra-day pattern of insulin sensitivity has been performed by using the dataset of 30 days of simulations for each subject.
By construction, each virtual subject behaves to a specific pattern of intra-day variability of insulin sensitivity, i.e. with three values of insulin sensitivity during the day (at breakfast, lunch and dinner), to which a random gaussian noise of 20% of its nominal value is added each day of the simulation.
However, in order to assess the daily pattern of insulin sensitivity among the 7 possible classes, firstly a model of the error in the estimation of S_I was needed, both for S_I^{SP} as well as S_I^{OMM}.

3.2.1 Estimation of intra-day pattern from S_I^{SP}

To do this, three different approaches have been used for the comparison between S_I^{SP} values, or between their "normalized" values obtained by dividing the daily S_I^{SP} by their daily maximum, by means of a statistical test. To do this we employed the z-test, which is a statistical test used to assess if the mean value of a distribution is significantly different from another with known values of their variances [16]. The level of significance (α) is set to 0.05.
The z-test calculates the standard score with the following formula:

$$z = \frac{(x_1 - x_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{3.5}$$

Where x_1 is the maximum S_I^{SP} of the day, or equal to one in the "normalized" case and x_2 one of the other two S_I^{SP} estimates, or their "normalized" values, σ_1^2 and σ_2^2 the variances of the two estimates and n_1 and n_2 the number of samples used for x_1

and x_2, that are always equal to 1 because all the estimates are scalar.
The null hypotesis H_0 is that x_1 and x_2 are similar, so, with α fixed to 0.05, if $-1.96 < z < 1.96$ H_0 is accepted and the value 'high' (h) is assigned to x_2 too.
On the other hand, if $z < -1.96$ or $z > 1.96$ H_0 is rejected and x_2 assumes the value 'low' (l) [23]. After the z-test has been calculated, each subject has for every day of the simulation a sequence of three number, that corresponds to 1 if the value assumed in the meal is 'high', or 0 if the value is 'low'.
When the intra-day pattern is found, the corresponding class is assigned according to the scheme defined in Fig. 2.2.
Hence in order to properly apply the z-test, the knowledge of the variance of S_I is fundamental for the determination of the pattern, because it influences the calculation of the z-value in Eq. (3.5).
In this thesis three different methods have been proposed to assess the variance of the estimated S_I^{SP}:

Method 1. The first method consists of calculating the variance from the Mean Squared Error (MSE) of the estimated S_I^{SP} ($S_I^{SP,est}$) and its prediction from the oral glucose minimal model ($S_I^{SP,pred}$):

$$MSE = \frac{\sum_{i=1}^{n} \left(S_I^{SP,est} - S_I^{SP,pred}\right)^2}{n-1} \tag{3.6}$$

The sample variance of the MSE is calculated within intervals of n elements spanning the entire axis and is fitted the model of the variance defined by the following equation:

$$var(S_I^{SP}) = s^2 = \gamma \cdot (S_I^{SP,pred})^2 + \delta \tag{3.7}$$

The identification of the parameters γ and δ of the model (3.7) was done by means of Weighted Least Squares (WLS) estimator:

$$\left[\gamma, \delta\right] = \left(G^T \Sigma_v^{-1} G\right) G^T \Sigma_v^{-1} s^2 \tag{3.8}$$

Where G is the matrix of the model in Equation (3.7) ($G = \begin{bmatrix} (S_{I,1}^{SP})^2 & 1 \\ ... & ... \\ (S_{I,n}^{SP})^2 & 1 \end{bmatrix}$), s^2 is the sample variance obtained from the MSE in Equation (3.6) for each

specific interval and Σ_v the matrix of the variance of the estimation error of s^2, calculated as [3]:

$$var(s^2) = \frac{2s^4}{(n-1)} \tag{3.9}$$

calculated for each specific interval.

Method 2. The second method consists, assuming the same description of variance as function of S_I^{SP}, of finding γ and δ so as to minimize the error in pattern classification between the ones calculated from S_I^{SP} and S_I^{OMM}.
This is done by iteratively randomly selecting combinations of γ and δ which are then put in the Equation (3.7) to calculate the variance used for the z-test for pattern classification using S_I^{SP}. Then, for each combination of γ and δ the classification error against S_I^{OMM} is calculated and the best combination is selected.
In this case the function implemented in Matlab is `fminsearch`, that finds the parameters that minimize the cost function, which is the classification error between S_I^{SP} and S_I^{OMM}.

Method 3. The method is applied on daily values of S_I^{SP} which are normalized against the maximum estimation of each day for each subject.
In this way the "normalized" S_I^{SP} values are within the region [0,1] and every day has a maximum that is equal to 1 and the other values of that particular day are between 0 and 1. Then the sample variance of the "normalized" S_I^{SP} is calculated in each 0.1-long interval from the MSE of the elements of that interval.

Finally, once the intra-day pattern of S_I^{SP}, for each day and subject, has been binary quantified using the z-test with the selected method for the calculation of the variance, each patient was associated to the corresponding class as previously defined in Fig. 2.2.

3.2.2 Validation

As in the regression analysis explained in Subsection 3.1.1, the classification of daily pattern with S_I^{SP} need to be validated against the one estimated with S_I^{OMM}.
The method used for the classification of the daily pattern of S_I^{OMM} is the same as

S_I^{SP}, i.e. employing the z-test methodology, but the variance is calculated as:

$$var(S_I^{OMM}) = \left(CV * S_I^{OMM}\right)^2 \tag{3.10}$$

Where CV is the coefficient of variation specific for each measure of S_I^{OMM}, representing the uncertainty in its estimation. After the correlation analysis, only S_I^{OMM} estimated with an acceptable precision ($CV < 50\%$) were included in the analysis. The validation was made by comparing classes assigned with S_I^{SP}, with all the methods proposed in Subsection 3.2.1 with those assigned with S_I^{OMM}.

3.3 Modeling inter-day Variability of intra-day pattern of Insulin Sensitivity

To model the inter-day variability of intra-day S_I pattern a stochastic approach, based on Markov Chains models, is employed. This is based on the fact that a patient can have a daily pattern of insulin sensitivity from which he/she can move based on behavioral changes. In particular, these models describe the probabilities of transition from a class to another based on the present state and not on the previous transitions, i.e. memoryless.
This represents a first step towards modeling S_I variability on the long-term, which can be tested whenever real datasets are available and/or further updated to account also for trends among days as well as seasonal variations.

