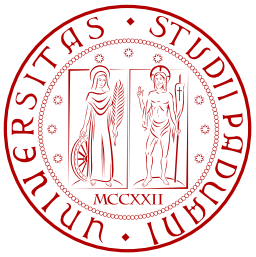

UNIVERSITY OF PADOVA

DEPARTMENT OF INFORMATION ENGINEERING

Ph.D. Course on Information Engineering

Curriculum: Bioengineering

Series: XXXVII

Mathematical Modeling of Glucose-Lactate Metabolism in Adolescents with and without Obesity during Oral Glucose Tolerance Tests

Course director:

Prof. Fabio Vandin

Advisor:

Prof. Chiara Dalla Man

Ph.D. candidate:

Jacopo Bonet

A thesis submitted

for the degree of

Philosophiæ Doctor (Ph.D.)

Typeset in LaTeX

Padua, Italy, 2024

Abstract

Before 1984, with the development of the so called *"Lactate-Shuttle theory"* by Brooks et.al, lactate was considered a waste product of anaerobic glycolysis (i.e. the first process of glucose utilization), generated during muscle fatigue and hypoxia state. Thereafter, several studies on this molecule were conducted, showing its key role in human metabolism.

From these studies, it emerged that lactate is always the final product of glycolysis, it can enter the mitochondria in which it is oxidized, and serves as a fuel in several human organs (e.g. heart, brain, liver and muscles), but also it has a signaling role, e.g. together with insulin and free fatty acids (FFAs).

In pathological states characterized by insulin resistance (i.e. the impaired insulin effect on the endogenous glucose production and utilization), such as obesity, diabetes and metabolic dysfunction-associated steatotic liver disease (MASLD), the high circulating glucose is primarily processed through glycolysis thus producing lactate. The high lactate levels observed in these pathological states contributes to the development of liver inflammation, ultimately leading to liver fibrosis and cirrhosis. Thus, lactate can be used as a marker in such diseases, in fact it was proposed as the molecule connecting obesity with insulin resistance state.

Mathematical modeling of lactate kinetic could be useful to estimate parameters related to its metabolism, enhancing the knowledge on this molecule and

processes in which it is involved.

Some minimal models (i.e. parsimonious models able to summarize the key characteristics of a biological process) were proposed to this aim. However, they have been validated during non-physiological tests, such as intravenous glucose tolerance tests (IVGTTs) or lactate infusions. This precluded the applicability of such models to more physiological post-prandial conditions. Also, to the best of our knowledge, none of the literature simulation models (i.e. models characterized by a large number of equations and parameters to fully implement the knowledge of the system under study) accounted for lactate kinetics.

That is why, in this project, the first aim was to set up a minimal model of the glucose-lactate kinetics during a 4h oral glucose tolerance test (OGTT), able to estimate parameters related to the insulin resistance state (i.e. liver and disposal insulin sensitivities, SI_L and SI_D, respectively) along with the glycolytic flux (i.e. lactate production rate, LPR) in a population of adolescents overweight/with obesity. Starting from the existing Single Tracer Oral Minimal Model (STOMM), proposed by Visentin et.al., we integrated a mono-compartmental description of the lactate kinetic, showing that this was the best choice for predicting both glucose and lactate during an OGTT. We also showed that model parameters estimated during a 4h-protocol well correlated with those obtained from a 3h-protocol. This allows to increase the usability of the model and alleviate the burden for the patient. Finally, the model was successfully applied to an independent dataset of lean/overweight/obese adolescents, showing that it was able to account for differences related to the obesity degree. Moreover, the difference observed in some model parameters in adults and adolescents were concordant with those reported in the literature studies. Particularly, the correlations obtained between model parameters and

subjects-specific anthropometric characteristics, along with the differences related to the age, opened the possibly to extend the minimal model to the non-linear mixed-effect modeling, in which such characteristics could be directly integrated into the model.

The second aim of this project was to develop a simulation model of lactate, together with glucose, insulin and C-peptide dynamics during an OGTT, extending the Type 2 Diabetes simulator (T2Ds) proposed by Visentin et.al. We showed that the integration of an adapted version of the selected minimal model of the lactate kinetic was the best choice also for this kind of model. The simulator was successfully employed to clone a real population of adolescents, showing that no statistically significant differences occurred between the virtual population and the real one for what concerns both substrates and hormones time courses, and key model parameter distributions. This simulator could be employed to *in-silico* test efficacy and safety of drugs targeting lactate metabolism, paralleling what it has been done with insulin treatments in the context of T2D and Type 1 Diabetes (T1D). As a matter of fact, T1D and T2D simulators have been used to simulate how exogenous insulin affects glucose metabolism, allowing to predict treatment outcomes, optimize dosing strategies, and improve patient management. Similarly, by simulating drug interactions with lactate metabolism, we could gain valuable insights into potential therapeutic effects prior to *in-vivo* testing.

Nevertheless, despite the interesting and promising results summarized above, we have to acknowledge some important limitations. The first is that the study only included lactate data from adolescents with/without obesity, which restricted the ability to compare lactate-related parameters with adults, but also with subjects affected by other pathological states such as diabetes. Secondly, although lactate metabolism is closely linked to other molecules like FFAs, this

study could not explore that connection due to the lack of such data. Finally, data from only a single OGTT were available, preventing the possibility to assess the presence of potential patterns in substrates and parameters throughout the day.

These will be the starting points for the future development of this projects. Additionally, incorporating lactate data from wearable sensors could further expand the models. Specifically, for what concerns the minimal model, the access to such data might broaden model applicability to real-life, free-living conditions. On the other hand, for what concerns the simulation model, the availability of such data could make the simulator a useful tool to improve the performance of control algorithms that, currently, rely solely on continuous glucose monitoring (CGM) and enhance decision-making processes in the management of diabetes and MASLD.

Sommario

Prima del 1984, con lo sviluppo della cosiddetta *"Lactate-Shuttle Theory"* proposta da Brooks et al., il lattato era considerato un prodotto di scarto della glicolisi anaerobica (ovvero il primo processo di utilizzazione del glucosio), generato durante l'affaticamento muscolare e in condizioni di ipossia. Successivamente, numerosi studi sono stati condotti su questa molecola, dimostrando il suo ruolo chiave nel metabolismo umano.
Da questi studi è emerso che il lattato è sempre il prodotto finale della glicolisi, può entrare nei mitocondri dove viene ossidato e funge da substrato metabolico in vari organi umani (cuore, cervello, fegato e muscoli), ma ha anche un ruolo di *signaling*, ad esempio con l'insulina e gli acidi grassi liberi (FFAs).
In stati patologici caratterizzati da resistenza all'insulina (ovvero l'effetto compromesso dell'insulina nel ridurre la produzione endogena di glucosio e nell'incrementare la sua utilizzazione), come obesità, diabete e malattia epatica associata a disfunzione metabolica (MASLD), le alte concentrazioni di glucosio circolante vengono principalmente processate attraverso la glicolisi, producendo così lattato. Gli elevati livelli di lattato osservati in questi stati patologici contribuiscono allo sviluppo di infiammazione epatica, portando infine a fibrosi epatica e cirrosi. Pertanto, il lattato può essere utilizzato come marcatore in tali malattie, infatti è stato proposto come la molecola che collega l'obesità allo stato di resistenza all'insulina.

Per questo motivo, la modellizzazione matematica della sua cinetica potrebbe essere utile per stimare parametri legati al suo metabolismo, migliorando la conoscenza su questa molecola e sui processi in cui è coinvolta.
A tale scopo, in letteratura sono stati proposti alcuni modelli minimi (cioè modelli parsimoniosi in grado di riassumere le caratteristiche chiave di un processo biologico). Tuttavia, questi sono stati validati durante test non fisiologici, come in test intravenosi di tolleranza al glucosio (IVGTT) o di infusioni di lattato. Questo ha impedito l'applicabilità di tali modelli a condizioni più fisiologiche come quelle post-prandiali. Inoltre, nessuno dei modelli di simulazione presenti in letteratura (cioè modelli caratterizzati da un maggior numero di equazioni e parametri per implementare completamente la conoscenza del sistema oggetto di studio) tengono conto della cinetica del lattato.
Per questo motivo, in questo progetto, il primo obiettivo è stato quello di sviluppare un modello minimo della cinetica glucosio-lattato durante un test di tolleranza orale al glucosio (OGTT) di 4 ore, in grado di stimare parametri relativi allo stato di resistenza all'insulina (cioè le sensibilità all'insulina epatica e di utilizzazione, rispettivamente SI_L e SI_D) insieme al flusso glicolitico (cioè il tasso di produzione di lattato, LPR) in una popolazione di adolescenti sovrappeso/con obesità. Partendo dal modello minimo orale con singolo tracciante (STOMM), proposto da Visentin et al., abbiamo integrato una descrizione mono compartimentale della cinetica del lattato, dimostrando che questa era la scelta migliore per predire sia il glucosio che il lattato durante un OGTT. Abbiamo anche dimostrato che i parametri del modello stimati durante un protocollo di 4 ore correlavano significativamente con quelli ottenuti da un protocollo di 3 ore. Ciò permette di aumentare l'usabilità del modello e alleviare l'onere per il paziente. Infine, il modello è stato applicato con successo a un dataset indipendente di adolescenti magri/sovrappeso/obesi, dimostra-

ndo che era in grado di tenere conto delle differenze legate al grado di obesità. Inoltre, in accordo con quanto riportato negli studi di letteratura, sono state osservate differenze in alcuni parametri del modello tra adulti e adolescenti. In particolare, le correlazioni ottenute tra i parametri del modello e le caratteristiche antropometriche specifiche dei soggetti, insieme alle differenze legate all'età, hanno aperto la possibilità di estendere il modello minimo alla modellizzazione non-lineare a effetti misti, in cui tali caratteristiche potrebbero essere direttamente integrate nel modello.

Il secondo obiettivo di questo progetto era quello di sviluppare un modello di simulazione della dinamica del lattato, assime alle dinamiche di glucosio, insulina e C-peptide durante un OGTT, estendendo il simulatore del diabete di tipo 2 (T2Ds) proposto da Visentin et al. Abbiamo dimostrato che l'integrazione di una versione adattata del modello minimo del lattato selezionato fosse la scelta migliore anche per questo tipo di modello. Il simulatore è stato impiegato con successo per clonare una popolazione reale di adolescenti, dimostrando che non ci sono differenze statisticamente significative tra la popolazione virtuale e quella reale per quanto riguarda l'andamento temporale sia dei substrati che degli ormoni, e le distribuzioni dei principali parametri del modello. Questo simulatore potrebbe essere utilizzato per testare *in-silico* farmaci il cui target sia il metabolismo del lattato, in modo simile a quanto fatto con i trattamenti insulinici nel T2D e nel diabete di tipo 1 (T1D). Infatti i simulatori del T1D e T2D sono stati impiegati per simulare come l'insulina esogena influenza il metabolismo del glucosio, consentendo di prevedere gli esiti del trattamento, ottimizzare le strategie di dosaggio e migliorare la gestione del paziente. Allo stesso modo, simulando le interazioni farmacologiche con il metabolismo del lattato, potremmo ottenere preziose informazioni sugli effetti terapeutici potenziali prima dei test *in-vivo*.

Tuttavia, nonostante i risultati interessanti e promettenti riassunti precedentemente, dobbiamo riconoscere alcune importanti limitazioni. La prima è che lo studio ha incluso solo dati di lattato da adolescenti con/senza obesità, il che ha limitato la possibilità di confrontare i parametri legati al lattato con gli adulti, ma anche con soggetti affetti da altre patologie come il diabete. In secondo luogo, sebbene il metabolismo del lattato sia strettamente legato ad altre molecole come gli FFAs, questo studio non ha potuto esplorare tale connessione a causa della mancanza di tali dati. Infine, essendo disponibili dati da un solo OGTT, non è stato possibile valutare la presenza di potenziali pattern giornalieri nè della concentrazione dei substrati coinvolti nè nei parametri del modello.

Superare queste limitazioni sarà l'obiettivo degli sviluppi futuri di questo progetto di tesi.

Inoltre, l'incorporazione di dati di lattato da sensori indossabili potrebbe ampliare ulteriormente i modelli. In particolare, per quanto riguarda il modello minimo, l'accesso a tali dati potrebbe estendere la sua applicabilità a condizioni di vita reale. D'altra parte, per quanto riguarda il modello di simulazione, la disponibilità di tali dati potrebbe rendere il simulatore uno strumento utile per migliorare le prestazioni degli algoritmi di controllo che attualmente si basano esclusivamente sul monitoraggio continuo del glucosio (CGM) e migliorare i sistemi di supporto alle decisioni nella gestione del diabete e MASLD.

.

Contents

List of Abbreviations

ADA American Diabetes Association.

AUC Area Under the Curve.

BIC Bayesian information criterion.

BMI body mass index.

BMIPCT body mass index percentile.

BW body weight.

CGM continuous glucose monitoring.

CV coefficient of variation.

DNL de novo lipogenesis.

FFA free fatty acid.

G6P glucose 6 phosphate.

GLUT glucose transporter.

IVGTT intravenous glucose tolerance test.

LPR lactate production rate.

MAP maximum a posteriori.

MASLD metabolic dysfunction-associated steatotic liver disease.

MCT monocarboxylate transporters.

MMTT mixed meal tolerance test.

OGMM Oral Glucose Minimal Model.

OGTT oral glucose tolerance test.

ROS reactive oxygen species.

STOMM Single Tracer Oral Glucose Minimal Model.

T2D type 2 diabetes.

TCA tricarboxylic acid cycle.

TOST two one-sided test.

WRSS weighted residuals sum of squares.

List of Figures

List of Tables

Chapter 1

Introduction and aim of the thesis

In this first chapter, an introduction on lactate and the metabolic processes in which it is involved is carried out. Alongside this introduction, to fully understand the scope and aims of this thesis, the attention will be focused on lactate kinetics, its role in pathological conditions, and the possibility to mathematically model its metabolism to better understand the mechanism in which it is involved.

1.1 Lactate as key molecule for human metabolism

Before 1984, with the development of the so called *"Lactate-Shuttle theory"* by Brooks et.al. [1][2][3][4], lactate was considered a waste product of anaerobic glycolysis (i.e.in absence or with reduced oxygen levels), generated during muscle fatigue and hypoxia state. The Cory cycle [5] was the only known lactate pathway, in which the lactate produced by the anaerobic glycolysis in the muscles, it is converted back into pyruvate in the liver and finally into glucose, as one of the main gluconeogenetic precursors [6].

Despite that, after the theory proposed by Brooks et.al [1][2][3][4], lactate now is considered a key molecule involved in several metabolic processes, also its

signaling role has emerged. First, the *"Lactate Shuttle theory"* claims that lactate is always the final product of glycolysis, both in aerobic and anaerobic conditions [3][7]. Secondly, this theory that accounts for the "intracellular" [1] and "cell to cell" lactate shuttle [4], shows that lactate can also enter the mitochondria [8], through transporters (i.e. monocarboxylate transporters, MCTs [9]) and be used as oxidative substrate.

Other theories were those proposed in [4][10], claiming that lactate can be used as an alternative metabolic fuel inside neurons, but also utilized by the heart, lungs, muscles during meal (Figure 1.1) and exercise conditions.

Lastly and briefly, lactate signaling role emerged [11][12], particularly in controlling its own catabolism, glycolysis and lipolysis (i.e. the metabolic reaction which generates free fatty acids, FFAs, from adipose tissues).

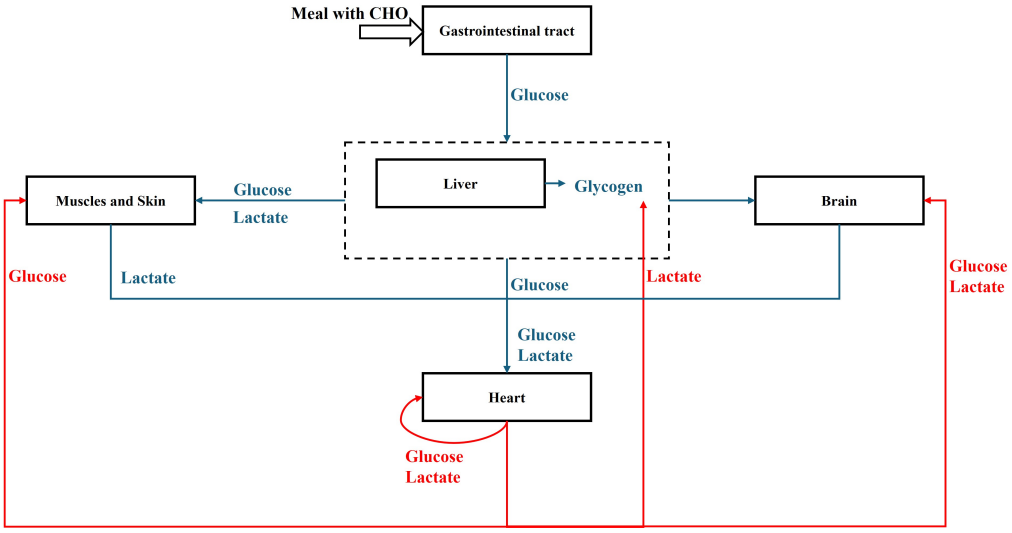

Figure 1.1: Glucose-Lactate metabolism during postprandial conditions (adapted from Brooks et.al [4]). Depending on liver glycogen levels, some dietary glucose bypasses the liver and enters the systemic circulation. Glucose is taken up by non-contracting muscles, particularly red and intermediate fibers. Glycolysis in these and other tissues leads to the release of lactate into the central venous circulation (blue lines), which is then taken up by the liver from the arterial circulation (red lines). In contrast, glucose from dietary intake or digestion in the gastro-intestinal tract can be directly taken up by the liver during the first circulatory pass for direct glycogen synthesis. Additionally, lactate released from various cells and tissues supports brain metabolism and functions such as glutamatergic signaling. Similarly, systemic lactate serves as an energy source for the heart.

