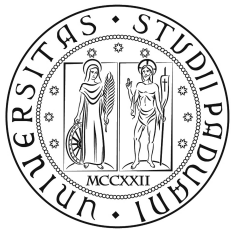


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

CORSO DI LAUREA MAGISTRALE IN
INGEGNERIA CHIMICA E DEI PROCESSI INDUSTRIALI

Tesi di Laurea Magistrale in
Ingegneria Chimica e dei Processi Industriali

Optimal design of feasible clinical tests for the identification of physiological models of type 1 diabetes mellitus

Relatore: Prof. Fabrizio Bezzo
Correlatore: Dr. Federico Galvanin

Laureando: DAVIDE PRADELLA

Anno accademico 2015/2016

Abstract

Type 1 diabetes mellitus is a disease characterised by inadequate insulin secretion, causing damaging imbalance to the glucoregulatory system. Cutting edge therapies involve the use of an artificial pancreas, that is a device capable of maintaining a physiological glucose concentration through accurate insulin infusions. Customisation of these devices to the individual needs may be accomplished provided that a precise identification of the patient is performed.

This work considers optimal model-based design of experiment techniques to devise a new clinical test for the identification of model parameters in subjects affected by T1DM. The pivotal feature required for this test is its implementation feasibility. In order to satisfy this requirement, a novel approach based on heuristic indices is introduced to assess several aspects that could impact on the practical applicability of the designed protocol. Among the criteria considered, a particular emphasis is put on assuring safety conditions for the test and on extracting the maximum information content for a precise estimation of the model parameters.

A robust clinical protocol is devised, which exhibits statistically satisfactory parameter estimation capabilities and complies with the requirements imposed by practicality.

Riassunto

Il diabete mellito di tipo 1 è una malattia autoimmune che causa la progressiva distruzione delle cellule beta del pancreas, da cui deriva l'incapacità di produrre insulina e quindi di regolare la glicemia. Pur riguardando una percentuale ridotta della popolazione diabetica, colpisce milioni di persone in tutto il mondo, generando una spesa economica di miliardi di dollari per la sua cura e il trattamento delle complicazioni che ne derivano.
Il trattamento tradizionale richiede la somministrazione giornaliera di molteplici dosi di insulina da parte del paziente. Questo metodo, oltre ad essere invasivo, non consente un controllo ottimale della glicemia ed è influenzato dall'abilità del paziente. La terapia più promettente in fase di sviluppo consiste nell'impiego di un pancreas artificiale, che consenta di sopperire all'incapacità del paziente di produrre insulina da sè. Il funzionamento di questo dispositivo sfrutta un algoritmo di controllo basato su un modello del paziente diabetico. Affinchè questa terapia risponda alle esigenze del singolo soggetto, è richiesta un'accurata identificazione dei parametri del modello mediante opportuni test clinici.
Questa tesi si propone di progettare test clinici che consentano un'identificazione ottimale dei parametri di modelli fisiologici relativi a soggetti diabetici, tali da soddisfare i requisiti di sicurezza richiesti per una loro implementazione clinica.
Il metodo impiegato nell'analisi è la progettazione di esperimenti basata su modello, che permette di ottenere la procedura ottimale da seguire nell'esecuzione del test sia in termini di campionamento sia di variabili di controllo. La progettazione del nuovo protocollo è volta a massimizzare la quantitità di informazione ottenibile attraverso un campionamento discreto, ottemperando ad alcuni vincoli imposti sulla sicurezza.
La capacità identificativa dei test viene valutata tramite alcune statistiche, relative alla stima dei parametri su due pazienti simulati, affetti da diabete in maniera più o meno severa. La valutazione complessiva dell'attuabilità del test avviene sulla base di una serie di indici euristici introdotti a tal proposito. Questi, oltre a considerare l'informazione fornita dal test, tengono in considerazione la sua sicurezza, invasività e durata.
L'analisi prende avvio dalla valutazione dei due test tradizionali usati in ambito clinico:

il test di tolleranza orale al glucosio e quello endovenoso. Partendo da tali basi, il lavoro si estende poi alla progettazione di nuovi protocolli con l'obiettivo principale di garantire una identificazione statisticamente soddisfacente del soggetto. I risultati forniscono importanti indicazioni, tra cui la migliore efficacia di un test orale per l'identificazione ma, al tempo stesso, la necessità di migliorare la sicurezza associata a tali protocolli. Al fine di garantire una maggiore affidabilità al test, a fronte dell'elevata variabilità osservata nelle riposte da soggetti differenti, si ricorre all'impiego di tecniche di progettazione robusta. Questo metodo simula una popolazione diabetica per trarre informazioni statistiche sulla variabilità media osservata nelle risposte, sulla base della quale viene progettato un nuovo protocollo. Il test ottenuto in tal modo, oltre a realizzare stime parametriche soddisfacenti, si dimostra efficace nell'ottemperare a tutti i requisiti richiesti per un'implementazione clinica, in modo particolare garantendo un sostanziale miglioramento del livello di sicurezza per il paziente.

Contents

Introduction

Type 1 diabetes mellitus (T1DM) is a disease affecting millions of people around the world and causing billions of dollars to be spent on its treatment and complications every year. This metabolic disorder is characterised by the lack of insulin in the organism, due to the autoimmune destruction of the pancreatic beta cells. As a result, high blood glucose concentrations are observed in the patient, which lead to unsafe conditions that, if prolonged over time, may cause severe damage.
The typical management of this disease requires the patient to perform multiple daily insulin injections, in order to maintain the normoglycaemia. There are several drawbacks related to this therapy, such as it is highly invasive and relies on the ability of the patient to keep his/her glycaemia under control.
The most promising therapy for the future is represented by the use of an artificial pancreas, that is a device capable of mimicking the action of the pancreas in insulin delivery. Such a device requires a control algorithm for the calculation of the insulin dosage based on the continuous glycaemia measurements. The availability of a physiological model, describing the glucoregulatory system of a diabetic patient, allows assessing the algorithm performance on a simulated subject. Alternatively, in model predictive control, the model itself is used in the algorithm, so as to customise the therapy to the individual.
A tailored therapy requires a precise identification of the model parameters of the patient by means of experimental measurements. Since the condition of the subject is usually unknown a priori, it is necessary to adopt optimal design of experiments techniques in order to devise tests, which comply with safety requirements. The goal of the thesis is to apply a model-based design of experiments approach to conceive clinical tests allowing for a statistically satisfactory identification of the parameters of a simple physiological model. Since the main focus is on the feasibility of the designed protocol, a novel approach based on five heuristic indices is introduced based on which the overall performance of the different designs may be compared.
The first chapter presents an overview of the disease and its global impact. Then, the chief clinical tests used for clinical and research purposes are discussed, along with the current therapy and the future perspectives on its management. Much relevance is

given to the review of the most important physiological models of T1DM that have been proposed over the years for the description of the glucose and insulin dynamics. Finally, the motivation for the use of a model-based design of experiments (MBDoE) procedure is explained, and the objectives of the thesis are outlined.
The second chapter deals with the mathematical part of the model-based design of experiments. Firstly, the Lynch and Bequette model is detailed, since it is the one selected for the simulation of the diabetic patients in this work. Secondly, the theory underlying MBDoE is presented, along with the in silico generation of experimental data and the parameter estimation procedure. Lastly, five heuristic indices are introduced, so as to evaluate the feasibility of the designed protocols.
Chapter three is about the preliminary analysis of the Lynch and Bequette model, by means of two different standard clinical tests. This allows to understand how the model, i.e. the simulated patient, reacts to different types of stimuli and how much information can be extracted with these two protocols.
In chapter four, the different designed protocols obtained with MBDoE are presented and their parameter estimation capability evaluated using two parametric sets, representing diabetic patients with different illness severity. Eventually, a comparison among the several designed tests is made in order to evaluate their overall performance.
In chapter five, a robust design of experiments approach is adopted to reduce the variability of the outputs. This method allows to devise a clinical protocol which improves the safety conditions when an unknown subject is to be identified.

Chapter 1

Diabetes mellitus

In this chapter an overview of diabetes mellitus is presented with a focus on the disease, its diagnosis and therapy. The different types of diabetes are analysed, with a particular emphasis on type 1, which is the main objective of this work. Then, the social and economic impact of diabetes are discussed and the main glucose tolerance tests utilised for its diagnosis are outlined. Much relevance is given to an overview of the proposed physiological models of type 1 diabetes and their role in state-of-the-art therapies, such as the use of an artificial pancreas. Going into detail, the importance of designing suitable clinical protocols for the identification of the parameters of physiological models is dealt with. Finally, the objectives of this thesis are pointed out along with a novel approach based on indices for the evaluation of the designed clinical protocols.

1.1 Introduction

Carbohydrates introduced in the body through food are processed in the digestion into their monomers: glucose, galactose and fructose. Both galactose and fructose are converted into glucose by the liver, in order to make them usable for the cells. Glucose constitutes the main source of energy for the organism and is carried to the various parts of the body where is needed through blood circulation. The level of glucose in blood is referred to as glycaemia and it is regulated by the glucoregulatory system through two hormones secreted by pancreas: insulin and glucagon.
Insulin is synthesised by beta cells, which are located in the islets of Langerhans in the pancreas. This hormone is responsible for some of the glucose to be directly used by some cells, such as cerebral cells and red blood cells, and for the excess to be stored as glycogen in the liver. On the opposite, glucagon is secreted by alpha cells in the pancreas and carries out the conversion of glycogen back to glucose when glycaemia decreases.

Diabetes mellitus is a metabolic disorder characterised by high glucose concentrations in blood due to insufficient insulin secretion, insulin effectiveness or both. Enduring deviations of blood glucose concentration from the physiological values (normoglycaemia) are denoted as hyperglycaemia, if they exceed the upper level of normoglycaemia, and hypoglycaemia, if they go beyond its lower limit. Persistent hyperglycaemic conditions can cause serious consequences mainly to the nerves and blood vessels. If the subject does not receive an adequate treatment, potentially deadly conditions may occur, such as cardiovascular diseases, amputation, blindness and kidney failure. Similarly, hypoglycemic conditions should be avoided since they can lead to shakiness, fatigue and eventually coma.

1.2 Diabetes classification

Three types of diabetes mellitus have been identified: type 1 and type 2, which regard almost the totality of the people affected, and gestational diabetes, which accounts for the remaining part.

1.2.1 Type 1 diabetes mellitus

Type 1 diabetes mellitus (T1DM), previously referred to as juvenile-onset or insulin dependent diabetes mellitus, is caused by the autoimmune destruction of pancreatic islet beta cells. Being incapable of producing the insulin by itself, the subject needs multiple daily external administrations for survival.
In the absence of insulin, the three main target tissues where glucose is normally delivered (liver, muscle and fat) are not capable of taking it up so that it flows, along with amino acids and fatty acids, in the bloodstream. Inability of glucose to enter into the cells causes the organism to use fatty acids for energy production; however, fatty acids utilisation produces ketones, which tend to accumulate in blood being non-volatile and, due to their acidity, lower pH.
Symptoms are increased urination and thirst, constant hunger, weight loss, recurrent blurred vision and fatigue.
T1DM typically afflicts children and young adults and no evidence of a correlation with body weight has been identified. The causes of this disease are still unknown, even though it is generally agreed that T1DM is related to a complex interaction between genes and environmental factors. In light of that, it is not preventable with the current knowledge. Nevertheless, a swift diagnosis and an adequate therapy allow the individual to conduct a decent life, even if life expectancy is reduced by more than 10 years [1].
T1DM is the target of this work because no prevention is possible and, although it

affects only 10% of the diabetic patients, it still represents a huge number of people.

1.2.2 Type 2 diabetes mellitus

Type 2 diabetes mellitus (T2DM), previously known as adult-onset or non-insulin dependent diabetes mellitus, derives from a relative insulin deficiency. Subjects affected by this illness show heterogeneous conditions ranging from a severe resistance to insulin but unaltered secretion capability to defects in insulin secretion [2]. The symptoms of T2DM may be similar to those of T1DM (excessive urination and thirst, increased hunger, vision changes), even though they are less evident, so that the diagnosis often occurs some years after the disease onset.
T2DM accounts for nearly 90% of the people affected by diabetes all over the world [3] and strong correlations have been highlighted between its onset and excess body weight and sedentary lifestyle.

1.2.3 Gestational diabetes mellitus

Gestational diabetes mellitus (GDM) is characterised by hyperglycaemic levels during pregnancy in an individual not previously affected by diabetes. Occurrence of this disease may lead to complications during pregnancy and delivery, moreover, it enhances the risk of incurring in type 2 diabetes mellitus in future life.

1.3 Global impact of diabetes mellitus

According to the World Health Organization [4], in 2012 diabetes was directly responsible for the death of 1.5 million people in the world, especially in low-income countries, and this number is expected to increase more than 50% in the next 10 years. In 2014, diabetes was estimated to affect 422 million adults around the world [5]. From Figure 1.1, it can be observed as the number of people affected in the period 1980-2014 has significantly increased in poor and developing countries (in orange and green). On the contrary, in western countries (in light blue), apart from the USA, prevention started earlier, hence the increase has slowed down.

These astounding figures explain why, in developed countries, 10% or more of the health budget is allocated to the treatment of diabetes and its complications. Direct medical costs associated with diabetes are related to its prevention and treatment. These include inpatient hospital care, outpatient and emergency care, medications and medical supplies such as injection devices and self-monitoring equipment.
Based on recent cost estimates, it has been evaluated the direct annual cost of diabetes to the world amounts to US$ 825 billion [7]. This figure has been estimated by the

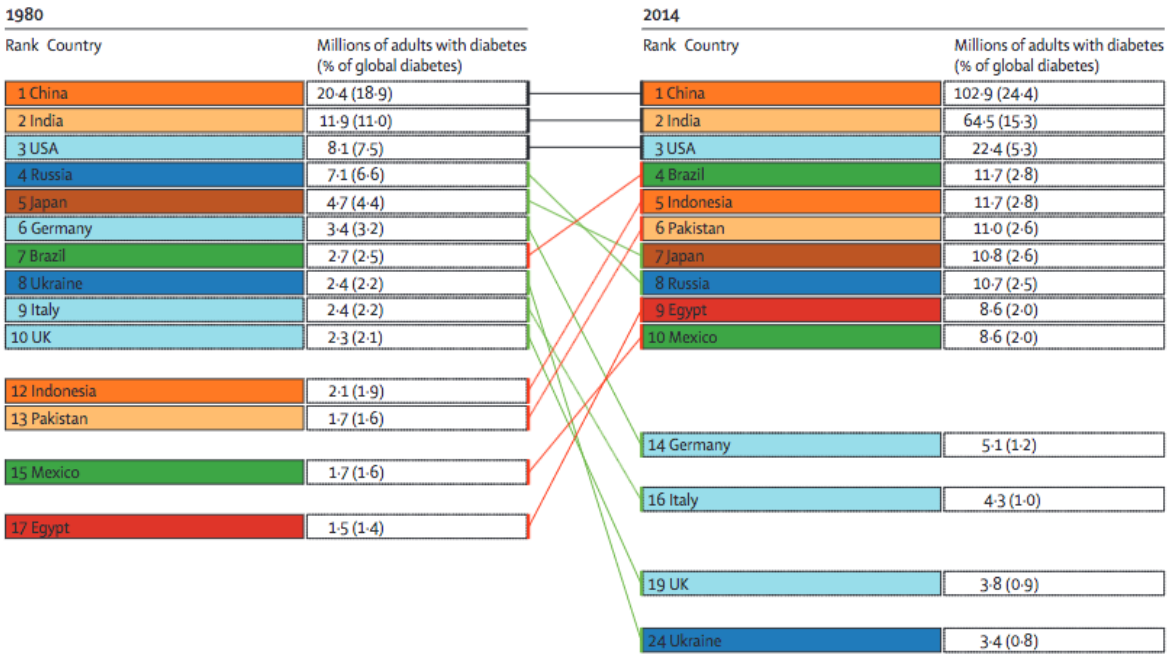


Figure 1.1. Countries with the largest number of adults with diabetes in 1980 and 2014 (colours for each country indicates its region) [6].

International Diabetes Federation to have more than tripled over the period 2003-2013 due to the increase in the number of people affected and rise in the per capita diabetes spending. The national expense for diabetes ranks these countries at the forefront: China ($ 127 billion), the USA ($ 105 billion), India ($ 73 billion) and Japan ($ 37 billion).

Nearly 60% of the global cost is borne by low-income and middle-income countries, where significant parts of medical expenses are generally paid out-of-the-pocket [7], which affects treatment adherence and causes serious financial hardships for patients and their families.

1.4 Standard clinical tests

Several standard clinical tests have been set up for the diagnosis of diabetes and some of them are also used for research purposes. They are generally based on a stimulus-response pattern, so the patient receives a glucose meal and glycaemia is measured at fixed times. The diagnosis is straightforward once the test results are available by comparison with standard validated values. A quick overview of the most established clinical tests is presented below.

- The Oral Glucose Tolerance Test (OGTT) is widely used for screening, since it is the one recommended by the World Health Organization [8]. The test is to be performed in the morning after the subject, in the 3 days before, has followed an

unrestricted diet with at least 150 g carbohydrates/d and an overnight fast prior to the test. A pre-test sample is collected to assess fasting values, then, at t=0 min, 75 g of glucose are administered in the form of a drink with 250-300 mL of water over the course of 5 min. For screening purposes, samples are taken at end of the test, which lasts 2 hours while, if the aim is model identification, more samples are collected in between.

- The Intravenous Glucose Tolerance Test (IVGTT) was proposed by Bergman [9] in 1981 to evaluate glucose and insulin sensitivity. The patient has to be consumed at least 300 g carbohydrates/d in the 3 days preceding the test, which is carried out in the morning following an overnight fast. Three 2-mL pre-test samples are taken 13, 8 and 3 min prior to the beginning of the test to monitor basal values, following which an intravenous glucose injection is performed at 300 mg/kg in an antecubital vein, starting at t=0 min and over a 60 s duration. During the test, 23 2-mL additional samples are collected at times: 2, 4, 6, 8, 10, 12, 14, 16, 19, 22, 27, 32, 42, 52, 62, 72, 82, 92, 102, 122, 142, 162 and 182 min. Since it is more invasive for the subject, it is mainly used for research purposes, for instance to detect insulin sensitivity in response to a variation in glycaemia.

- The Postprandial Glucose Test is used both for screening and evaluating the effectiveness of a treatment in diabetic patients. It is executed after the subject has eaten a meal containing at least 100 g of carbohydrates and has fasted for 2 h with a variable sampling schedule.

- The Euglycemic Hyperinsulinemic Clamp, discussed by De Fronzo et al. [10], is a test used to evaluate the insulin resistance while injecting an amount of glucose sufficient to compensate the rise in insulin. The test lasts 2 h and the glucose infusion profile applied in the last 30 minutes quantifies insulin sensitivity.

The OGTT is currently the most employed test for diabetes diagnosis because it has been widely validated and, like the postprandial, is an oral test, which means it is less invasive for the subject. Nevertheless, the choice of the optimal test for research purposes is not trivial and it is worth being carefully analysed. In fact, it must take into account many different aspects, such as:

- low stress for the subject,

- maximisation of the information obtainable,

- inherent safety throughout the execution.

1.5 Type 1 diabetes mellitus management

Safety conditions for the patient affected by T1DM are ensured by maintaining its normoglycaemic values through exogenous injections of insulin to prevent hyperglycaemic conditions. Dosage and timing of the insulin administration are crucial to avoid also the risk of incurring in hypoglycaemic values. In the past, the approach consisted of instructing the subject to check the glycaemia during the day through self-monitoring of blood glucose. Blood samples were collected from the fingertips and glycaemia was read on a specific device. Multiple insulin injections are required during the day, usually long-lasting effect insulin is administered once a day and additionally short-acting insulin before meals. This therapy is not expensive but it is invasive, due to the exogenous injections, and does not allow an optimal glycemic control. Despite this technique is still used, thanks to technological improvement more effective methods have emerged allowing for continuous glucose monitoring. Frequent measurement permits, on the one hand, to study the dynamics of glucose-insulin in the body and, on the other hand, to enhance the treatment through an ad hoc therapy for the patient. Moreover, continuous glucose monitoring is a breakthrough because it paves the way to the advent of an artificial pancreas, that is a device capable of reproducing the function of this organ.
In the last 40 years, much work has been done to develop an artificial closed-loop pancreas able to adjust the insulin infusion in response to glucose changes. In fact, this kind of device would be less invasive for the subject, being either wearable or implantable, and automate patient's action. The first generation of artificial pancreas, called *Biostator*, was based on two manipulated variables (glucose and insulin infusion) and programmed to operate minimising hyperglycaemia while avoiding hypoglycaemia [11]. Nowadays, the frontier is to develop portable devices (wearable artificial pancreas, WAP), which continuously monitor glucose and provide insulin whenever required. WAPs typically consist of three parts, as displayed in Figure 1.2:

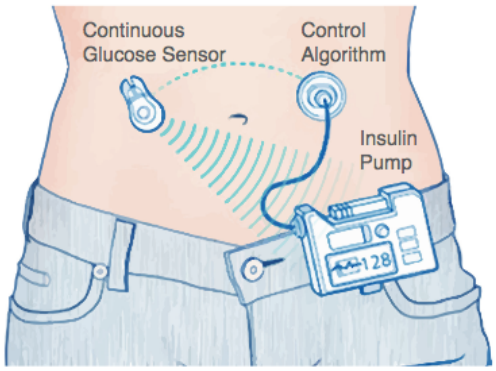


Figure 1.2. Wearable artificial pancreas components [12].

- a continuous glucose sensor, which detects the glycaemia at regular intervals;

- a controller, to regulate the insulin infusion rate through the use of control algorithms;

- a continuous insulin pump for its administration, via a cannula inserted under the skin.