3.3.1 Markov Chains models

Markov Chains are discrete time stochastic processes (i.e. collection of random variables): in fact at each time t, where $t = 1 \dots n$, the chain belongs to a countable set S called "Space State", defined as $S = \left\{S_1, S_2, \dots, S_n\right\}$ [10].
At each time t the model is in a state S_t and the probability that it assumes the state S_{t+1} at the successive iteration $t + 1$ is called p. Since the transition from a state to another does not depend on the path made before time t, the system is called "memoryless", and has the following property:

$$P\Big[X(t+1) = S_{t+1}|X(t) = S_t, X(t-1) = S_{t-1}, \ldots, X(0) = S_0\Big] = P\Big[X(t+1) = S_{t+1}|X(t) = S_t\Big] \quad (3.11)$$

Moreover we considered, at least in this preliminary approach, homogeneous Markov Chains, i.e. since transition probabilities p do not depend on time:

$$P\Big[X(t+1) = S_{t+1}|X(t) = S_t] = P\Big[X(1) = S_1|X(0) = S_0\Big] = p \quad (3.12)$$

Thus Markov Chain can be represented by a transition matrix, that summarizes all the informations described above:

$$P = \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{bmatrix} \quad (3.13)$$

Since it is a stochastic matrix, each element is non negative and each line sums up to 1.
Markov Chains have been used for the modeling of inter-day variability because they benefit of some important properties for modeling insulin sensitivity variations among days:

- They are able to model an extended time horizon of events, helping to model S_I variability in T1D patients' lives in the long period.

- Differently from other stochastic models they assume that events can reoccur: it is very common infact that a subject belong to the same class for more than one day.

- During each cycle the patient may make a transition from one state to another. In fact inter-day variability of a subject is based on the transitions that he/she does from a class to another one every day.

- In the end, Markov Chains assume that states are mutually exclusive: this means that in one day a patient belongs to one and only one class.

3.3.2 Estimation of Markov transition matrix from S_I^{SP}

In this thesis we aim to estimate the transitions that each subject makes among the seven possible classes during 90 days of simulation.
Of note, three different scenarios have been taken into account for the classification of intra-day pattern of S_I^{SP} used in the calculation of transition matrices, also including the case of missing S_I^{SP} estimates during one day:

- The first one is that the day has no missing S_I^{SP} estimates: in this case the class is assigned based on the corresponding pattern;
- The second case concerns the presence of one missing S_I^{SP} estimate in the day. In this case there are two possible solutions:
  - If both the other two estimates are classified as 'high' (h), then 4 possible classes are considered for that specific day:
    - The missing value is set to 'low' (l) and the other two estimates remain 'high'(h);
    - The missing value is set to 'high' (h) and the other two estimates remain 'high'(h);
    - The missing value is set to 'high' (h) and one of the other two values is modified from 'high' (h) to 'low' (l);
    - The missing value is set to 'high' (h) and both other estimates are set as 'low' (l).
  - If only one other estimate is 'high' (h), then 3 possible classes are considered for that specific day:
    - The missing value is set to 'low' (l) and the estimate 'high' (h) is not modified;
    - The missing value is set to 'high' (h) and the estimate 'high' (h) is not modified;
    - The missing value is set to 'high' (h) and the estimate is modified from 'high' (h) to 'low' (l).
- The third case is when more than one S_I^{SP} estimate is missing for that day. In this case there are not enough informations to assign a pattern for that day, which is considered missing.

After the classes have been calculated, the transition probabilities p_{ij} are calculated by normalizing the quantity of passages from i to j against the total number of transitions to j. Transition matrices built to model the inter-day variability of insulin sensitivity share the same structure with the one theorized in the formula (3.13): the structure consists of a square matrix for each subject, with dimension equal to 7, i.e. the classes of the daily pattern of insulin sensitivity. In the diagonal the probability of staying in the same class after a day was represented, while the other elements of the matrix corresponded to the probability of transition from a class to another, according to the number of row and column.

3.3.3 Validation

The validation of the methodology proposed in Subsection 3.3.2 is assessed by comparing, for each subject, the results obtained against the ones randomly generated and incorporated into the UVA/Padova T1D simulator, used to generate the 90 days database.
In order to assess a metric for the comparison between the transition matrix estimated with the methodology based on S_I^{SP} and the one simulated with the UVA/Padova T1D simulator, we exploit one of the properties of the transition matrix itself, i.e. that each row sum up to 1. The error (%) committed by modeling the transition matrix has been calculated in two ways:

1. The first method measures the similarity of the total transition matrix: it takes into account all the rows. In this case the error (%) is calculated as:

$$err_{tot}(\%) = \left(\frac{1}{7} \sum_{i=1}^{7} \frac{1}{max\big(p_i + \hat{p}_i, 1\big)} \sum_{j=1}^{7} |p_{ij} - \hat{p_{ij}}| \right) \cdot 100 \tag{3.14}$$

2. The second method measures the similarity of the effective transition matrices, i.e. accounting only for the rows of the original matrix which are different from zero, known by construction. The error (%) in this case is:

$$err_{eff}(\%) = \left(\frac{1}{k} \sum_{i=1}^{7} \frac{1}{p_i + \hat{p}_i} \sum_{j=1}^{7} |p_{ij} - \hat{p_{ij}}| \right) \cdot 100 \tag{3.15}$$

Where k is the number of rows different from zero in the original matrix, p_{ij} and $\hat{p_{ij}}$ are the transition probabilities of going from class i to class j in the original and estimated matrices respectively, p_i nand $\hat{p}_i$ are the sum of the probabilities of i in the original and estimated matrices respectively.

Chapter 4

Results

In this Chapter results obtained with the methods described in Chapter 3 are presented. In particular, the first section is about the regression and the correlation analysis between the new index here proposed of insulin sensitivity from minimally invasive technologies (S_I^{SP}) and the reference one based on the oral glucose minimal model (S_I^{OMM}).
The second one deals with the estimation of intra-day pattern of insulin sensitivity from S_I^{SP} and its validation against the one estimated from S_I^{OMM}.
The last section is about the modeling of inter-day variability of intra-day pattern of insulin sensitivity based on Markov Chains models and its validation against the Markov Chain model built for each subject and incorporated into the UVA/Padova T1D simulator.

4.1 Regression analysis and Validation of S_I^{SP} against S_I^{OMM}

As it was explained in Subsection 3.1.1, the validation of the estimation of insulin sensitivity based on minimally-invasive technologies is made against the oral minimal model one.
The analysis has been performed using the one day scenario where three meals (breakfast, lunch and dinner) have been performed with patient-specific insulin sensitivity varying between the day based on patient-specific intra-day pattern of insulin sensitivity.
In the figure (Fig. 4.1) estimates of S_I^{SP} are plotted against S_I^{OMM} with different colors for breakfast lunch and dinner, and the regression made for each meal is represented. Moreover, in the same figure, linear regression calculated by considering all the estimates is represented with the red line.