1.2 Lactate metabolism in pathological conditions

Monitoring lactate is proven to be important in several metabolic diseases, particularly in metabolic dysfunction associated steatotic liver disease (MASLD) [13][14][15][16], as well as in conditions such as obesity and diabetes [17][18]. Individuals with obesity and prediabetes often exhibit a high degree of insulin resistance, which results in increased glucose production and reduced uptake in muscles due to a diminished insulin effect [19]. This scenario leads to an excess of circulating glucose, which is used as a metabolic substrate. While some glucose is stored as glycogen, it cannot be completely converted into it [20], thus, glucose is processed through glycolysis, ultimately producing lactate. Lactate is either metabolized or released into the bloodstream, serving as

a key marker of glycolysis. The lactate is subsequently converted into pyruvate and acetyl-CoA, entering the tricarboxylic acid cycle, TCA, (i.e. the final path of the oxidative metabolism) and hepatic de novo lipogenesis, DNL, (i.e. the process of generation of new fat by the liver), which contributes to palmitate formation [13][21][22]. This second process ultimately leads to inflammation and liver fibrosis [23] (Figure 1.2).

Also, in states of obesity and insulin resistance, an increased flux of FFA from adipose tissue and enhanced DNL driven by insulin resistance led to hepatic fat accumulation, a hallmark of MASLD [21][24]. As MASLD progresses, it exacerbates overall insulin resistance, contributing to cardiometabolic complications such as diabetes [25][26][27].

Thus, lactate emerges as a crucial molecule in these pathophysiological states, acting as a link between obesity and insulin resistance [28]. By understanding lactate kinetics and modeling drugs that affect its metabolism (e.g., [15][16]), we can potentially develop innovative methodologies to improve our understanding of these metabolic conditions.

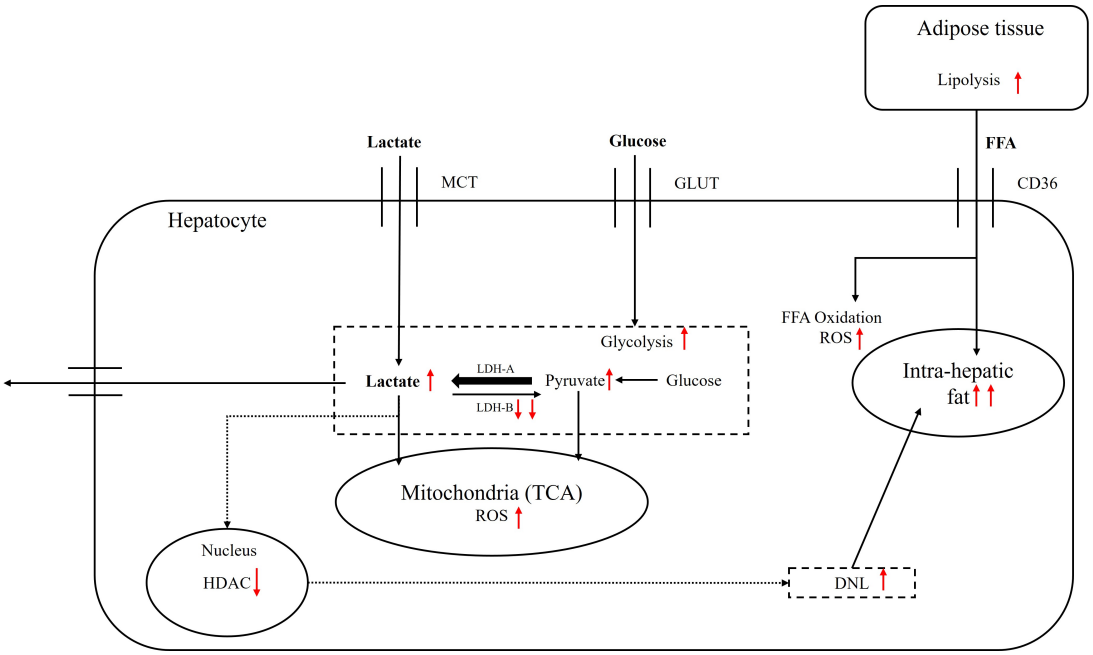

Figure 1.2: Hepatocytes glucose-lactate-FFA metabolism in MASLD, adapted from Q. Lu et.al. [15]. Glucose and lactate enter the liver hepatocytes through MCT and GLUT respectively. Glycolysis is upregulated, leading to higher levels of pyruvate in both plasma and liver. Pyruvate is subsequently converted to oxaloacetate (in the TCA) or to lactate, both of which are increased in MASLD. The LDH-B enzyme (i.e. responsible of the lactate to pyruvate conversion) activity in reduced, leading to high lactate levels [16]. Elevated lactate boosts the expression of lipogenic enzymes due to decreased nuclear HDAC activity, increasing DNL. Increased lipolysis in the adipose tissue raises plasma FFA levels. Fatty acid oxidation is also heightened, not just in the mitochondria, but also in the endoplasmic reticulum. Oxidation in the latter results in more ROS production, which causes inflammation and liver damage. The increased glycolysis and fatty acid oxidation, which elevate acetyl-CoA levels and enhance the TCA cycle, along with the reduced activity of the mitochondrial respiratory chain, lead to further ROS generation.

1.3 Modeling lactate kinetics

Mathematical modeling of lactate kinetic can provide insight into its metabolism and relationship with glucose and insulin, helping to elucidate the mechanisms causing pathological states such as insulin resistance and diabetes.

1.3.1 Literature lactate kinetic models

Mathematical models have been developed to describe lactate kinetics during exercise and intravenous glucose tolerance test (IVGTT). For instance, Durand et al. [29] applied several lactate kinetic models, previously developed in [30][31][32], to examine lactate responses across various exercise conditions. In another study by Wahl et.al.[33], the authors investigated lactate transport into and out of red blood cells under different warm-up protocols and before and after a 30-second maximal-effort exercise. Additionally, a comprehensive model of whole-body fuel homeostasis during exercise conditions was developed in [34], and later updated in [35]. This model incorporated oxygen consumption and epinephrine dynamics, both directly tied to relative exercise intensity, to modulate hormone and glucose responses, including lactate kinetics in response to varying exercise intensities and modalities. For what concern the IVGTT, the first mathematical model was developed by Watanabe et.al [36] in 1995, which consisted of the lactate kinetic integration into the glucose minimal model proposed in [37], with the aim to predict both glucose and lactate plasma concentrations during this protocol. The type of tests used in that study consisted of an insulin-modified frequently sampled intravenous glucose tolerance test (INS-MOD FSIGT) and in a FSIGT. The first test was used to evaluate the impact of insulin on lactate kinetic, and compared with the results of

the FSIGT, in which only glucose was infused showing that insulin infusion increases plasma lactate concentration, but the lactate production was limited by the rate of glucose uptake (i.e. insulin resistance state may have profound effect on lactate kinetics). More recently, in 2012, Stefanovski et.al [38] developed a simple mathematical model to describe the glucokinase activity (i.e. one of the key glycolytic enzymes) in the liver and to predict lactate plasma concentrations during a tolbutamide-modified FSIGT, to assess if insulin can acutely influence lactate kinetics. It is worth noting that, differently from [36], in this work authors used glucose as the model forcing function (i.e. assumed to be known and without error). Stefanovski and coworkers concluded that, during a FSIGT, the lactate production depends only on glucose, independently from dynamical changes in insulin concentrations. Despite that, since the authors found correlations between lactate related parameters and the disposition index (DI, an index of beta-cell function accounting for prevailing insulin resistance) they suggested that *"in subjects with reduced DI, lactate appearance is diminished similar to the results in those with diabetes"*, again underlying that insulin resistance state impact lactate appearance. Lastly, in 2024, Romeres et.al. [39] described lactate kinetic after a primed infusion of labelled lactate and identified differences concerning lactate clearance rate in healthy and T1D adults during exercise under eu/hyperglycemic conditions.

1.3.2 Gap in the literature

Although lactate kinetics has been widely studied under exercise conditions or IVGTT, to the best of our knowledge, it has not been examined during an oral glucose tolerance test (OGTT) simultaneously with glucose kinetic. It is also worth noting that the IVGTT creates a non-physiological condition in the glucose regulatory system, by delivering glucose directly into the bloodstream,

bypassing the gastrointestinal tract. In contrast, the OGTT consisted of an oral glucose ingestion, typically 75g, and the subsequent measurement of plasma glucose and insulin concentrations. This type of test was used in several studies, and OGTT data have been used for the development of the Oral Glucose Minimal Model (OGMM) [40], to assess the overall insulin sensitivity (SI), and also in [41] to develop the Single Tracer Oral Minimal Model, STOMM, which was used to estimate the liver and disposal insulin sensitivity indexes (SI_L, SI_D respectively) from labeled and unlabeled plasma glucose concentrations. The development of a minimal model considering both glucose and lactate kinetics during an OGTT, would be useful to estimate parameters related to its metabolism and to assess the glycolitic flux under physiological conditions.

Moreover, none of the published metabolic simulators (e.g. [42][43][44][45][46]) account for lactate kinetic. The development of a simulator of the lactate kinetics and the subsequent possibility to conduct *in-silico* trials on the glucose-lactate system could be of great help for the research in this field for several reasons. Firstly, these tools enable researchers and clinicians to predict and analyze the pharmacokinetic and pharmacodynamic behavior of new drugs in a controlled virtual environment before proceeding to human trials. This approach enhances safety by identifying potential side effects early on, but also optimizes the drug development process, potentially reducing costs and time to market. Furthermore, a simulator that accounts for lactate kinetics can provide valuable insights into metabolic disorders and conditions where lactate metabolism is disrupted, such as sepsis [47], diabetes, or MASLD.

For example, such a simulator could aid in understanding how different drugs interact with lactate production and clearance under various pathological states (e.g. [15][16]), offering a platform to tailor drug dosages and their combinations more effectively.

Moreover, developing a lactate kinetics simulator can set-up the basis for the development of new algorithms for real-time monitoring and treatment adjustments, similar to the closed-loop system and basal insulin treatment proposed for type 1 (T1D) [48][49][50] and, more recently, for type 2 diabetes (T2D) [51]. The impact of such advancements could be significant, not only in the management of chronic conditions but also in acute care settings where rapid adjustments of lactate levels are critical.

1.4 Aim and structure of the thesis

Lactate has an emerging metabolic role, involved in several metabolic processes, also linked to obesity, diabetes, and MASLD.

The possibility to model its kinetic during physiological perturbation and extract some parameters related to glycolytic process will be a key to better understand the body response and assess the impact of this molecule in pathological conditions. As previously discussed, the existing mathematical models can describe lactate kinetics, but during non-physiological conditions, such as IVGTT. Moreover, no simulation model of lactate has been proposed so far. The aim of this thesis is to fulfill this gap.

This is why, in the first part of the thesis, "minimal models" (i.e. parsimonious models able to capture the key aspects of a metabolic process) are presented, along with the methodology of selecting the best performing one. The impact of protocol duration on the estimated parameters and fluxes along with the analysis of the differences between adolescents with/without obesity are also assessed. The second part of this thesis describes the development and use of a "simulation model" (i.e. a model with a huge number of equations and parameters) that is a comprehensive description attempting to fully implement

the body knowledge about the system allowing to perform simulation and to conduct *in-silico* trials in specific populations.

The work carried out in this project is organized as follows.

In Chapter 2, the two datasets used in this study are described. The first one, collected at the Kansas University, consists in data collected in 23 subjects during a 3h/4h-75g OGTT, and the second one, collected at the Yale University, consists in data collectd in 24 subjects from 3h-75g OGTT.

In Chapter 3, the parameters estimation techniques and the methods to select the best performing minimal model are discussed. In the first part of the chapter, all the procedures and assumptions that allowed model identification are described and motivated, whereas the second part deals with the procedure for model assessment. Finally, in the last part, the results of model selection phase, along with model applications, are presented. In Chapter 4, the simulation model application and results are described in detail. For the reader convenience, further details on the integration of an adapted version of the minimal model into the simulator are reported in Appendix A.

Finally, in Chapter 5, the discussion and future development are presented.

Chapter 2

Datasets

In this chapter, the two datasets used in this thesis project for models development, use and simulation are presented. The first dataset (Dataset 1) was provided by the Kansas University (Kansas City, KS, USA) and consisted of plasma glucose, lactate and insulin measurements collected during 3h/4h-75g OGTT in twenty-three adolescents with obesity. The second Dataset (Dataset 2) was provided by Yale University (New Heaven, CT, USA) and consisted of plasma glucose, insulin and lactate measurements collected during a 3h-75g OGTT in twenty-four adolescents with/without obesity. The first dataset was used to develop the minimal model, and to evaluate if the parameters and fluxes estimated from a 4h-protocol were comparable with those obtained from a shorter one (3h). The second dataset, which contained also lean subjects, was used to assess whether model parameters and fluxes were different in adolescents with and without obesity. Finally, data of subjects for which C-peptide measurement were also available (i.e. forty-four subjects) were used to develop the glucose, insulin, C-peptide and lactate adolescents' simulator.

2.1 Dataset 1

The first dataset was collected at the Kansas University and the study was supported by NIH, R01MD015974.

2.1.1 Subjects

Twenty-three adolescents with (Mean±SD) body weight (BW) of 99±22 kg, and age of 15±2 years were enrolled. The body mass index (BMI) was calculated as the weight in kilograms divided by the height in square meters (33±7 kg/m^2), and the body mass index percentile (BMIPCT) was assessed using age and sex specific BMI reference data, to classify subjects into lean/overweight/obese [52]. Particularly, subjects with BMIPCT < 85th percentile were classified as leans, with 85th $\leq$ BMIPCT < 95th percentile were classified as overweight, obese otherwise. The inclusion criteria was BMIPCT $\geq$ 85th percentile (i.e. only subjects with obesity/overweight were included). Twenty-two subjects were with obesity, while one subjects was overweight. Subjects baseline characteristics are reported in Table 2.1.

2.1.2 Protocol

After an overnight fast, a 75g-4h OGTT was performed and glucose, insulin, and lactate plasma concentrations were collected. In 4 subjects, the study was stopped earlier, at 180min, because glucose levels were falling below 70mg/dL, one subject was excluded from the study since insulin was not available between 60-120min, while other two subjects had not lactate measurements after 3h (in other words 16 subjects were studied for >3h). Glucose (Figure 2.1, panel A) and lactate (panel C) concentrations were measured by using a YSI2700-STAT-Analyzer (Yellow Springs Instruments, Yellow Springs,

Table 2.1: Dataset 1 subjects baseline characteristics

Baseline	**Mean**$\pm SD$**/Count**
Sex (Male/Female)	17/6
Age (years)	15±2
Body weight (kg)	99±22
Body mass index (kg/m^2)	33±7
Lean/overweight/obese	0/1/22
Basal glucose (mg/dL)	94±7
NFG/IFG	19/4
2-h glucose (mg/dL)	122±23
NGT/IGT	21/2
Basal insulin (µU/mL)	35±26
Basal lactate (mg/dL)	9.73±2.79

Mean±SD/count baseline characteristics of the 23 subjects under study. NFG/IFG: normal/impaired fasting glucose; NGT/IGT: normal/impaired glucose tolerance.

Ohio) at t=[0 10 20 30] and every 15 minutes thereafter. With this sampling frequency, 16 subjects had a total of 36 samples (18 for glucose and 18 for lactate), while the other 4 subjects had a total of 28 samples (14 for glucose and 14 for lactate) after the glucose ingestion. Insulin measurements (Figure 2.1, panel B) were collected at t=[0 10 20 30] and every 30 minute thereafter and measured by using a radioimmunoassay (RIA; Linco, St. Charles, Missouri).

In accordance with the American Diabetes Association (ADA) criteria [53], adolescents were classified into normal/impaired glucose tolerance (NGT/IGT) depending on 2h-glucose plasma concentration (IGT if 2h-glucose$\geq$140mg/dL, NGT otherwise), and into normal/impaired fasting glucose (NFG/IFG) depending on fasting plasma glucose (IFG if fasting glucose $\geq$100mg/dL, NFG otherwise). Twenty-one subjects were NGT, while two were IGT. The two IGT subjects were also IFG, while other two were IFG but NGT. The study was approved by the Human Investigation Committee of Kansas University Medical Center, where it was performed. Children's parents provided written informed consent to participate in the study.