Crucial for a WAP is to have available a reliable physiological model of the glucoregulatory system, i.e. representing both glucose and insulin dynamics, for the controller to be able to predict at each time the optimal insulin infusion rate. While the first devices based on a closed-loop configuration have appeared on market, much efforts are still addressed to find robust glucose sensors and a reliable control algorithm. The major challenges remaining are related to sensor calibration errors, unknown meal disturbances, physical activity effect and insulin-glucose sensitivity variability [13].
A major approach in the development of WAP control algorithms is represented by the use of model predictive control (MPC) strategies, which consist in applying a model to predict the future output of blood glucose while optimising the future control moves, namely the insulin infusion. Typically, the objective function to be minimised is the difference between the desired set point and the blood glucose value estimated by the model over the prediction time horizon [14]. This fact highlights the necessity of identifying a reliable model in order to assure optimal performance to the artificial pancreas.

1.6 T1DM models

An accurate description of the glucose-insulin dynamics in the body is desirable, in order to make predictions on the behaviour of the subject to possible disturbances, such as meals or physical exercise.
There are two main categories of mathematical models for the prediction of blood glucose concentration: data-driven models and knowledge-driven models. The former are empirical models based on a black box approach, that is they consider only input-output data neglecting any physiology. These models are advantageous over knowledge-driven ones because of their structural simplicity and low time required for their development. However, data-driven models lack of any insights on glucose and insulin levels in tissues and organs, since they do not consider any physiological information. On the other hand, knowledge-driven models take into account the physiological interactions between glucose and insulin in the glucoregulatory system. Therefore, the effort required to develop these models is significantly higher, as they usually involve systems of differential and algebraic equations. The knowledge-driven models can be further divided into lumped or semiempirical models and comprehensive models. The former

models entail a limited number of dynamic equations, in which various organs or tissues in the body are lumped into one or more compartments [15]. The latter models are more complex and require more time for development than lumped models, which will be examined in detail in the next section.

1.6.1 Lumped physiological models

Bolie was a pioneer in the development of physiological models of the glucoregulatory system. In 1961, he proposed a linear model based on two differential ordinary differential equations representing glucose and insulin concentrations, based on his condensed theory. According to this theory, the liver, pancreas and peripheral tissues communicate with each other by means of a single compartment through which glucose and insulin concentration changes are distributed rapidly and uniformly [15]. The limitations of this model are connected with the omission of the action of the kidneys and intravascular-extravascular differences in concentrations of insulin and glucose.
The most widely known and used model was proposed in 1981 by Bergman et al. [9] and is generally referred to as *minimal model*. It was originally developed to measure insulin sensitivity and glucose effectiveness in an IVGTT, but it has gained popularity due to its relative simplicity. In fact, it uses only three differential equations and a few parameters to describe the effect of insulin on the regulation of glucose. It is a compartmental model composed of two submodels: one to account for the glucose disappearance and the other for the insulin kinetics, as shown in Figure 1.3.

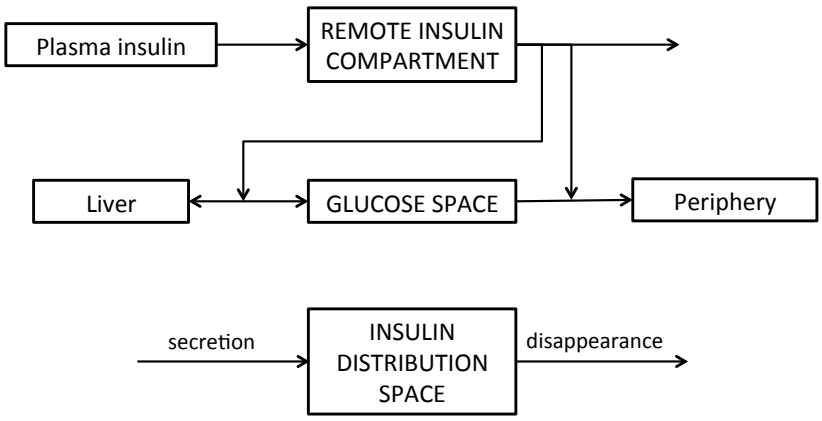


Figure 1.3. *Schematic of compartments in the minimal model: glucose disappearance (above) and insulin kinetics (below).*

Glucose disappearance is characterised by assuming a remote insulin compartment, where plasma insulin enters and can act effectively on the dispersal of glucose into the periphery and the liver and on the inhibition of hepatic glucose production. Postprandial insulin secretion can be distinguished in a first-phase release, which begins

within two minutes of the meal consumption and lasts 10-15 minutes and a following second-phase release, which is responsible for restoring of normoglycaemia. Therefore, first-phase insulin release is represented as a bolus accessing the plasma compartment at the time of the glucose injection, whereas the second-phase insulin secretion is described by the third differential equation of the model. The minimal model presents two strong limitations: firstly, it does not allow representing an exogenous insulin injection and, secondly, no description of the rate of glucose appearance after a meal is permitted. These limitations make the model not suitable for describing the typical insulin treatment in a T1DM individual.

A more comprehensive model was proposed by Lynch and Bequette (LB model) in 2002 [16], which is in turn derived from Fisher's model (1991). Fisher's model is a modification of the minimal model, where the three differential equations are adjusted to take into account glucose absorption after a meal. In addition, since it has been adapted to T1DM, it omits the term of insulin secretion, which is replaced by an exogenous insulin infusion term. The LB model adds a fourth differential equation to link blood and subcutaneous glucose concentrations accounting for a first order delay between them. A schematic of the LB model is presented in Figure 1.4 with the possible manipulable inputs, i.e. glucose meal and insulin infusion, and measurable outputs, i.e. glucose and insulin concentrations.

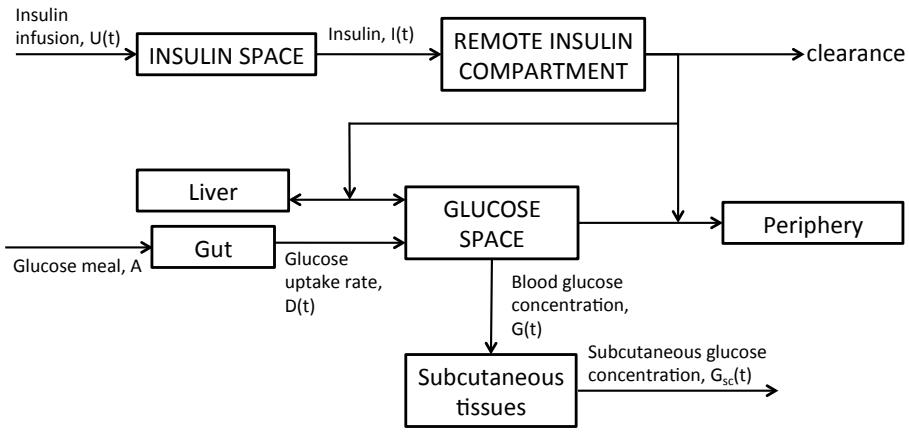


Figure 1.4. *Schematic of Lynch-Bequette model with inputs and outputs.*

Discrete results were achieved by Lynch and Bequette in the application of the model to model predictive control, even though lack of representation during food assumption and nocturnal periods was observed. In 2006, Van Herpe et al. proposed a modified minimal model whose extended time validity makes it appropriate to control blood glucose concentration in intensive care unit patients [17]. At the same time, Roy et al. [18] developed an extended minimal model which considers also free fatty acids, which play a relevant role in altering the glucoregulatory system and thus are represented

with an additional differential equation.
Several clinical and in vitro studies have shown that pancreatic insulin secretion follows rapid small amplitude and slower large amplitude oscillatory behaviour depending on the glucose infusion [15]. This fact has been considered by Sturis et al. [19] in developing a new physiological model of six ordinary differential equations and five nonlinear equations in the attempt to represent the slow oscillations of insulin and glucose. In 2000, Tolic et al. [20] introduced two models of insulin receptor dynamics in the model of Sturis et al. Later, Bennett et al. recognised that there were three variables in the model of Sturis et al. to represent the delay of plasma insulin to inhibit hepatic glucose production which had no clinical meaning. Therefore, they simplified the original model reducing the 6 ODEs to 3 by introducing an explicit time delay [21].
In 2004, Hovorka and Wilinska developed a more detailed model, known as Hovorka-Wilinska model [22], which permits a better description of the glucoregulatory system, taking into account also the absorption of subcutaneous short-acting insulin. The model includes three separate subsystems: one for glucose, one for insulin and one for insulin action and entails an increased complexity due to the higher number of differential equations and model parameters. Nevertheless, it has been proved to be effective for the characterisation of glucose homeostasis.
One of the latest models has been proposed by Dalla Man et al. in 2007 [23] to describe accurately glucose assumption through an oral meal. The model consists of a glucose and an insulin subsystem. The glucose subsystem is divided into two compartments: the first one accounting for glucose masses in plasma and rapidly equilibrating tissues and the second one in the slowly equilibrating tissues. The insulin subsystem is characterized by a two-compartment model: one considering insulin mass in plasma and the other in liver. The model was derived using data on normal subjects, thus additional modifications are required for its extension to T1DM. Moreover, the fact that it is a mean model must taken under consideration because it could affect individual prediction capabilities.

1.7 Design of clinical tests for model identification

The importance of having a reliable model of T1DM, when it comes to the design of an artificial pancreas, has already been stressed above, so now the relation between physiological models and clinical tests can be discussed.
Standard clinical tests are mainly used for the diagnosis of diabetes so that, based on their results, it is possible to state if the patient is ill or healthy. Nevertheless, they are of primary importance also for research purposes, in order to determine the parameters of physiological models, that is to identify the subject.

In the preliminary model analysis, the aim is, given a set of candidate models and some variables to be measured, to determine whether it is possible to design an experiment that allows to univocally identify the parameters of these models (identifiability) and to distinguish the different model structures from one another (distinguishability). Structural identifiability is a condition that guarantees that the same input, applied to the same model and at the same initial conditions, but using different parameter vectors, results in predicted output trajectories not overlying. Distinguishability, instead, ensures that two different models with different parameter vectors and with the same input and initial conditions generate two predicted output trajectories not overlying [24]. Once verified that the proposed model is endowed with structural identifiability and distinguishability, the identification of model parameters can be examined. Unique parameter identification depends on several factors: the structure of the model, the type of data collected in the experiment, sampling times at which data are collected and the level of input perturbation (oral/intravenous glucose and pattern of subcutaneous/intravenous insulin administrations) induced during the clinical test [15]. An accurate estimation of model parameters results in a model tailored on the individual subject, which can provide enhanced care therapy, for instance realising a customised artificial pancreas.

When performing a clinical test, an excitation pattern consisting in a glucose administration is applied to the subject and, depending on this solicitation and on how the response is detected, the parameter estimation may be more or less precise. This issue is not new, since in 1981 Bergman [9] had already pointed it out, stating that the IVGTT he had analysed could possibly not represent the best stimulus pattern for the estimation of metabolic parameters. Bergman suggested that using different temporal patterns for the glucose and insulin administration could improve the process. In fact, in 1987 he showed that the injection of insulin some time after glucose in the subject leads to an increased precision in the estimation of the parameters [25].

More recently, the possibility to obtain better parameter estimation using different tests has emerged. A schematic of the general procedure to be applied is shown in Figure 1.5. Initially, the test is designed, subject to the contraints to be met, making use of model-based design of experiments (MBDoE), then the experiment is carried out and model parameters are identified. Finally, statistical analysis is used to determine the accuracy of the estimated parameters.

Applying MBDoE techniques allows to design alternative clinical tests, derived from the standard ones, which can produce relevant improvement on parameter estimation while meeting also certain constraints. Adopting the Hovorka-Wilinska model, Galvanin et al. [26] have shown that modifications of some standard clinical tests, like the OGTT and the postprandial, can guarantee more accurate parameter estimation.

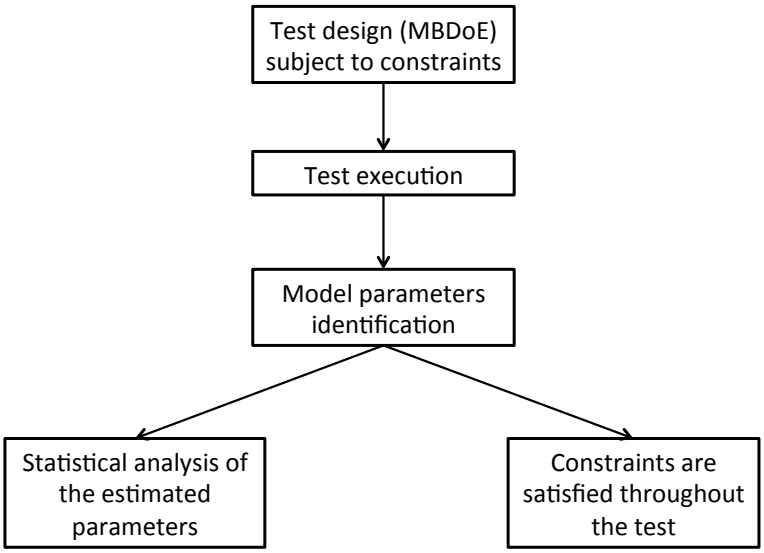


Figure 1.5. *Steps to parameters identification using MBDoE procedures.*

In this work three variables were optimised subject to safety constraints on the individual: carbohydrates content of the meal, insulin infusion rate and sampling times. It was proved that a statistically satisfactory parameter estimation for the single subject is achievable through modified clinical protocols, even though some limitations were encountered, such as excessive length (from 10 to 14h) and a difficult optimisation of exogenous insulin administration. Galvanin et al. [27] have also investigated the use of an advanced MBDoE technique that combines the online model-based design of experiments with the backoff-based MBDoE in the design of clinical tests. This approach allows to use experimental data as soon as they are available to adjust the design and takes into account the uncertainty associated with model parameters in order to design an optimally informative and safe test. Although parameter estimation achieved using this method is satisfactory, there could be some issues related to the duration of the test and restoration of basal values. Therefore, the need for more viable tests is the starting point for further investigations on the topic.

In 2014, Laguna et al. [28] have enquired an identification method based on interval analysis, which takes into account both variability of the subject and model imprecisions. The leading criterium of the procedure is the minimisation of a composite cost index consisting of a glucose envelope width and an Hausdorff distance-based error with respect to the envelope. Experimental tests showed that this kind of approach has good prediction capability in average and better identification performances when the maximum intra-patient variability is stimulated.

The use of continuous glucose monitoring systems (CGMSs) could ease the design of optimal clinical tests, even though noise in the measurements and formulation of

the test itself may hinder model identification in this type of approach, as shown by Galvanin et al. [29].

1.8 Objectives

The objective of this thesis is to devise feasible clinical protocols allowing for a precise identification of type 1 diabetes models suited to the characteristics of the individual subject, while meeting constraints related to:

- safety conditions for the subject must be guaranteed, assuring normoglyceamic values throughout the test and steady basal glucose concentration at the end of it;
- invasivity on the patient (type of glucose and insulin admnistration) must be minimised during execution, in order to make the protocol preferable among other options;
- ease of implementation, which means the test must be not excessively time demanding and not too complex to execute, in terms of sampling and glucose and insulin administration;
- robustness is required, so as to always assure parameter estimation is possible;
- the information content achievable must be maximised in order to identify the model parameters for the single subject with the greatest possible precision, so as to endow it with very reliable prediction capability.

All these requirements cannot be met using standard clinical tests because they have not been designed specifically for this purpose. Nevertheless, based on their structure a suitable test is to be devised which, regardless of the subject, is capable of assuring an optimal parameter estimation so as to design feasible customised care solutions.
In order to account for all the features above, a novel approach will be introduced in this work to simplify the choice of the best protocol among the several options available. This approach consists in the definition of some heuristic indices, which are related to the contraints listed above and allow a quick evaluation of the suitability of a designed clinical test for parameter identification.

Chapter 2

Model-based design of experiments

This chapter deals with the theoretical aspects of model-based design of experiments. After a preliminary discussion on the MBDoE procedure, the physiological model of Lynch and Bequette is described in detail. Then, much relevance is given to the mathematical description of the theory behind the MBDoE approach and the several options available in the software where it is embedded. Finally, a novel methodology based on heuristic indices is presented in order to evaluate the designed tests.

2.1 Introduction

State-of-the-art devices for T1DM management, such as the WAPs, require the availability of reliable physiological models for the description of glucose dynamics. Since each subject is characterised by its own physiology, a precise identification of its model parameters needs to be carried out in order to optimise the performance of the control algorithm, which is responsible for maintaining normogylcaemia. Parameter identification relies on the execution of clinical tests to acquire as much information as possible to achieve this purpose.
The most commonly used clinical protocols have already been discussed in § 1.4; however, they were not specifically designed with this aim, so their sampling schedule is often inappropriate. To overcome this drawback, it is useful to adopt a model-based design of experiments approach, which allows to design new or modified clinical protocols to accomplish the identification task.
MBDoE is a sequential procedure, which consists of three steps:

1. experiment design: based on the preliminary knowledge of the system, the optimal sampling schedule is sought while acting on some manipulable variables. The objective is to maximise the information content derivable subject to some constraints related to the safety of the patient;

2. experiment execution: experimental data are collected carrying out the experiment with the optimal conditions obtained in the previous step;

3. parameter identification: based on experimental data and using some estimators, the identification of the unknown model parameters is performed.

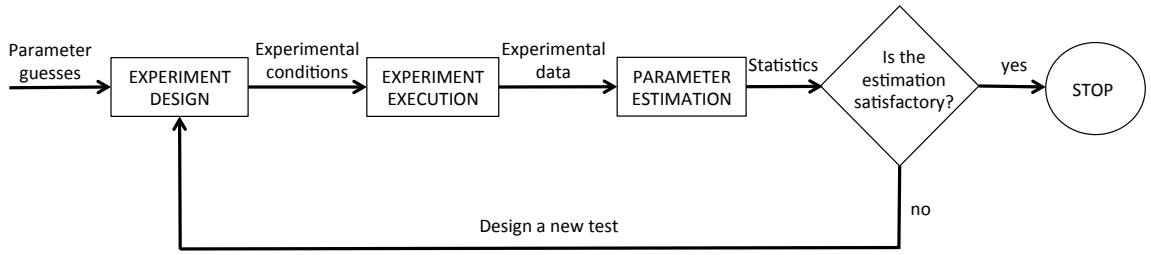


Figure 2.1. MBDoE procedure for parameter estimation [26].

2.2 Lynch-Bequette model

In this work the Lynch and Bequette model [16] (also referred to as LB model) is selected because of its ease in representing the glucoregulatory system compared to alternative more recent models. Although lumped physiological models, such as the LB model, are not able to represent the dynamics of glucose at various organ or tissue levels, they entail a lesser complexity, also in terms of number of parameters, compared to comprehensive models. In fact, it consists of only 4 differential equations, which are displayed below:

$$\frac{dG(t)}{dt} = -\theta_1 G(t) - X(t)[G(t) + G_b(t)] + D(t) \tag{2.1}$$

$$\frac{dX(t)}{dt} = -\theta_2 X(t) + \theta_3 I(t) \tag{2.2}$$

$$\frac{dI(t)}{dt} = -\theta_4[I(t) + I_b] + \frac{U(t)}{V_i} \tag{2.3}$$

$$\frac{dG_{sc}(t)}{dt} = \frac{G(t) - G_{sc}(t)}{5} - R_{ut} \tag{2.4}$$

A description of the model variables and constants is given in Table 2.1 and 2.2, respectively.

As pointed out in § 1.6, equations (2.1)-(2.3) are taken from Fisher's model, whereas equation (2.4) relates subcutaneous and blood glucose values considering a 5 minutes delay between them. Examining the system in detail: equation (2.1) describes the dynamics of blood glucose concentration, equations (2.2) and (2.3) characterise the

dynamics of insulin, using a variable proportional to insulin in the remote compartment (X) and plasma insulin concentration (I).

Table 2.1. Lynch-Bequette model variables.

Variable	Description
$G(t)$	Difference between blood glucose concentration and its basal value [mg/dL]
$D(t)$	Glucose intake velocity after a meal [mg/dL/min]
$U(t)$	Exogenous insulin infusion rate [mU/min]
$X(t)$	Variable proportional to insulin in the remote compartment [mU/L]
$G_{sc}(t)$	Difference between subcutaneous glucose concentration and its basal value [mg/dL]
$I(t)$	Difference between plasma insulin concentration and its basal value [mU/L]

Table 2.2. Lynch-Bequette model constants.

Constant	Description
V_i	Insulin distribution volume [L]
R_{ut}	Glucose usage velocity [mg/dL/min]
G_b	Glucose basal value [mg/dL]
I_b	Insulin basal value [mU/min]

The inputs to the model are represented by the glucose meal D(t) and the exogenous insulin infusion rate U(t). The model, as presented by equations 2.1-2.4, describes an OGTT, where the meal-absorption dynamics, differently from Fisher, is not modelled through the exponential form $D(t) = A \cdot exp(-0.05t)$ but using the more accurate expression proposed by Hovorka and colleagues [22]:

$$D(t) = \frac{1000 \cdot A \cdot t \cdot exp(-t/t_{max,G})}{t^2_{max,G} \cdot BW \cdot V_g} \tag{2.5}$$

where A is the carbohydrate content of the meal expressed in grams and $t_{max,G}$ represents the time-of-maximum appearance rate of glucose in the accessible glucose compartment. BW and V_g are factors related to the characteristics of the subject: the first one, accounts for its body weight and the second one, assesses the available glucose distribution volume.
In this analysis, the following values for the subject have been considered:

$$BW = 70\,\text{kg} \quad t_{max,G} = 20\,\text{min} \quad V_g = 1.6\,\text{dL/kg} \tag{2.6}$$

This choice represents a lean subject, with the same glucose distribution volume reported by Hovorka et al. [22] and a sensible value for the time-to-maximum of glucose

uptake.
On the contrary, in the IVGTT, glucose is administrated intravenously, thus the exponential term *D(t)* in equation 2.1 is substituted with the glucose injection, U_v, while the rest remains unchanged.
The monitored variable is the total glucose concentration, which is given by:

$$G_{tot}(t) = G(t) + G_b \tag{2.7}$$

The values of G_b, I_b and V_i are chosen according to Lynch and Bequette [16]:

$$G_b = 80\,\text{mg/dL} \quad I_b = 15\,\text{mU/L} \quad V_i = 12\,\text{L} \tag{2.8}$$

whereas for R_{ut} the value proposed by Sorensen [30] is selected:

$$R_{ut} = 0.75\,\text{mg/(dL} \cdot \text{min)} \tag{2.9}$$

θ_1-θ_4 represent the model parameters, which allow for the identification of the single subject metabolic behaviour. A description of the parameters is presented in Table 2.3:

Table 2.3. Lynch-Bequette model parameters.