	α	β	**R**	p-value
Breakfast	1.1125	-1.1363	0.8385	7.22e-20
Lunch	0.90649	0.005234	0.9256	2.79e-42
Dinner	0.96279	0.34757	0.8996	9.02e-34

Table 4.1: Values of the parameters identified for each meal with the one-day scenario.

As previously reported, S_I^{OMM} values estimated with a CV higher than 50% are not considered in the regression analysis.

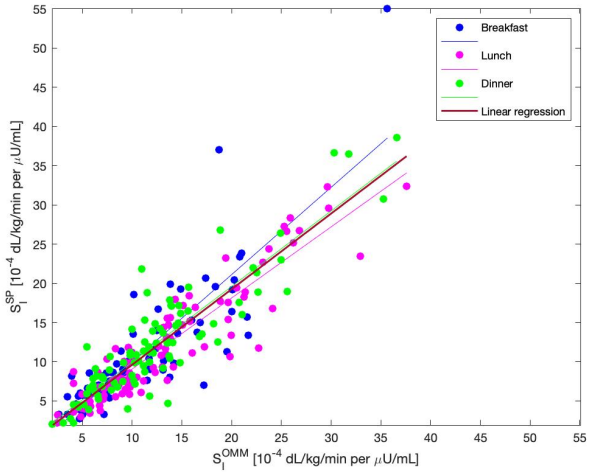


Figure 4.1: S_I^{SP} vs S_I^{OMM} from the one day scenario. Each meal has a different color: breakfast is blue, lunch pink and dinner green. Regression analysis is represented with a red line.

The parameters calculated from the linear regression analysis performed with the dataset of one day of observations will be used in the whole study. In addition, we also performed the same analysis using the 30-days simulation scenario where patient-specific insulin sensitivity values where also allowed to randomly vary within ±20% of their nominal value among days. As can be observed in Table 4.2 parameters identified with the 1 day and the 30 days scenarios are similar:

	α	β	**R**	p-value
1 day	0.96398	-0.024904	0.8846	1.99e-87
30 days	0.93102	0.52968	0.8788	0

Table 4.2: Values of the parameters identified with the one-day and 30-days scenarios.

In fact as it is evident from the next figure (Fig. 4.2), the parameters and the coefficient of correlation R are similar to which obtained with the dataset of 30 days

of observations.

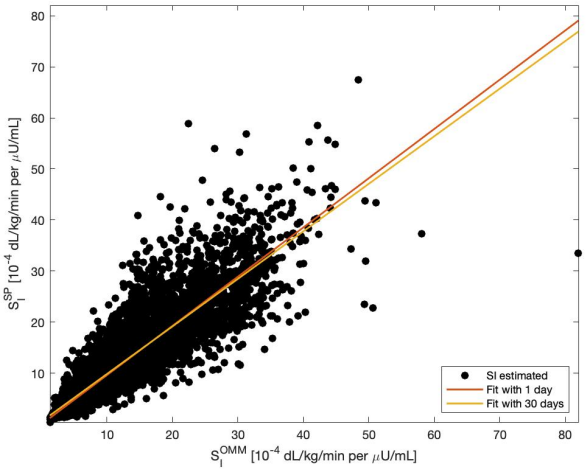


Figure 4.2: S_I^{SP} vs S_I^{OMM} from the 30-days simulation scenario (black dots) and regression analysis (yellow line). In addition regression analysis obtained from the one-day simulation scenario (red line).

It is important to note that the error between S_I^{SP} and S_I^{OMM} tends to increase with increasing values of S_I: this is a qualitative evaluation that confirms the law expressed in Equation (3.7), because the higher S_I the higher the error in S_I^{SP} estimation.

From the Table 4.2 one deduces a high correlation between S_I^{SP} and S_I^{OMM} ($R = 0.8846$), but also the linear regression obtained with the same process on the dataset based on 30 days of simulation (the yellow line in Figure 4.2), reflects a high correlation ($R = 0.8788$) with similar regression coefficients.
Since it is known that CGM sensors are affected by noise, the same analysis is performed by using plasma glucose data, whose measurements are more accurate than CGM by construction (random gaussian noise with zero mean and coefficient of variation equal to 2%). The following figures represent the regression analysis between S_I^{SP} calculated from plasma glucose vs S_I^{OMM}:

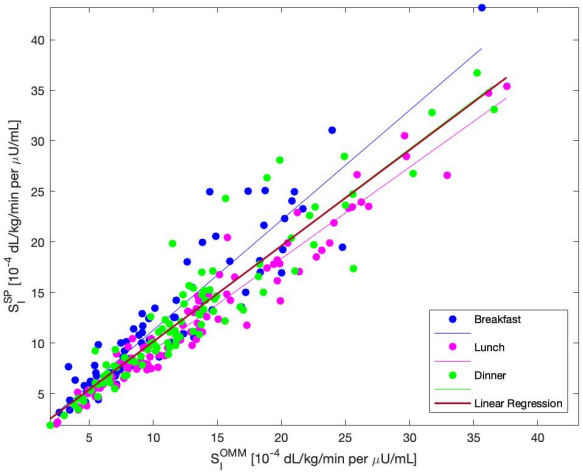


Figure 4.3: S_I^{SP} vs S_I^{OMM} from the one day scenario. Each meal has a different color: breakfast is blue, lunch pink and dinner green. Regression analysis is represented with a red line.

The parameters obtained by performing linear regression with these data are:

	α	β	**R**	p-value
Breakfast	1.0858	0.4447	0.9298	1.16e-31
Lunch	0.90335	0.28957	0.9757	9.21e-66
Dinner	0.9524	0.64584	0.9372	7.54e-44

Table 4.3: Values of the parameters identified for each meal with the one-day scenario.

As it expected and evident from Table 4.3, the coefficient of correlation improves by using glucose in the calculation of S_I^{SP}. As before, these parameters can be used also in the analysis done with the dataset of 30 days of simulation, based on results showed in the Table 4.4:

	α	β	**R**	p-value
1 day	0.94655	0.69528	0.9384	6.34e-123
30 days	0.93541	0.69151	0.9282	0

Table 4.4: Values of the parameters identified with the one-day and 30-days scenarios.