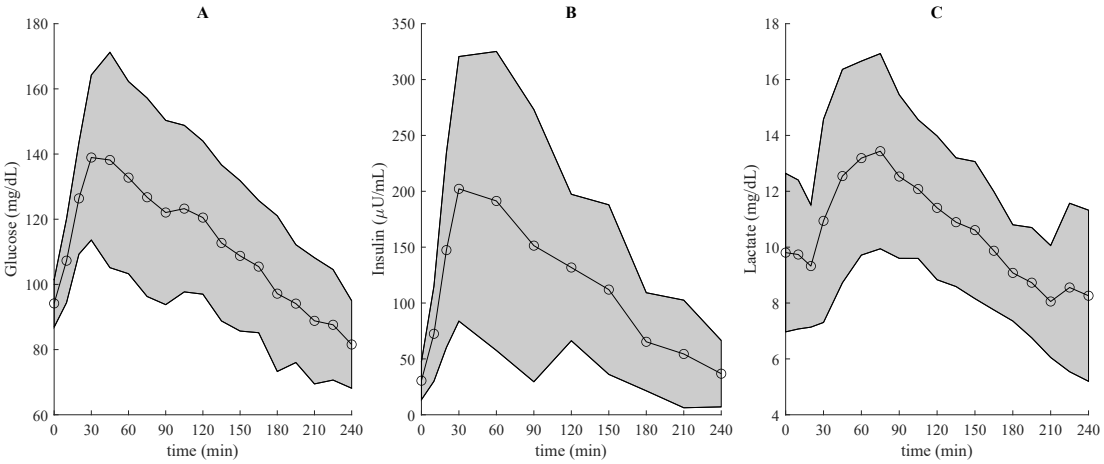

Figure 2.1: Mean (continuous line with circles) and mean±SD intervals (grey areas) of plasma glucose (panel A), insulin (panel B) and lactate concentrations (panel C).

2.2 Dataset 2

The dataset was collected at Yale University and the study was supported in part by MIUR (Italian Minister for Education) under the initiative "Departments of Excellence" (Law 232/2016) and in part by NIH R01DK114504 and R01MD015974

2.2.1 Subjects

Twenty-four subjects were enrolled to determine the pathogenesis of obesity in youth (NCT03454828) and characterized by BW of 83±27kg, age of 16±3 years. The BMI was calculated as for Dataset 1 (29±9 kg/m^2), and BMIPCT was used classify subjects into lean/overweight/obese [52]. Fifteen adolescents were overweight/with obesity (four were overweight, eleven were with obesity) and nine were lean. Subjects baseline characteristics are reported in Table 2.2.

Table 2.2: Dataset 2 subjects baseline characteristics

Baseline Mean±*SD*/Count	**All subjects N=24**	**Lean N=9**	**Overweight/Obese N=15**
Sex (Male/Female)	9/15	2/7	7/8
Age (years)	16±3	16±2	16±3
Body weight (kg)	83±27	61±7	97±26*
Body mass index (kg/m^2)	29±9	21±2	34±8*
Lean/overweight/obese	9/4/11	/	/
Basal glucose (mg/dL)	86±6	83±6	88±6
NFG/IFG	24/0	9/0	15/0
2-h glucose (mg/dL)	95±20	107±21	88±16*
NGT/IGT	24/0	9/0	15/0
Basal insulin (µU/mL)	13±6	8±3	15±5*
Basal lactate (mg/dL)	8.83±2.43	8.20±1.71	9.10±2.80

Mean±SD/count baseline characteristics of the 24 subjects under study (second column), and of lean (third column) and overweight/obese (fourth column). * indicates statistically significant difference between lean vs overweight/obese using T-test or Mann-Whitney test depending on variable distribution. NFG/IFG: normal/impaired fasting glucose; NGT/IGT: normal/impaired glucose tolerance.

2.2.2 Protocol

After an overnight fast, subjects underwent a 3h-75g OGTT. Plasma measurements of glucose, insulin and lactate were collected at t=-15, 0, 10, 20, 30, 60, 90, 120, 150,180 min after the glucose ingestion. In particular, plasma glucose (Figure 2.1, panel A) was determined with an YSI 2700 Analyzer (YSI Inc), plasma insulin (Figure 2.2, panel B) concentrations were measured by Radioimmunoassay (Linco), while a colorimetric method was used to measure plasma concentrations of lactate (kit LC2389; Randox) during the study (Figure 2.2, panel C). With this sampling frequency, the subjects had a total of 18 samples (9 for glucose and 9 for lactate) after the glucose ingestion.

Similarly to Dataset 1 and in accordance with the American Diabetes Association criteria [53], adolescents were classified into NGT/IGT and into NFG/IFG. All the subjects were NGT and NFG.

Yale internal review board (IRB) approved the study (HIC20000239), while children's parents provided written informed consent to participate in the study.

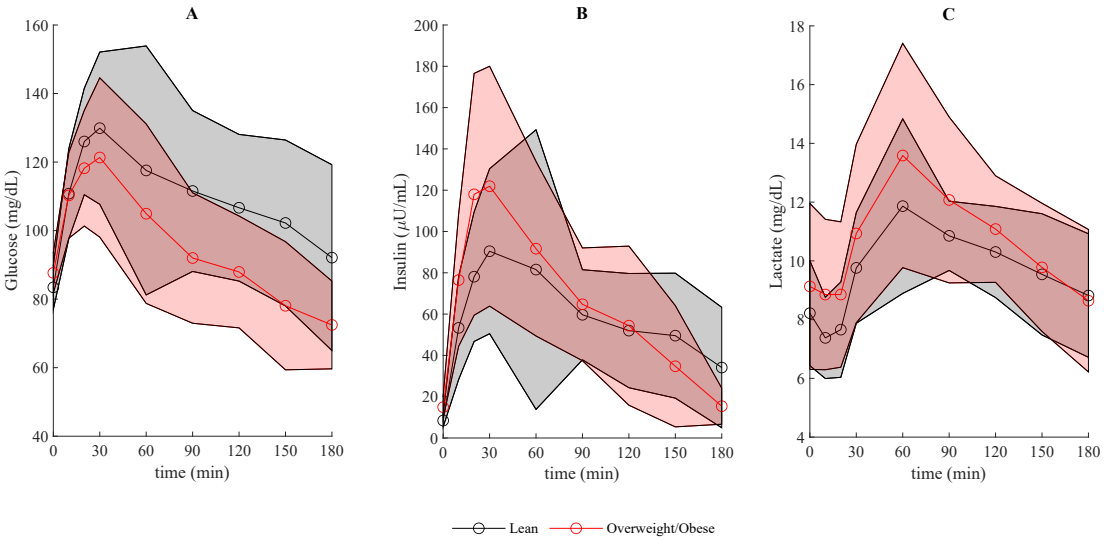

Figure 2.2: Mean (continuous line with circles) and mean±SD intervals (shaded areas) of plasma glucose (panel A), insulin (panel B) and lactate (panel C) concentrations in lean (black) and overweight/obese (red).

Chapter 3

The Oral Glucose-Lactate Minimal Model

[1]As mentioned in Chapter 1, the first aim of this thesis was to develop a parsimonious (i.e. minimal) model to estimate key parameters related to glycolysis during an OGTT.

Some mathematical models were proposed in the literature to describe lactate kinetics [36][38] during IVGTT, but none during a more physiological one, like OGTT. In this chapter, the glucose-lactate oral minimal model is presented.

First, some literature models were tested, and new mathematical models were developed on Dataset 1. The best one was selected based on standard criteria (see Chapter 3.2.3). Dataset 1 was also used to determine the impact of protocol duration on the estimated model parameters and on the glycolytic flux (i.e. lactate production rate (LPR)). The second dataset was used to assess the model performance on a different cohort of adolescents which contained also lean subjects.

Finally, the impact of obesity degree and age on model parameters in all the

[1]This chapter contains material published in [54][55].

available subjects is discussed.

3.1 Models

In this paragraph, all the literature and new tested models are presented. Since the protocol was an OGTT, the models shared the same description of meal glucose rate of appearance (R_a), already described in the OGMM [40]. R_a was modelled as a piece-wise linear function:

$$R_a(\alpha,t) = \begin{cases} \alpha_{i-1} + \frac{\alpha_i - \alpha_{i-1}}{t_i - t_{i-1}} \cdot (t - t_{i-1}) & \text{for } t_{i-1} \leq t \leq t_i \text{ with } i = 1, ..., 5 \\ \alpha_5 \cdot e^{\frac{-(t-t_5)}{T}} & t > t_5 \end{cases} \tag{3.1}$$

with α denoting [α_1, α_2, ..., α_5] and t_5=180min and T=60min, assuming that:

$$\int_0^\infty R_a(\alpha,t)\, dt = \frac{D \cdot f}{BW} \tag{3.2}$$

where D is the oral glucose ingested dose (typically 75g for an OGTT in adults), f the fraction of ingested glucose that appears in the plasma (f=0.9, as in [40]) and BW the subject body weight. This constraint provided an additional relationship among the unknown parameter α_i, thus reducing the number of R_a parameters by one.

Despite sharing the R_a description, the tested models presented in the following were based on different assumptions, in accordance with literature evidence or already available models. Further details on the common assumptions used during model identification are reported in Paragraph 3.2.2.

3.1.1 Model 1

Model 1 was inspired by that proposed by Stefanovski et.al. [38] to describe lactate kinetics during an IVGTT. However, since the purpose of this minimal

model was to predict both glucose and lactate data, at variance to the published study, glucose was not used as model forcing function (i.e. assumed to be known without error), but its kinetic description was integrated into the model. To do this, as a first step, we assumed that all the glucose, regardless from insulin concentration, is converted into glucose 6 phosphate (G6P), if above G_b, and then finally into lactate. Model equations consisted in three ordinary differential equations:

$$\begin{cases} \dot{Q}_G(t) = -k_{PG} \cdot [Q_G(t) - V_G \cdot G_b] + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{Q}_P(t) = -k_{LP} \cdot Q_P(t) + k_{PG} \cdot [Q_G(t) - V_G \cdot G_b] \cdot f_{switch} & Q_P(0) = 0 \\ \dot{Q}_L(t) = -k_L \cdot [Q_L(t) - Q_{Lsb}(t)] + 2k_{LP} \cdot Q_P(t) \cdot f_{switch} & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.3}$$

where Q_G , Q_P and Q_L (mg/kg) are the glucose, intermediate metabolite (i.e. G6P) and lactate masses respectively, while G_b (mg/dL), L_b (mg/dL) and V_G (dL/kg), V_L (dL/kg), are the glucose and lactate basal plasma concentrations and volumes respectively. Q_{Lsb} and f_{switch} were defined as:

$$\begin{cases} Q_{Lsb} = (t \cdot \frac{(L_M - L_b)}{t_M} + L_b) \cdot V_L \\ f_{switch} = \begin{cases} 1 \text{ if } G(t) \geq G_b \\ 0 \text{ if } G(t) < G_b \end{cases} \end{cases} \tag{3.4}$$

where Q_{Lsb} is the "sliding baseline" for lactate, L_M is the lactate concentration at the end of the experiment (i.e. 180/240min) and t_M=180 or 240min depending on protocol duration.

In other words, the model assumes that glucose is firstly converted into glucose-6-phosphate with a rate constant k_{PG} (min^{-1}) only if glucose plasma concen-

tration is above its basal value. Glucose-6-phosphate is then converted into lactate with a rate constant k_{LP} (min^{-1}), the factor 2 account for the fact that one molecule of G6P in converted into two molecules of lactate. Lactate is then cleared from the plasma with a rate constant k_L (min^{-1}) but depending on Q_{Lsb}. It is worth noting that, in this model, no insulin effect on glucose metabolism was considered, as in [38]. A schematic representation of Model 1 is reported in Figure 3.1.

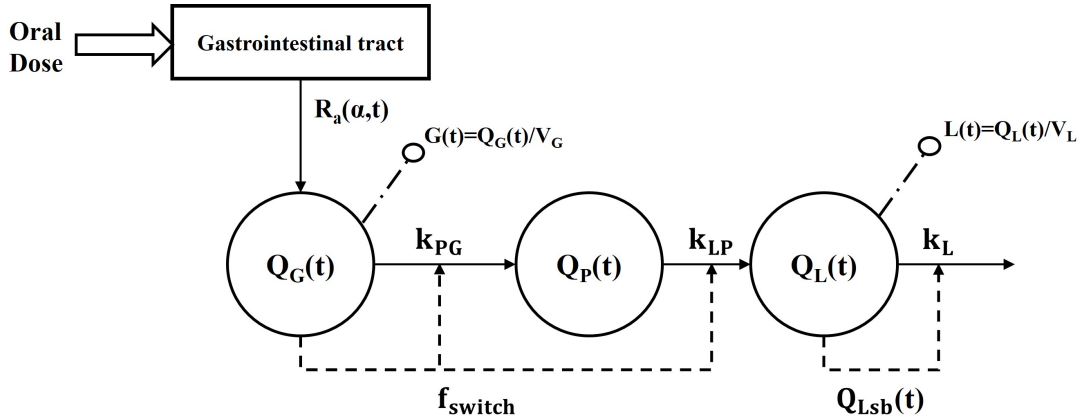

Figure 3.1: Schematic representation of Stefanovski et.al. model [38] adapted to describe both glucose and lactate plasma concentration during an OGTT. Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. Q_G , Q_P and Q_L are glucose, G6P and lactate plasma mass respectively; G(t) and L(t) are the glucose and lactate plasma concentrations respectively; $R_a(\alpha,t)$ is the glucose rate of appearance described as in Dalla Man et.al. [40]; k_{PG} is the rate constant of glucose to G6P conversion; k_{LP} is the rate constant of G6P to lactate conversion; k_L is the rate constant of lactate plasma clearance; f_{switch} is the control action of glucose on both glucose-G6P and G6P-lactate conversions; Q_{Lsb} is the sliding basal lactate concentration; V_G and V_L are the glucose and lactate plasma volumes respectively.

3.1.2 Model 2

Model 2, at variance of 1, also considered the insulin effect on glucose metabolism, with the same mathematical description of the OGMM [40], while the other model equations remained the same of 3.1.1. Model equations are:

$$\begin{cases} \dot{Q}_G(t) = -[p_1 + X(t)] \cdot Q_G(t) + p_1 \cdot G_b \cdot V_G + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}(t) = -p_2 \cdot X(t) + p_2 \cdot SI' \cdot [I(t) - I_b] & X(0) = 0 \\ \dot{Q}_P(t) = -k_{LP} \cdot Q_P(t) + [p_1 + X(t)] \cdot [Q_G(t) - V_G G_b] f_{switch} & Q_P(0) = 0 \\ \dot{Q}_L(t) = -k_L \cdot [Q_L(t) - Q_{Lsb}(t)] + 2k_{LP} \cdot Q_P(t) \cdot f_{switch} & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.5}$$

where p_1 (min^{-1}) is the fractional glucose effectiveness, representing the ability of glucose to promote its disposal and inhibit its production; the net insulin action on both glucose production and disposal, X(t), is assumed to be delayed with respect to plasma insulin, with a rate constant p_2 (min^{-1}); SI' (1/min per μU/mL) is the fractional insulin sensitivity (i.e. per unit distribution volume) and I_b (μU/mL) the basal insulin concentration. Total insulin sensitivity, SI (i.e. the ability of insulin to inhibit glucose production and enhance glucose utilization, dL/kg/min per μU/mL) was defined as:

$$SI = SI' \cdot V_G \tag{3.6}$$

It is worth noting that p_1 depends on SI':

$$p_1 = \frac{GEZI}{V_G} + I_b \cdot SI' \tag{3.7}$$

where GEZI (dL/kg/min) is the glucose effectiveness at zero insulin.

It is worth noting that, like for Model 1, we still assumed that all the glucose is converted first into G6P, if above G_b, and then into lactate, but, here, we integrated also the insulin control on glucose metabolism like in [40]. A schematic representation of Model 2 is reported in Figure 3.2.