Parameter	Description
θ_1	Rate of blood glucose disappearance into liver or periphery [min^{-1}]
θ_2	Insulin disapperance velocity from the remote compartment [min^{-1}]
θ_3	Insulin apperance velocity in the remote compartment [min^{-1}]
θ_4	Insulin disappearance velocity from the insulin space [min^{-1}]

θ_1 is also known as glucose effectiveness (S_G) and quantifies the capacity of the glucoregulatory system to reduce blood glucose concentration after a meal. Therefore, this parameter is very relevant for T1DM patients. The ratio θ_3/θ_2 represents the sensitivity to insulin (S_I), that is the effectiveness of insulin in restoring glucose basal values. These two parameters are less important in patients affected by T1DM, as they do not produce insulin on their own.
The design of a feasible clinical protocol for parameter identification represents the ultimate objective of the present work.
The established LB model can be used to evaluate standard clinical tests but it also proves useful to design enhanced modified protocols, where the two manipulable variables (i.e. glucose administration and insulin infusion) allow maximising the obtainable information content.

2.3 Optimal model-based design of experiments

This section deals with the theory behind optimal model-based design of experiments. Firstly, a mathematical description of the general MBDoE problem is provided with the variables and constraints involved. Secondly, the dynamic optimisation strategy is tackled with an overview on the main optimisation criteria available.

2.3.1 Mathematical model and constraints

Dynamic models, such as the Lynch and Bequette model, can be represented using a general form consisting of a system of differential and algebraic equations (DAEs):

$$\begin{cases} \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \mathbf{w}, \hat{\boldsymbol{\theta}}, t) = 0 \\ \hat{\mathbf{y}}(t) = \mathbf{h}(\mathbf{x}(t)) \end{cases} \tag{2.10}$$

where $\mathbf{x}(t) \in \Re^{N_x}$ is the vector of time-dependent state variables, $\mathbf{u}(t) \in \Re^{N_u}$ and $\mathbf{w} \in \Re^{N_w}$ are the time-dependent and time-invariant vectors control variables, respectively, $\hat{\boldsymbol{\theta}} \in \Re^{N_\theta}$ is the set of unknown model parameters to be estimated, t is the time, $\hat{\mathbf{y}} \in \Re^{N_y}$ is the vector of output responses estimated by the model.
Specifically, in the LB model, insulin infusion $U(t)$ represents the time-dependent control variable and glucose meal A the time-invariant control variable, while the variable to be measured is the total glucose concentration $G_{tot}(t)$.
In order to solve system (2.10), a set of initial conditions, i.e. values of the state variables at time $t = t_0$, is required in the form:

$$\begin{cases} \mathbf{f}(\dot{\mathbf{x}}(t_0), \mathbf{x}(t_0), \mathbf{u}(t_0), \mathbf{w}, \hat{\boldsymbol{\theta}}, t_0) = 0 \\ \hat{\mathbf{y}}(t_0) = \mathbf{h}(\mathbf{x}(t_0)) \end{cases} \tag{2.11}$$

In physiological models, state variables cannot assume any value but need to be constrained to assure safety conditions to the patient throughout the experiment. Therefore, in the design procedure the following safety end-point constraints are imposed on:

- lower bound of glucose concentration (LB), which is considered an "hard" constraint, that is not to be violated, in order to prevent hypoglycemic conditions. This is mathematically expressed as:

$$LB - G_{tot}(t) \leq 0 \rightarrow G_{tot}(t) \geq LB \tag{2.12}$$

but it is implemented in the software as:

$$\frac{dV_1}{dt} = max(0, LB - G_{tot}(t)) \leq 0 \tag{2.13}$$

where V_1 represents the amount of the contraint violation. The first entry of the function *max* represents the maximum violation allowed, while the second entry defines the function on which the condition applies.

- upper bound of glucose concentration (UB), which can be treated either as a "hard" or a "soft" constraint, since a small violation (V_2) is much more tolerable by the organism. Similarly to equation 2.13, this is implemented in the software as:

$$\frac{dV_2}{dt} = max(0, G_{tot}(t) - UB) \leq 0 \tag{2.14}$$

- final glucose concentration, which must be restored within a certain range of the initial basal value; a 10% deviation from the initial value has been allowed at the end of the test.

- derivative of the final glucose concentration (Γ):

$$\Gamma = \frac{dG}{dt} \tag{2.15}$$

which is an additional requirement to guarantee that the basal glucose value is not only reached but also maintained at the end of the test, in order for the patient to be safely discharged. Due to its definition, this constraint is compelled to be as close as possible to zero. Typical values used in the design phase for lower and upper bounds of Γ are -0.01 and 0.01, respectively.

In this work, a glucose concentration of 170 mg/dL is selected as upper bound and 60 mg/dL as lower bound.
MBDoE aims at reducing the model parameters uncertainty region by optimising the experimental settings, which are grouped in the design vector $\boldsymbol{\varphi} \in \Re^{n_\varphi}$:

$$\boldsymbol{\varphi} = \{\mathbf{y}_0, \mathbf{u}(t), \mathbf{w}, \mathbf{t}^{sp}, \tau\} \tag{2.16}$$

where $\mathbf{y}_0$ is the set of initial conditions on the measured variables, $\mathbf{t}^{sp} = [t_1, ..., t_{n_{sp}}]^T$ is the vector of sampling times, which defines the times at which the output variables are measured and τ is the duration of the experiment.
In order to facilitate the solution of the dynamic optimisation problem, some control vector parametrisation techniques can be adopted to approximate the set of control variables $\mathbf{u}(t)$ over a number of intervals, as shown by Vassiliadis and colleagues [31].

For instance, in this work a piecewise constant approximation is used for exogenous insulin infusion, which means that the variables requiring optimisation are:

- the $(n_{sw}-1)$ times at which the insulin profile changes in value, which are collected in the vector $\mathbf{t}_{sw}$ of switching times;
- the n_{sw} values that the control variable assumes before and after a switching time, which are collected in the vector $\mathbf{z}_{sw}$ [32].

Furthermore, due to practical limitations or safety reasons, constraints on maximum values of insulin infusion rate and on upper and lower duration of switching intervals are imposed in the design stage.

2.3.2 Objective function and optimisation criteria

The aim of MBDoE is the maximisation of the information content derivable from the designed test, so as to achieve a statistically satisfactory parameter identification by optimising the design vector $\boldsymbol{\varphi}$. In particular, it is the function ψ of the variance-covariance matrix that needs to be minimised acting on the elements of the design vector:

$$\boldsymbol{\varphi}_{opt} = arg\,min\{\psi[\mathbf{V}_\theta(\boldsymbol{\theta}, \boldsymbol{\varphi})]\} \tag{2.17}$$

For a single experiment, the variance-covariance matrix of model parameters $\mathbf{V}_\theta$ is defined as:

$$\mathbf{V}_\theta(\boldsymbol{\theta}, \boldsymbol{\varphi}) = [\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{H}_\theta(\boldsymbol{\theta}, \boldsymbol{\varphi})]^{-1} \tag{2.18}$$

where $\mathbf{H}_\theta$ represents the Fisher information matrix (FIM) and $\boldsymbol{\Sigma}_\theta$ the preliminary variance-covariance matrix of model parameters.
For a dynamic system with discrete sampling the FIM, according to the formulation of Zullo [33], assumes the form:

$$\mathbf{H}_\theta(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{k=1}^{n_{sp}} \sum_{i=1}^{N_y} \sum_{j=1}^{N_y} s_{ij} \mathbf{Q}_i^T \mathbf{Q}_j + \mathbf{H}_\theta^0 \tag{2.19}$$

where s_{ij} is the ij-th element of the N_yxN_y inverse matrix of measurement errors, $\mathbf{Q}_i$ is the n_{sp}xN_θ dynamic sensitivity matrix of the i-th measured response, whose elements take the form:

$$\mathbf{Q}_i = \left[\frac{\partial \hat{y}_{il}}{\partial \theta_m}\right] \quad l = 1, ..., n_{sp} \quad m = 1, ..., N_\theta \tag{2.20}$$

and $\mathbf{H}_\theta^0$ is the prior dynamic information matrix.
Several design criteria have been proposed for the maximisation of the information matrix. The most known ones are the so called "alphabetical criteria", which are also available in *gPROMS*®:

- A-optimality: minimises the trace of the variance-covariance matrix $\mathbf{V}_\theta$:

$$\psi_A(\mathbf{V}_\theta) = \frac{1}{N_\theta} \sum_{i=1}^{N_\theta} (\nu_{ii}) \tag{2.21}$$

where ν_{ii} represents the elements on the diagonal of the variance-covariance matrix, which are summed because of the definition of the trace. In geometrical terms, this criterion corresponds to minimising the dimensions of the hyperrectangle within which the confidence ellipsoid of model parameters can be inscribed, as shown in Figure 2.2 (in 2D).

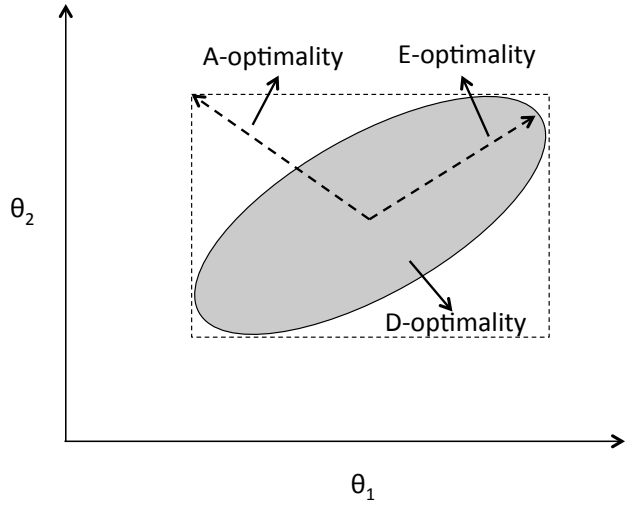


Figure 2.2. Graphical interpretation of the different design criteria used in MBDoE in a 2D case.

- D-optimality: minimises the determinant of the variance-covariance matrix:

$$\psi_D(\mathbf{V}_\theta) = det(\mathbf{V}_\theta)^{\frac{1}{N_\theta}} \tag{2.22}$$

This is also known as the minimum volume criterion because it minimises the volume of the confidence ellipsoid. In 2D this is the equivalent of minimising the area of the ellipse (Figure 2.2).

- E-optimality: minimises the largest eigenvalue of the variance-covariance matrix:

$$\psi_E(\mathbf{V}_\theta) = \max_{k=1,\ldots,N_\theta} \lambda_k(\mathbf{V}_\theta) \tag{2.23}$$

The eigenvalues of the variance-covariance matrix represent the lengths of the minor and major axes of the confidence ellipsoid. Therefore, by minimising the largest eigen-

value, the design forces the confidence ellipsoid to become as spherical as possible. In 2D it means the ellipse is forced as close as possible to a circumference (Figure 2.2).

2.4 Experimental data generation

Once the optimal design vector $\boldsymbol{\varphi}_{opt}$ has been obtained, the following step of the MB-DoE consists in the experiment execution. At this stage, experimental data need to be collected from the designed test. However, this is not actually possible because the test is still in its design phase. For this reason, it cannot be performed *in vivo* because safety and effectiveness have not been proved and the optimal conditions obtained from the model in the design phase could jeopardise the subject in reality. In addition, at this point data retrieval from a real patient would significantly augment the overall time required for the sequential design procedure.
Therefore, experimental data are generated *in silico* using the same physiological model but with different parameters (parametric mismatch). This procedure consists in the following steps:

1. sensible perturbations are assigned to the initial parameters used for the experiment design, generating a new set $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4]^T$, which is assumed to represent the real physiology of the patient. This set is used to simulate the glucoregulatory response of that subject;

2. glucose concentrations collected at the sampling times obtained from the designed test constitute the experimental measurements of the simulated patient;

3. some stochastic and ergodic disturbance factors are applied to these measurements, in order to simulate the experimental noise, that is the random errors that make each repetition of the experiment different from the others. Assuming that systematic errors are avoided, the stochastic component can be regarded as normally distributed with zero mean. The general variance model has the form:

$$\sigma_i^2 = \omega_i^2(\hat{y}_i^2)^{\gamma_i} \tag{2.24}$$

where ω_i is the standard deviation of the i-th measured response and γ_i the heteroschedastic factor. In this work a constant relative variance of 0.033 is considered according to Galvanin and colleagues [26], thus $\gamma = 1$ and equation 2.24 becomes:

$$\sigma_i^2 = \omega_i^2(\hat{y}_i^2) = (0.033G_{tot})^2 \tag{2.25}$$

2.5 Parameter estimation

The parameter estimation step consists in the identification of the model parameters in a statistically sound way based on the experimental data collected or generated *in silico*. The optimal result of this procedure is a vector that minimises the distance between experimental data $\mathbf{y}$ and model predicted responses $\hat{\mathbf{y}}$.

2.5.1 Estimators

The most widely used estimators for the evaluation of the model parameters of a dynamic model are:

- least squares (LS): is the simplest one and is defined as:

$$\Phi^{LS}(\mathbf{y}) = \sum_{i=1}^{n_{sp}} \left[(\mathbf{y}_i - \hat{\mathbf{y}}_i)^T (\mathbf{y}_i - \hat{\mathbf{y}}_i) \right] \tag{2.26}$$

- weighted least squares (WLS): takes into account the variance-covariance matrix of measurement errors:

$$\Phi^{WLS}(\mathbf{y}, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_{n_{sp}}) = \sum_{i=1}^{n_{sp}} \left[(\mathbf{y}_i - \hat{\mathbf{y}}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - \hat{\mathbf{y}}_i) \right] \tag{2.27}$$

  As suggested by Galvanin [32], the problem associated with LS and WLS is that they only provide an estimate of model parameters $\hat{\theta}$ but no a-posteriori statistics about their precision.

- maximum likelyhood (ML): considers a probability distribution that allows to derive the a-posteriori variance-covariance matrix of model parameters, which makes this approach suitable for MBDoE. For independent and normally distributed measurement errors, the ML estimator can be expressed according to Bard [34] as:

$$\Phi^{ML}(\mathbf{y}, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_N) = 2\pi^{\frac{N_y}{2}} \prod_{i=1}^{n_{sp}} |\boldsymbol{\Sigma}_{\boldsymbol{i}}|^{-\frac{1}{2}} exp\left\{ -\frac{1}{2} \sum_{i=1}^{n_{sp}} \left[(\mathbf{y}_i - \hat{\mathbf{y}}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - \hat{\mathbf{y}}_i) \right] \right\} \tag{2.28}$$

  This is also the estimator that *gPROMS*® uses for the evaluation of model parameters.

2.5.2 Assessment of parameter estimates

Parameter estimates need to be evaluated in terms of precision and accuracy. Precision of the estimate corresponds to enclosing the set of parameters to a restricted confidence region, whereas accuracy expresses how much close to the true value (which is unknown) the estimate is.
Precision of the estimates is evaluated by means of confidence intervals, which are defined as:

$$K_i(\alpha) = \bar{t}\Big(\frac{1-\alpha}{2}, n_{sp} - N_\theta\Big) \cdot \sqrt{\nu_{ii}} \qquad i = 1, ..., N_\theta \tag{2.29}$$

where $\bar{t}$ represents the upper critical value of a t-distribution with $(n_{sp} - N_\theta)$ degrees of freedom and ν_{ii} the variance of the i-th parameter, i.e. element of the variance-covariance matrix. A α% confidence region indicates that if the experiment is repeated and the parameters are estimated from the new experimental data, their values will lie in this confidence region with α% probability. Therefore, the confidence intervals define the bounds of the confidence ellipsoid:

$$[\hat{\theta}_i - K_i(\alpha); \hat{\theta}_i + K_i(\alpha)] \qquad i = 1, ..., N_\theta \tag{2.30}$$

where $\hat{\theta}_i$ is the i-th estimated parameter.
The t-values express the percentage accuracy of the estimated parameters with respect to the 95% confidence intervals, i.e. $\bar{t}(0.95, n_{sp} - N_\theta)$, which are given by:

$$\bar{t_i} = \frac{\hat{\theta}_i}{K_i(0.95)} \qquad i = 1, ..., N_\theta \tag{2.31}$$

These values need to be compared with the reference t-value obtained from a Student t-distribution with confidence α and $(n_{sp} - N_\theta)$ degrees of freedom. A t-value larger than the reference one means that the parameter has been accurately estimated, otherwise a poor estimate has been accomplished.
Accuracy cannot be evaluated directly since the true value of the parameters remains unknown, instead, the lack-of-fit test is used to assess that the residuals, i.e. difference between measured output and predicted response, are actually minimised. It consists in a χ^2-test considering the sum of weighted residuals:

$$SWR = \sum_{i=1}^{N} [(\mathbf{y}_i - \hat{\mathbf{y}}_i)^T \Sigma_i^{-1} (\mathbf{y}_i - \hat{\mathbf{y}}_i)] \tag{2.32}$$

The fitting of the experimental data is adequate and the model represents a good description of the system if:

$$SWR < \chi^2_{ref} \tag{2.33}$$

where χ^2_{ref} is a reference χ^2 distribution with $(n_{sp} - N_\theta)$ degrees of freedom.

2.6 Preliminary analysis of the model

Prior to the design of experiments procedure, an a-priori evaluation of the physiological model can be performed so as to get some insights on its structure and potential. In the present work, two preliminary studies are conducted:

- sensitivity analysis,
- information content analysis.

2.6.1 Sensitivity analysis

Sensitivity analysis consists in the assessment of how some variations on the inputs to the model reflect on the predicted outputs. In the first place, it is a qualitative analysis aimed at detecting the temporal effect of the variation on some variables, since a dynamic model is involved. This kind of observation allows, in the case of a parametric model, to determine which parameters can be more easily identified (for a specific set of experimental settings) and which, instead, could require a reparametrisation to facilitate the process. The responses predicted by the model $\hat{\mathbf{y}}$ to a change in the parametric set $\boldsymbol{\theta}$ are collected in the N_yxN_θ matrix of local sensitivities:

$$\mathbf{Q}(t) = \begin{bmatrix} q_{1,1}(t) & \dots & q_{1,N_\theta}(t) \\ \vdots & \ddots & \vdots \\ q_{N_y,1}(t) & \dots & q_{N_y,N_\theta(t)} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \hat{y}_1(t)}{\partial \theta_1} & \cdots & \dfrac{\partial \hat{y}_1(t)}{\partial \theta_{N_\theta}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \hat{y}_{N_y}(t)}{\partial \theta_1} & \cdots & \dfrac{\partial \hat{y}_{N_y}(t)}{\partial \theta_{N_\theta}} \end{bmatrix} \tag{2.34}$$

2.6.2 Information content analysis

Information content analysis in terms of trace of the FIM can be used for an evaluation of complex systems with a large number of parameters. In this case, the dynamic profile of the trace highlights time periods when the maximum information can potentially be obtained through sampling.
Alternatively, a different evaluation of the information content can be performed by comparing continuous and discrete sampling. Continuous cumulative information content I_c derivable in the entire test can be computed through the integral of the trace of the FIM:

$$I_c = \int_{t=0}^{\tau} tr[\mathbf{H}_\theta(t)]dt \tag{2.35}$$

while discrete information content I_d is given by considering only the information obtained at sampling times:

$$I_d = \sum_{i=1}^{t_{sp}} tr[\mathbf{H}_\theta(t)]\Delta t \tag{2.36}$$

Beyond the gap in value between the two strategies, this analysis can provide interesting insights on the best time allocation of the samples.

2.7 Software

The entire simulation work presented in this thesis from modelling to experiment design and parameter estimation is carried out using *gPROMS® Model builder 4.1.0* [35]-[36], which is an advanced modelling software by *Process Systems Enterprise Ltd.*
gPROMS® is an equation-oriented modelling system used for building, validating and executing first-principles models. It allows for the simultaneous resolution of large systems of algebraic and differential equations. Its peculiarity is that it can be applied to a wide variety of problems: from steady-state simulation to dynamic simulation and from parameter estimation to model-based design of experiments.
Generation of the noise applied to glucose measurements is made using the function *normrnd* of *MATLAB® R2013a* by *MathWorks®* [37], which produces arrays from the normal distribution with the specified mean and variance.

2.8 Indices

The main objective of this work is to devise a feasible clinical test for the identification of T1DM. So far, the attention has been focused on the experiment design procedure, aimed at estimating the model parameters with the greatest precision.
In order to discriminate among the possible tests the one offering the best performance, some criteria for the evaluation must be defined. These criteria must take into account, beyond the information content derivable, the assurance of safety conditions throughout the execution and feasibility factors for a successful application.
In order to fulfil these requirements a set of 5 heuristic indices is proposed as the framework for the test design procedure. All indices are normalised, i.e. defined on a scale ranging from 0 to 1, where the boundary values represent the least and most desirable one, respectively. This allows a quick comparison among the tests, both to identify possible improvement factors and to evaluate the best protocol.
Each index is individually presented and discussed in the following sections.