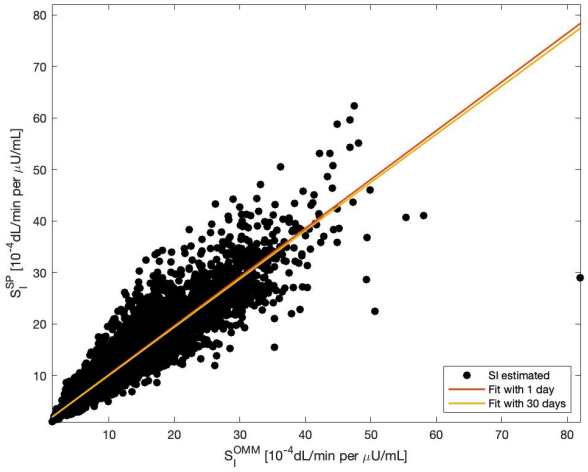


Figure 4.4: Prediction with 1 day and 30 days parameters on 30 days dataset.

By performing regression analysis with parameters identified with the 30-days dataset, the coefficient of correlation R is equal to 0.9282, that is a bit worse than that calculated with 1-day simulation (see Table 4.4), but it is still higher than that calculated by using CGM in the formula.

4.2 Classification of intra-day pattern of Insulin Sensitivity

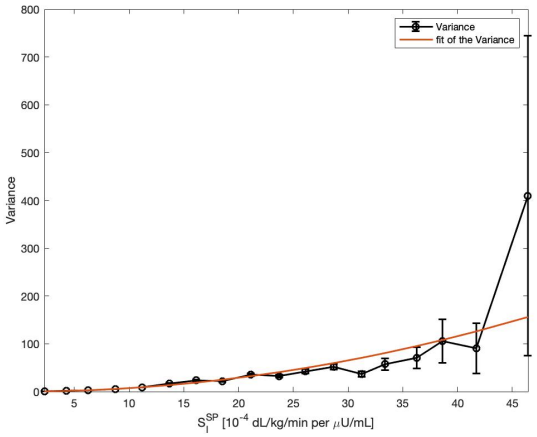
As previously described in Subsection 3.2.1, intra-day pattern of insulin sensitivity was calculated by performing a statistical analysis. However, to do so, the uncertainty on insulin sensitivity estimation, as defined by its variance, is required.

4.2.1 Measurement error of S_I^{SP}

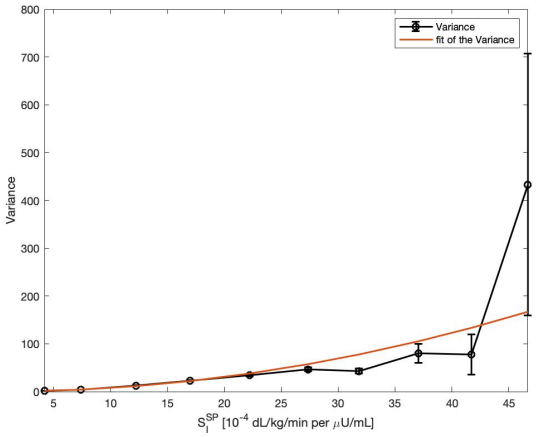
Here below are reported the different methods for the calculation of the variance, i.e. the uncertainty, in S_I^{SP} estimates and a comparison between the results is done:

Method 1. The first method consists of fitting the law for the variance reported in (3.7) as estimated from Mean Squared Error for different intervals of S_I^{SP}, and by means of Weighted Least Squares estimator.
Parameters are estimated by fitting the entire range of S_I^{SP}. Model parameters estimation (γ, δ) was performed using different amplitudes for the intervals where the variance was computed from the MSE. In particular, the smaller is the interval the more precise is the fit, as it is clear in Figure 4.5.

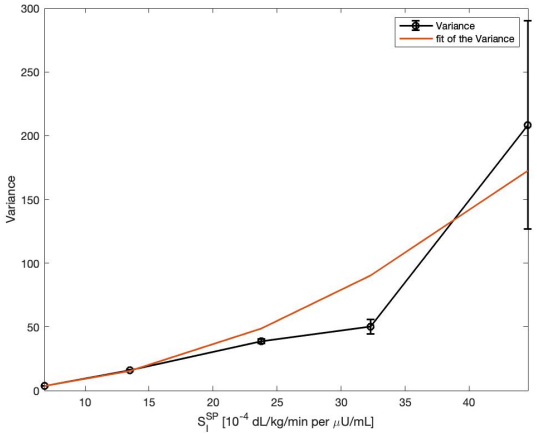
Errorbars give a visual information about the uncertainty on the calculation of the variance, using the formula reported in Equation (3.9).



(a) Fit of the variance with intervals of length 2.5.



(b) Fit of the variance with intervals of length 5.



(c) Fit of the variance with intervals of length 10.

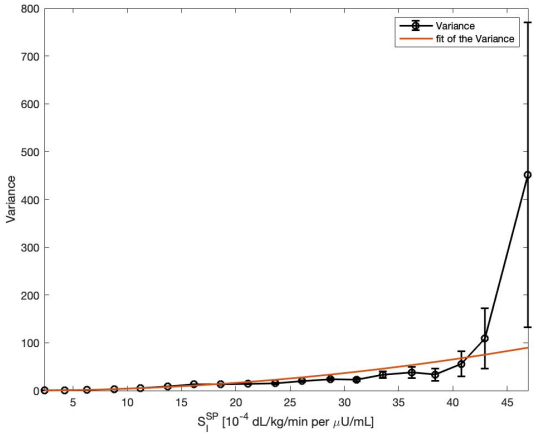
Figure 4.5: Variance calculated in intervals and its fit.

From Figure 4.5 it is evident that an unique fit for the whole range of values of S_I^{SP} doesn't fit well high values. In fact, also with intervals 2.5 long, from the value of about 22.5 the fit doesn't go through the central element of the set. Moreover, high values of S_I^{SP}, which have a bigger probability of error than the lower, are not well described. By using intervals of greater length, the situation is even worse.
The parameters obtained in the three cases are summarized in the following table:

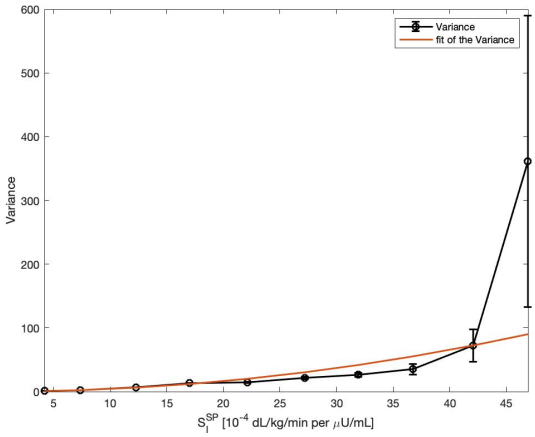
	Int = 2.5	**Int = 5**	**Int = 10**
γ	0.0724	0.0767	0.0869
δ	0.0712	0.0233	-0.4179

Table 4.5: Parameters of the fit of variance with intervals of different length.