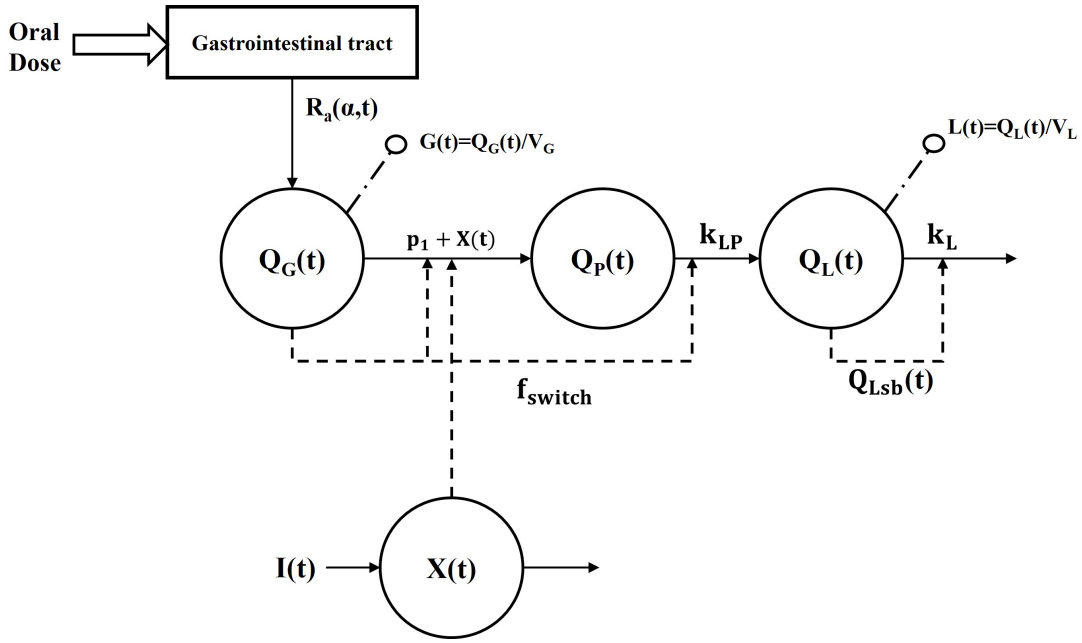

Figure 3.2: Schematic representation of Stefanovski et.al. model [38] adapted to describe both glucose and lactate plasma concentrations during an OGTT, combined with the OGMM of Dalla Man et.al. [40] Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. Q_G , Q_P and Q_L are glucose, G6P and lactate plasma mass respectively; G(t), L(t) and I(t) are the plasma glucose, lactate and insulin concentrations respectively; X(t) is the delayed insulin effect on both glucose production and utilization; $R_a(\alpha,t)$ is the glucose rate of appearance; p_1 is the fractional glucose effectiveness; k_{LP} is the rate constant of G6P to lactate conversion; k_L is the rate constant of lactate plasma clearance; f_{switch} is the control action of glucose on both glucose-G6P and G6P-lactate conversions; Q_{Lsb} is the sliding basal lactate concentration; V_G and V_L are the glucose and lactate plasma volumes respectively.

3.1.3 Model 3

Model 3 is an extension of the one proposed in [36], integrating the R_a description of [40] in it. Model equations are the following:

$$\begin{cases} \dot{Q}_G(t) = -[p_1 + X(t)] \cdot Q_G(t) + p_1 \cdot G_b \cdot V_G + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}(t) = -p_2 \cdot X(t) + p_2 \cdot SI' \cdot [I(t) - I_b] & X(0) = 0 \\ \dot{Q}_P(t) = -(k_P + k_{LP}) \cdot Q_P(t) + 2 \cdot fr \cdot [p_1 + X(t)] \cdot Q_G(t) + k_{PL} \cdot Q_L(t) & Q_P(0) = Q_{Pb} \\ \dot{Q}_L(t) = -(k_L + k_{PL}) \cdot Q_L(t) + k_{LP} \cdot Q_P(t) & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.8}$$

Differently from the previous models, Model 3 assumed that only a fraction of glucose is converted into pyruvate (Q_P), *fr* (%), while the other (1-*fr*) follows other metabolic paths.

In addition, the model considers the possible exchange between the pyruvate and lactate compartments, through the rate parameters k_{LP} and k_{PL} (min^{-1}). Both pyruvate and lactate can be irreversibly cleared from the plasma, with rate constant k_P and k_L (min^{-1}). The model was not *a priori identifiable* (see Paragraph 3.2.1), thus, a reparameterization was performed as in the original work [36], leading to the following model equations:

$$\begin{cases} \dot{Q}_G(t) = -[p_1 + X(t)] \cdot Q_G(t) + p_1 \cdot G_b \cdot V_G + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}(t) = -p_2 \cdot X(t) + p_2 \cdot SI' \cdot [I(t) - I_b] & X(0) = 0 \\ \ddot{Q}_L(t) = -A_1 \cdot \dot{Q}_L(t) - A_2 \cdot Q_L(t) + A_3 \cdot [p_1 + X(t)] \cdot Q_G(t) & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.9}$$

With:

$$\begin{cases} A_1 = k_L + k_{LP} + k_P + k_{PL} \\ A_2 = k_L \cdot k_{LP} + k_P \cdot k_{PL} + k_L \cdot k_P \\ A_3 = 2fr k_{LP} \end{cases} \tag{3.10}$$

A schematic representation of Model 3 is reported in Figure 3.3.

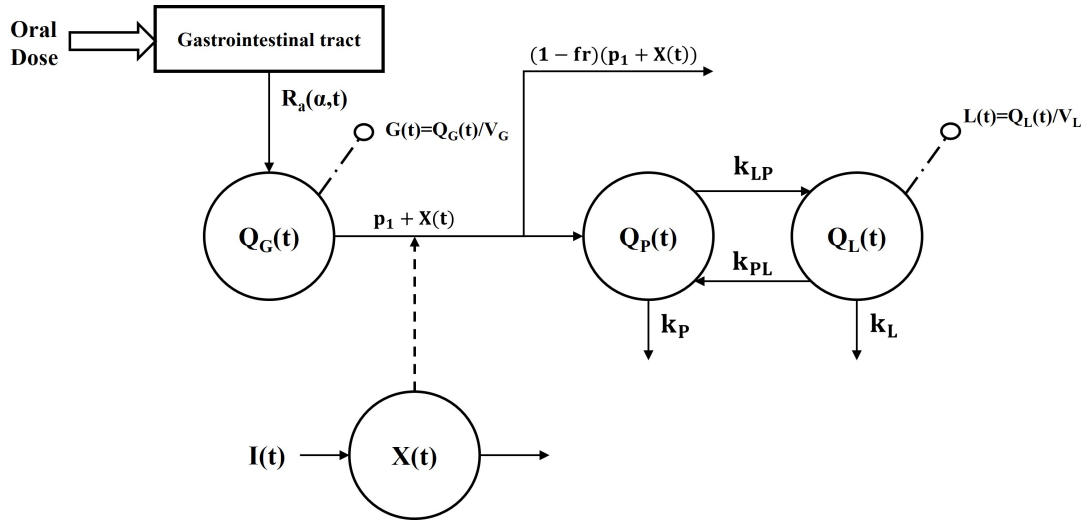

Figure 3.3: Schematic representation of Watanabe et.al. model [36] which was adapted for the OGTT setting by integration the of the glucose rate of appearance as in Dalla Man et.al. [40]. Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. Q_G , Q_P and Q_L are glucose, pyruvate and lactate plasma mass respectively; G(t), L(t) and I(t) are the plasma glucose, lactate and insulin concentrations respectively; X(t) is the delayed insulin effect on both glucose production and utilization; $R_a(\alpha,t)$ is the glucose rate of appearance; *fr* is the faction of glucose converted into lactate; p_1 is the fractional glucose effectiveness; k_{LP} and k_{PL} are the rate constant of pyruvate-lactate and lactate-pyruvate conversion respectively; k_P and k_L are the rate constant of pyruvate and lactate plasma clearance respectively; V_G and V_L are the glucose and lactate plasma volumes respectively.

3.1.4 Model 4

Differently from Model 2 and 3, in which X(t) represented the net insulin effect on both glucose production and utilization, these two components were considered separately in Model 4. To do that, the model was based on the STOMM [41], which was originally identified on glucose tracer data, making it possible the distinction of the insulin effect into its two components. The model equations are the following:

$$\begin{cases} \dot{Q}_G(t) = -[p_{1D} + X_D(t)] \cdot Q_G(t) + EGP(t) + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}_D(t) = -p_{2D} \cdot X_D(t) + p_{2D} \cdot SI'_D \cdot [I(t) - I_b] & X_D(0) = 0 \\ \dot{Q}_L(t) = -k_L \cdot Q_L(t) + 2 \cdot fr \cdot [p_{1D} + X_D(t)] \cdot Q_G(t) & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.11}$$

where, differently from Model 3 and 2, p_{2D} (min^{-1}) and SI'_D (1/min per μU/mL) represented, respectively, the rate constant of the disposal insulin effect, X_D (t), and the fractional disposal insulin sensitivity (i.e. per unit distribution volume). It is worth noting that p_{1D} (fractional disposal glucose effectiveness) relationship with SI'_D was the same as the one reported in Eq. 3.7. The lactate production rate (LPR), that is the best proxy of glycolytic flux, was then defined as:

$$LPR(t) = 2 \cdot fr \cdot [p_{1D} + X_D(t)] \cdot Q_G(t) \tag{3.12}$$

The model incorporates a description of the endogenous glucose production (EGP) as function of plasma glucose concentration and insulin action:

$$\begin{cases} EGP(t) = EGP_b - p_3 \cdot [G(t) - G_b] - X_L(t) - X_{der}(t) \\ EGP_b = V_G \cdot p_{1D} \cdot G_b \\ \dot{X}_L(t) = -p_{2P} \cdot X_L(t) + p_{2P} \cdot X_P(t) & X_L(0) = 0 \\ \dot{X}_P(t) = -p_{2P} \cdot X_P(t) + p_{2P} \cdot p_4 \cdot [I(t) - I_b] & X_P(0) = 0 \\ X_{der}(t) = \begin{cases} k_{GR} \cdot \dot{G}(t) \text{ if } \dot{Q}_G(t) > 0 \\ 0 \text{ if } \dot{Q}_G(t) \leq 0 \end{cases} \end{cases} \tag{3.13}$$

X_L represents the insulin effect on EGP and, delayed with respect to plasma insulin by a second order model with rate constant p_{2P} (min^{-1}), X_{der} is the control action of glucose rate of change, through k_{GR} (dL/kg), on EGP, and p_3 (dL/kg/min) represents the sensitivity to glucose.

This model structure guarantees to distinguish the insulin action on glucose production from that on glucose utilization, and to estimate the liver and disposal insulin sensitivities, SI_L and SI_D (dL/kg/min per μU/mL), calculated from model parameters as:

$$\begin{cases} SI_L = \frac{p_4}{G_b} \\ SI_D = SI'_D \cdot V_G \end{cases} \tag{3.14}$$

A schematic representation of Model 4 is reported in Figure 3.4

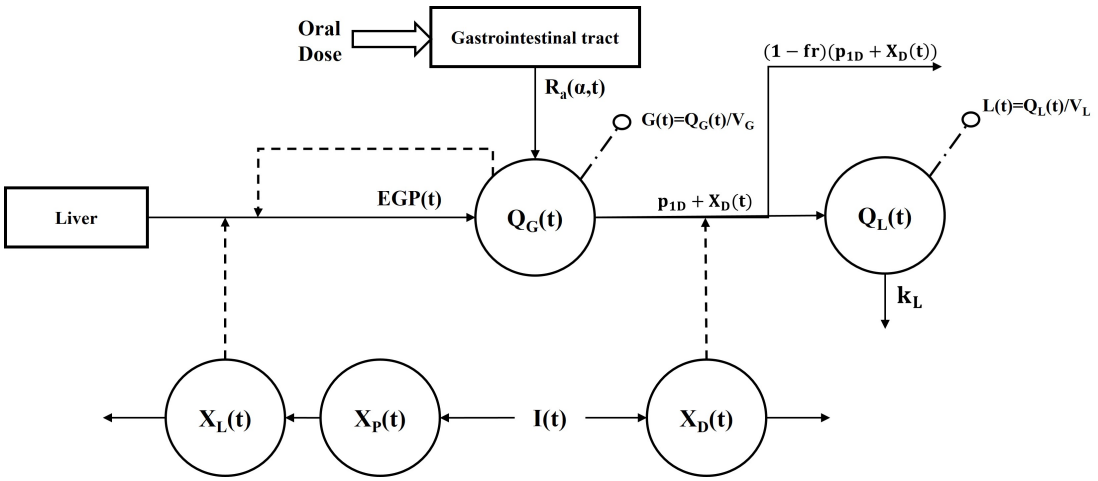

Figure 3.4: Schematic representation of the STOMM [41] adapted to integrate lactate kinetic description. Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. Q_G, and Q_L are glucose and lactate plasma mass respectively; G(t), L(t) and I(t) are the plasma glucose, lactate and insulin concentrations respectively; $X_D(t)$, $X_P(t)$ and $X_L(t)$ are the delayed insulin effect on glucose disposal, production and liver respectively; $R_a(\alpha,t)$ is the glucose rate of appearance; EGP(t) is the endogenous glucose production; p_{1D} is the fractional disposal glucose effectiveness; *fr* is the fraction of glucose converted into lactate; k_L is the rate constant of lactate plasma clearance; V_G and V_L are the glucose and lactate plasma volumes respectively.

3.1.5 Model 5

The structure of Model 5 was the same of Model 4, but a control action on lactate compartment influx was added, following a Langmuir type of control (i.e. lactate controls on glycolysis as proposed in [11]):

$$\begin{cases} \dot{Q}_G(t) = -[p_{1D} + X_D(t)] \cdot Q_G(t) + EGP(t) + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}_D(t) = -p_{2D} \cdot X_D(t) + p_{2D} \cdot SI'_D \cdot [I(t) - I_b] & X_D(0) = 0 \\ \dot{Q}_L(t) = \begin{cases} -k_L \cdot Q_L(t) + 2 \cdot (1 - \frac{L(t)}{\beta}) \cdot [p_{1D} + X_D(t)] \cdot Q_G(t) \text{ if } L(t) < \beta \\ -k_L \cdot Q_L(t) \text{ if } L(t) \geq \beta \end{cases} & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.15}$$

In other words, depending on lactate plasma concentration, if it overcomes

a threshold, β (mg/dL), glucose is no longer converted into lactate, and it follows other metabolic paths. EGP description is the same as reported in Eq. 3.13. A schematic representation of Model 5 is reported in Figure 3.5

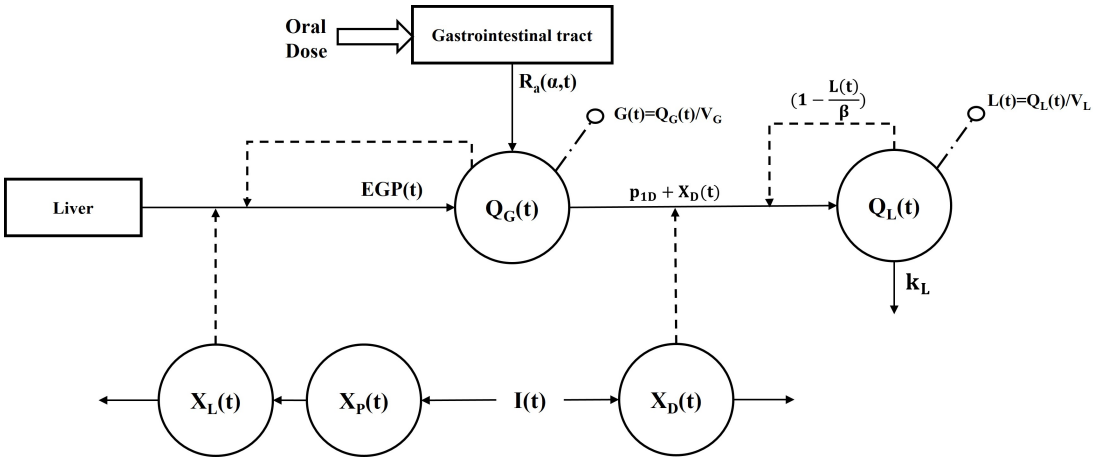

Figure 3.5: Schematic representation of the STOMM [41] adapted to integrate lactate kinetic description with an additional control by lactate plasma concentration on its own production. Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with white circle indicate the measured plasma concentrations. Q_G, and Q_L are glucose and lactate plasma mass respectively G(t), L(t) and I(t) are the plasma glucose, lactate and insulin concentrations respectivel; $X_D(t)$, $X_P(t)$ and $X_L(t)$ are the delayed insulin effect on glucose disposal, production and liver respectively; $R_a(\alpha,t)$ is the glucose rate of appearance; EGP(t) is the endogenous glucose production; p_{1D} is the fractional disposal glucose effectiveness; β is the threshold above which glucose is no more converted into lactate; k_L is the rate constant of lactate plasma clearance¸ V_G and V_L are the glucose and lactate plasma volumes respectively.

3.1.6 Model 6

Model 6 integrated the lactate-pyruvate kinetic description of Model 3, with the glucose, EGP and insulin effects as Model 4. As already discussed, the pyruvate-lactate two-compartment model was reparametrized, leading to:

$$\begin{cases} \dot{Q}_G(t) = -[p_{1D} + X_D(t)] \cdot Q_G(t) + EGP(t) + R_a(\alpha,t) & Q_G(0) = V_G G_b \\ \dot{X}_D(t) = -p_{2D} \cdot X_D(t) + p_{2D} \cdot SI'_D \cdot [I(t) - I_b] & X_D(0) = 0 \\ \ddot{Q}_L(t) = -A_1 \cdot \dot{Q}_L(t) - A_2 \cdot Q_L(t) + A_3 \cdot [p_{1D} + X_D(t)] \cdot Q_G(t) & Q_L(0) = V_L L_b \\ G(t) = \frac{Q_G(t)}{V_G} \\ L(t) = \frac{Q_L(t)}{V_L} \end{cases} \tag{3.16}$$

with A1, A2 and A3 the same as Eq. 3.10. A schematic representation of Model 6 is reported in Figure 3.6.