2.8.1 Information content index

It quantifies the information content derivable from a test by means of the discrete information content (I_d), defined by equation 2.36, achieved during the sampling procedure, by comparing it with the maximum value ($I_{d,max}$) obtained among all the tests analysed at a given parametric set. Mathematically, it is expressed as:

$$\gamma = \frac{I_d}{I_{d,max}} \tag{2.37}$$

2.8.2 Invasivity index

It is an attempt to quantify the impact of the glucose and insulin administration strategies on the subject. Compared to the other indices, it is more tricky to define because it depends on the tolerance of the individual.
A basic difference between the OGTT and the IVGTT lies in glucose administration: for the first it is oral, while for the second it is intravenous.
In terms of insulin administration, several pathways exist: the most used is subcutaneous delivery because it allows continuous infusion and it is less invasive (since it involves the subcutaneous layer) than intravenous injection, which is normally utilised just for surgery or complications. A third preferable route is oral delivery, which is a non-invasive method but it still requires studies due to insulin lack of permeability in the intestinal epithelium [38].
Invasivity index has been assessed through the following equation:

$$\alpha = 1 - \frac{\alpha_{glucose} + \alpha_{insulin}}{2} \tag{2.38}$$

where the factors $\alpha_{glucose}$ and $\alpha_{insulin}$ may assume the values: 0 for oral administration, 0.5 for subcutaneous (only for insulin administration) and 1 for intravenous.

2.8.3 Safety index

This index takes into account the violations of glucose upper and lower bounds observed during the entire duration of the test in terms of areas. As already pointed out, upper and lower bound violations are to be treated differently because of their different impact on the subject. Thus, the total violation β' is calculated using a weighting factor:

$$\beta' = V_1 + 0.5V_2 \tag{2.39}$$

where V_1 and V_2 represent the violations of glucose lower and upper bounds, respectively.

However, since the sought value should lie in the range from 0 to 1, applying normalisation the safety index β becomes:

$$\beta = 1 - \frac{\beta'}{\beta_{max}} \tag{2.40}$$

where β' is the total violation of the test considered and β_{max} the maximum violation observed over all the tests analysed at a given parametric set.

2.8.4 Sample index

This index accounts for the overall number of samples which are collected during the test execution. A maximum of 60 samples (n_{max}) collected is considered, assuming to apply a discrete sampling schedule to the OGTT (whose duration is 2 h) with a minimum time between measurements of 2 min. Therefore, the index results:

$$\pi = 1 - \frac{n_{sp}}{n_{max}} \tag{2.41}$$

2.8.5 Duration index

This index considers the time horizon for the execution of the test. As a reference, a maximum duration of 12 h (t_{max}), corresponding to the length of a day hospital, is assumed for its completion, that is taking into account the time required to restore and maintain the glucose basal value of the subject. Therefore, the index is simply defined as:

$$\tau = 1 - \frac{t_{test}}{t_{max}} \tag{2.42}$$

Chapter 3

Preliminary analysis of the model

This chapter deals with the preliminary evaluation of the Lynch and Bequette model by means of two clinical tests. Firstly, the two reference tests adopted in this work are described, in terms of schedule and inputs to the model. Secondly, sensitivity analysis is applied to the model to investigate how it responds to small variations in the input of glucose and in model parameters using the two standard protocols. Finally, a comparison is made of the cumulative information content derivable from the model using the standard tests with discrete and continuous sampling.

3.1 Standard clinical tests

In § 1.4 the structure of the main standard clinical tests has already been discussed. The present work is concerned with only two of them: the Oral Glucose Tolerance Test and the Intravenous Glucose Tolerance Test, which are considered the framework for the development of optimised protocols.
As outlined by the WHO, the OGTT entails only two glucose samples, which are clearly insufficient for model identification purposes. Because of this and the unfeasibility to restore glucose basal value within the 2 h duration (using 2 U of insulin), a different protocol for the OGTT has been considered. This modified protocol, proposed by Dalla Man and coworkers [39], consists in a 5 h test using the same glucose administration but collecting 11 samples at times: 0, 10, 20, 30, 60, 90, 120, 150, 180, 240 and 300 min. From now on, this test will be referred to as *standard OGTT*. In the case of the OGTT, the LB model, described in § 2.2, has two control variables: glucose administration and insulin infusion. For the standard OGTT, the glucose meal amounts to 75 g and the insulin infusion rate is kept constant at 16.666 mU/min [16], which corresponds to a total of 5 U.
The schedule of the IVGTT is maintained unchanged, as described in § 1.4, with 23 samples collected at the sampling times prescribed by Bergman [9] over the 182 min

duration of the test. The so established protocol will be referred to as *standard IVGTT*. Similarly to the OGTT, there are still two control variables in the model: glucose injection and insulin infusion. For the IVGTT, the term representing the glucose meal *D(t)* in equation 2.1 of the LB model is replaced by the glucose injection U_v. Following the guidelines of this test, a glucose injection of 300 mg/kg over a 1 min interval is performed at time t=0, which means the term U_v corresponds to:

$$U_v = \frac{300\,\text{mg/kg} \cdot \cancel{BW}}{1\,\text{min} \cdot V_g \cdot \cancel{BW}} = \frac{300\,\text{mg/kg}}{1\,\text{min} \cdot 1.6\,\text{dL/kg}} = 187.5 \frac{\text{mg}}{\text{dL} \cdot \text{min}} \tag{3.1}$$

In the same way as the OGTT, insulin infusion rate is maintained constant at 16.666 mU/min so, due to the shorter duration of the IVGTT, the total amount of insulin is reduced to 3033 mU.
The initial set of parameters $\boldsymbol{\theta_0}$ for the model has been chosen according to Lynch and Bequette, as shown in Table 3.1.

Table 3.1. Initial set of parameters $\boldsymbol{\theta_0}$ for the LB model.

Parameter	Value [min^{-1}]
θ_1	0.028735
θ_2	0.028344
θ_3	$5.035 \cdot 10^{-5}$
θ_4	5/54

3.2 Sensitivity analysis

3.2.1 Sensitivity to glucose administration

Firstly, sensitivity analysis has been carried out on the control variable regarding glucose meal/injection. The input of glucose A expressed in grams has been increased of 10% in the standard OGTT (Figure 3.1) and, similarly, the glucose injection has been augmented of 10% in the standard IVGTT (Figure 3.2). In both cases, the output variable of interest is the glucose concentration G_{tot}, because it is the measured variable in the experiments to be designed and, more importantly, it is the variable that needs to be controlled in the patient.
Figure 3.1 and 3.2 show a first important difference between the two standard protocols on the glucose profile due to the type of glucose administration. In the standard OGTT, the oral glucose assumption is modelled with the exponential equation 2.5 and this results in a delayed smooth peak. On the contrary, in the standard IVGTT the intravenous glucose injection produces an immediate spike.

Additionally, in both cases the effect of the glucose change in input is concentrated on the spike of the curve and its proximity and softens when moving away from it.
As a further consideration, the amount of insulin administered is insufficient for both the standard protocols to prevent violations of glucose upper bound, even though basal glucose values are restored at the end of them.

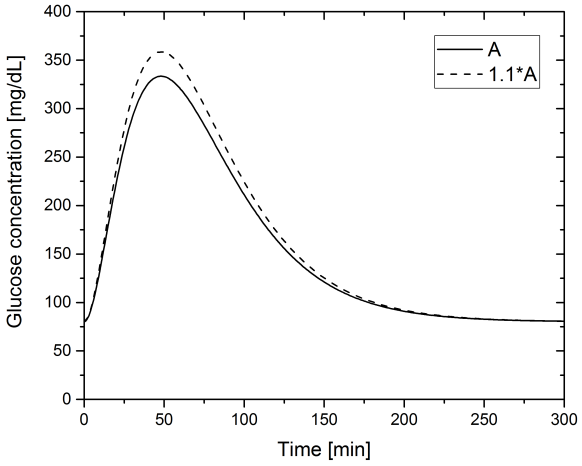


Figure 3.1. Sensitivity of glucose concentration G_{tot} to a 10% increase in the glucose meal A in a standard OGTT.

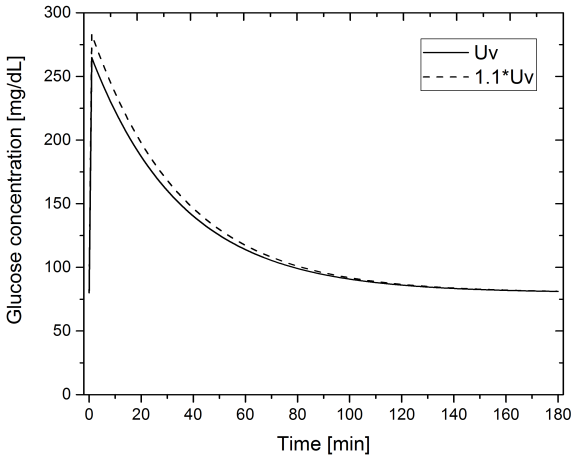


Figure 3.2. Sensitivity of glucose concentration G_{tot} to a 10% increase in the glucose injection U_v in a standard IVGTT.

3.2.2 Sensitivity to model parameters

In the second place, sensitivities of glucose concentration to model parameters have been individually tested upon a 10% increase in their values. Sensitivity matrix is

calculated according to formula 2.34, whose elements, considering only one output (i.e. G_{tot}), can be expressed as:

$$q_i = \frac{G_{tot}(\theta_i') - G_{tot}(\theta_i)}{\theta_i' - \theta_i} \cdot \theta_i \quad i = 1, \ldots, N_\theta \tag{3.2}$$

where θ_i and θ_i' represent the original and perturbed sets of parameters, respectively, and $G_{tot}(\theta_i')$ and $G_{tot}(\theta_i)$ the corresponding responses of the model.
The results are presented in Figure 3.3 for the standard OGTT and 3.4 for the standard IVGTT.

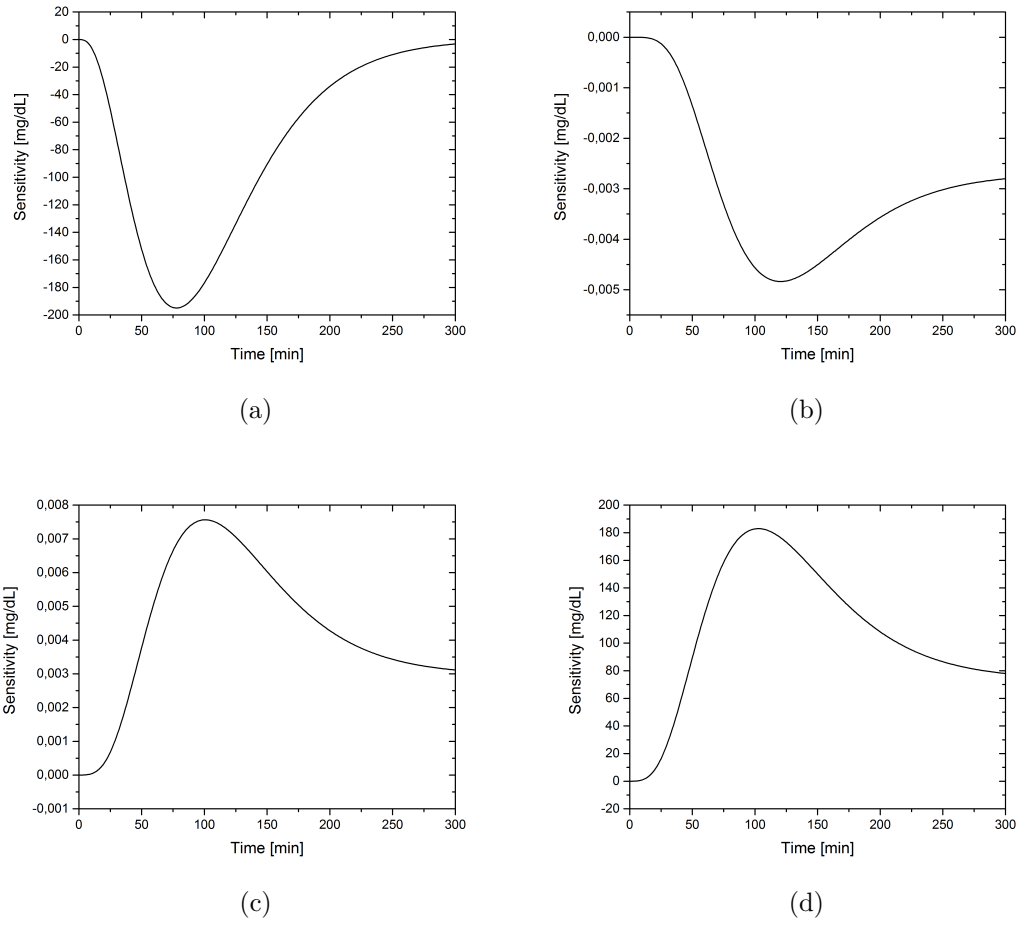


Figure 3.3. *Standard OGTT sensitivity analysis to a 10% increase in model parameters: θ_1 (a), θ_2 (b), θ_3 (c) and θ_4 (d).*

Peaks exhibited in sensitivity indicate zones where the larger deviations due to the parameter change are observed, hence where the maximum information content of that specific test can be extracted.
For the standard OGTT, glucose sensitivities to the parameters show spikes distributed in the times ranging from 75 to 125 min. Looking back at Figure 3.1, it can be observed that this temporal distribution corresponds to the period following the maximum in

glucose concentration occurring after a meal. It is worth noting that glucose sensitivities to θ_2 and θ_3 are 4 orders of magnitude lower than those of the other parameters due to fact that they are not closely related to glucose dynamics.

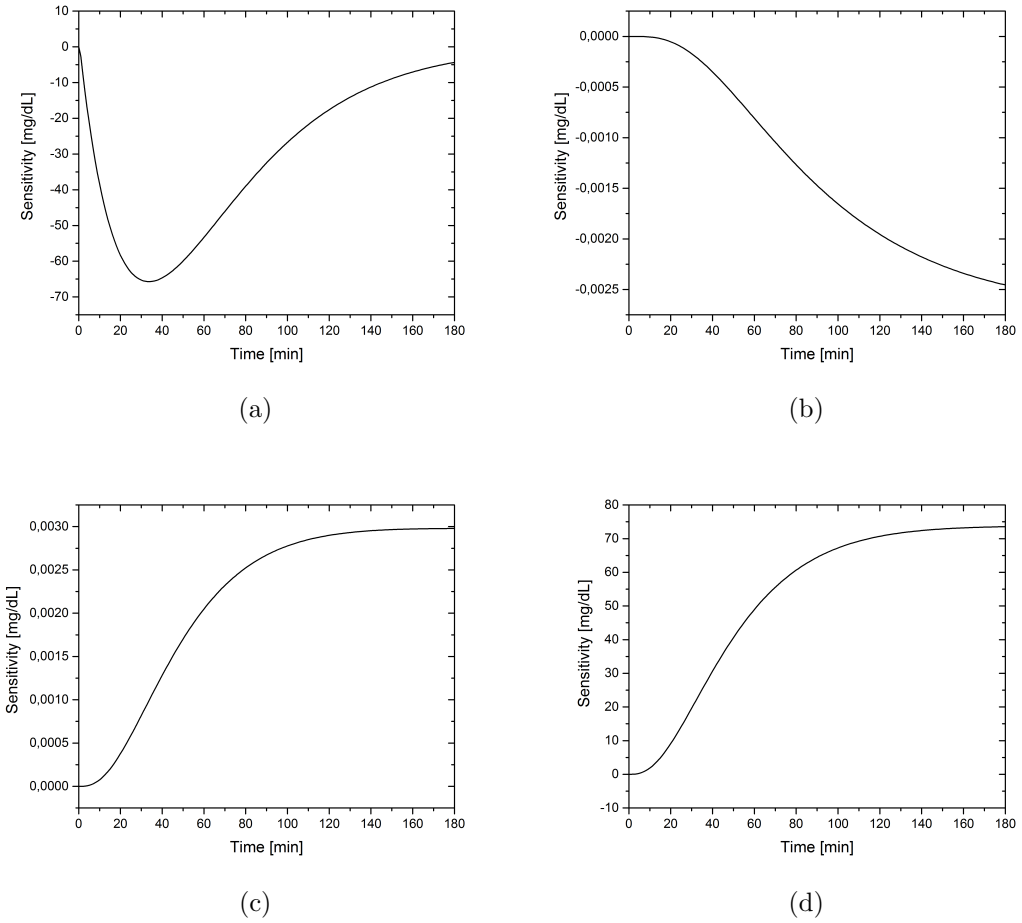


Figure 3.4. *Standard IVGTT sensitivity analysis to a 10% increase in model parameters: θ_1 (a), θ_2 (b), θ_3 (c) and θ_4 (d).*

Figure 3.4 shows glucose sensitivities to model parameters for a standard IVGTT. A peak due to glucose injection is observed around 30 min for θ_1 (a), which is related to glucose effectiveness, and thus is anticipated compared to the OGTT because of the faster response induced by the injection. Similarly to the OGTT, glucose sensitivities to θ_2 and θ_3 (b, c) are 4 orders of magnitude lower than the other parameters. Apart from θ_1, sensitivities to the other parameters present an asymptotic behaviour, which makes it harder to locate the best sampling zone but also indicates that the information may theoretically be extracted over a longer period.

Finally, it is interesting to note that the highest sensitivity values in the standard OGTT, i.e. for θ_1 and θ_3, are more than double of the same sensitivities in the standard IVGTT. This fact can be interpreted as a hint of the higher potentiality of the OGTT regarding the information content derivable.

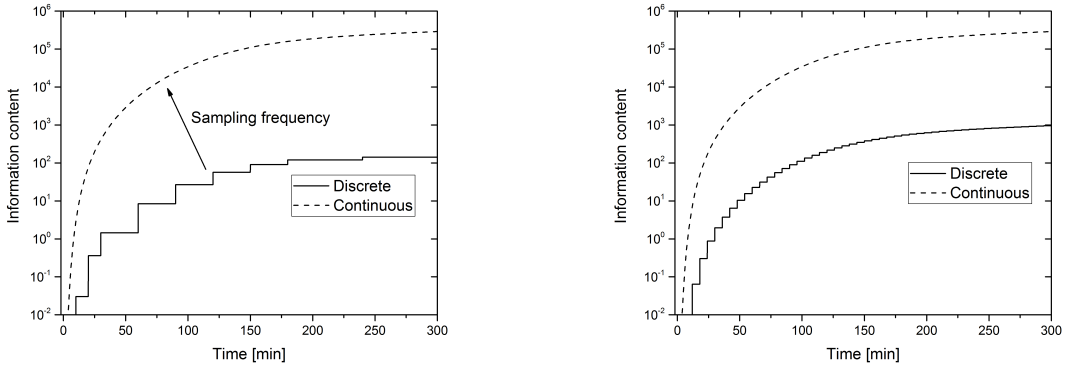
3.3 Information content analysis

The information content theoretically derivable through the standard OGTT and IVGTT has been inquired comparing the continuous and discrete information, which are computed using equations 2.35 and 2.36. Additionally, in order to evaluate the effect of an increase in the number of samples, the effect of the sampling schedule of the standard tests has been compared to 50 time-equispaced samples.
Final data of the cumulative information content from the standard tests with regular and augmented sampling are presented in Table 3.2, while graphical comparisons of the cumulative information content are plotted in Figure 3.5-3.6.

Table 3.2. Cumulative information content of standard OGTT and IVGTT: continuous versus discrete sampling.

Test		Discrete sampling	Continuous sampling
OGTT	11 samples	150.06	$2.89 \cdot 10^5$
	50 samples	963.94	$2.89 \cdot 10^5$
IVGTT	23 samples	125.23	$8.66 \cdot 10^4$
	50 samples	475.93	$8.66 \cdot 10^4$

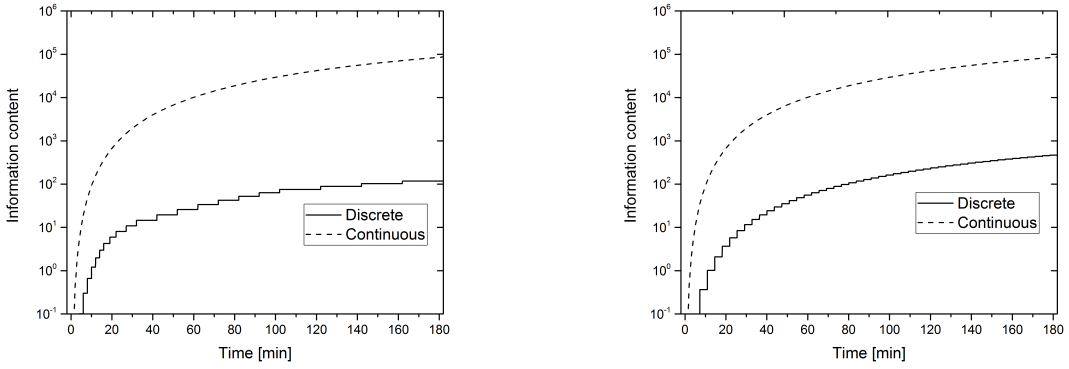
From Table 3.2, it is clear that the OGTT compared to the IVGTT allows for an higher information content to be reached both in the continuous and discrete sampling. The reason for this lies in the different type of stimulus induced on the patient: the OGTT uses a meal which is more physiological than an injection. In addition, the OGTT induces an higher violation of glucose upper bound, that is responsible for a greater information gain during the test.
The information content significantly decreases (3 orders of magnitude) from the continuous to the discrete sampling, while, as it is easily deducible, increasing the sampling frequency a greater discrete information can be achieved. Theoretically, by increasing the number of samples to infinite the information content achieveable reaches the same realised with a continuous sampling strategy.
Figure 3.5 shows that, in the OGTT, the rise in the sampling frequency (b), especially in the first 100 min where the height of the steps is higher with the standard sampling (a), is responsible for the final higher value.