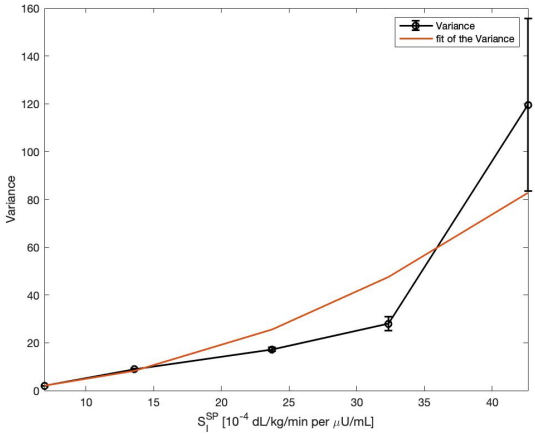
The same method is applied to results obtained by using plasma glucose instead of CGM. The plot of the variances calculated in intervals of different lengths are:



(a) Fit of the variance with intervals of length 2.5.



(b) Fit of the variance with intervals of length 5.



(c) Fit of the variance with intervals of length 10.

Figure 4.6: Variance calculated in intervals and its fit.

Parameters identified by fitting the variances are:

	Int = 2.5	**Int = 5**	**Int = 10**
γ	0.0408	0.0409	0.0456
δ	-0.0696	0.1133	-0.1045

Table 4.6: Parameters of the fit of variance with intervals of different length.

From the comparison between results obtained with CGM and plasma glucose, it is evident that estimates obtained by using plasma glucose are more precise. In fact, on the one hand the parameter γ, that is an index of the coefficient of variation of the estimates, is lower when using glucose. On the other hand δ is about the same in all the cases considered.

Method 2. The second method consists of finding the error in pattern classification, using the z-test, between S_I^{SP} and S_I^{OMM}. In this case the function of Matlab `fminsearch` is used and the parameters γ and δ from the curve defined by fitting the variance in intervals of 2.5-length, as presented in Table 4.5, are set as starting parameters and best parameters obtained were: γ=0.0208 and δ=-0.0111. To verify that they're not local minima but really the best parameters, several attempts have been done with different starting parameters and different ranges of upper and lower bounds.
The same technique is applied to results obtained by using glucose in Equation (3.2) and in this case, parameters obtained by applying `fminsearch` were: γ=0.0135 and δ=0.4501.
As showed in Method 1, by using plasma glucose data instead of CGM, γ is lower, thus indicating a lower uncertainty of estimations.

Method 3. The third method calculates the variance from "normalized" S_I values, i.e. on the S_I values normalized by the maximum of each day for each specific subject. The "normalized" S_I values, both for S_I^{SP} and S_I^{OMM}, are reported in the following plot:

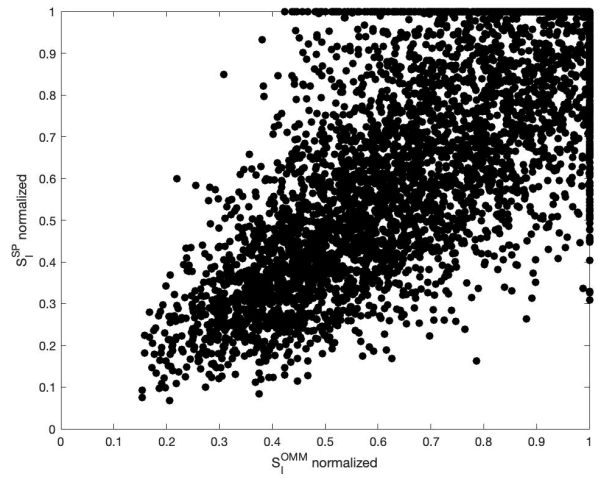


Figure 4.7: Plot of S_I^{SP} vs S_I^{OMM} with "normalized" values, i.e. obtained by normalizing, for each day and each subject, the S_I values for the daily maximum S_I.

All the markers in $S_I^{SP} = 1$ and $S_I^{OMM} = 1$ correspond to the maximum values of each day, for SI_I^{SP} and SI_I^{OMM} respectively.
Then the variance is calculated and the starcaise function calculated as the average MSE of the "normalized" S_I^{SP} vs S_I^{OMM} values in 0.1 intervals (Fig. 4.8):

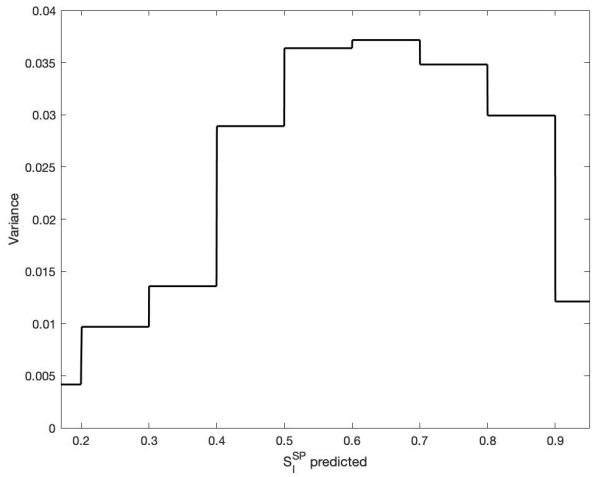


Figure 4.8: Variance of "normalized" S_I^{SP}.

As it was anticipated in the Subsection 3.2.1, sample variance of "normalized" S_I^{SP} is obtained from the MSE in intervals of length of 0.1.
Moreover it's evident from the plot that in this case it does not seem like a parabolic curve, so it can not be fitted by Equation (3.7).
On the other hand, values obtained by normalizing S_I^{SP} calculated with glucose are:

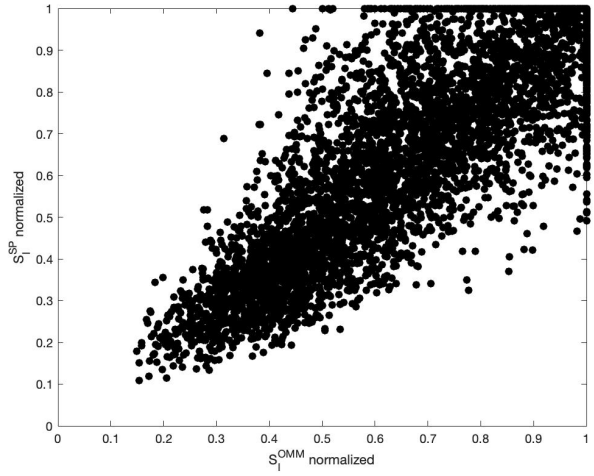


Figure 4.9: Plot of S_I^{SP} vs S_I^{OMM} with "normalized" values using plasma glucose instead of CGM.