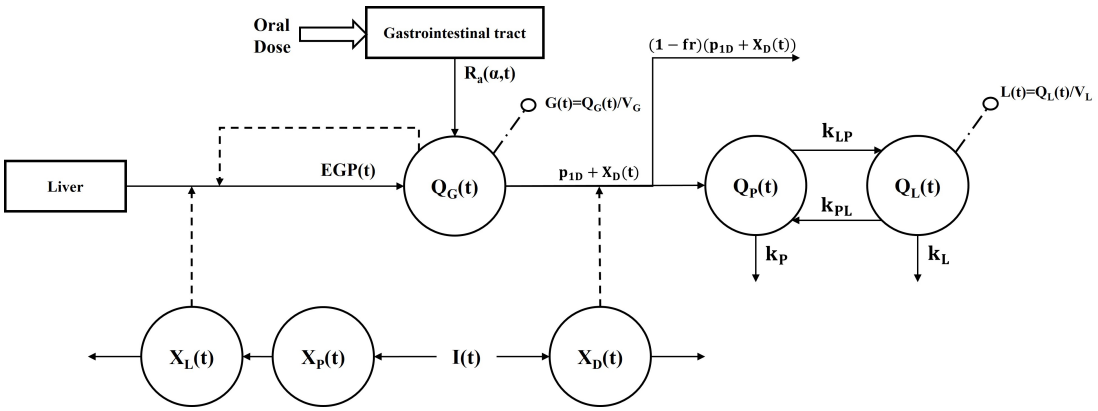

Figure 3.6: Schematic representation of the STOMM [41] combined with lactate kinetic description as proposed by Watanabe et.al. [36]. Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with white circle indicate the measured plasma concentrations. Q_G, Q_P and Q_L are glucose, pyruvate and lactate plasma mass respectively; G(t), L(t) and I(t) are the plasma glucose, lactate and insulin concentrations respectively; $X_D(t)$, $X_P(t)$ and $X_L(t)$ are the delayed insulin effect on glucose disposal, production and liver respectively; $R_a(\alpha,t)$ is the glucose rate of appearance; EGP(t) is the endogenous glucose production; *fr* is the fraction of glucose converted into lactate; p_{1D} is the disposal fractional glucose effectiveness; k_{LP} and k_{PL} are the rate constant of pyruvate-lactate and lactate-pyruvate conversion respectively; k_P and k_L are the rate constant of pyruvate and lactate plasma clearance respectively; V_G and V_L are the glucose and lactate plasma volumes respectively.

3.2 Model identification and assessment

In this paragraph, it is described how the models were identified, and all the assumptions used to make them identifible. In the *A priori identifiability* section, the *a priori identifiability* of the models is discussed; in *Parameter estimation* section, the main assumptions and method used for parameters estimation (*a posteriori identifiability*) are presented. In the *Model assessment* section, the methods used to evaluate model performance along with the criteria used to select the best performing one are reported. Finally, in the *Statistical Analysis*, the statistical tests used to compare models and to assess parameter correlations are described.

3.2.1 A priori identifiability

The first step in the model development was the assessment of its *a priori identifiablity*. This was necessary to exclude non-identifiable models, and eventually reparametrized the model to make it identifiable. In fact, *a priori identifiability* tests if, from a theoretical point of view, given a model structure and an input-output configuration it is possible to determine the model parameters from the data. In particular, a model is globally identifiable if the model parameters can be uniquely determined from the available data, while it is locally identifiable if a finite number of solutions exists. In the other case the model is non-identifiable, i.e. an infinite number of solutions for model parameters exists.

The assessment of model *a priori identifiablity* was done using the DAISY software [56], which implements an algorithm based on differential algebra.

Particularly, the models described in paragraph 3.1 were not *a priori identifiable*, thus several assumptions were done as previously proposed in [40][41]. In fact,

to ensure the *a priori identifiability*, we assumed that both glucose and lactate volume of distribution (V_G and V_L respectively) were equal to the population value of 1.45 dL/kg as in [40][41]. Other than this, for the models which were based on [40] and [41], p_1 or p_{1D} parameters were described as in Equation 3.7. Finally, as already presented in 3.1.3 and 3.1.6, beyond fixing both V_G and V_L, the models were reparametrized as in the original work [36] to make them identifiable.

3.2.2 Parameter estimation

The model parameters were estimated by using the Bayesian maximum a posteriori (MAP) estimator, with the primary objective of enhancing model *a posteriori identifiability*.

A Gaussian Bayesian prior was selectively applied to the parameters for whom literature values were available, such as the one in [40][41] for adult subjects. In particular this information, defined as mean (μ_p) and covariance (Σ_p), was incorporated into the estimation process (Equation 3.18 and Table 3.1).

Regarding the other parameters, we applied non-informative priors, allowing them to be fully quantified from the data. This helped the estimation process, while keeping the freedom of the model parameters to adapt to the specific population, in line with literature that highlights possible differences between populations [57][58].

Specifically, for Model 1, no *a priori* knowledge was available, so parameters were identified without the use of *a priori* information, relying solely on the data. In contrast, for Models 2 and 3 (based on OGMM [40]), a Gaussian prior was applied to the parameters GEZI and p_2, with means set equal to the one reported in [40] and standard deviations assumed to be 20% of their respec-

tive means, as done in previous works [59]. Similarly, for Models 4-6 (based on STOMM [41]), Gaussian priors were applied to the parameters p_{2D}, $GEZI_D$, p_{2P}, p_3 and k_{GR}, with means from STOMM, and standard deviations set to 20% of the means. Notice that, at variance with [41], here the meal was unlabeled. Despite that, Models 4-6 still allowed for the segregation of glucose production and disposal by leveraging lactate measurements, which reflected glucose utilization.

Finally, for parameters where no *a priori* knowledge was assumed, non informative priors were used, setting variances to infinity and means to zero.

All these assumptions were used to thus define the mean model parameters vector (μ_p) and parameters covariance matrix (Σ_p), see Table 3.1 for further details.

Table 3.1: Model parameter *a priori* knowledge from the literature

Model	Parameters	μ_p	SD	Reference
1	No assumptions			
2-3	GEZI	0.036 mg/kg/min	0.007 mg/kg/min	[40]
	p_2	0.01 min^{-1}	0.002 min^{-1}	
4-5-6	p_{2D}	0.025 min^{-1}	0.005 min^{-1}	[41]
	$GEZI_D$	0.021 mg/kg/min	0.004 mg/kg/min	
	p_{2P}	0.020 min^{-1}	0.004 min^{-1}	
	p_3	0.017 dL/kg/min	0.003 dL/kg/min	
	k_{GR}	0.121 dL/kg	0.024 dL/kg	

Mean population values (μ_p) and standard deviations (SD) used in each tested minimal model. Notice that for these parameters, the standard deviations were assumed to be 20% of their respective μ. Parameters not reported in this table were assumed to have a mean value equal to zero and infinite variance.

Each model was identified on both plasma glucose and lactate data, simultaneously. For both, measurement errors were assumed to be independent, Gaussian, with zero mean and SD dependent on glucose and lactate concen-

tration respectively using a constant coefficient of variation (CV) of 2%, as in [40][60]. This assumption was used to define the measurement error covariance matrices (Σ_G and Σ_L):

$$\Sigma_G = \begin{bmatrix} (0.02G_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (0.02G_{NG})^2 \end{bmatrix}$$
$$\Sigma_L = \begin{bmatrix} (0.02L_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (0.02L_{NL})^2 \end{bmatrix} \tag{3.17}$$

where G and L indicate glucose and lactate plasma concentrations respectively, N_G and N_L are the glucose and lactate number of samples respectively, while insulin was used as model forcing function so assumed to be known without error. In all the cases, following the MAP approach and the hypothesis previously described on data measurement error and prior distributions, the objective function (OBJ) to be minimized was the following:

$$\begin{aligned} OBJ(\hat{p}) = (Y_G - \hat{Y}_G)^T \cdot \Sigma_G^{-1} \cdot (Y_G - \hat{Y}_G) + (Y_L - \hat{Y}_L)^T \cdot \Sigma_L^{-1} \cdot (Y_L - \hat{Y}_L) \\ + (\hat{p} - \mu_p)^T \cdot \Sigma_p^{-1} \cdot (\hat{p} - \mu_p) \end{aligned} \tag{3.18}$$

where Y and $\hat{Y}$ were the experimental data and the model prediction vectors, respectively, with subscripts G and L indicating the glucose and lactate plasma concentrations; Σ_G and Σ_L were the covariance matrices of glucose and lactate measurement errors (Equation 3.17); $\hat{p}$ was the parameter vector to be estimated; μ_p its expected value and Σ_p its covariance matrix.
Numerical identification was performed in Matlab R2022b [61], using the *ode45()* solver to integrate model equations and the *lqnonolin()* built-in function to determine model parameters minimizing the equation 3.18.

3.2.3 Model assessment

The model assessment consisted in testing if the model is suitable for the purpose, i.e. to well describe the data with the minimum number of physiologically meaningful and precisely estimated model parameters. To do this, the following criteria needs to be met: the normality [62] and randomness [63] of the weighed residuals; the ability of the model to fit the data, assessed by means of the weighted residuals sum of squares (WRSS), as well as visual inspection; the ability to provide precisely estimated parameters (*a posteriori identifiability*) assessed by means of the median [25th-75th] percentile of percentage coefficient of variation (CV%) and the percentage of subjects with at least one parameter estimated with CV%>100% (N_{CV}). The models which passed the previous tests were compared by means of a parsimony criterion, i.e. the Bayesian information criterion (BIC), defined as:

$$\mathrm{BIC} = 2ln(\mathrm{WRSS}) + \mathrm{M}ln(\mathrm{N_G} + \mathrm{N_L}) \tag{3.19}$$

where M was the the number of model parameters, while $\mathrm{N_G}$ and $\mathrm{N_L}$ were the number of glucose and lactate data points, respectively.
Once the best model was selected, it was used to evaluate: firstly, the impact of protocol duration on parameter estimates and, secondly, to determine if the obesity degree has an impact on the parameter estimates..

Impact of protocol duration

As already reported, the glucose-lactate minimal model was developed using subjects of Dataset 1, which underwent a 4h OGTT.
To assess the impact of the protocol duration, the selected model was identified on the same subjects but using only the first 3h of the OGTT.
The key parameters and LPR time course estimated from the 4h-OGTT were

compared with those obtained with the 3h-OGTT. The results of this analysis are reported in 3.3.2.

Model use in individuals with and without obesity

After the evaluation of the protocol duration impact, the glucose-lactate minimal model was applied to Dataset 2, which contained also lean subjects (i.e. 9 subjects without obesity and 15 subjects overweight/with obesity). This allowed us to compare model parameters and fluxes in subjects with and without obesity. These results are presented in paragraph 3.3.3.

3.2.4 Statistical analysis

Data and results are reported as median [25th -75th], if not normally distributed, or mean±SD, if normally distributed. Paired and unpaired comparisons were done by using Student-T test/Wilcoxon test, or unpaired T test/ Mann Whitney test respectively, depending on distributions (assessed using Shapiro-Wilk test [64]). Two groups non-significant comparisons were further analyzed using equivalence and non-inferiority tests, which aim to statistically demonstrate that, within a specified reasonable range of values (called the equivalence or non-inferiority margin), the groups are either practically equivalent or one is not inferior to the other. Null (H0) and alternative (H1) hypothesis, in case of Equivalence and Non-inferiority tests, are:

$$\begin{cases} \text{Equivalence} & \text{H0} : |\mu_1 - \mu_2| > \Psi_{\text{INF}}; \text{H1} : |\mu_1 - \mu_2| \leq \Psi_{\text{SUP}} \\ \text{Non-Inferiority} & \text{H0} : \mu_1 - \mu_2 < -\Psi_{\text{INF}}; \text{H1} : \mu_1 - \mu_2 \geq -\Psi_{\text{INF}} \end{cases} \tag{3.20}$$

where μ_1 and μ_2 are the mean parameter values of the two groups, Ψ_{INF} and Ψ_{SUP} are the equivalence margins of the equivalence test, while Ψ_{INF} is the non-inferiority margin of the non-inferiority test. Here, when performing the

equivalence analysis, instead of defining the margins as expected differences, we used the log-transformed two one-sided test (TOST) method [65][66], where the margins were defined as the ratio between values. For paired comparisons, we used a narrow margin of 0.7 to 1.3, corresponding to a ±30% range from the no effect (ratio = 1). This margin was chosen because we assumed that, in paired samples, the overall variability is low, which allows for a smaller margin of equivalence. In contrast, for unpaired samples, we defined the margins using a more flexible approach based on the pooled coefficient of variation (CV), where the margin was set to 1 ± 0.5 * CV and the CV was derived from the pooled standard deviation and mean. This approach accounts for the larger variability in unpaired samples, where the groups may differ more inherently, leading to a wider margin of equivalence. This also applies to the non-inferiority test, except for defining a single margin (i.e. non-inferiority margin) instead of two. Correlation analysis was performed using Pearson or Spearman correlation coefficients depending on the linearity/nonlinearity of the relationship. Significance level was assumed as α=0.05, while for two-sided and one-sided tests, α was set equal to 0.05 and 0.025 respectively. The analysis was conducted in Matlab R2022b [61] and in R (version 4.3.2) [67].

3.3 Results

In this paragraph, the results of the selection of the best minimal model, the impact of protocol duration and of obesity degree on model parameters are shown.
In 3.3.1 the minimal model selection is described. Particularly, the technique described in 3.2.3 were used to select the best performing model.
In 3.3.2 the impact of the protocol duration is discussed. In particular, the LPR time courses and correlation analysis between parameters estimated using the 4h OGTT were compared to the ones estimated using the 3h OGTT.
In 3.3.3 the application of the selected minimal model to Dataset 2 is reported, along with the comparisons between subjects with/without obesity in terms of data and estimated LPR AUCs and model parameters. Also, the comparison with literature adults' parameters is discussed, where possible.

3.3.1 Model selection

In this section the model selection phase, using the techniques reported in 3.2.3 is described. Only subjects of Dataset 1 for which data were available for more than 3h were used to develop the model (i.e. 16 subjects). For each model, in Table 3.2 and Figures 3.7-3.8 are reported the results of the model assessment procedure. After this phase, the selected minimal model was applied to all the available subjects of Dataset 1 (Figure 3.9).
Model 1 was discarded due to poor randomness and normality of the residuals, with 37% of the subjects failing both tests. It is also worth noting that Model 1 was characterized by the worst WRSS, probably due to the absence in the model of an insulin effect on glucose kinetics, leading to unsatisfactory glucose and lactate predictions.

Model 5 was characterized by the same problem of Model 1, again with 37% of the subjects failing the normality test on the residuals. It is worth noting that Model 5 had also a poor precision of the estimated parameters (i.e. 81% of the subjects had at least one parameter estimated with CV%>100%). Specifically, the β parameter (representing lactate's control over its own production) was difficult to estimate, probably because lactate levels did not increase enough during the test compared to exercise conditions or after a lactate infusion, making it harder to determine such a threshold.

Similarly, Model 2 was discarded since it was the worst in terms of randomness of the residuals compared to the remaining models. To note also that, in 25% of the subjects, SI was estimated with poor precision, and this issue likely arose from considering a total insulin effect like in the OGMM, without distinguishing between insulin effect on glucose production and disposal, with the second one being more impactful on lactate production during post-prandial conditions.

The remaining three models (i.e. Model 3, 4 and 6), which were based on the STOMM structure, were not different in terms of WRSS (P=0.35). Nevertheless, both Model 3 and Model 6 showed unsatisfactory *a posteriori identifiability* since 56% and 50% of the subjects had at least one parameter with CV%>100% respectively. Both Model 3 and 6 shared the same two-compartment description of lactate kinetic based on [36], and this reduced the precision in some of the model parameter estimates.

Model 4 was selected as the best one since it provides good model performance, both in terms of model fit and *a posteriori identifiability*.

Table 3.2: Minimal Model selection

Model	Run Test	Normality Test	WRSS	CV%	N_{CV}	BIC
1	63%	63%	1095 [441-1755]	7 [6-8]%	13%	38.9 [37.3-39.9]
2	69%	88%	490 [291-941]	9 [7-11]%	25%	47.8 [46.5-48.8]
3	75%	82%	388 [294-895]	8 [7-19]%	56%	51.0 [46.9-51.7]
4	**94**%	**82**%	**349 [262-371]**	**13 [11-15]**%	**6**%	**61.4 [59.9-61.6]**
5	88%	63%	326 [254-600]	14 [12-16]%	81%	61.5 [59.9-62.0]
6	88%	88%	322 [266-535]	13 [12-16]%	50%	68.5 [65.8-68.8]

Data are reported as median and [25th-75th] percentile or percentages. The bold values highlight the best-selected model. WRSS: weighted residual sum of squares; N_{CV}: percentage of subjects with at least one parameters estimated with CV>100%; BIC: Bayesian information criterion.