(a) (b)

Figure 3.5. Cumulative information content from the OGTT with standard, i.e. 11 samples (a), and increased sampling, i.e. 50 samples (b), compared to the continuous one.

On the contrary, Figure 3.6 displays how in the IVGTT the overall discrete information content enhancement is less evident than the OGTT due to the better temporal distribution of the samples, which results in a more uniform increase of the information at each step.



(a) (b)

Figure 3.6. Cumulative information content from the IVGTT with standard, i.e. 23 samples (a), and increased sampling, i.e. 50 samples, compared to the continuous one.

Chapter 4

Designed clinical tests

This chapter represents the core of this thesis because it deals with the design of clinical tests for the identification of physiological models complying with some safety and feasibility requirements.
Firstly, the assumptions made in the design phase and in the following parameter estimation task are discussed. Secondly, the inadequacy of the two standard protocols presented in Chapter 3 is shown and the designed clinical tests are individually addressed, both in terms of design and parameter estimation capability. Finally, a comparison is made among the designed tests to assess their overall performance by means of the heuristic indices previously introduced.

4.1 Introduction

In this chapter the design of clinical protocols using a model-based design of experiments approach is tackled. The objective is to devise an optimal test procedure for the identification of model parameters in subjects affected by T1DM.
Based on the Lynch and Bequette model, described in § 2.2, the following tests have been designed and will be discussed:

- 5h-11s oral test: an oral test with one meal, a duration of 5 h and 11 samples taken;
- 5h-22s oral test: similar to the previous one, but with 22 samples collected;
- 5h-22s-2m oral test: an oral test with two meals, a duration of 5 h and 22 samples taken;
- 3h-23s intravenous test: an intravenous test with one glucose injection, a duration of 3 h and 23 samples collected;

- 3h-23s-2i intravenous test: similar to the previous one, but with two glucose injections.

The reason why these tests are presented will be pointed out in the chapter as soon as they are introduced.
Beyond the settings applied for their design and schedule obtained for the execution, the tests will be characterised based on their identification capabilities. For this analysis, two different parametric sets have been chosen: one involving a moderate and the other a severe change with respect to the design parametric set.

4.1.1 Parametric sets

The values of the model parameters previously reported in Table 3.1 refer to a healthy subject. Taking those parameters as a reference, the set has been normalised and from now on it will be indicated as $\mathbf{\Theta_0} = [1, 1, 1, 1]^T$. This set has been considered as the starting point for the design procedure of all the tests presented in this work because the purpose is not just to identify a population of diabetic subjects. On the contrary, the protocol should ideally be able to identify every possible individual, assuming that no a-priori information on the severity of his/her illness is available.
In addition, two different parametric sets are taken into account for assessing the parameter estimation performance of the designed tests. Set $\mathbf{\Theta_1}$ considers a slightly diabetic patient, while set $\mathbf{\Theta_2}$ describes a subject with a more severe condition. Both sets are normalised with respect to the parameters of the healthy subject and reported in Table 4.1.

Table 4.1. Parametric sets used for the parameter estimation task.

Parametric set	Values [min^{-1}]
$\mathbf{\Theta_1}$	$[0.8, 1.2, 0.8, 1.2]^T$
$\mathbf{\Theta_2}$	$[0.6, 1.4, 0.6, 1.4]^T$

Two aspects have been examined for the choice of these two parametric sets:

- sensible deviations from the healthy subject set, so as to reproduce a hypotetical T1DM patient;
- different degrees of deviation, i.e. a moderate and a severe change, in order to compare and contrast the parameter estimation capability of different designed tests.

Typical values of the parameters have been found in literature [9] for lean subjects with a low glucose tolerance. In terms of parameter values, this primarily reflects in

a decrease of the first and third parameter and an increase of the fourth one. In fact, the first parameter represents the rate of blood glucose disappearance into liver or periphery, also known as glucose effectiveness. This capacity is highly compromised in diabetic patients. The third parameter represents insulin appearance velocity in the remote compartment, where it is effective in accelerating glucose disappearance. Typically, the value of this parameter in diabetic patients is significantly lower than healthy subjects due to the deficiency of their glucoregulatory system. The rate of insulin disappearance from the insulin space, represented by the fourth parameter, undergoes a substantial increment, since there is no endogenous insulin production in a T1DM patient. The second parameter has been somehow overestimated in set $\boldsymbol{\Theta_2}$, in order to investigate the identification performance of the designed protocols.

4.1.2 Design assumptions

In all the tests performed, it is assumed to measure blood glucose concentration. The number of samples collected is not subject to optimisation, but only their temporal allocation. A minimum distance of 5 min between measurements is assigned to the oral tests, according to Galvanin et al. [26]. In the intravenous tests, the distance is reduced to 2 min as in the standard IVGTT, because of the shorter duration and different kind of stimulus induced in the subject.
A minimum duration of 5 min is selected for the insulin infusion intervals. However, a higher value is often imposed so as to obtain a protocol easier to be implemented, hence reducing the complexity of execution.
The lower bound is a "hard" constraint, thus its violation is not allowed, while the upper bound is a "soft" constraint, hence small violations are sometimes tolerated. As regards the other end-point constraints, the final glucose concentration needs to be restored within a $\pm$ 5% range of the basal value (i.e. 80 mg/dL). Additionally, the derivative of the final glucose concentration needs to be as close as possible to zero, so as to assure the conservation of the basal value.
Finally, a calculation of the typical insulin requirement per meal is performed, in order to allow for a comparison with the design results. For adult T1DM patients, an insulin value of 0.54 U/kg/d has been found in literature [40], which corresponds to 37.8 U/d for a 70 kg person. It is assumed that the daily calorie need for a person is equivalent to 2000 kcal/d and 55% is introduced as carbohydrates, as recommended by the Food and Agriculture Organization of the United Nations [41]. This is equivalent to an intake of 1100 kcal/d from carbohydrates. Glucose and sugars in general provide 3.75 kcal/g [41], hence approximately 295 g of carbohydrates are daily required:

$$\frac{1100\,\text{kcal/d}}{3.75\,\text{kcal/g}} = 293.3\,\text{g/d} \approx 295\,\text{g/d} \tag{4.1}$$

Based on this value, the amount of insulin required by the patient in a standard OGTT can be computed through the following proportion:

$$37.8\,\mathrm{U/d} : 295\,\mathrm{g/d} = x : 75\,\mathrm{g} \quad \rightarrow \quad x = 9.61\,\mathrm{U} \tag{4.2}$$

Similarly, in a standard IVGTT, which entails a 21 g glucose administration, the estimated insulin needed corresponds to 2.69 U.

4.2 Standard OGTT

This test has already been presented in Chapter 3 and analysed with some preliminary sensitivity studies on the glucose assumption and on the model parameters. It consists of a 5 h test during which 11 blood samples are taken at the times reported in Table 4.2.

Table 4.2. Sampling times for the standard OGTT.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 10, 20, 30, 60, 90, 120, 150, 180, 240, 300]

Insulin is perfused at a constant rate of 16.666 mU/min, as stated by Lynch and Bequette [16].
This protocol has not been obtained through a design of experiments procedure, hence it is not optimised. Therefore, it is interesting to enquire how informative this protocol is.
Figure 4.1 shows the glucose profile of the healthy subject (i.e. set $\boldsymbol{\Theta_0}$) and the one obtained after parameter identification with set $\boldsymbol{\Theta_1}$. It is worth noting the great mismatch between the two curves, especially after the peak of glucose has been reached on. This gap is due to the influence of the parameters on the response that, in the case of a diabetic patient, causes a significant increment of glycemia.
In terms of safety, it is clear how this protocol jeopardises the patient. In fact, despite the upper bound is not a hard constraint, glucose concentration of the subject heavily exceeds this limit for about two hours.
Safety could be improved through an optimisation of the insulin infusion profile. In particular, it can be noticed that the current amount of insulin administered (i.e. 5 U) is far below the reference value computed in Equation 4.2 for such a glucose meal.

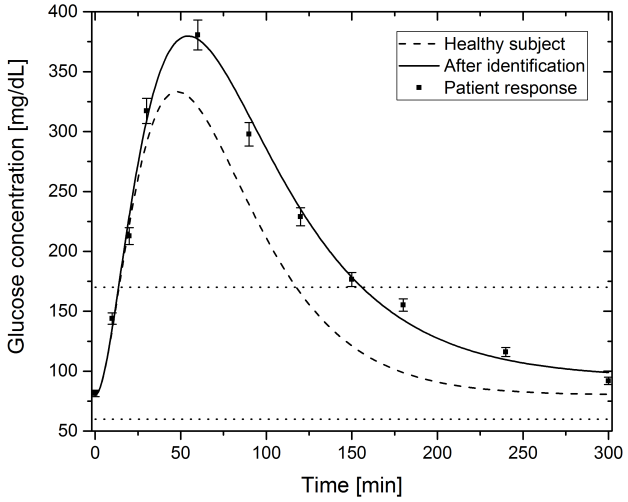


Figure 4.1. Glucose profiles of the standard OGTT for the healthy subject (i.e. with set $\mathbf{\Theta_0}$) and after identification with set $\mathbf{\Theta_1}$.

From Table 4.2, it is evident that the estimated parameters are far from the true values. In effect, this non-optimised protocol provides statistically unsatisfactory estimates, as it is proved by the t-values, which are all well below the reference one.

Table 4.2. Parameter estimation statistics for the standard OGTT with set $\mathbf{\Theta_1}$ (asterisks denote t values failing the t-test).

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.895)
Θ_1	0.7775	1	0.700	1.111*
Θ_2	0.6650	1	6.480	0.103*
Θ_3	3.9925	1	3694	0.001*
Θ_4	1.0212	1	20.13	0.051*

Table 4.3 highlights that also the fitting of the experimental data is inadequate, thus this test cannot be used for identification purposes with the current schedule.

Table 4.3. Lack of fit test for the standard OGTT with set $\mathbf{\Theta_1}$ (asterisk indicates a bad fit).

Weighted residual	χ^2-value (95%)
27.3241*	14.0671

4.3 Standard IVGTT

This test has a duration of 182 min and entails the collection of 23 blood samples with the schedule reported in Table 4.4.

Table 4.4. Sampling times for the standard IVGTT.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[2, 4, 6, 8, 10, 12, 14, 16, 19, 22, 27, 32, 42, 52, 62, 72, 82, 92, 102, 122, 142, 162, 182]

Similarly to the standard OGTT, the insulin infusion rate is constant and equal to 16.666 mU/min. This results in an overall insulin infusion slightly greater (i.e. 3.03 U) than the value estimated for such a glucose assumption.
Figure 4.2 shows the glucose profiles with set $\mathbf{\Theta_0}$ and after identification with set $\mathbf{\Theta_1}$. A sharp spike is observed with a lower maximum than the standard OGTT and a relatively reduced violation of the upper bound.

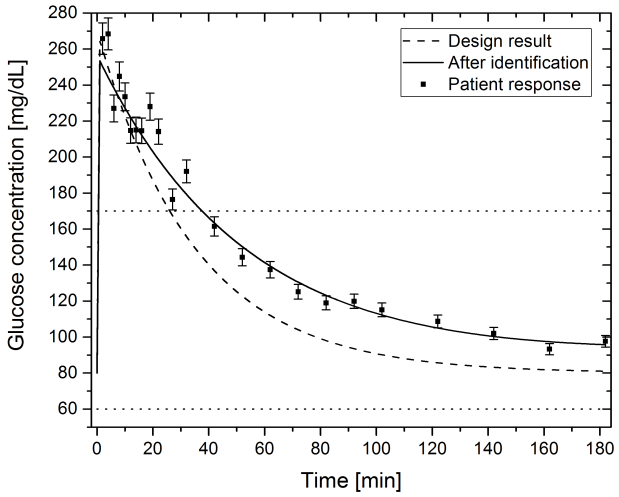


Figure 4.2. Glucose profiles of the standard OGTT for the healthy subject (i.e. with set $\mathbf{\Theta_0}$) and after identification with set $\mathbf{\Theta_1}$.

From an a-priori evaluation, as a consequence of the higher number of samples collected, an enhanced identification capability would be expected compared to the standard OGTT. However, in Table 3.2 it has already been pointed out that the standard OGTT provides a higher information content despite its lower number of samples. This fact reflects on the parameter estimates reported in Table 4.5. Only the first parameter can be estimated satisfactorily, while no estimate is obtained for the second

one. The other two parameters present completely unreliable values, as it is showed by the huge confidence intervals associated with them.

Table 4.5. Parameter estimation statistics for the standard IVGTT with set $\boldsymbol{\Theta_1}$ (asterisks denote t values failing the t-test).

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.725)
Θ_1	0.6336	1	0.111	5.707
Θ_2	0	1	-	-
Θ_3	0.3507	1	143.5	0.002*
Θ_4	1.0559	1	214.06	0.044*

Finally, Table 4.6 shows that a bad fit of the experimental data is still obtained.

Table 4.6. Lack of fit test for the standard IVGTT with set $\boldsymbol{\Theta_1}$ (asterisk indicates a bad fit).

Weighted residual	χ^2-value (95%)
48.4254*	31.4104

4.4 5h-11s oral test

This first designed protocol aims at optimising the standard OGTT applying a MBDoE approach. It maintains the same duration and number of samples, trying to extract as much information as possible while controlling the insulin infusion rate.

4.4.1 Experiment design

The sampling times adopted for the standard OGTT are not the result of an optimisation, hence they do not allow to maximise the information content achievable during the test. Therefore, it makes sense to apply the design of experiments to devise the best temporal allocation of the samples to be collected, using insulin infusion as the control variable. Table 4.7 reports the settings imposed in the design phase to the manipulated variable.

Table 4.7. Controlled variable settings for the 5h-11s oral test design of experiment.

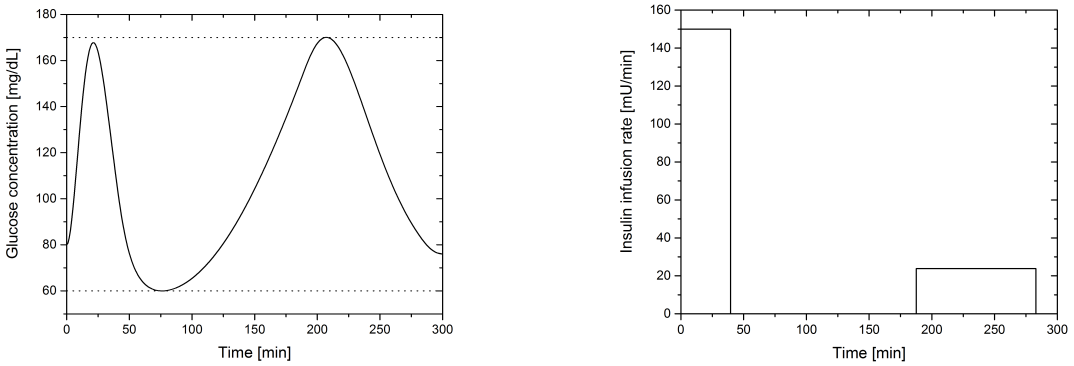
Insulin settings	Value
Control type	piecewise-constant
Intervals number	4
Interval lower bound [min]	5
Upper bound [mU/min]	[150,115,115,115]

Furthermore, the test is required to be safe for the subject, thus the end-point constraints displayed in Table 4.8 have been enforced in order to ensure that kind of design. On the one hand, the constraints on the upper and lower bound violations aim at assuring harmless conditions during the execution. On the other hand, the constraints on the final glucose concentration and its derivative force to restore the basal values of the patient prior to the test.

Table 4.8. End-point constraints for the 5h-11s oral test design of experiments.

Constrained variable	Lower bound	Upper bound
Lower bound violation [mg · min/dL]	-0.1	0.1
Upper bound violation [mg · min/dL]	-0.1	0.1
G_{tot} [mg/dL]	76.0	84.0
Γ [mg/(dL · min)]	-0.01	0.01

The result of the MBDoE procedure is shown in Figure 4.3, in terms of glucose profile obtained (a) and insulin infusion schedule (b). Firstly, it is clear how the design allows to enclose the glucose profile between the lower and upper bound, thus maximising the safety of the patient, at least with this parametric set. However, it could endanger a different subject, i.e. having a highly different parametric set. Secondly, a simple insulin administration profile is obtained, since insulin needs to be perfused only in two of the four control intervals.



(a) (b)

Figure 4.3. 5h-11s oral test: glucose profile from the MBDoE procedure (a) and insulin administration profile (b).

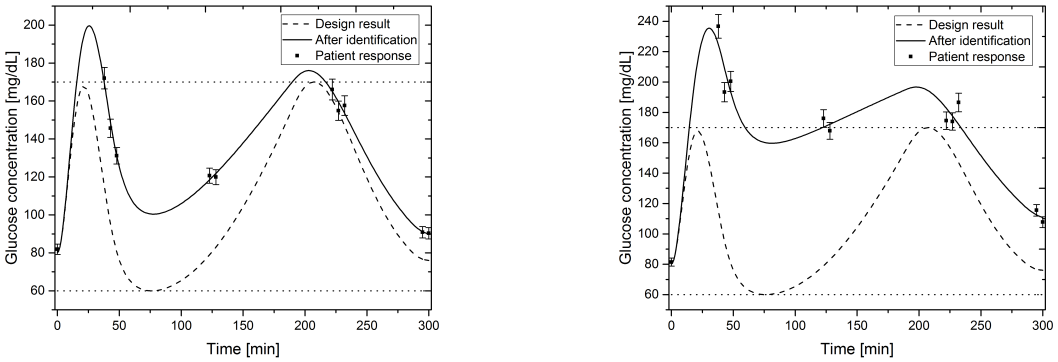
Table 4.9 collects the detailed results of the optimisation. The first thing to notice is the information content improvement accomplished thanks to the optimisation, since index I_d increases from 150 to 268.
Unfortunately, the main issue of this test is related to the excessive insulin infusion rate required in its initial part. This value is beyond those reported in literature for the available insulin infusion devices. Therefore, this test is only meant to show the enhancements permitted by a MBDoE approach over the standard OGTT.

Table 4.9. Optimised design variables and discrete information content for the 5h-11s oral test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 38, 43, 48, 123, 128, 222, 227, 232, 295, 300]
U [mU/min]	[149.99, 0, 23.82, 0]
t_U [min]	[39.52, 148.01, 95.46, 17.01]
I_d	268.00

4.4.2 Parameter estimation

Parameter identification has been performed using sets $\mathbf{\Theta_1}$ and $\mathbf{\Theta_2}$. Looking at Figure 4.4, a significant mismatch is evident between the designed and the identified profiles, which increases as the gap in the corresponding parametric sets rises. This behaviour, on the one hand, moves the subject away from the lower bound but, on the other hand, causes a modest violation of the upper bound.



(a) (b)

Figure 4.4. Glucose profiles from the MBDoE procedure and after identification with set $\boldsymbol{\Theta_1}$ (a) and $\boldsymbol{\Theta_2}$ (b) for the 5h-11s oral test.

The estimates of the parameters are statistically satisfactory for both the sets used, as shown in Table 4.10 and 4.11. This means that the optimisation strategy adopted is actually effective in devising a protocol capable of increasing the information content achievable.

Table 4.10. Parameter estimation statistics for the 5h-11s oral test with set $\boldsymbol{\Theta_1}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.895)
Θ_1	0.6469	1	0.1836	3.524
Θ_2	1.3089	1	0.1799	7.274
Θ_3	0.7439	1	0.0637	11.68
Θ_4	1.0630	1	0.1703	6.242

Table 4.11. Parameter estimation statistics for the 5h-11s oral test with set $\boldsymbol{\Theta_2}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.895)
Θ_1	0.4716	1	0.1814	2.600
Θ_2	1.4866	1	0.2784	5.340
Θ_3	0.5517	1	0.0860	6.419
Θ_4	1.1534	1	0.3461	3.332

4.5 5h-22s oral test

This test attempts to maintain the positive aspects observed in the 5h-11s oral test, such as the good estimation capabilities, while getting rid of the unfeasible insulin

infusion rate used in that test.

4.5.1 Experiment design

In this optimisation, the number of samples is risen to 22, so as to furtherly increase the information derivable. This is a sensible value because it is close to the number of samples taken in the standard IVGTT, but distributed over a longer time period.
Similarly to the 5h-11s oral test, only insulin infusion is considered as control variable in the design. However, the maximum on the upper bound is fixed at 100 mU/min, according to Lynch and Bequette [16], and the control intervals reduced to 3, as shown in Table 4.12.

Table 4.12. Controlled variable settings for the 5h-22s oral test design of experiment.

Insulin settings	Value
Control type	piecewise-constant
Intervals number	3
Interval lower bound [min]	30
Upper bound [mU/min]	[100, 90, 90]

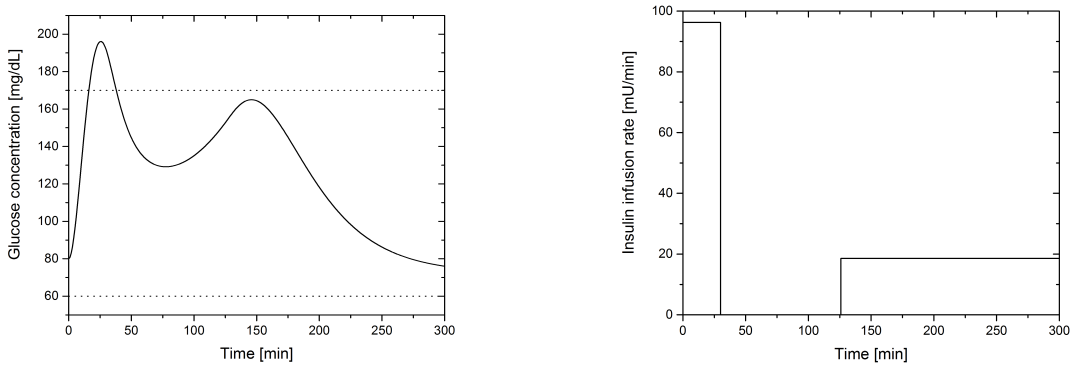
The end-point constraints are partially relaxed. In fact, due to the lower insulin infusion rate imposed, a small violation of the upper bound needs to be permitted (Table 4.13), for the optimisation to converge. Additionally, the constraint on the derivative of the final glucose concentration is loosened, so as to keep the duration of the test unchanged.