From Figure 4.9 it is clear than in this case there is a bigger correlation between S_I^{SP} and S_I^{OMM}: in fact the cloud of estimates is less wide than the cloud of data points in the case of S_I^{SP} calculated with CGM. The variance obtained is the following:

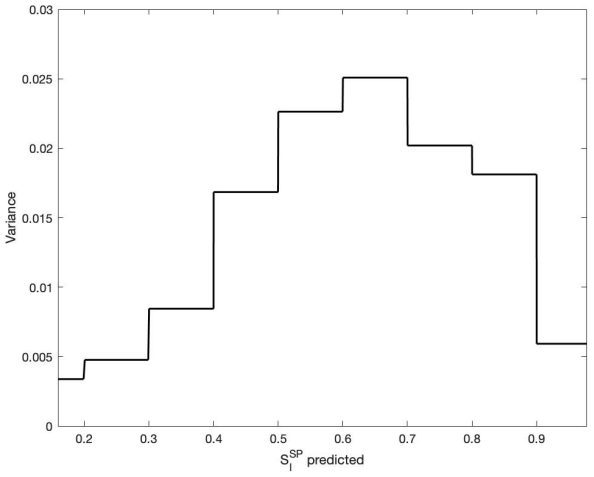


Figure 4.10: Variance of "normalized" S_I^{SP} using plasma glucose instead of CGM.

4.2.2 Estimation of intra-day pattern of S_I^{SP} and validation

First of all, before comparing the performance of intra-day pattern classification using the z-test exploiting each method for the calculation of the variance for S_I^{SP}, the reference intra-day pattern of S_I^{OMM} must be calculated using the z-test with the variance defined in Equation (3.10).

Below, the reference classes assigned with S_I^{OMM}, are:

Class	**N° of days/classes**
1	378
2	199
3	269
4	128
5	401
6	358
7	273

Table 4.7: Number of days assigned to the seven classes with S_I^{OMM}.

A total number of classes assigned with S_I^{OMM} equal to 2006 has been calculated due to elimination of those days with unreliable S_I^{OMM} or S_I^{SP} estimates.
Depending on the method used to calculate variance with S_I^{SP}, the z-test gave three different sets of the pattern, which were compared against the reference one calculated from S_I^{OMM}.
Each class is compared to the corrispondent one represented in Table 4.7, and the amount of equal assignments between S_I^{SP} and S_I^{OMM} is:

Class	**Method 1**	**Method 2**	**Method 3**
1	336	195	358
2	81	101	101
3	125	128	131
4	36	85	80
5	115	225	179
6	44	192	130
7	120	222	118

Table 4.8: Number of days assigned to the same class with S_I^{SP} and S_I^{OMM}.

In the first column are reported the classes assigned by calculating the z-test with the variance defined by parameters estimated using the Method 1. In the second column using the variance from the parameters estimated with the Method 2. Lastly, in the third column the variance of "normalized" S_I^{SP} and represented in Figure 4.8 is considered for the z-test.
It is interesting to compare results: as partially expected, the best case is the one that uses Method 2, in which about the 57% of the classes are the same as S_I^{OMM}. Also in the third case results are good (about 55% of the classes). The worst results are obtained in the first case, where only the 43% of classes was matched.

It is important to underline another aspect: with the first technique, class 1 is equal to classes assigned with S_I^{OMM} in the 89% of cases, that is the highest value among the three. On the other hand, class 6 is right only in the 12% of cases, that is the worst result.
Classification is performed also with results obtained by using glucose in the Equation (3.2). The reference classes assigned with the oral glucose minimal model are slightly different since a lower number of unreliable S_I^{SP} values were removed from the analysis:

Class	**N° of days/classes**
1	386
2	205
3	301
4	144
5	415
6	375
7	318

Table 4.9: Number of days assigned to the seven classes with S_I^{OMM}.

From Table 4.9 it is evident that in this case the total number of classes assigned is 2144. The amount of matched classes is represented in Table 4.10:

Class	Method 1	Method 2	Method 3
1	337	239	322
2	114	118	130
3	173	171	217
4	85	112	122
5	170	233	323
6	102	204	269
7	234	286	243

Table 4.10: Number of days assigned to the same class with S_I^{SP} and S_I^{OMM}.

With the first method the 56% of the same classes assigned with S_I^{OMM} are decided. With the second one the 63% and with the third one the 76%, thus showing a great improvement with respect to results obtained with CGM data.
After the comparison, the best methods appear to be the second and the third, and parameters and results obtained with these two methods will be used in the last part of the thesis for modeling the inter-day variability of intra-day pattern of S_I.

4.3 Modeling inter-day variability from intra-day pattern of Insulin Sensitivity

After the intra-day pattern of insulin sensitivity is estimated, results obtained are used to assess inter-day variability. In particular, Method 2 and 3 for the calculation of the error on S_I^{SP} are used with the parameters calculated with the dataset of 30 days of simulations.
In this Section the dataset of 90 days of simulation is used, where the daily pattern of S_I^{SP} has been estimated for each day and subject, as explained in Section 3.3. Once the daily pattern is calculated using Method 2 and 3 for each day of simulation in each specific subject, classes are assigned. The second step is the calculation of transition probabilities from a class to another with Markov models, as explained in Subsection 3.3.1. Two examples of transition matrices obtained with the methods 2 and 3 in a representative subject are:

- **Case 1.**

CLASSES	1	2	3	4	5	6	7
1	0.2197	0.0833	0	0	0	0	0.6250
2	0.3333	0	0	0	0	0	0.6667
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0.2581	0.0161	0	0	0	0	0.7258

(a) Estimated transition matrix from Method 2 of a representative subject.

CLASSES	1	2	3	4	5	6	7
1	0.2353	0.1176	0	0	0	0	0.6471
2	0	0.3333	0.1667	0	0	0	0.5000
3	0	0	0.3333	0	0	0	0.6667
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	1
6	0	0	0	0	0	0	0
7	0.2097	0.0323	0.0161	0	0.0161	0	0.7258

(b) Estimated transition from Method 3 of a representative subject.

Figure 4.11: Transition matrices of a representative subject.

While the correspondent original transition matrix of the representative subject is:

CLASSES	1	2	3	4	5	6	7
1	0.2500	0	0	0	0	0	0.7500
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0.2143	0	0	0	0	0	0.7857

Figure 4.12: Original transition matrix of a representative subject.