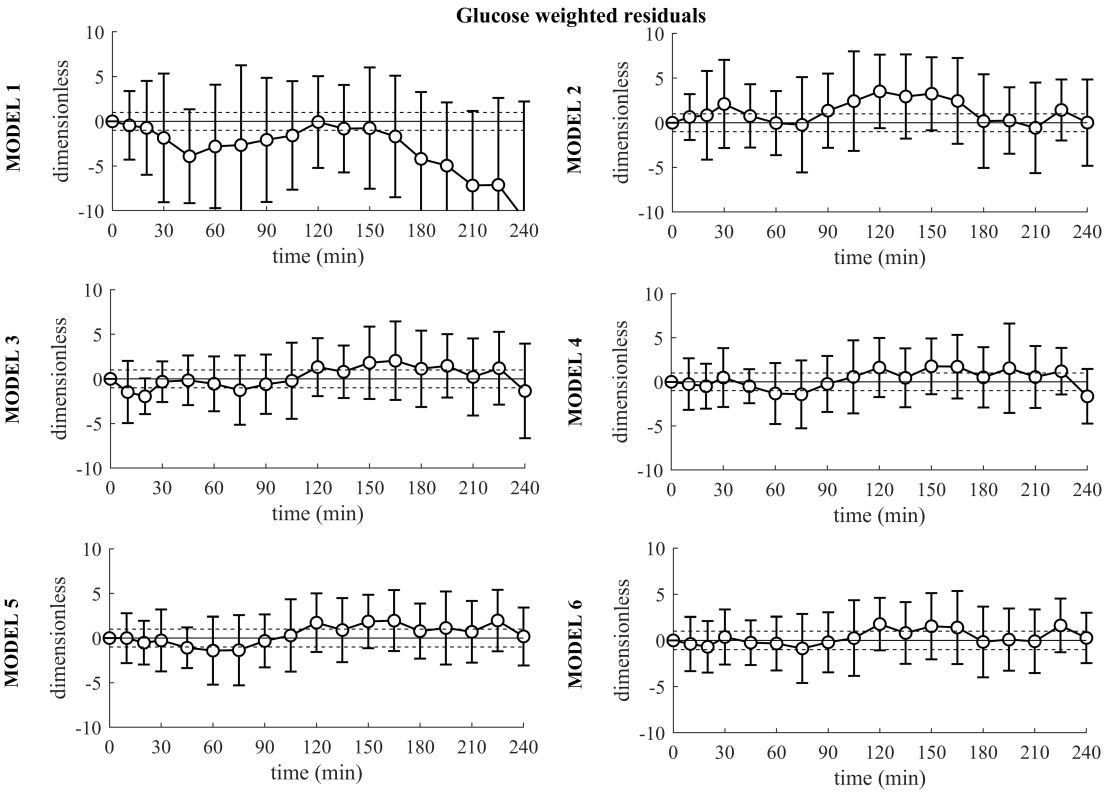

Figure 3.7: Mean±SD glucose weighted residuals (circle with vertical bars) for each tested model. Dashed horizontal lines indicate [-1: 1] range

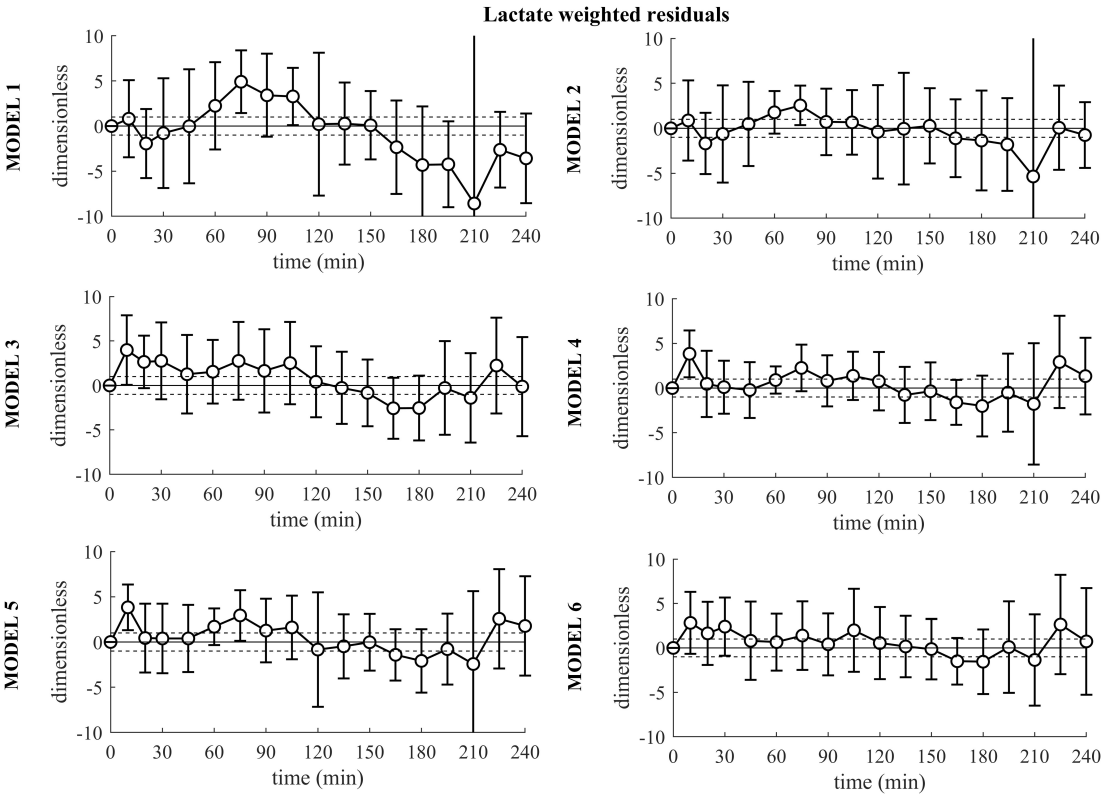

Figure 3.8: Mean±SD lactate weighted residuals (circle with vertical bars) for each tested model. Dashed horizontal lines indicate [-1: 1] range.

Model 4 was then applied to all the available subjects of Dataset 1, providing the glucose and lactate predictions reported in Figure 3.9, while model estimated parameters are reported in Table 3.3.

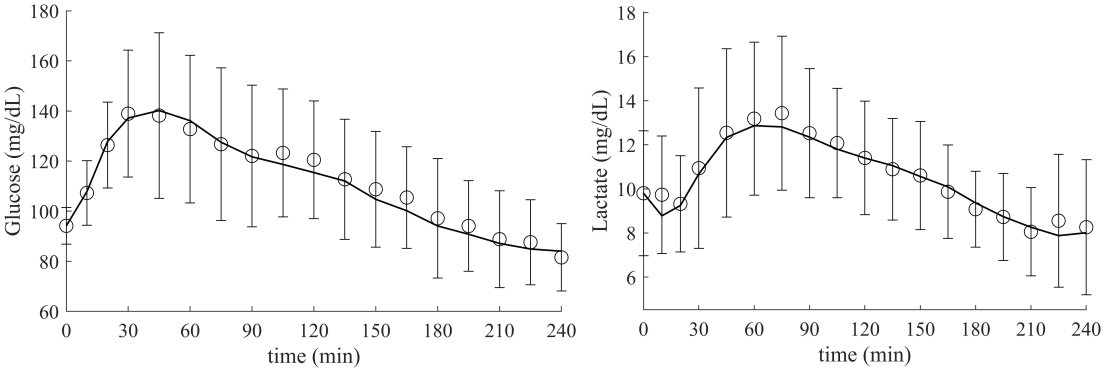

Figure 3.9: Mean±SD experimental data (circles and vertical bars) vs. mean model predictions (continuous lines) for glucose (left) and lactate (right).

Since Dataset 1 contains also data collected during a 3h-OGTT, while the

model was developed using only the data from a 4h-OGTT, in the next paragraph, the impact of protocol duration on model parameter estimates is discussed.

Table 3.3: Glucose-Lactate Minimal model parameter estimates (Dataset 1)

Parameter	Estimate	Precision
Glucose rate of appearance		
α_1 (mg/kg/min)	4 [4-6]	7 [6-9]%
α_3 (mg/kg/min)	3 [2-4]	10 [8-15]%
α_4 (mg/kg/min)	4 [3-4]	2 [2-3]%
α_5 (mg/kg/min)	2 [2-3]	10 [5-13]%
Lactate kinetic		
fr (%)	26 [17-37]	16 [13-19]%
k_L (min^{-1})	0.058 [0.044-0.087]	8 [7-9]%
Glucose disposal		
p_{2D} (min^{-1})	0.031 [0.022-0.038]	10 [8-15]%
SI_D ($dL/kg/min/\mu U/mL\ x10^{-5}$)	5.1 [2.0-12]	21 [18-28]%
$GEZI_D$ (dL/kg/min)	0.021 [0.018-0.026]	13 [10-15]%
Endogenous glucose production		
p_3 (dL/kg/min)	0.017 [0.016-0.020]	14 [12-15]%
p_{2P} (min^{-1})	0.021 [0.015-0.024]	12 [10-15]%
p_4 ($mg/kg/min/\mu U/mL$)	0.040 [0.023-0.053]	12 [8-15]%
k_{GR} (dL/kg)	0.120 [0.118-0.121]	14 [13-14]%

Median [25th -75th] of estimated model parameters (second column) and their precision (third column).

3.3.2 Impact of protocol duration

As already discussed, since 16 subjects of Dataset 1 were studied for 4h, the parameters and LPR time courses were compared with their 3h estimated counterparts. It is also to note that the 3h protocol is the standard one for the OGTTs, so an assessment of the model ability to perform well even with protocol duration shorter than 4h was necessary.

Figure 3.10 shows the comparison between the estimated LPR (Equation 3.12) considering the entire OGTT duration with the estimation provided using only the first 180min of the OGTT. The two were comparable both in terms of mean±SD time courses (Fig 3.10) and AUC: 340.4 [256.3-572.2] and 337.9 [297.4-468.4]) mg/kg, P=0.53, in the 4h-OGTT and 3h-OGTT, respectively. AUC was calculated between 0-180min for both. The equivalence test was performed as described previously (see Paragraph 3.2.4), using the estimates of the 4h protocol as reference group and obtaining a P = 0.0005.

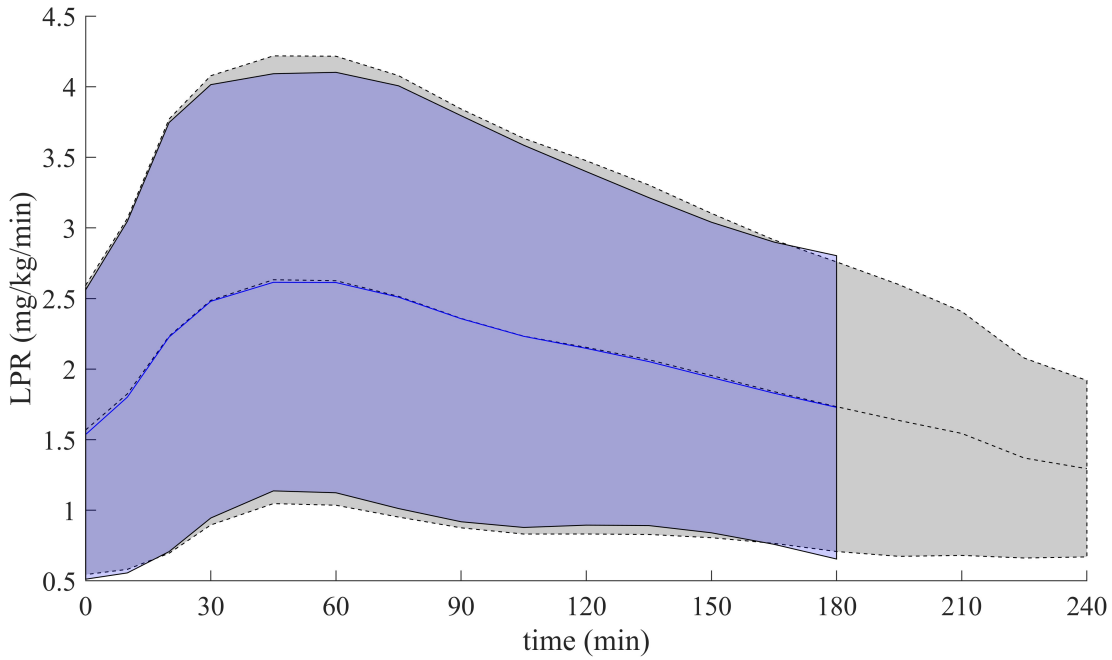

Figure 3.10: Mean±SD LPR time courses estimated using a protocol of 240min (grey dashed line and grey area) vs a protocol of 180min (blue continuous line and purple area).

Concerning model parameters estimates, results can be found in Table 3.4. As previously pointed out, the comparisons were made only between parameters SI_L, SI_D, *fr* and k_L, for which no *a priori* knowledge was assumed. No statistically significant differences were found between the estimates obtained using the two protocols (Wilcoxon test, P=0.9, 0.8, 0.6, 0.6 ; equivalence test, P= 0.0116, 0.0412, 0.0063, 0.0006 for for SI_L, SI_D, *fr* and k_L respectively), concluding that the observed difference is small enough to be considered practically irrelevant within the context of the equivalence margin (see Paragraph 3.2.4). Moreover, the correlations between the estimates were always statistically significant and very good. Interestingly, the SI_L estimates had the lowest correlation coefficient (ρ=0.58) between the two protocols.

Table 3.4: Key minimal model parameters comparisons using 4h and 3h OGTT protocol

Parameter	240min estimate	180min estimate	Correlation
SI_L (dL/kg/min/μU/mL x10^{-4})	4.2 [2.9-5.6]	4.2 [3.3-5.9]	0.58 (P=0.02)
SI_D (dL/kg/min/μU/mL x10^{-5})	5.1 [1.8-12]	3.9 [1.7-11]	0.93 (P<$10^{-}4$)
fr (%)	25 [17-37]	30 [17-45]	0.92 (P<10^{-4})
k_L (min^{-1})	0.06 [0.04-0.09]	0.06 [0.05-0.09]	0.95 (P<10^{-4})
LPR-AUC (mg/kg)	340.4 [256.3-572.2]	337.9 [297.4-468.4]	0.96 (P<10^{-4})

Key glucose-lactate minimal model parameters with their median [25th -75th] estimation during 240min (first column) and 180min (second column) protocols. In the last column the Pearson correlation coefficient between the estimates and its P-value (P) are reported. No statistically significant differences (assessed using Wilcoxon test) emerged between the estimates obtained with the two protocols. Note that, for all the variables, the equivalence test yielded to significant results.

3.3.3 Impact of obesity degree and age

After the model selection phase and the assessment of the protocol duration impact, the minimal model was applied to a different dataset (Dataset 2), characterized by a shorted protocol duration (i.e. 180min OGTT), and containing twenty-four subjects, some of which were lean adolescents. This allowed us to firstly assess the model performance on a new data set and evaluate its ability to describe glucose and lactate data across different populations and, secondly, to understand if obesity shows any effect on glucose-lactate kinetics as assessed from model parameters.

The model successfully predicted both glucose and lactate time courses (Figure 3.11), providing the parameter estimates reported in Table 3.5. WRSS was 150 [116-272] in lean subjects and 171 [132-220] in subjects overweight/with obesity respectively (Mann-Whitney test P=0.86, equivalence P= 0.047), confirming the ability of the model to describe the system in different populations. In Table 3.5, the parameters estimated in the Database 2 are reported.

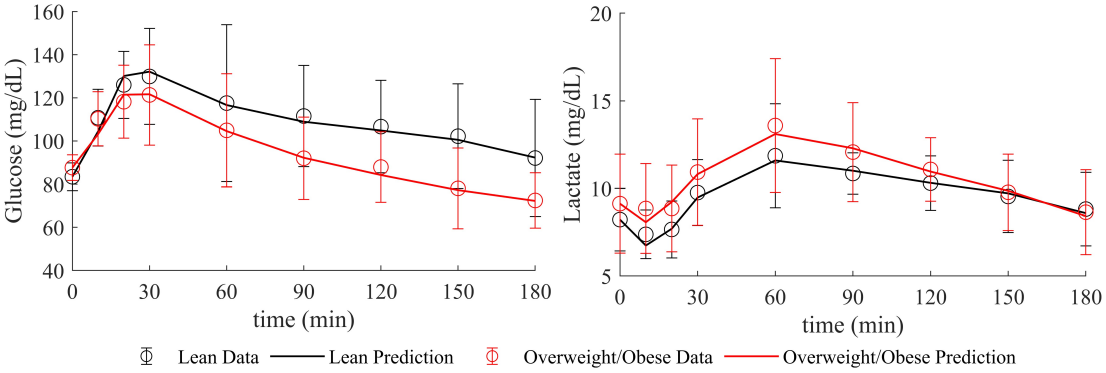

Figure 3.11: Mean±SD experimental data (circles and vertical bars) vs. mean model predictions (continuous lines) for glucose (left) and lactate (right) in lean (black) and overweight/obese adolescents (red).

Table 3.5: Minimal model parameter estimates (Dataset 2)

Parameter	Estimate	Precision
Glucose rate of appearance		
α_1 (mg/kg/min)	6 [4-9]	6 [5-8] %
α_3 (mg/kg/min)	6 [4-7]	5 [4-7]%
α_4 (mg/kg/min)	6 [4-7]	3 [3-5]%
α_5 (mg/kg/min)	3 [2-5]	9 [6-13]%
Lactate kinetic		
fr (%)	22.5 [12.6-54.1]	18 [13-25]%
k_L (min^{-1})	0.076 [0.053-0.117]	13 [11-16]%
Glucose disposal		
p_{2D} (min^{-1})	0.023 [0.015-0.036]	10 [8-13] %
SI_D ($dL/kg/min/\mu U/mL\ x10^{-4}$)	6.9 [1.7-11.1]	14 [8-22]%
$GEZI_D$ (dL/kg/min)	0.025 [0.010-0.030]	10 [8-19]%
Endogenous glucose production		
p_3 (dL/kg/min)	0.017 [0.016-0.018]	14 [13-15]%
p_{2P} (min^{-1})	0.022 [0.020-0.027]	12 [10-13]%
p_4 ($mg/kg/min/\mu U/mL$)	0.055 [0.047-0.064]	11 [9-12]%
k_{GR} (dL/kg)	0.121 [0.119-0.122]	14 [14-15]%

Median [25th -75th] of estimated model parameters (second column) and their precision (third column).