Table 4.13. End-point constraints for the 5h-22s oral test design of experiment.

Constrained variable	Lower bound	Upper bound
Lower bound violation [mg · min/dL]	-0.1	0.1
Upper bound violation [mg · min/dL]	-0.1	500
G_{tot}[mg/dL]	76.0	84.0
Γ [mg/(dL · min)]	-0.1	0.1

The glucose and insulin profiles obtained applying the MBDoE procedure are shown in Figure 4.4. It can be noticed that allowing for a small violation of the upper bound prevents glucose concentration from hitting the lower bound. Insulin infusion profile is

even simpler than the 5h-11s oral test and also the amount of insulin perfused decreases from 8.2 U to 6.1 U.



(a) (b)

Figure 4.4. 5h-22s oral test: glucose profile from the MBDoE procedure (a) and insulin administration profile (b).

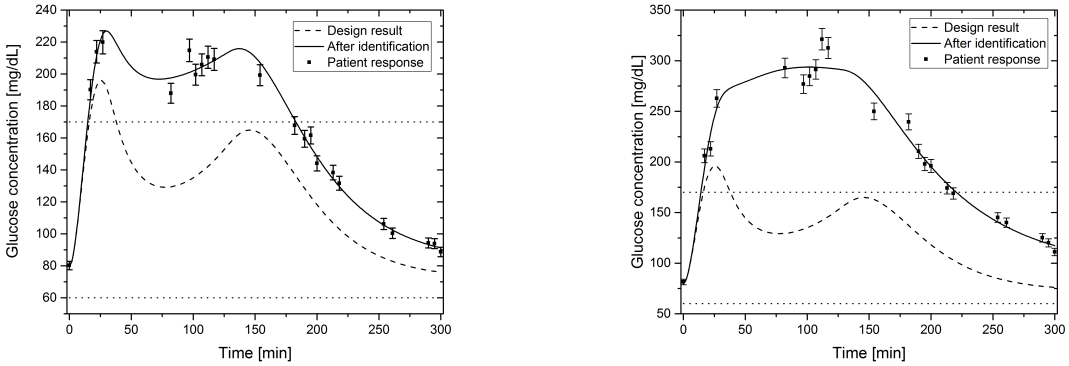
The settings obtained for the test, as the sampling times, control variable values and switching times are displayed in Table 4.14. Furthermore, the discrete information content exhibits a huge increment as a consequence of the rise in samples.

Table 4.14. Optimised design variables and discrete information content for the 5h-22s oral test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 17, 22, 27, 82, 97, 102, 107, 112, 117, 154, 182, 190, 195, 200, 213, 218, 254, 261, 290, 295, 300]
U [mU/min]	[96.31, 0, 18.59]
t_U [min]	[30, 95.83, 174.17]
I_d	455.57

4.5.2 Parameter estimation

The effect of the parametric mismatch is relevant (Figure 4.5), especially for set $\mathbf{\Theta_2}$ (b), which shows a remarkable violation of the upper bound. This represents a potentially harmful condition for the patient, hence it needs to be handled if the test is to be applied on a clinical scale.
Additionally, it should be observed that in both cases it is difficult to obtain a stable end-point glucose concentration, i.e. a zero derivative.



(a) (b)

Figure 4.5. Glucose profiles from the MBDoE procedure and after parameter identification with set $\boldsymbol{\Theta_1}$ (a) and $\boldsymbol{\Theta_2}$ (b) for the 5h-22s oral test.

The effect of the samples rise is numerically evident also by the estimates achieved, presented in Table 4.15 and 4.16. In fact, the t-values associated with the estimated parameters are greater than those reported for the 5h-11s oral test for both sets. This means higher precision in the identification, as it is also proved by the reduction of the confidence intervals.

Table 4.15. Parameter estimation statistics for the 5h-22s oral test with set $\boldsymbol{\Theta_1}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.734)
Θ_1	0.7738	1	0.0964	8.031
Θ_2	1.2316	1	0.1680	10.05
Θ_3	0.8117	1	0.1360	5.969
Θ_4	1.1566	1	0.0877	13.19

Table 4.16. Parameter estimation statistics for the 5h-22s oral test with set $\boldsymbol{\Theta_2}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.734)
Θ_1	0.6038	1	0.1024	5.897
Θ_2	1.4844	1	0.2348	6.322
Θ_3	0.6533	1	0.1935	3.376
Θ_4	1.4164	1	0.2156	6.569

4.6 5h-22s-2m oral test

The objective of this design is to avoid or at least reduce the violation of the glucose upper bound, so as to increase the safety of the test. This can be accomplished, for instance, by administering multiple glucose meals with a reduced amount of glucose in each one.

4.6.1 Experiment design

The design of this protocol departs from the positive results obtained from the oral tests previously devised, in terms of parameter estimation performance. The number of samples taken is maintained at 22, so as to possibly reproduce the high information content previously achieved. The oral administration of glucose is kept as well, but it is splitted into two different meals of reduced quantity, which are to be optimised. Therefore, in addition to the insulin infusion rate, two other control variables can be adjusted: the amount of glucose of each meal and the meal times. The amount of glucose per meal is enclosed between 10 and 30 g, while the meal times are chosen so as the first one occurs at t=0 min and the second one can move in the range from 100 to 150 min. Differently from insulin infusion, these two variables are time-invariant. The settings used for the manipulated variables in the design are reported in Table 4.17.

Table 4.17. Controlled variable settings for the 5h-22s-2m oral test design of experiment.

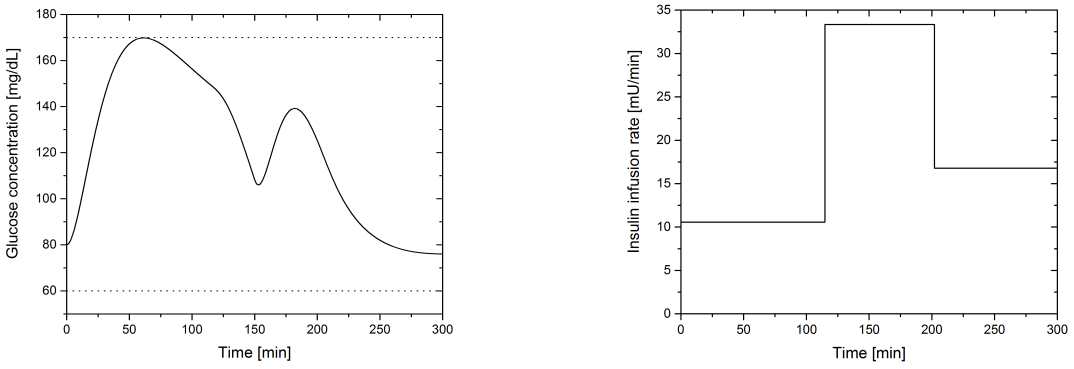
	Insulin infusion	Glucose amount	Meal times
Control type	piecewise-constant	time-invariant	time-invariant
Intervals number	3	1	1
Interval lower bound	20 min	-	-
Lower-Upper bounds	0-50 mU/min	10-30 g	[0-1, 100-150] min

Since the aim is to improve the safety of the protocol, stricter limits have been imposed on the design not tolerating any violation at all (Table 4.18).

Table 4.18. End-point constraints for the 5h-22s-2m oral test design of experiments.

Constrained variable	Lower bound	Upper bound
Lower bound violation [mg · min/dL]	-0.01	0.01
Upper bound violation [mg · min/dL]	-0.01	0.01
G_{tot} [mg/dL]	76.0	84.0
Γ [mg/(dL · min)]	-0.01	0.01

The design result, shown in Figure 4.6, allows to avoid any violation of the upper bound and, at the same time, to assure a higher stabilisation of the final glucose concentration.



(a) (b)

Figure 4.6. 5h-22s-2m oral test: glucose profile from the MBDoE procedure (a) and insulin administration profile (b).

The two peaks that are observed are due to the two meals, which also produce a different type of insulin infusion profile from those observed in the previous tests. In fact, the maximum insulin infusion rate is located in the central interval because of the second meal.

Table 4.19 exhibits the detailed results of the optimisation. It can be observed that the second meal has the effect of concentrating the samples in the last 100 min of the test. The overall amount of glucose administered to the patient has been reduced to 51 g, even though the total amount of insulin perfused has been scarcely decreased to 5.8 U.

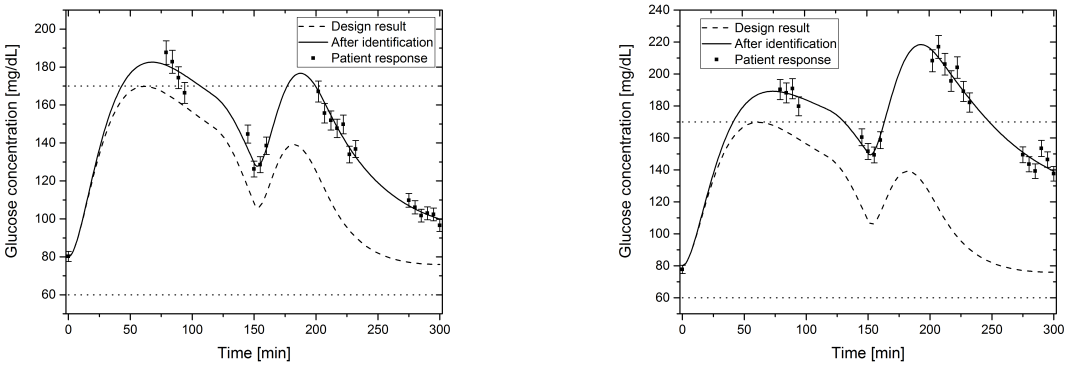
The discrete information content derived at the design conditions is only slightly reduced compared to the 5h-22s oral test, hence the parameter identification should still be satisfactory.

Table 4.19. Optimised design variables and discrete information content for the 5h-22s-2m oral test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 79, 84, 89, 94, 145, 150, 155, 160, 202, 207, 212, 217, 222, 227, 232, 275, 280, 285, 290, 295, 300]
U [mU/min]	[10.57, 33.34, 16.78]
t_U [min]	[114.87, 87.20, 97.94]
Meals [g]	[21.09, 30.00]
t_{meals} [min]	[0, 150]
I_d	431.56

4.6.2 Parameter estimation

From Figure 4.7, the result of the identification with set $\mathbf{\Theta_1}$ (a) by means of the 5h-22s-2m oral test appears to be quite good in terms of safety. A small violation of the upper bound is observed and glucose final concentration is not too far from its basal value. Unfortunately, the identification performed with set $\mathbf{\Theta_2}$ (b), emphasises the limitation of this protocol. In fact, the final glucose concentration is too distant from the basal value and the final derivative is not stable yet, thus making the test unsuitable for clinical applications.



(a) (b)

Figure 4.7. Glucose profiles from the MBDoE procedure and after identification with set $\mathbf{\Theta_1}$ (a) and $\mathbf{\Theta_2}$ (b) for the 5h-22s-2m oral test.

Parameter estimates are satisfactory with both sets (Table 4.20 and 4.21), even though not as good as those realised in the 5h-22s oral test, as it can be observed comparing the t-values.

Table 4.20. Parameter estimation statistics for the 5h-22s-2m oral test with set $\boldsymbol{\Theta_1}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.734)
Θ_1	0.8067	1	0.1744	4.626
Θ_2	1.0835	1	0.2929	3.699
Θ_3	0.7431	1	0.1169	6.354
Θ_4	1.1978	1	0.1576	7.600

Table 4.21. Parameter estimation statistics for the 5h-22s-2m oral test with set $\boldsymbol{\Theta_2}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.734)
Θ_1	0.8856	1	0.2417	3.665
Θ_2	1.0343	1	0.3032	3.411
Θ_3	0.6769	1	0.2008	3.371
Θ_4	1.9089	1	0.3921	4.868

4.7 3h-23s intravenous test

Similarly to the optimisation of the standard OGTT performed in the 5h-11s oral test, this design aims at detecting the performance of an optimised standard IVGTT.

4.7.1 Experiment design

Considering the same duration and number of samples of the standard IVGTT, a MBDoE approach is applied in order to identify the optimal sampling schedule. The control variable for the optimisation is still the insulin infusion, which can be modified over three time intervals (Table 4.22).
Since safety is a major concern, it is chosen to inject only two-thirds of the glucose used in the standard IVGTT, i.e. 14 g. As a consequence, the glucose injection rate reduces to 125 mg/dL/min. This is supposed to induce a considerable reduction in the initial glucose peak, thus reducing the risk of an upper bound violation. However, since the violation cannot be avoided completely, a small tolerance is allowed in order for convergence to be reached, as shown in Table 4.23.

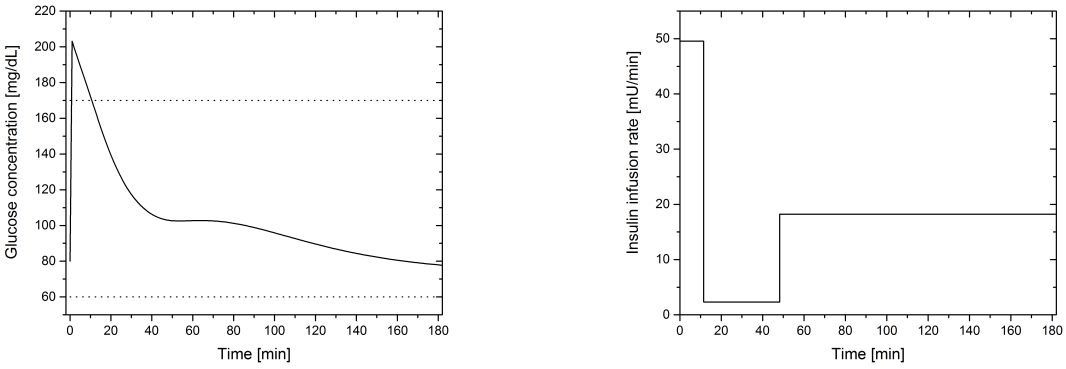
Table 4.22. Controlled variable settings for the 3h-23s intravenous test design of experiment.

Insulin settings	Value
Control type	piecewise-constant
Intervals number	3
Interval lower bound [min]	10
Lower-Upper bounds [mU/min]	0-50

Table 4.23. End-point constraints for the 3h-23s intravenous test design of experiment.

Constrained variable	Lower bound	Upper bound
Lower bound violation [mg · min/dL]	-0.01	0.01
Upper bound violation [mg · min/dL]	-0.01	200
G_{tot} [mg/dL]	76.0	84.0
Γ [mg/(dL · min)]	-0.1	0.1

The reduction in the glucose injection results in a sharp decrease of the initial spike in the glucose concentration profile (Figure 4.8). Nevertheless, the upper bound violation is still present, even though negligible. The advantage of this protocol consists in the good stabilisation of the final glucose concentration profile permitted by the short stimulus induced on the subject.



(a) (b)

Figure 4.8. 3h-23s intravenous test: glucose profile from the MBDoE procedure and after parameter identification with set $\boldsymbol{\Theta_1}$ (a) and insulin administration profile (b).

As a result of the optimisation the time of the initial samples is postponed and there is a higher concentration of samples in the final part of the test (Table 4.24).

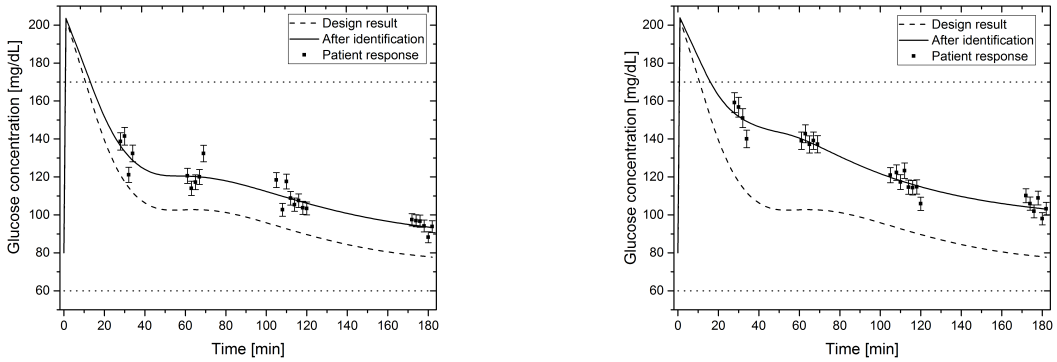
In addition, there is a good increase in the discrete information content derived with respect to the standard IVGTT, which should result in an improved parameter estimation capability. However, it can be noticed that the information achievable with this optimised protocol remains below the one obtainable from the 5h-11s oral test. This fact confirms what was previously observed in the preliminary analysis about the higher information content of the standard OGTT over the standard IVGTT.

Table 4.24. Optimised design variables and discrete information content for the 3h-23s intravenous test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[28, 30, 32, 34, 61, 63, 65, 67, 69, 105, 108, 110, 112 114, 116, 118, 120, 172, 174, 176, 178, 180, 182]
U [mU/min]	[49.57, 2.31, 18.23]
t_U [min]	[11.45, 36.83, 133.72]
I_d	208.08

4.7.2 Parameter estimation

Figure 4.9 shows that the parametric mismatch still causes a significant gap between the designed test on the healthy subject and the profile after identification with diabetic sets. Nevertheless, due to the different type of stimulus induced by the glucose injection, the violation of the upper bound remains quite limited regardless of the parametric set.



(a) (b)

Figure 4.9. Glucose profiles from the MBDoE procedure and after identification with set $\mathbf{\Theta_1}$ (a) and $\mathbf{\Theta_2}$ (b) for the 3h-23s intravenous test.

The designed test exhibits an acceptable parameter estimation capability with set

$\mathbf{\Theta_1}$, although the second parameter is on the border line of acceptability (Table 4.25). The limitations of this protocol clearly emerge when the parameter estimation with set $\mathbf{\Theta_2}$ is performed and statistically unsatisfactory estimates are obtained for the second and third parameters (Table 4.26). Additionally, also the fourth parameter presents too much a large confidence interval value, which makes the identification highly unreliable.

Table 4.25. Parameter estimation statistics for the 3h-23s intravenous test with set $\mathbf{\Theta_1}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.729)
Θ_1	0.7886	1	0.1149	6.864
Θ_2	1.1464	1	0.6457	1.775
Θ_3	0.8355	1	0.2726	3.065
Θ_4	1.2037	1	0.1629	7.391

Table 4.26. Parameter estimation statistics for the 3h-23s intravenous test with set $\mathbf{\Theta_2}$ (asterisks denote t-values failing the t-test).

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.729)
Θ_1	0.6557	1	0.1713	3.827
Θ_2	3.6051	1	5.344	0.675*
Θ_3	1.0708	1	1.7	0.630*
Θ_4	1.8098	1	0.7203	2.513

4.8 3h-23s-2i intravenous test

The purpose of this design is to improve safety and identification performance of the 3h-23s intravenous test using two glucose injections.

4.8.1 Experiment design

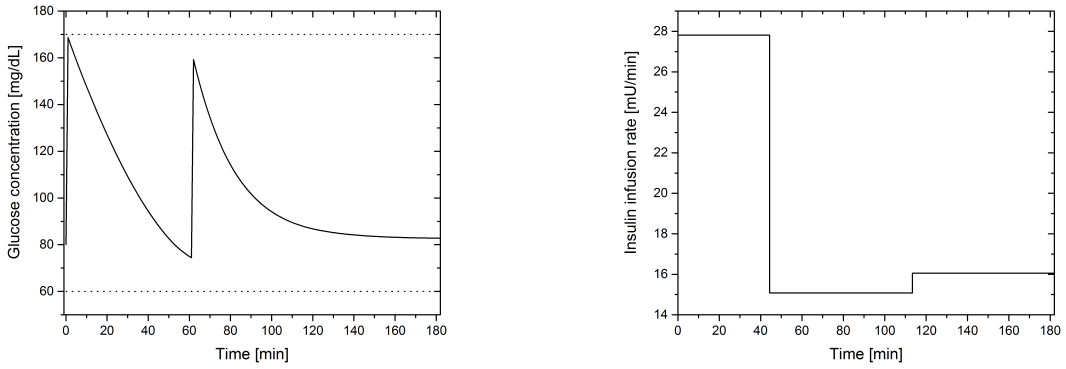
The optimisation of the 23 samples is carried out adjusting three control variables: the insulin infusion rate, the amount of glucose injected and the injection times.
Insulin infusion is manipulated over three intervals and narrower bounds, since less glucose per injection is assumed by the patient. Manipulation of the amount of glucose injected, along with no upper bound violation tolerated, allows to enclose the glucose profile between the two bounds. The first glucose injection is forced to occur in the first minute of the test, while the second one can move in the range from 40 to 100 min,

as shown in Table 4.27.
As regards the other end-point constraints imposed on the design, they are the same used in the 5h-22s-2m oral test (Table 4.18).

Table 4.27. Controlled variable settings for the 3h-23s-2i intravenous test design of experiment.