These are two examples of transition matrices estimated with S_I^{SP} with the two different methods. In the diagonals there is the probability that the subject stays in the same class for two consecutive days. The elements outside from the disgonal represent the probability that the subject goes from the class i to the class j, indicated by rows and columns respectively.
For a comparison, the same process is applied using plasma glucose instead of CGM. In this case, the transition matrices are:

CLASSES	1	2	3	4	5	6	7
1	0.1818	0.1818	0	0	0	0	0.6364
2	0.3636	0	0	0	0	0	0.6364
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0.0746	0.1343	0	0	0	0	0.7910

(a) Estimated transition matrix from Method 2 of a representative subject.

CLASSES	1	2	3	4	5	6	7
1	0.1818	0.1818	0	0	0	0	0.6364
2	0.3636	0	0	0	0	0	0.6364
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0.0746	0.1343	0	0	0	0	0.7910

(b) Estimated transition from Method 3 of a representative subject.

Figure 4.13: Transition matrices of a representative subject, using plasma glucose data.

- **Case 2.**

CLASSES	1	2	3	4	5	6	7
1	0.2000	0	0.1000	0.6500	0.0500	0	0
2	1	0	0	0	0	0	0
3	0.3333	0	0	0.1111	0.5556	0	0
4	0.2264	0.0189	0.0943	0.6604	0	0	0
5	0.1167	0	0.3333	0.5000	0	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0

(a) Estimated transition matrix from Method 2 of a representative subject.

CLASSES	1	2	3	4	5	6	7
1	0	0.2000	0.6000	0	0	0	0.2000
2	0	0.1250	0.1250	0.7500	0	0	0
3	0.1176	0.1176	0.1176	0.2941	0.0588	0.1176	0.1765
4	0.0577	0.0769	0.1731	0.6923	0	0	0
5	0	0	0	1	0	0	0
6	0	0	0	1	0	0	0
7	0	0.2500	0.5000	0.2500	0	0	0

(b) Estimated transition from Method 3 of a representative subject.

Figure 4.14: Transition matrices of a representative subject.

While the correspondent original transition matrix of the representative subject is:

CLASSES	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0.1358	0.7778	0.0864	0	0
4	0	0	0.0959	0.9041	0	0	0
5	0	0	0.2481	0.5865	0.1654	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0

Figure 4.15: Original transition matrix of a representative subject.

As before, for a comparison, the same process is applied using plasma glucose instead of CGM. In this case, the transition matrices are:

CLASSES	1	2	3	4	5	6	7
1	0.0909	0	0.0909	0.8182	0	0	0
2	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0
4	0.0282	0	0.0845	0.8873	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0

(a) Estimated transition matrix from Method 2 of a representative subject.

CLASSES	1	2	3	4	5	6	7
1	0	0	0	0	0	1	0
2	0	0	0	1	0	0	0
3	0	0	0	0.1250	0	0.8750	0
4	0.0141	0.0141	0.0845	0.8873	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0.1250	0.8750	0	0	0
7	0	0	0	0	0	0	0

(b) Estimated transition from Method 3 of a representative subject.

Figure 4.16: Transition matrices of a representative subject, using plasma glucose data.

After the transition matrices are calculated, the similarity between the estimated and the orginal ones is calculated. To have a reliable evaluation of the error committed in the estimation two metrics are used, as reported in Subsection 3.3.3. The first method evaluates the similarity of total matrices, calculated with the formula (3.14). The second one assesses the similarity of matrices by using only the rows where the original matrix is different from 0, as defined in the formula (3.15). By calculating the error (%) in the two cases presented before, results obtained are:

Case	**Equation**	**CGM**		**Glucose**	
		Method 2	Method 3	Method 2	Method 3
1.	err_{tot}	15.5%	34.6%	0%	0.2%
	err_{eff}	8.5%	42.1%	0%	1.1%
2.	err_{tot}	47.1%	73.6%	43.2%	70%
	err_{eff}	43.3%	38.5%	67.6%	63.44%

Table 4.11: Error (%) committed in Case 1 and Case 2.

In general, the mean error (%) obtained by applying the two methods for the calculation of S_I^{SP}estimation error, both with CGM and glucose are represented in Table 4.12:

Equation	**CGM**		**Glucose**	
	Method 2	Method 3	Method 2	Method 3
err_{tot}	54% (18,03)	72% (20,5)	33% (22,3)	39% (23,8)
err_{eff}	59% (28,5)	64% (24,6)	64% (40,3)	62% (37)

Table 4.12: Error (%) and (SD) committed by modeling inter-day variability.

From Table 4.12 it is clear that the error committed by modeling inter-day variability with Markov matrices calculated from CGM data is high. In general, better results are obtained by using the formula (3.14), and using plasma glucose instead of CGM, hence this highlights the importance of recalibration of CGM data to improve the signal to noise ratio.

Chapter 5

Conclusions

Diabetes Mellitus is a major public healthcare problem worldwide. Since a real treatment does not exists, reasearches are focused on finding techniques to control the level of glycemia and mantain it in the normal range, to avoid/limit long- and short-terms complications. In particular patients with Type 1 Diabetes (T1D), who are insulin-dependent, face everyday with the issue of calculating the exact output of insulin to administer, based on their glucose levels, as well as the amount of carbohydrates ingested, health state etc.

A key parameter for the optimization of patient-specific therapies for T1D patients is represented by insulin sensitivity. However such parameter is known to vary within (intra-) and between (inter-) day, hence assessing its variability is of fondamental importance. The aim of this thesis is to model intra-day and inter-day variability of insulin sensitivity in T1D patients every day life. To do so an in silico database was used exploiting the last version of the UVA/Padova simulator [22], which includes several important features like intra-day variability of insulin sensitivity and dawn phenomenon, to which a random model generating inter-day variability of intra-day pattern was added for the purpose of this work.

In this thesis three different scenarios, exploiting the 100 virtual adult population with T1D, have been simulated for three different purposes: the first one is 1-day scenario, which was used to develop and test a new formula for S_I estimation from minimally-invasive technologies, i.e. CGM sensors and insulin pumps, and validate it agaist the oral minimal model S_I (S_I^{OMM}), estimated from plasma glucose and insulin measurements. The second one consisted of 30 days of simulation, which was used to assess intra-day variability of the new S_I^{SP} proposed by using statistical methods for pattern classification, and the last one consisted of 90 days which was used to model inter-day variability of intra-day pattern of insulin sensitivity by estimating a Makov Chain model describing the probability of transition between the estimated intra-day pattern of S_I^{SP}.

Data obtained from CGM sensors and CSII pumps were used to develop and test a

formula for S_I^{SP} estimation [19]. Results show a good correlation between S_I^{SP} and S_I^{OMM}, of about 0.88. This methodology has been used to calculate the intra-day variability of S_I^{SP}, allowing to assign a daily pattern of S_I^{SP} to each subject by means of statistical methods. However, to do so, the uncertainty related to S_I^{SP} estimation, quantified by the variance of its measurement error, was needed.