Then, all the available subjects with/without obesity were compared in terms of substrate time courses and estimated model parameters (Table 3.6). First, no statistically significant differences emerged on glucose and lactate over-basal Area Under the Curve (AUC). Moreover, both the equivalence and non-inferiority test null hypothesis were accepted, assuming lean subjects as the reference group. Nevertheless, insulin over-basal AUC was significantly higher in subjects with obesity compared to leans (P=0.034), while the LPR over-basal AUC was significantly lower in the first group compared to the second (P=0.006), as reported in Table 3.6. Interestingly, the total glucose, lactate and LPR AUCs were not significantly different (but only for LPR AUC the equivalence was significant with P= 0.045), while insulin AUC was, again, higher in subjects with obesity (P=0.007). These results were confirmed by the

differences in model parameters. In fact SI_D and SI_L were lower in subjects with obesity compared to lean, with SI_D equal to 1.2 [0.3-2.8] vs 5.9 [1.2-9.1] dL/kg/min/μU/mL x10^{-4}, P=0.036, respectively; and SI_L equal to 4.7 [3.7-6.5] vs 6.9 [6.2-10.0] dL/kg/min/μU/mL x10^{-4}, P=0.003, respectively.

These results highlight that in subjects with obesity, despite the higher insulin exposure, the insulin resistance state limited the capability of releasing lactate during post-prandial conditions. However, subjects with obesity showed no differences in total LPR AUC with respect to lean, meaning that, the basal glycolytic metabolism is enhanced in those subjects, as already observed in [68]. However, the limited sample size may have hindered the detection of differences in lactate-related parameters between the two groups. In fact, a post-hoc sample size analysis indicated that a total of 218 subjects would be required to achieve a statistical power of 0.8 for the parameter k_L. Nonetheless, the non-inferiority test for this parameter was significant (P=0.001), demonstrating that, using the specific margins defined in 3.2.4, lean individuals have a fractional lactate clearance rate that is at least not inferior to that of individuals with obesity.

Secondly, we found correlations between lactate-related model parameters and some anthropometric characteristics. In particular, in subjects with obesity, we found that SI_D was significantly correlated with both BMI (ρ= -0.38, P=0.031), and *fr* (ρ= -0.46, P=0.009). Although no significant correlation appeared between basal lactate concentration and BMI, *fr* was positively associated with BMI (ρ= 0.40, P=0.025). These results extended what was previously obtained in [55]. Nevertheless, in lean subjects, only the association between SI_D and *fr* was was close to be significant (ρ= -0.68, P=0.050), while the associations between SI_D and *fr* with BMI were not (P=0.249 and 0.177, respectively).

Finally, although the *a priori* knowledge used to identify the model was taken

from studies performed in adults, it was able to account for differences between adults and adolescents (Table 3.7).

Table 3.6: Minimal model parameter estimates in adolescents with/without obesity

Overbasal-AUC	**Lean**	**Obese**	**P**
Glucose(mg/dL/min x10^3)	3.4 [2.8-8.2]	3.6 [1.7-5.3]	0.353
Insulin (pmol/L min x10^4)	9.4 [6.5-12.2]	12.5 [9.4-22.8]	**0.034**
Lactate (mg/dL min x10^2)	3.6 [2.5-5.6]	2.9 [2.0-5.0]	0.603
LPR (mg/kg)	127.3 [104.5-185.1]	75.8 [42.3-113.1]	**0.006**
Key model parameters			
EGP_b (mg/kg/min)	2.8 [2.4-3.7]	2.3 [1.8-3.4]	0.353
SI_L (dL/kg/min/μU/mL x10^{-4})	6.9 [6.2-10.0]	4.7 [3.7-6.5]	**0.003**
SI_D (dL/kg/min/μU/mL x10^{-4})	5.9 [1.2-9.1]	1.2 [0.3-2.8]	**0.036**
fr(%)	13.3 [12.3-93.5]	24.6 [16.2-38.3]	0.603
k_L (min^{-1})	0.068 [0.054-0.157]	0.059 [0.042-0.082]	0.161

Median [25th -75th] percentile of substrate and LPR over-basal AUCs, along with key model parameters in adolescent with/without obesity. P is the p-value of the comparisons conducted using Mann-Whitney test. Note that, for all the comparisons in which no statistical differences were found, the null hypothesis of the equivalence test was always accepted. Only for EGP_b and k_L, the non-inferiority was significant (P= 0.002, 0.001).

In the study of [41], SI_D and SI_L were estimated (mean±SE) equal to 9.93±2.18 and 5.48±0.54 10^{-4} dL/kg/min/μU/mL in healthy, and 5.34±6.17 and 5.41±3.55 10^{-4} dL/kg/min/μU/mL in prediabetic adults.

In our study, adolescents with obesity showed SI_D equal to 2.96±0.81 10^{-4} dL/kg/min/μU/mL, while SI_L equal to 5.02±0.42 10^{-4} dL/kg/min/μU/mL. This was in agreement with [57], which showed a higher insulin resistance state in adolescents with obesity compared to adults.

Similarly, EGP_b (Equation 3.13) was estimated higher in adolescents with/without obesity, compared to healthy and prediabetic adults (3.01±0.42, 3.06±0.53, 2.01±0.05, and 1.75±0.34 mg/kg/min, respectively). This also was in line with [69] and [70], where authors used tracer techniques to determine the basal EGP.

It is worth noting that no statistical test were performed to compare adults and

adolescents since adults data were not available in our study.

Table 3.7: Minimal model parameter estimates in adolescents and adults

Parameter	Adolescents without obesity	Adolescents with obesity	Healthy adults [41]	Adults with prediabetes [41]	Adolescents [70]
BMI (kg/m^2)	21±0.8	35±1*	25±1	31±1	24±2
EGP_b (mg/kg/min)	3.06±0.53	3.01±0.42	2.01±0.05	1.75±0.34	3.12±0.16
SI_L ($dL/kg/min/\mu U/mL\ x10^{-4}$)	8.08±0.75	5.02±0.42*	5.48±0.54	5.41±3.55	NA
SI_D ($dL/kg/min/\mu U/mL\ x10^{-4}$)	12.65±7.97	2.96±0.81*	9.93±2.18	5.34±6.17	NA

Data are reported as mean±SE to be consistent with [41][70] The first two column report parameters of the subjects used in this thesis project, while the remaining column of the literature values reported in [41][70]. Asterisk indicates statistically significant differences between adolescents with/without obesity. NA: not available

Chapter 4

The Oral Glucose-Lactate Simulation Model

[1] While in Chapter 3 the oral glucose-lactate minimal model was presented, in this chapter a simulation model of the glucose-lactate system is proposed. The model is based on an already existing model of glucose-insulin-C-peptide dynamics, in which the minimal model of lactate kinetic and glucose-lactate interaction are incorporated. The main purpose of such model was to simulate glucose, lactate, insulin, and C-peptide data during an OGTT by using a large number of equations and parameters to fully implement the knowledge of the system under study. Furthermore, to make the simulator potentially useful to test drug affecting lactate metabolism, one also needs to equip it with an *in-silico* population able to match the behavior of a real one, similarly to what done for the T2D simulator, initially developed to match early stage type 2 diabetic population and then extended to reproduce data of insulin-naïve and advance stage type 2 individuals [44][71].

[1] This chapter contains material published in [71][72]

4.1 Models

The existing models of insulin, C-peptide and glucose subsystems [42][43] [44] are presented first in Paragraphs 4.1.1 and 4.1.2. Then, since the model used to describe glucose kinetics is more complex than that implemented in the minimal model, the lactate model selected in Chapter 3 needs to be adapted to be plugged into the simulator (see Paragraph 4.1.3 and Appendix A for further details).

4.1.1 Insulin and C-peptide subsystem

In this paragraph, the insulin and C-peptide secretion and kinetics models used in [44] are described. The insulin and C-peptide secretion was modelled as the sum of three components: the static (ISR_s, pmol/min), the dynamic (ISR_d, pmol/min) and the basal (ISR_b, pmol/min) insulin secretions:

$$\begin{cases} ISR(t) = ISR_s(t) + ISR_d(t) + ISR_b & ISR(0) = ISR_b = CP_b \cdot k_{01} \cdot V_C \\ \dot{ISR}_s(t) = -\alpha \cdot [ISR_s(t) - V_c \cdot \Phi_s \cdot (G(t) - G_b)] & ISR_s(0) = 0 \\ \dot{ISR}_d(t) = \begin{cases} V_c \cdot \Phi_d \cdot \dot{G}(t) \text{ if } \dot{G}(t) \geq 0 \\ 0 \text{ if } \dot{G}(t) < 0 \end{cases} & ISR_d(0) = 0 \end{cases} \tag{4.1}$$

where CP_b (pmol/L) is the basal plasma C-peptide concentration; k_{01} (min^{-1}) is the rate parameter calculated as proposed in [73]; α (min^{-1}) is the delay between glucose and insulin secretion; Φ_s (10^{-9} min^{-1}) is the β-cell responsivity to glucose; V_C is the C-peptide volume of distribution (L); Φ_d (10^{-9}) is the β-cell responsivity to glucose rate of change; G (mmol/L) is the plasma glucose concentration with G_b its basal value; $\dot{G}$ (mmol/L/min) is the glucose rate of change. The C-peptide kinetics is described by a two compartment model:

$$\begin{cases} \dot{CP}_1(t) = -(k_{01} + k_{21}) \cdot CP_1(t) + k_{12} \cdot CP_2(t) + \frac{ISR(t)}{V_C} & CP_1(0) = CP_b \\ \dot{CP}_2(t) = -k_{12} \cdot CP_2(t) + k_{21} \cdot CP_1(t) & CP_2(0) = \frac{k_{21} \cdot CP_b}{k_{12}} \end{cases} \tag{4.2}$$

where k_{01}, k_{21}, k_{12} (min^{-1}) are the rate parameters calculated based on [73]; CP_1 and CP_2 (pmol/L) are the C-peptide concentrations in the plasma and peripheral compartments respectively; CP_b, is the plasma C-peptide concentration. Insulin kinetics is described by a three compartment model:

$$\begin{cases} \dot{I}_L(t) = -(m_1 + m_3(t)) \cdot I_L(t) + m_2 \cdot I_P(t) + \frac{ISR(t)}{BW} & I_L(0) = \frac{I_{pb} \cdot m_2 + \frac{ISR_b}{BW}}{m_1 + m_3(0)} \\ \dot{I}_P(t) = -(m_2 + m_4 + m_5) \cdot I_P(t) + m_1 \cdot I_L(t) + m_6 \cdot I_{EV}(t) & I_P(0) = I_{Pb} = I_b \cdot V_I \\ \dot{I}_{EV}(t) = -m_6 \cdot I_{EV}(t) + m_5 \cdot I_P(t) & I_{EV}(0) = \frac{I_{pb} \cdot m_5}{m_6} \\ I(t) = \frac{I_P(t)}{V_I} & I(0) = I_b \\ m_3(t) = \frac{HE(t) \cdot m_1}{1 - HE(t)} \text{ with } HE(t) = -a_G \cdot G(t) + a_{0G} \text{ s.t. } HE(t) \epsilon [0, 0.99] & \end{cases} \tag{4.3}$$

where m_1, m_2 (fixed to population value of 0.268 min^{-1} [74]), m_3, m_4, m_5, m_6 (min^{-1}) are rate parameters; I_L, I_P, and I_{EV} (pmol/kg) are the liver, plasma, and extra-vascular insulin mass respectively; V_I (L/kg) the plasma insulin volume; I(t) is the plasma insulin concentration (pmol/L); HE (dimensionless) is the hepatic insulin extraction; a_G (dL/mg) is a parameter governing the glucose control on HE; a_{0G} (dimensionless) is the extrapolated HE at zero glucose. A schematic representation of the insulin and C-peptide subsystem is reported in Figure 4.1

4.1.2 Glucose subsystem

In this paragraph, the glucose kinetic model and the mathematical descriptions of the glucose rate of appearance (R_a), endogenous glucose production (EGP) and glucose utilization (U) are presented. Differently from the minimal

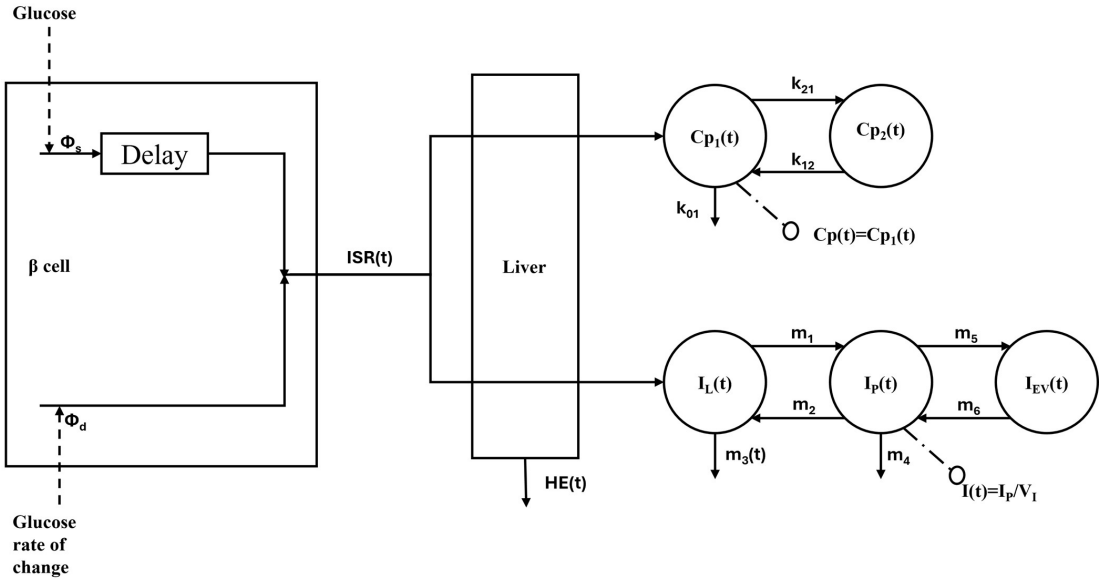

Figure 4.1: Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. The β cell rectangle represents the insulin secretion, ISR(t), model described in Equation 4.1, with Φ_s and Φ_d the β cell responsivities to glucose and glucose rate of change respectively. The C-peptide kinetic model is described in Equation 4.2, with k_{12}, k_{21} and k_{01} the C-peptide rate constant of plasma-periphery, periphery-plasma and plasma clearance respectively. The Insulin kinetic model is described in Equation 4.3, with m_1 and m_2 the rate constant of liver-plasma, plasma-liver exchange respectively; m_5 and m_6 are the rate constant of plasma-extra vascular and extra vascular-plasma exchange respectively; m_3 and m_4 are the liver and plasma disappearance rate respectively; HE(t) is the hepatic extraction.

models, in which glucose kinetics was described by using a single compartment, here a second compartment, representing glucose mass in the tissue, was added. The model equations are the following:

$$\begin{cases} \dot{Q}_{GP}(t) = -E(t) - U_{ii}(t) - k_1 \cdot Q_{GP}(t) + k_2 \cdot Q_{GT}(t) + EGP(t) + R_a(t) & Q_{GP}(0) = G_b \cdot V_G \\ \dot{Q}_{GT}(t) = -U_{id}(t) - k_2 \cdot Q_{GT}(t) + k_1 \cdot Q_{GP}(t) & Q_{GT}(0) = G_{Tb} \cdot V_G \\ G(t) = \frac{Q_{GP}(t)}{V_G} \\ E(t) = \begin{cases} k_{e1} \cdot Q_{GP}(t) \text{ if } Q_{GP}(t) > k_{e2} \\ 0 \text{ if } Q_{GP}(t) \leq k_{e2} \end{cases} \end{cases} \tag{4.4}$$

where Q_{GP} (mg/kg) is the plasma glucose mass; Q_{GT} (mg/kg) is the tissue glucose mass; EGP (mg/kg/min) is the endogenous glucose production; R_a (mg/kg/min) is the glucose rate of appearance; U_{ii} and U_{id} (mg/kg/min) are the glucose insulin independent and insulin dependent utilizations respectively; k_1 and k_2 (min^{-1}) are rate parameters of plasma-tissue exchange; E (mg/kg/min) is the renal excretion; k_{e1} (min^{-1}) and k_{e2} (mg/kg) are the glomerular filtration rate and renal threshold of glucose respectively (both fixed to population values of 0.0005 min^{-1} and 339 mg/kg based on [42]); V_G (dL/kg) is the plasma glucose volume.