	Insulin infusion	Glucose injection	Injection times
Control type	piecewise-constant	time-invariant	time-invariant
Intervals number	3	1	1
Interval lower bound	10 min	-	-
Lower-Upper bounds	0-40 mU/min	50-120 mg/(dL·min)	[1-1, 40-100] min

Figure 4.10 confirms that, similarly to the 3h-23s intravenous test, the type of stimulus generated by an injection permits to stabilise the glucose profile and restore the basal value in a simple way.
The design procedure yields two glucose injections, approximately of the same amount, at a distance of an hour (Table 4.28).
At the same time, it can be noticed that the increment realised in the discrete information content compared to the previously designed intravenous test is not so significant as expected.



(a) (b)

Figure 4.10. 3h-23s-2i intravenous test: glucose profile from the MBDoE procedure (a) and insulin administration profile (b).

Table 4.28. Optimised design variables and discrete information content for the 3h-23s-2i intravenous test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[24, 26, 28, 30, 39, 59, 61, 74, 101, 103, 105, 107, 109, 111, 113, 115, 118, 172, 174, 176, 178, 180, 182]
Glucose injections [mg/(dL·min)]	[90.52, 87.52]
$t_{injections}$ [min]	[0, 61]
U [mU/min]	[27.81, 15.08, 16.06]
t_U [min]	[44.35, 69.06, 68.59]
I_d	269.51

4.8.2 Parameter estimation

Figure 4.10 displays the good safety achieved with this test in the identification of different diabetic subjects, as it is proved by the negligible violations of the upper bound observed. End-point glucose concentration is not very close to the basal value, but, quite a flat end-profile is obtained.

Nevertheless, the poor identification performance of this new protocol can already be perceived by looking at the fact that there is only slight improvement in the t-values after identification with set $\mathbf{\Theta_1}$ (Table 4.29) compared to the 3h-23s intravenous test. In fact, from Table 4.30, it can be seen that this test fails in the estimation of the second and third parameters with set $\mathbf{\Theta_2}$.

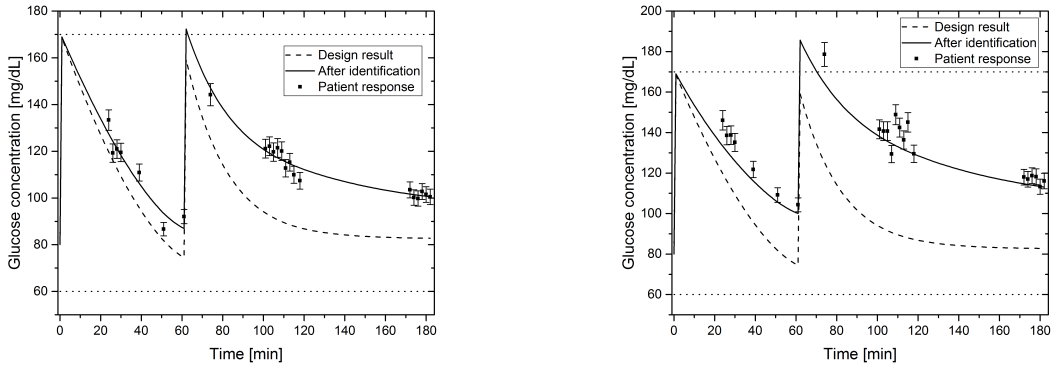


Figure 4.11. Glucose profiles from the MBDoE procedure and after identification with set $\mathbf{\Theta_1}$ (a) and $\mathbf{\Theta_2}$ (b) for the 3h-23s-2i intravenous test.

Table 4.29. Parameter estimation statistics for the 3h-23s-2i intravenous test with set $\boldsymbol{\Theta_1}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.729)
Θ_1	0.8080	1	0.1889	4.278
Θ_2	1.5252	1	0.6223	2.451
Θ_3	1.1467	1	0.4516	2.539
Θ_4	1.1796	1	0.1390	8.486

Table 4.30. Parameter estimation statistics for the 3h-23s-2i intravenous test with set $\boldsymbol{\Theta_2}$ (asterisks denote t-values failing the t-test).

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.729)
Θ_1	0.7275	1	0.2684	2.710
Θ_2	1.3618	1	2.1570	0.631*
Θ_3	0.7910	1	1.2280	0.644*
Θ_4	1.3981	1	0.4021	3.477

4.9 Evaluation of the designed protocols

In this section, some considerations on the information content derived at different conditions are pointed out and an overall assessment of the designed protocols is made.

4.9.1 Parametric bias on the information content

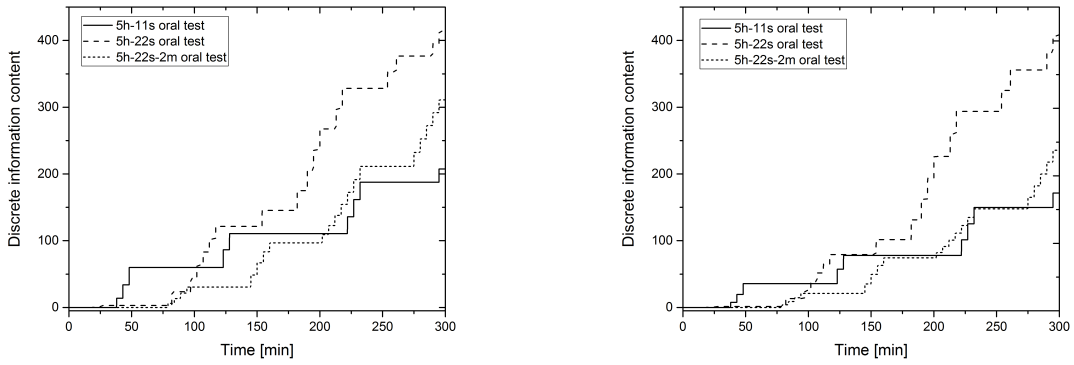
The glucose profiles previously examined for different parametric sets have shown that the model parameters have a strong influence on the glucose response. To put it another way, it means that each patient has his/her own parameters, which are initially unknown and produce a different response to the same stimulus. The information content derivable in a test is highly dependent on the parametric set of the subject considered, hence it is not sensible to compare different tests based on distinct parametric sets.

Therefore, it is chosen to compare different tests only at the same parametric set. However, an issue could be raised regarding the fact that the choice of the specific set on which to base the comparison could affect the corresponding result. With respect to this, an empirical analysis has shown that for each test the following equality holds,

regardless of the parametric set considered:

$$log\left(\frac{I_d}{I_c}\right) = constant \tag{4.3}$$

where I_d and I_c represent the discrete and continuous information content realised with the same parametric set, respectively. The use of the logarithm is due to the two/three orders of magnitude difference between the discrete and continuous information.
In practical terms, Equation 4.3 means that the fraction of discrete information derived over the continuous information potentially achievable is the same for a specific test for every parametric set examined. This implies that if a test is found to be more informative than another for a certain parametric set then, also considering other parametric sets, it will remain so. This can be visualised from the cumulative information content graphs reported in Figure 4.12 for sets $\boldsymbol{\Theta_1}$ and $\boldsymbol{\Theta_2}$. The 5h-22s oral test is shown to be the most informative for both sets and the same hierarchy is maintained also for the other two tests.



(a) (b)

Figure 4.12. Comparison of the cumulative information content derivable from three designed protocols considering set $\boldsymbol{\Theta_1}$ (a) and $\boldsymbol{\Theta_2}$ (b).

4.9.2 Overall assessment

Beyond the capability to identify the model parameters in a statistically sound way, the purpose of this work is to provide some criteria based on which to evaluate the feasibility of the designed tests. In fact, the optimally devised protocol should be able to identify an unknown subject while complying with the safety requirements needed for its clinical implementation.
A novel approach is hereby introduced, which takes into account five distinct indices, whose mathematical formulation has already been presented in § 2.8. These heuristic indices allow to draw a comparison between different tests considering the factors that

could impact the most on their applicability.

In light of the previous considerations, it is chosen to compare the tests on the basis of set $\mathbf{\Theta_2}$. Computation of the indices is performed considering safety and information realised after the identification with that set. Therefore, only the tests that allow a statistically satisfactory parameter estimation are taken into account, so as to avoid deceiving results because of the great parametric mismatch. As a result of this, only three designed protocols are displayed in Figure 4.13. The radar chart is a useful tool

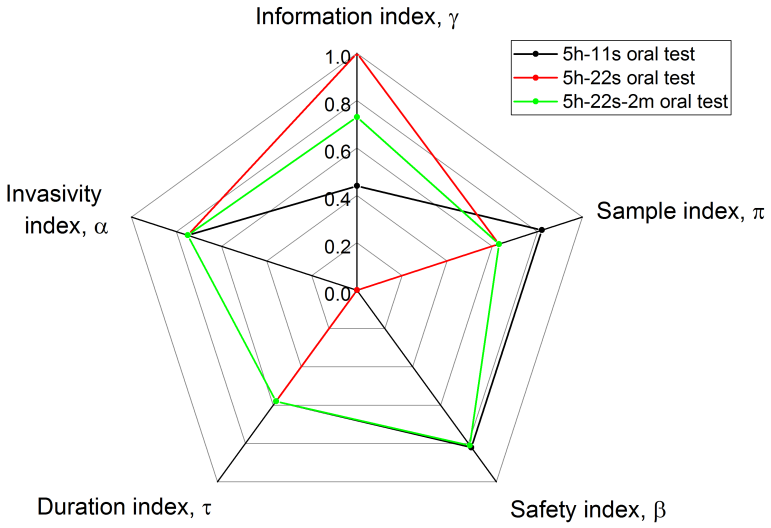


Figure 4.13. Indices radar chart for the designed protocols that allow for a statistically satisfactory parameter estimation with set $\mathbf{\Theta_2}$.

for a quick evaluation of the most suitable clinical protocol based on the different indices introduced. In fact, the larger the area of the polygon, the greater the suitability of the designed test.

Firstly, the oral tests have been shown to outperform the intravenous ones for identification purposes, because they produce a more physiological type of stimulus on the subject. As a matter of fact, only oral tests are represented in the radar chart.

Secondly, it can be easily noticed the rise in the information content that can be accomplished through an increase in the number of samples, i.e. a decrease in π. Nevertheless, this fact has a direct impact on the safety of the protocol: the 5h-22s oral test presents the highest information index but, at the same time, its safety reaches a minimum.

The 5h-22s-2m represents a good compromise between high information content and good safety, however, it has been shown to be difficult to restore basal glucose values at the end of the test (Figure 4.7b) due to the second meal.

It is clear that none of these protocols completely meets the requirements needed for a clinical application, even though valuable information has been derived. The major

issue associated with these tests is represented by the rapid loss of their safety due to the parametric mismatch. As a final consideration, some additional improvements on safety are necessary, in order to guarantee a feasible implementation of the most informative protocols.

Chapter 5

Robust design

This chapter deals with the design of a test that ensures the compliance of safety requirements over a wide range of variability of the model parameters.
Firstly, the strategy adopted for the implementation of a robust design is presented.
Secondly, the effectiveness of such an approach in the design of a new protocol is shown, in terms of safety enhancement and parameter estimation performance.

5.1 Introduction

In previous analyses, it has been shown that subject's responses are highly dependent on the parameter values. In other words, this means that a great variability in the outputs is observed moving in the domain of uncertainty of the model parameters. The main drawback associated with this fact consists in the large deviations of the glucose profile from the design result, which can cause significant violations of the upper bound.
In order to prevent the uncertainty of model parameters from leading to unsafe operating conditions, robust design of experiment techniques may be implemented. A successful application of this approach allows to design tests which are insensitive to the starting value of the parameters. The method consists in the use of stochastic simulations, which reproduce the physiological behaviour of a large group of diabetic patients. The results are then applied to devise tests with stricter safety requirements, so as to minimise the effect of parameter uncertainty when performed on unknown subjects.

5.2 Variability assessment

The effect of the uncertainty of the model parameters on the glucose output is tested on a population of 500 individuals affected by T1DM, adopting the 5 h-22 s oral test.

The set of parameters of each of them is assigned randomly using a uniform distribution within a certain range of variability. A uniform distribution is chosen because each parameter is equally likely to be selected for the simulated patient. With respect to the normalised set of parameters considered in the design stage for the healthy subject, i.e. $\mathbf{\Theta_0}$, the variability range assigned in the uniform distribution is the following:

$$[0.43 - 0.73, 1.03 - 1.20, 0.21 - 0.40, 1.40 - 1.60]$$

This choice has been made according to data reported in literature by Bergman [9] for lean diabetic subjects.
Glucose profiles resulting from the 500 simulated patients are analysed by means of some statistics, so as to assess their gap from the upper threshold. In particular, an expected value approach is chosen, which considers the a priori uncertainty in the model parameters by assuming that the parametric set to be used belongs to a population of subjects, whose distribution is known, i.e. a uniform distribution over the range specified above [42]. This method consists in selecting the maximum glucose value of each profile and computing the mean and standard deviation over the entire population taken into account.
Standard deviation (SD) is then used for designing robust tests, in which the upper bound violation constraint is reformulated as:

$$\frac{dV_1}{dt} = max(0, G_{tot}(t) - UB + 2 \cdot SD) \tag{5.1}$$

A value of 19.53 for the standard deviation has been found but it has been chosen to round it off to 20, so that the upper bound in the new design lowers from 170 to 130 mg/dL.

5.3 Robust oral test

The 5h-22s oral test has been detected the most informative among the several designs presented in Chapter 4. Therefore, it is assumed as a reference for the development of a clinical protocol which combines a satifactory identification capability with an excellent safety.

5.3.1 Experiment design

An initial design is performed controlling both insulin infusion and the amount of glucose of the meal, so as to be able to enforce the stricter constraint on the upper threshold (Equation 5.1). Insulin infusion is controlled as for the 5h-22s oral test (Ta-

ble 4.12), while the glucose meal can move in the range from 30 to 60 g. The end-point constraints imposed on the design are similar to the 5h-22s-2m oral test (Table 4.18), except for the glucose derivative constraint, which is relaxed in the range from -0.1 to 0.1. Relaxation of this constraint is required because otherwise the good excitation pattern previously obtained is lost and, as a result, the parameter identification performance is dramatically reduced.
A value close to 40 g is found for the meal from this preliminary design, so it is rounded off to 40. Then, a new design is carried out fixing the meal at 40 g and controlling only the insulin infusion with the same settings used before.
This designed protocol will be referred to as *robust oral test*. The results of this test, i.e. sampling times and insulin infusion schedule, are reported in Table 5.1.

Table 5.1. Optimised design variables and discrete information content for the robust oral test.

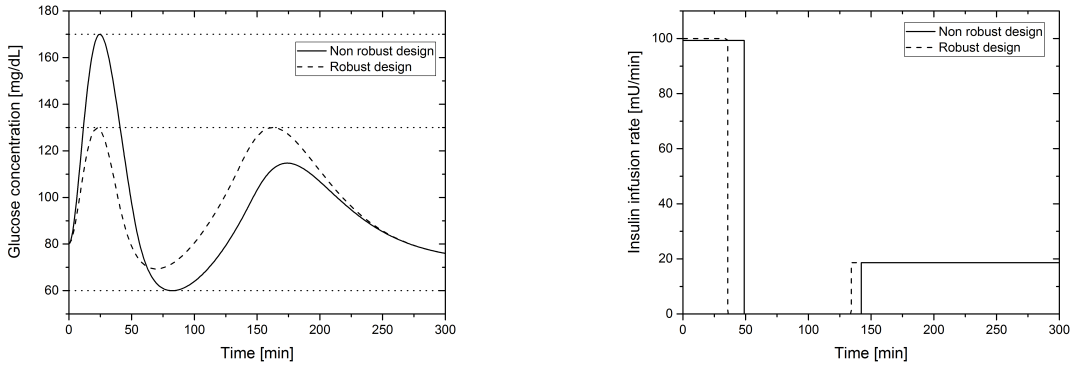
Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 37, 42, 47, 52, 57, 62, 107, 112, 117, 122, 180, 185, 190, 195, 200, 275, 280, 285, 290, 295, 300]
U [mU/min]	[100, 0, 18.62]
t_U [min]	[35.78, 98.39, 165.83]
I_d	413.29

The high value found for the discrete information content, which is slightly lower than the 5h-22s oral test, indicates that good parameter estimates can possibly be achieved with this new protocol.
In addition to good identification performance, the objective of the robust design is primarily that of guaranteeing safety conditions to the unknown patient. At this purpose, the robust oral test can be compared to an identical design obtained at the same conditions but using the original upper bound constraint, i.e. 170 mg/dL. Initially, another preliminary design is performed adjusting both insulin infusion and glucose administration, which now varies in the range from 30 to 70 g. A value of 62.23 g is obtained, thus it is rounded off to 62 g.
Fixing the glucose meal at this value, a new design is carried out in which only the insulin infusion is controlled as before. The results of the new protocol, which will be referred to as *non robust oral test*, are presented in Table 5.2.

Table 5.2. Optimised design variables and discrete information content for the non robust oral test.

Variable	Value
$\mathbf{t_{sp}}$ [min]	[0, 44, 49, 54, 59, 98, 103, 108, 113, 124, 131, 171, 181, 191, 196, 201, 222, 251, 285, 290, 295, 300]
U [mU/min]	[99.30, 0, 18.65]
t_U [min]	[48.93, 93.26, 157.82]
I_d	490.03

It can be noticed that the non robust oral test is even more informative than the 5h-22s oral test (Table 4.14), because it induces a higher stimulus in the subject pushing the glucose profile towards both bounds (Figure 5.1a). Similarly, the robust oral test, thanks to the reduced glucose meal, presents a similar behaviour but allows to maintain the glucose profile below the new upper bound imposed by the robust design. As regards insulin infusion, Figure 5.1b shows that the two profiles are approximately equivalent. A slightly longer insulin infusion is required in the non robust oral test, since a greater glucose meal is applied.



(a) (b)

Figure 5.1. Comparison between robust and non robust oral tests: glucose profile from the MBDoE procedure (a) and insulin administration profile (b).

5.3.2 Parameter estimation

The effectiveness of the robust design can be shown performing a parameter estimation task with set $\mathbf{\Theta_2}$. Figure 5.2 shows the glucose profiles after identification with that set for the robust and the non robust oral tests. Experimental measurements are not displayed in the graph for the sake of clarity. It can be observed that the robust oral test succeeds in maintaining the glucose profile within the upper bound, even though

set $\mathbf{\Theta_2}$ does not belong to the population of subjects simulated in the robust design, for the reasons pointed out in § 4.1.1.

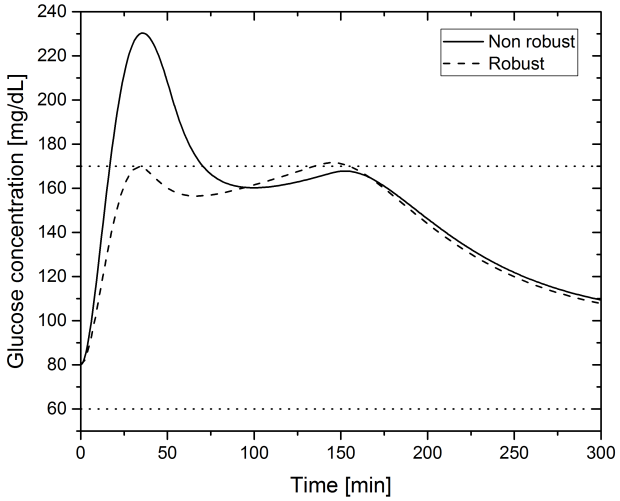


Figure 5.2. Glucose profiles of the robust and non robust oral tests after identification with set $\mathbf{\Theta_2}$.

The robust protocol devised does not guarantee that no violation of the upper threshold is obtained for any choice of the parametric set. Nevertheless, since an expected value approach has been used, it ensures that, at least in a neighbourohood of the variability range taken into account in the robust design, the violation is limited. In terms of parameter estimation performance, both tests permit to identify the model parameters with set $\mathbf{\Theta_2}$ in a statistically sound way. A light improvement in the estimates is observed in the non robust oral test, due to the higher information content realised. However, since the focus is now on safety, only the robust oral test estimates are reported in Table 5.3.

Table 5.3. Parameter estimation statistics for the robust oral test with set $\mathbf{\Theta_2}$.

Parameter	Final value	Initial guess	Confidence interval 95%	95% t-value (Ref. 1.734)
Θ_1	0.6999	1	0.1347	5.196
Θ_2	1.3043	1	0.1519	8.589
Θ_3	0.6193	1	0.0961	6.447
Θ_4	1.5739	1	0.2424	6.493

A comparison between these estimates and those achieved with the 5h-22s oral test (Table 4.16) shows that the effect of the loss of information content due to safety

enforcement is not so evident. In fact, the t-values prove that, for the second and third parameter, better estimates are achieved using the robust oral test.

5.4 Final considerations

The non robust oral test seems to grant an enhancement of the information content compared to the 5h-22s oral test. However, since the primary concern of the robust design is to ensure safety conditions to the subject, the non robust oral test is excluded from further analyses.
Figure 5.3 presents a comparison based on the 5 heuristic indices between the 5h-22s oral test and the robust oral test.

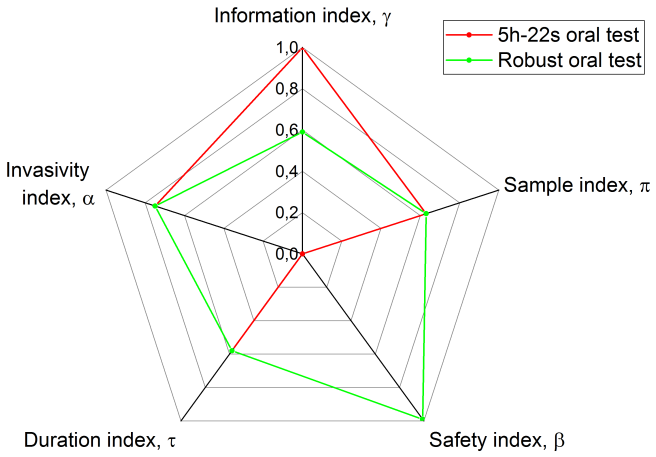


Figure 5.3. Indices radar chart: comparison between the 5h-22s oral test and the robust oral test after parameter estimation with set $\boldsymbol{\Theta_2}$.