Three different methods have been used to calculate its variance. In particular, in the first method the variance was calculated by fitting a model as the error between S_I^{SP} vs S_I^{OMM}, in the second one parameters of the curve that describes the variance were identified by minimizing the error in daily pattern classification between S_I^{SP} vs S_I^{OMM} and the third method calculated the variance from the "normalized" S_I^{SP} estimates vs S_I^{OMM}.

The methods which gave the best outcomes resulted to be the second and the third one. Thanks to these methods, the 57% and the 55% of classes were correctly assigned with respect to the pattern defined by S_I^{OMM}, respectively.

In the end, parameters obtained have been used to assess inter-day variability of the intra-day pattern of insulin sensitivity. This has been done by estimating the Markov Chains model used to generate the 90-days simulation scenario for each subject. Results showed that the error (%) committed in the estimation of the original Markov Chain model was still high (around 50%) when using CGM data without any recalibraion algorithm.

In fact, since CGM sensors are known to be affected by higher measurement error than plasma glucose measurements, results obtained with S_I^{SP} have been also calculated by using plasma glucose instead of CGM data. In fact it has been demonstrated that, by using glucose, the correlation between S_I^{SP} vs S_I^{OMM} was higher (R=0.93), the classes correctly assigned with respect to S_I^{OMM} with the second and the third method have increased to the 63% and 76% respectively, and error (%) committed by constructing the transition matrices improved resulting around 30-40%.

In this thesis an in-depth analysis has been made to the process of estimation of inter-day variability of insulin sensitivity, starting from the estimation of S_I^{SP} and its validation, going through the classification of intra-day pattern of insulin sensitivity and ending with the modeling of inter-day variability of intra-day pattern of insulin sensitivity with Markov models.

Results showed in this thesis represent step toward the development of models which can automatically adapt T1D therapy following behavioral changes occurring in daily lif conditions of patients with T1D.

Bibliography

[1] Bergman RN, Ider YZ, Bowden CR, Cobelli C, "Quantitative estimation of insulin sensitivity." Am J Physiol 1979;236: 667-677.

[2] Breda E, Cavaghan MK, Toffolo G, Polonsky KS, Cobelli C, "Oral glucose tolerance test minimal model indexes of β-cell function and insulin sensitivity." Diabetes 2001;50:150-158.

[3] Casella G, Berger RL, "Statistical Inference" Thompson learning 2002, p. 331.

[4] Cobelli C, Dalla Man C, Toffolo G, Basu R, Vella A, Rizza R, "The oral minimal model method." Diabetes 2014;63:1203-1213.

[5] Cobelli C, Dalla Man C, Sparacino G, Magni L, De Nicolao G, Kovatchev BP, "Diabetes: models, signals and control." IEEE Rev Biomed Eng. 2009;2:54-96

[6] Dalla Man C, "Modeling and Control of Biological Systems." Slides of the course.

[7] Dalla Man C, Caumo A, Basu R, Rizza R, Toffolo G, Cobelli C, "Minimal model estimation of glucose absorption and insulin sensitivity from oral test: validation with a tracer method." Am J Physiol Endocrinol Metab 2004, 287:637-643.

[8] Dalla Man C, Yarasheski KE, Caumo A, et al., "Insulin sensitivity by oral glucose minimal models: validation against clamp." Am J Physiol Endocrinol Metab 2005, 289:954-959.

[9] DeFronzo RA,Tobin JD, Andres R, "Glucose clamp technique: a method for quantifying insulin secretion and resistance." Am J Physiol 1979; 237:214-223.

[10] Facchinetti A, "Clinical Engineering and Health Technology Assessment." Slides of the course.

[11] Kovatchev BP, Breton M, Dalla Man C, Cobelli C, "In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes." J Diabetes Sci Tech 2009;3:44-55.

[12] Meneghetti L, Dassau E, Doyle FJ 3rd, Del Favero S, "Machine Learning-based anomaly detection algorithms to alert patients using Sensor Augmented Pump of infusion site failures." J Diabetes Sci Tech 2021;9:1932296821997854.

[13] Mitra M, "Pancreatic Cancer and its treatment." Journal of GastroIntestinal Hepatic Surgery 1 2018; Figures.

[14] Pacini G, Mari A, "OGTT e IVGTT: due test a confronto per la valutazione dell'insulino-sensibilità e della funzione beta cellulare." G It Diabetol Metab 2007;27.

[15] Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N et al., "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045:Results from the International Diabetes Federation Diabetes Atlas, 9^{th} edition." Diabetes Res Clin Pract 2019;157:107843.

[16] Sapkota A, "Z-Test- Definition, Formula, Examples, Uses, Z-Test Vs T-Test." microbenotes.com/z-test/ 2021.

[17] Sargent J, "Technology will set you free." Nature milestones 2021.

[18] Schiavon M, Dalla Man C, Cobelli C, "Physiology-based Run-to-Run adaptation of insulin to Carbohydrate Ratio improve Type 1 Diabetes therapy:Results from an In Silico study." 2019 American Control Conference (ACC) 2019; 4124-4129.

[19] Schiavon M, Dalla Man C, Kudva YC, Basu A, Cobelli C, "Quantitative estimation of insulin sensitivity in type 1 diabetic subjects wearing a Sensor Augmented Insulin Pump." Diabetes Care 2014;37:1216-1223.

[20] Stanfield CL, "Fisiologia", EdiSES 2017.

[21] Vettoretti M, Cappon G, Facchinetti A, Sparacino G, "Advanced diabetes management using Artificial Intelligence and Continuous Glucose Monitoring sensors." Sensors (Basel) 2020;20:3870.

[22] Visentin R, Campos-Náñez E, Schiavon M, Lv D, Vettoretti M, Breton M, Kovatchev BP, Dalla Man C, Cobellli C, "The UVA/Padova Type 1 Diabetes simulator goes from single meal to single day." J Diabetes Sci Tech 2018;12:273-281.

[23] Visentin R, Dalla Man C, Kudva YC, Basu A, Cobelli C, "Circadian variability of insulin sensitivity: physiological input for in silico artificial pancreas." Diabetes Technology and Therapeutics 2015;17:1-7.

[24] *Insulin pen*, bbc.com/news/health, Figures.

[25] "dexcom.com"; Figures.

[26] "Il diabete in Itlalia." ISTAT, Comunicato Stampa 2017; Figures.

[27] "diabetesatlas.org", 2021

[28] "who.int/health-topics/diabetes"

[29] "diabetes.org.uk: Metformin and diabetes."