The first one has been shown to be the most informative among all the designed protocols, while the latter exhibits the highest safety. Taking merely the areas of the two polygons into account, the robust oral test clearly outperforms the 5h-22s oral test in an overall assessment. As a consideration related to this fact, it is important to highlight the effectiveness of the implementation of a robust approach in the design of clinical tests.
Additionally, it should be clear that the more information is desired, the less safety is achieved. This fact has clearly emerged from the previous comparison between the robust and the non robust oral tests. Broadly speaking, the implementation of stricter safety requirements in the design phase, due to the robust approach, causes a reduction of the degrees of freedom available. As a result, a lesser stimulus can be induced on the patient, at least considering the same duration for the test, thus a lower information

content can be accomplished.

As a final consideration, the research of an optimal clinical protocol for parameter identification seems to be more the result of a compromise on the highest information achievable at the maximum safety, rather than a simultaneous maximisation of both of them.

Beyond safety conditions over the entire duration of the test, the optimal protocol is also required to restore basal glucose values at the end of it. Nevertheless, this requirement is not simple to be satisfied, since its direct implementation in the MBDoE would result in a further decrease of the degrees of freedom available, hence a much more complex design. Therefore, a different procedure has been followed, in order to assure the compliance of this condition, based on the results obtained at the end of the test.

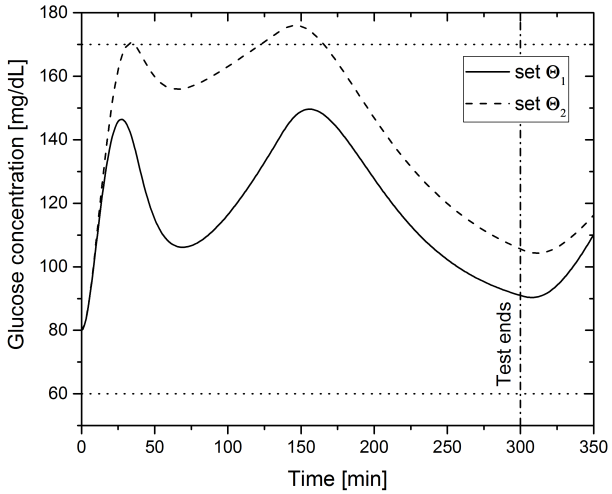


Figure 5.4. Glucose profiles of the robust oral test with sets $\mathbf{\Theta_1}$ and $\mathbf{\Theta_2}$ over a 350 min duration (test ends at 300 min).

Figure 5.4 shows the behaviour of glucose concentrations with sets $\mathbf{\Theta_1}$ and $\mathbf{\Theta_2}$ once the robust oral test terminates, i.e. at 300 min. As previously observed in Figure 5.2, glucose profile is not flat at the end of the test, due to the parametric mismatch. However, suspending insulin infusion at 300 min, a stabilisation of glucose concentration is achieved after a while, followed by a rapid increase due to the lack of insulin administration.

This fact is confirmed by looking at the values of the glucose derivative reported in Table 5.4: after 10 min in both cases the derivatives fall in the range imposed in the design, i.e. $-0.1 < \Gamma < 0.1$.

Table 5.4. Glucose derivative and glucose concentration values at the end of the robust oral test and 10 min after with sets $\boldsymbol{\Theta_1}$ and $\boldsymbol{\Theta_2}$.

Parametric set	Γ @ 300 min [mg/(dL · min)]	Γ @ 310 min [mg/(dL · min)]	G_{tot} @ 300 min [mg/dL]	G_{tot} @ 310 min [mg/dL]
$\boldsymbol{\Theta_1}$	-0.13	0.06	90.97	90.38
$\boldsymbol{\Theta_2}$	-0.20	-0.03	105.60	104.30

As for the final glucose values achieved, it should be considered that they are not too far from the basal value, i.e. 80 mg/dL, so they can be easily restored under the supervision of the medical staff or, alternatively, with the patient's knowledge of himself/herself.

Conclusions

The purpose of this thesis was to devise a feasible clinical protocol for the identification of physiological models of T1DM. The Lynch-Bequette model has been assumed as the reference model in this analysis, which has been entirely conducted by means of the software $gPROMS^{®}$.

The investigation has departed from the tests currently used in clinical facilities for diabetes diagnosis and research purposes: the standard OGTT and the standard IVGTT. A preliminary analysis has provided some valuable knowledge on them, such as the more physiological stimulus induced by an oral test and the higher information content derivable from it.

The chief part of the work has been the design of new clinical tests applying a MBDoE approach. Firstly, the poor performance of the two standard tests has been demonstrated, motivating the search of reliable alternatives. Secondly, five protocols have been enquired, primarily in terms of parameter estimation capability, using two distinct parametric sets representing different degrees of diabetes severity.

The design procedure applied to the standard tests has shown considerable improvements in both cases. However, at the same time, it has highlighted the superiority of an oral test over an intravenous one for parameter estimation purposes. In fact, the 5h-11s oral test succeeds in this task, while the 3h-23s intravenous test fails, due to its poor stimulus. This hurdle cannot be overcome even using two glucose injections (3h-23s-2i intravenous test), although excellent safety standards are exhibited during the test.

Focusing on the oral tests, the 5h-22s oral test has shown the most satisfactory parameter estimation performance among the designed protocols, along with a simple insulin infusion profile. The main issue related to this protocol is its unsatisfactory safety, especially in the identification of patients with severe parametric deviations from the healthy subject. As a countermeasure to this problem, a different protocol has been examined based on two glucose meals of reduced quantity. The resulting 5h-22s-2m oral test has only partially met the safety requirements desired for a clinical implementation and, at the same time, it has shown some difficulties in restoring the glucose basal value at the end of it.

It has been noticed that there is a strong dependence of the model outputs on the parameters, which causes significant deviations from the designed profiles, putting sometimes the subject at risk. In turn, there is also a bias of the information content due to the parametric set, which needs to be considered when comparing different tests. In light of the constant ratio between the discrete and continuous information content for the same test, regardless of the parametric set, an assessment of the most informative protocol is allowed.

The evaluation of the overall feasibility requirements of the tests has been based on five indices. The invasivity index cannot be improved with the kind of tests taken into account, duration can hardly be reduced for identification purposes, while the number of samples has been maintained at a sensible value, since it is directly linked to the information derivable. However, the MBDoE procedure has shown its effectiveness in maximising the information content while enforcing constraints on the safety.

The lack of a reliable protocol has triggered further investigation, which has entailed the application of a robust design approach. Starting from the most informative and least safe protocol (i.e. 5h-22s oral test), this method has allowed to design a highly reliable clinical protocol (robust oral test). This test is endowed with high robustness, that is its safety is quite insensitive to the parameter values, which is a major concern when an unknown subject is to be identified. In addition, it assures a statistically satisfactory parameter estimation, even though some information content is inevitably lost. As a final consideration, this protocol seems suitable because it has a simple insulin infusion profile and glucose values close to basal are restored at the end of it.

Future developments may involve the application of a multi-objective design strategy, where information content and safety are both optimised at the same time. Additionally, it could be interesting to take into consideration oral insulin administrations, so as to minimise the invasivity on the subject.

Notations

Symbols

A	glucose meal	[g]
BW	body weight	[kg]
D	glucose intake velocity	[mg/dL/min]
G	glucose concentration	[mg/dL]
G_b	basal glucose concentration	[mg/dL]
G_{sc}	glucose concentration in the subcutaneous layer	[mg/dL]
G_{tot}	total glucose concentration	[mg/dL]
I	difference between blood and basal insulin	[mU/L]
I_b	basal insulin concentration	[mU/L]
I_c	continuous information content	[-]
I_d	discrete information content	[-]
$I_{d,max}$	maximum discrete information content among all the tests	[-]
K_i	confidence interval	[-]
LB	lower glucose bound	[mg/dL]
N_θ	number of model parameters	[-]
n_{max}	maximum number of samples allowed in a test	[-]
n_{sp}	number of blood samples taken in a test	[-]
n_{sw}	number of switching levels of a control variable	[-]
N_u	number of manipulated inputs	[-]
N_w	number of time invariant controls	[-]
N_x	number of state variables	[-]
N_y	number of measured variables	[-]
N_θ	number of model parameters	[-]
q_i	i-th element of the dynamic sensitivity matrix	[-]
R_{ut}	glucose usage velocity in tissues	[mg/dL/min]
s_{ij}	ij-th element of the inverse matrix of Σ	[-]

S_G	glucose effectiveness	[min^{-1}]
S_I	insulin sensitivity	[mU·min/L]
t	time	[min]
$\bar{t}$	t-value	[-]
t_0	initial time	[min]
t_{max}	maximum test duration	[min]
$t_{max,G}$	time-to-maximum of glucose uptake	[min]
t_{sp}	glucose sampling time	[min]
t_{test}	test duration	[min]
U	exogenous insulin infusion rate	[mU/min]
U_v	intravenous glucose injection rate	[mg/dL/min]
UB	upper glucose bound	[mg/dL]
V_1	glucose lower bound violation	[mg·min/dL]
V_2	glucose upper bound violation	[mg·min/dL]
V_g	glucose distribution volume	[dL/kg]
V_i	insulin distribution volume	[L]
X	insulin concentration in the remote compartment	[mU/L]

Greek letters

α	invasivity index	[-]
$\alpha_{glucose}$	glucose invasivity index	[-]
$\alpha_{insulin}$	insulin invasivity index	[-]
$\alpha\%$	confidence region	[-]
β	safety index	[-]
β'	total bound violation of a test	[mg·min/dL]
β_{max}	maximum bound violation observed in the OGTT	[mg·min/dL]
γ	information content index	[-]
Γ	time derivative of glucose concentration at the end of a test	[mg/dL/min]
γ_i	heteroschedastic factor	[-]
θ_i	i-th model parameter	[-]
θ_i'	i-th perturbated parameter	[-]
λ_k	k-th eigenvalue	[-]
ν_{ii}	element on the diagonal of the variance-covariance matrix	[-]
π	sample index	[-]

σ_i	standard deviation of the i-th response $\hat{y}_i$	[-]
τ	test duration index	[-]
Φ^{LS}	objective function for LS parameter estimation	[-]
Φ^{ML}	objective function for ML parameter estimation	[-]
Φ^{WLS}	objective function for WLS parameter estimation	[-]
χ^2	chi-square	[-]

Vectors and matrices

χ^2_{ref}	reference chi-square	[-]
Ψ	$\mathbf{V}_\theta$ measurement function	[-]
ω_i	standard deviation of the i-th measured response	[-]
$\mathbf{H}_\theta$	Fisher information matrix	$[N_\theta \mathrm{x} N_\theta]$
$\mathbf{H}^0_\theta$	preliminary Fisher information matrix	$[N_\theta \mathrm{x} N_\theta]$
$\mathbf{Q}$	sensitivity matrix	$[n_{sp} \mathrm{x} N_\theta]$
$\mathbf{t_{sp}}$	vector of sampling times	$[n_{sp}]$
$\mathbf{t_{sw}}$	vector of switching times	$[n_{sw}+1]$
$\mathbf{u}$	vector of time-dependent control variables	$[N_u]$
$\mathbf{V}_\theta$	variance-covariance matrix of model parameters	$[N_\theta \mathrm{x} N_\theta]$
$\mathbf{w}$	vector of time-invariant control variables	$[N_w]$
$\mathbf{x}$	vector of state variables	$[N_x]$
$\mathbf{y}$	measurements vector	$[N_y]$
$\hat{\mathbf{y}}$	vector of estimated responses	$[N_y]$
$\mathbf{z_{sw}}$	vector of switching levels	$[n_{sw}]$
$\hat{\mathbf{y}}$	vector of estimated values of model parameters	$[N_\theta]$
$\theta_\mathbf{0}$	vector of initial model parameters	$[N_\theta]$
$\mathbf{\Theta_0}$	vector of normalised initial model parameters	$[N_\theta]$
$\mathbf{\Theta_1}$	first vector of normalised model parameters for the subject	$[N_\theta]$
$\mathbf{\Theta_2}$	second vector of normalised model parameters for the subject	$[N_\theta]$
$\mathbf{\Sigma_i}$	measurement errors variance-covariance matrix in the i-th experiment	$[N_y \mathrm{x} N_y]$
$\mathbf{\Sigma}_\theta$	prior variance-covariance matrix of model parameters	$[N_\theta \mathrm{x} N_\theta]$
$\boldsymbol{\varphi}$	design vector	$[n_\varphi]$
$\boldsymbol{\varphi_{opt}}$	optimal design vector	$[n_\varphi]$

Acronyms

CGMS	continuous glucose monitoring system
GDM	gestational diabetes mellitus
IVGTT	intravenous glucose tolerance test
LB	Lynch and Bequette model
LS	least squares
MBDoE	model-based design of experiments
ML	maximum likelyhood
MPC	model predictive control
OGTT	oral glucose tolerance test
SWR	sum of weighted residuals
T1DM	type 1 diabetes mellitus
T2DM	type 2 diabetes mellitus
WAP	wearable artificial pancreas
WHO	World Health Organization
WLS	weighted least squares

Bibliography

[1] Livingstone, S. J.; Levin, D.; Looker, H. C. et al. *Estimated Life Expectancy in a Scottish Cohort With Type 1 Diabetes, 2008-2010.* JAMA. 2015; 313(1):37-44. doi:10.1001/jama.2014.16425.

[2] Gardner, D. G. and Shoback, D. editors. *Greenspan's basic and clinical endocrinology.* 9th ed. San Francisco: McGraw Hill; 2011.

[3] World Health Organization, Geneva, 1999. *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus.* [online]. Available from: http://apps.who.int/iris/handle/10665/66040 [Accessed 30 December 2015].

[4] World Health Organization. Geneva, 2014. *Global health estimates: deaths by cause, age, sex and country, 2000-2012.* [Online]. Available from: http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html [Accessed 29 December 2015].

[5] World Health Organisation. Geneva, 2014. *Global status report on noncommunicable diseases 2014.* [Online]. Available from: http://www.who.int/nmh/publications/ncd-status-report-2014/en/ [Accessed 29 December 2015].

[6] NCD Risk Factor Collaboration (NCD-RisC). *Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants.* Lancet 2016; published online April 7. http://dx.doi.org/10.1016/S0140-6736(16)00618-8.

[7] Seuring, T.; Archangelidi, O.; Suhrcke, M. *The economic costs of type 2 diabetes: a global systematic review.* Pharmacoeconomics 2015; 33:811–31.

[8] World Health Organization, Geneva, 1999. *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus.* [online]. Available from: http://apps.who.int/iris/handle/10665/66040 [Accessed 30 December 2015].

[9] Bergman, R. N.; Phillips, L. S.; Cobelli, C. *Physiologic evaluation of factors controlling glucose tolerance in man.* J. Clin. Invest. 1981; 68, 1456-1467.

[10] De Fronzo, R. A.; Jordan, D. T.; Aundreb, A. *Glucose clamp technique: a method for quantifying insulin secretion and resistance.* Am. J. Physiol. 1979; 237, E214-E223.

[11] Bequette, B. W. *A critical assessment of algorithms and challenges in the development of a closed-loop artificial pancreas.* Diabetes Technol. Ther., 2005; 28-47.

[12] http://www.medicine.tcd.ie/physiology/assets/docs12_13/lecturenotes/AG/BMS/Artificial_Pancreas.pdf [Accessed 21 August 2016].

[13] Bequette, B. W.; Baysal, N.; Howsmon, D.; Cameron, F. *Steps towards a closed-loop artificial pancreas.* 41st Annual northeast bioengineering conference, 2015. doi: 10.1109/NEBEC.2015.7117081

[14] Bequette, B. W. *Algorithms for a closed-loop artificial pancreas: the case for model predictive control.* J Diabetes Sci. Technol. 2013; 7(6): 1632-1643.

[15] Balakrishnan, N. P.; Rangaiah, G. P.; Samavedham, L. *Review and analysis of blood glucose (BG) models for type 1 diabetic patients.* Ind. Eng. Chem. Res. 2011; 50(21):12041-12066.

[16] Lynch, S. M.; Bequette, B. W. *Model predictive control of blood glucose in type 1 diabetics using subcutaneous glucose measurements.* Proceedings of the American Control Conference, 2002; 4039-4043.

[17] Van Herpe, T.; Pluymers, B.; Espinoza, M.; Van der Berghe, G.; De Moor, B. *A minimal model for glycemia control in critically ill patients.* EMBS '06. 28th Annual Internation Conference of the IEEE, Engineering in Medicine and Biology Society, August 30 to September 3, 2006; pp 5432-5435.

[18] Roy, A.; Parker, R. S. *Dynamic modelling of free fatty acid, glucose, and insulin: an extended "minimal model".* Diabetes Technol. Ther. 2006, 8 (6), 617-626.

[19] Sturis, J.; Polonsky, K. S.; Mosekilde, E.; Van Cauter, E. *Computer model for mechanisms underlying ultradian oscillations of insulin and glucose.* Am. J. Physiol. Endocrinol. Metab. 1991, 260 (5), E801-809.

[20] Tolic, I. M.; Mosekilde, E.; Sturis, J. *Modelling the insulin-glucose feedback system: the significance of pulsatile insulin secretion.* J. Theor. Biol. 2000, 207 (3), 361-375.

[21] Bennett, D. L.; Gourley, S. A. *Asymptotic properties of a delay differential equation model for the interaction of glucose with plasma and interstitial insulin.* Appl. Math. Comput. 2004, 151 (1), 189-207.

[22] Hovorka, R.; Canonico, V.; Chassin, L. J.; Haueter, U.; Massi-Benedetti, M.; Federici, M.O.; Pieber, T. R.; Schaller, H. C.; Schaupp, L.; Vering, T.; Wilinska, M. E. *Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes.* Physiol. Meas. 2004; 25, 905-920.

[23] Dalla Man, C.; Rizza, R. A.; Cobelli, C. *Meal simulation model of the glucose-insulin system.* IEEE transactions on biomediacal engineering, 2007; vol. 54, no. 10, 1740-1749.

[24] Asprey, S. P.; Macchietto, S. *Statistical tools for optimal dynamic model building.* Comput. Chem. Eng. 2000; 24(2-7):1261-1267.

[25] Yang, Y. J.; Youn, J. H.; Bergman, R. N. *Modified protocols improve insulin sensitivity estimation using the minimal model.* Am. J. Physiol. 1987; 253, E595-E602.

[26] Galvanin, F.; Barolo, M.; Macchietto, S.; Bezzo, F. *Optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus.* Ind. Eng. Chem. Res. 2009; 48, 1989-2002.

[27] Galvanin, F.; Barolo, M.; Macchietto, S.; Bezzo, F. *Optimal design of clinical tests for the identification of physiological models of type 1 diabetes in the presence of model mismatch.* Med. Biol. Eng. Comput. 2011; 49(3):263-277.

[28] Laguna, A. J.; Rossetti, P; Ampudia-Blasco, F. J.; Vehí, J.; Bondia, J. *Experimental blood glucose interval identification of patients with type 1 diabetes.* Journal of Process Control 2014; 24, 171-181.

[29] Galvanin, F.; Barolo, M; Bezzo, F. *On the use of continuous glucose monitoring systems to design optimal clinical tests for the identification of type 1 diabetes mellitus.* Comput Methods Programs Biomed 2013; 109(2):157-170.

[30] Sorensen, J. T. *A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes.* Ph.D. thesis Dept. Chem. Eng. MIT, Cambridge, 1985.

[31] Vassiliadis, V. S.; Sargent, R. W. H.; Pantelides, C. C. *Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints.* Industrial and Engineering Chemistry Research 1994; 33(9):2111-2122.

[32] Galvanin, F. *Optimal model-based design of experiment in dynamic systems: novel techiniques and unconventional applications.* Ph.D. Thesis, DIPIC, Università di Padova, 2010.

[33] Zullo, L. *Computer aided design of experiments. An engineering approach.* Ph.D. Thesis, The University of London, U.K.; 1991.

[34] Bard, Y. *Nonlinear parameter estimation.* Academic Press. New York; 1974. p. 64

[35] Process Systems Enterprise. *gPROMS Introductory User Guide.* Process Systems Enterprise Ltd.: London; 2004.

[36] Process Systems Enterprise. *gPROMS Advanced User Guide.* Process Systems Enterprise Ltd.: London; 2004.

[37] The MathWorks. *MATLAB Getting Started Guide.* The MathWorks Inc.: Natick, USA; 2008.

[38] Fonte, P.; Araujo, F.; Reis, S.; Sarmento, B. *Oral insulin delivery: how far are we?.* J. Diabetes Sci. Technol. 2013; 7(2), 520-531.

[39] Dalla Man, C; Campioni, M.; Polonsky, K. S.; Basu, R.; Rizza, R. A.; Toffolo, G.; Cobelli, C. *Two-hour seven-sample oral glucose tolerance test and meal protocol.* Diabetes 2005; 54 (11), 3265-3273.

[40] Hirsch, I. B.; Bode, B. W.; Garg, S.; Lane, W.S.; Sussman, A.; Hu, P. et al. *Continuous subcutaneous insulin infusion (CSII) of insulin aspart versus multiple daily injection of insulin aspart/insulin glargine in type 1 diabetic patients previously treated with CSII.* Diabetes Care 2005; 28(3): 533-538.

[41] Food and Agriculture Organization of the United Nations. *Food energy-methods of analysis and conversion factors.* Report of a technical workshop, Rome; 2003.

[42] Asprey, S. P.; Macchietto, S. *Designing robust optimal dynamic experiments.* J. Process Control 2002; 12(4): 545-556.