The gastro-intestinal tract is modelled by three-compartments:

$$\begin{cases} Q_{sto}(t) = Q_{sto1}(t) + Q_{sto2}(t) & Q_{sto}(0) = 0 \\ \dot{Q}_{sto1}(t) = -k_{max} \cdot Q_{sto1}(t) + Dose \cdot \delta(t) & Q_{sto1}(0) = 0 \\ \dot{Q}_{sto1}(t) = -k_{empt}(Q_{sto}) \cdot Q_{sto2}(t) + k_{max} \cdot Q_{sto1}(t) & Q_{sto2}(0) = 0 \\ \dot{Q}_{gut}(t) = -k_{abs} \cdot Q_{gut}(t) + k_{empt}(Q_{sto}) \cdot Q_{sto2}(t) & Q_{gut}(0) = 0 \\ R_a(t) = \frac{f \cdot k_{abs} \cdot Q_{gut}(t)}{BW} \end{cases} \tag{4.5}$$

where Q_{sto} is the amount of glucose (mg) in the stomach, Q_{sto1} and Q_{sto2} (mg) are the solid and liquid phase; Dose (mg) is the amount of ingested glucose; Q_{gut} (mg) is the amount of glucose in the intestine; f (dimensionless) is the percentage of ingested glucose that appeared in the plasma (set to 0.9); BW (kg) is the body weight; R_a (mg/kg/min) is the glucose rate of appearance. The parameter k_{empt} depends on Q_{sto} as described in the following:

$$k_{empt}(Q_{sto}) = k_{min} + \frac{k_{max} - k_{min}}{2} \cdot \{\tanh[\alpha \cdot (Q_{sto} - b \cdot Dose)] - \tanh[\beta \cdot (Q_{sto} - c \cdot Dose)] + 2\} \tag{4.6}$$

in other words, k_{empt} is maximum, equal to k_{max}, when the stomach contains the amount of ingested glucose (i.e. Dose), then, it decreases with a rate α to

its minimum, k_{min}, and it recovers back to k_{max} with a rate β. Parameters b and c are the percentages of dose for which k_{empt} decreases/increases at (k_{max}-k_{min})/2, respectively.

The endogenous glucose production is modelled as:

$$\begin{cases} EGP(t) = k_{p1} - k_{p2}Q_{GP}(t) - k_{p3}X_L(t) - k_{p4}I_L(t) & EGP(0) = EGP_b \\ \dot{I'}(t) = -k_i \cdot [I'(t) - I(t)] & I'(0) = I_b \\ \dot{X}_L(t) = -k_i \cdot [X_L(t) - I'(t)] & X_L(0) = I_b \\ k_{p1} = EGP_b + k_{p2}Q_{GP}(0) - k_{p3}X_L(0) - k_{p4}I_L(0) & \end{cases} \tag{4.7}$$

where k_{p1} (mg/kg/min) is the extrapolated EGP (mg/kg/min) at zero glucose and insulin, calculated to guarantee the steady state condition at time 0min; k_{p2} (min^{-1}) is the hepatic glucose effectiveness; k_{p3} (mg/kg/min per pmol/L) is the hepatic insulin sensitivity; k_{p4} (mg/kg/min per pmol/kg) is the portal insulin sensitivity; Q_{GP} (mg/kg) is the plasma glucose mass; I'(pmol/L) is a delayed plasma insulin concentration; k_i (min^{-1}) is the rate of delayed insulin action; X_L (pmol/L) is the insulin action on EGP; $I_L(t)$ is the liver insulin mass as already defined in Equation 4.3.

Glucose utilization is the sum of two components, the insulin dependent (U_{id}) and insulin independent (U_{ii}) glucose utilization:

$$\begin{cases} U_{ii}(t) = F_{cns} \\ U_{id}(t) = \frac{V_{m0} + V_{mx} \cdot X(t) \cdot (1 + r1 \cdot risk) \cdot Q_{GT}(t)}{K_{m0} + Q_{GT}(t)} \\ \dot{X}(t) = -p_{2U} \cdot X(t) + p_{2U} \cdot [I(t) - I_b] \quad X(0) = 0 \end{cases} \tag{4.8}$$

where F_{cns} (mg/kg/min) is the glucose utilization by the central nervous system and is fixed to 1 based on [43]; V_{mx} (mg/kg/min per pmol/L) is the insulin sensitivity on glucose utilization; Q_{GT} (mg/kg) is the tissue glucose mass; X (pmol/L) is the insulin effect on glucose utilization; p_{2U} (min^{-1}) is the rate

of insulin action on glucose utilization; K_{m0} (mg/kg) is the glucose mass in the Michaelis-Menten relation; *risk* is the function that modulates the insulin sensitivity parameter (V_{mx}) such that:

$$\begin{cases} \text{risk} = \begin{cases} 0 \text{ if } G(t) > G_b \\ 10[f(G)]^2 \text{ if } G_{low} \leq G(t) \leq G_b \\ 10[f(G_{low})]^2 \text{ if } G(t) < G_{low} \end{cases} \\ f(G) = [\log(G)]^{r1} - [\log(G_b)]^{r2} \end{cases} \tag{4.9}$$

where G is the plasma glucose concentration (mg/dL); G_b (mg/dL) is the basal plasma glucose and G_{low} the lower threshold assumed to be equal to 60 mg/dL; r1 and r2 (dimensionless) are the parameters of risk function [43]. A schematic representation of the Glucose subsystem is reported in Figure 4.2

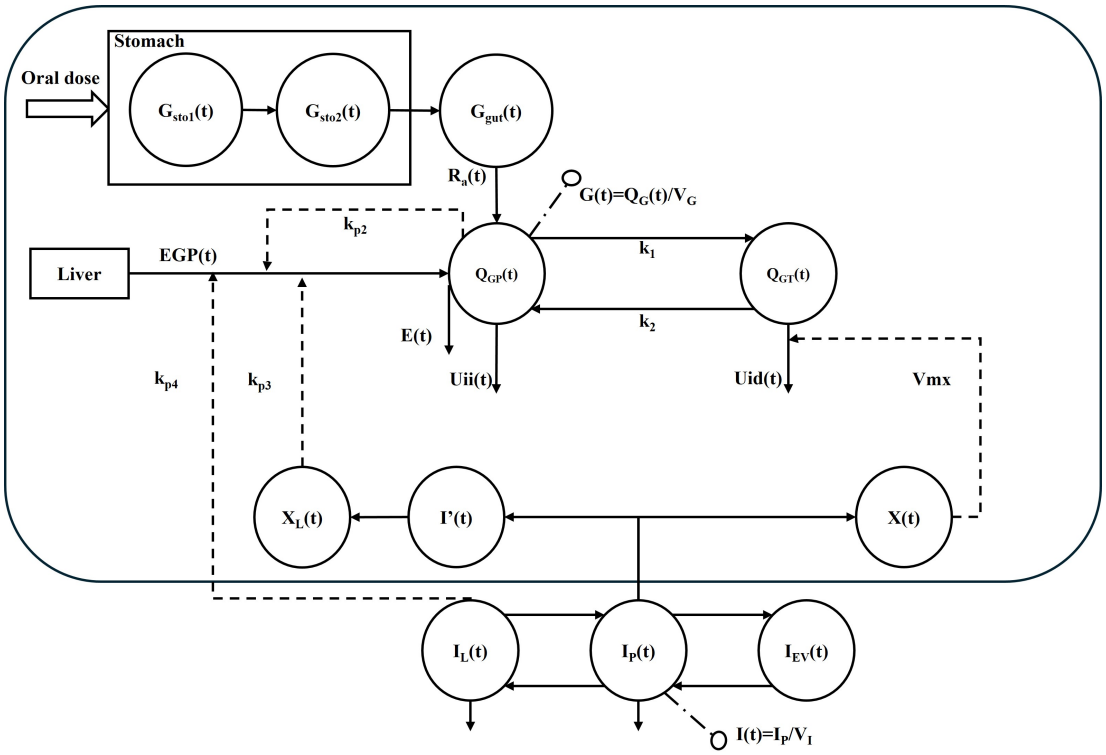

Figure 4.2: Schematic representation of the Glucose subsystem (squared) with $I_P(t)$, $I_L(t)$ and $I_{EV}(t)$ used to describe the delayed insulin I'(t), $X_L(t)$ and X(t). Continuous lines represent fluxes, dashed lines represent control actions, while dashed-dotted lines with circle indicate the measured plasma concentrations. Glucose kinetics is described in Equation 4.4, with Q_{GP} and Q_{GT} the glucose masses in the plasma and tissue respectively, E(t) the renal excretion; k_1 and k_2 are the exchanges rate parameters between plasma and tissue. Q_{sto1}, Q_{sto2}, Q_{gut} and R_a described the solid and liquid phase, the amount of glucose in the intestine and glucose rate of appearance respectively (Equation 4.5). The endogenous glucose production (EGP(t)) is described in Equation 4.7, with k_{p2}, k_{p3} and k_{p4} the hepatic glucose effectiveness (glucose control on its own production), hepatic insulin sensitivity ($X_L(t)$ control) and portal insulin sensitivity (I_L control) respectively. Finally, U_{ii} and U_{id} are the glucose insulin-independent and insulin-dependent utilizations; X(t) is the delayed insulin action of U_{id}; V_{mx} is the disposal insulin sensitivity (Equation 4.8). V_G and V_I are the glucose and insulin plasma volume respectively.

4.1.3 Lactate subsystem

The model of the glucose subsystem included in the simulator differs from that of the minimal model (see Paragraph 4.1.2). Consequently, the lactate kinetics described in Paragraph 3.1.4 needed some adjustment to be integrated into the simulator. For clarity, we provide below a detailed description of the model incorporated into the simulator, while the remaining steps of model development and selection are outlined in Appendix A. In particular, similarly to the hypothesis made for the minimal model (see Paragraph 3.1.4), the key assumption was that only a fraction of the total glucose utilization is used to generate lactate, while the other followed other metabolic paths, and V_L (lactate plasma volume, dL/kg) was assumed to be equal to glucose plasma volume (V_G, dL/kg). Lactate kinetic model is the following:

$$\begin{cases} \dot{Q}_{LP}(t) = -k_L \cdot Q_{LP}(t) + 2fr[U_{ii}(t) + U_{id}(t)] \quad Q_{LP}(0) = L_b \cdot V_L = L_b \cdot V_G \\ L(t) = \frac{Q_{LP}(t)}{V_L} = \frac{Q_{LP}(t)}{V_G} \end{cases} \tag{4.10}$$

where Q_{LP} is the lactate plasma mass (mg/kg); k_L (min^{-1}) is the rate constant of lactate utilization; *fr* (dimensionless) is the fraction of the total glucose utilization that generates lactate; L (mg/dL) is the lactate plasma concentration with L_b its basal value. Note that, this model structure results to be equivalent to the one of the selected minimal model (see Paragraph 3.1.4), in which the $p_{1D}Q_G(t)$ and $X_D(t)Q_G(t)$ are substituted by $U_{ii}(t)$ and $U_{id}(t)$, respectively. A schematic representation of the model is reported in Figure 4.3.

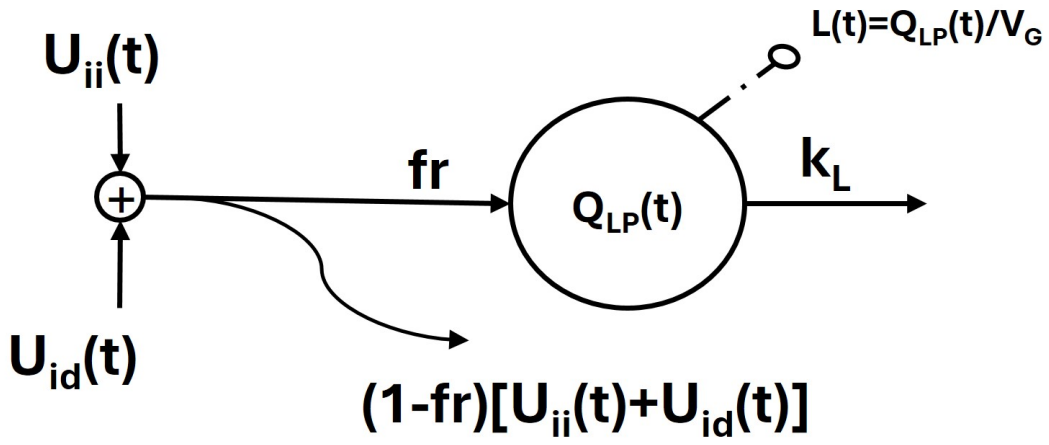

Figure 4.3: Schematic representation of lactate kinetic model. Continuous lines represent fluxes, while dashed-dotted lines with circle indicate the measured plasma concentrations. U_{ii} and U_{id} are the glucose insulin independent and insulin dependent utilizations respectively (Equation 4.8); *fr* is the fraction of the total glucose utilization that generates lactate; Q_{LP} is the lactate mass in the plasma; L is the lactate plasma concentration; k_L is the rate constant of plasma lactate utilization V_G is the glucose plasma volume.

4.2 *In-silico* cloning of a target population

One of the primary aims of current research is providing a tool for testing therapies, drugs, or other interventions in pathophysiological states using computer simulation. This involves employing mathematical models to generate simulated data that are indistinguishable from the real ones, prior to moving to *in-vivo* testing. The last aim of this thesis is to provide a tool to cloning real glucose, insulin, C-peptide and lactate data of a real population of pediatric subjects and use the generated virtual patients to possibly test what already described before. In this paragraph, the procedure developed to generate an *in-silico* population starting from a real one, along with the method to extract the most 100 representatives *in-silico* subjects are described, which consisted of four main steps [71] (Figure 4.4).

Model identification

If only the average clinical data of a target population are available, one can resort to the identification procedure described in Appendix A.2.1 for estimating the average of simulation model parameters, as done in [71].

Parameter distribution update

The estimated average model parameters is used to update the mean vector of the parameter distribution. At variance with what reported in [71], where only the insulin sensitivity and secretion indexes were modified, in this thesis, the entire average-parameters vector was updated. This approach was essential for generating a population of adolescents, as they differ significantly in age, body weight, and other factors from the predominantly healthy and T2D adult subjects (see Chapter 4.3.2).

In-silico subjects generation

Once the parameter distribution was updated, the desired number of parameter vectors was randomly sampled from it and Mahalanobis distance was used to exclude extreme values, setting the threshold of exclusion to 95th percentile [75]. Moreover, as already done in [71], other constraints were added to exclude unrealistic parameter combinations as reported in Table 4.1. For the gastro-intestinal tract model, parameters b and d are fractions, bounded to be lower than 1, while k_{max} was assumed to be higher than k_{min} (see Equation 4.6). For k_{abs} and the parameters related to hepatic exctraction and insulin PK model, the parameter maximum values was set to the 99th percentile of the distribution determined in [44]. In addition to the constraint already adopted in [71], *fr* was bounded to be in the range 0-1, since it was the fraction of the total glucose utilization which generates lactate. Once the virtual subject parameters were extracted from the distribution, the simulator was employed to simulate the same experimental protocol used to identify the model. In our case, a 3h-75g OGTT was simulated, and glucose, insulin, C-peptide and lactate plasma concentration were sampled. The choice of a 3h instead of the 4h OGTT with sampling time t= [0 10 20 30 60 90 120 150 180] min, was made mainly because not all the subjects were studied for more than 3h (i.e. Dataset 2). Nevertheless, other experimental conditions can be simulated, like different types of meals or therapies.

In-silico subjects extraction

It is worth noting that, although the generated subjects were obtained by randomly sampling the updated parameter distribution, this did not automatically imply that the time courses of glucose, insulin, C-peptide and lactate were representative of those measured in real population (see Chapter 4.3.1).

Table 4.1: Model parameter constraints

Process	Criteria
Gastro-intestinal-tract	k_{abs}<1, k_{max}>k_{min}, b<1, d<1, b>d
Hepatic extraction	$HE_b \epsilon$[0.3-0.99]
Insulin PK parameters	m_4<1.5, m_5<1, m_1<1
Lactate PK parameters	0<$fr \leq 1$

A method which guarantees the extraction of the 100 most representative virtual subjects has been proposed [71]. The procedure is described in the following. One first creates a large number of *in-silico* patients (typically 3000) with the method described above. Then, an initial reduction in the number of valid subjects is done based on the similarity between the single subject curves and mean±SD experimental data using the discrete Fréchet distance as proposed in [76] for polygonal curves. Particularly, the discrete Fréchet distance $d_F(P, Q)$ between two polygonal curves $P = (p_1, p_2, \ldots, p_R)$ and $Q = (q_1, q_2, \ldots, q_S)$, with p_i and q_j points in the *x-y* space, is defined as:

$$\begin{cases} \delta(p_i, q_j) = ||p_i - q_j|| \\ d_F(P, Q) = \min_L \max_{(i,j)\in L} \delta(p_i, q_j) \end{cases} \tag{4.11}$$

where $\delta(p_i, q_j)$ is the Euclidian distance between two points, and L is the coupling defined as:

$$L = [(i_1, j_1), (i_2, j_2), ...(i_w, j_w)] \text{ with } 1=i_1 \leq i_2 \leq i_w = R \text{ and } 1=j_1 \leq j_2 \leq j_w = S \tag{4.12}$$

i.e. a sequence of pairs of indexes (i,j) that represent a mapping between the points of the two curve P an Q. Each pair (i,j) indicates that the points p_i of P is being matched with the point q_j of Q, with R and S the number of points of

P and Q respectively. The choice of using such measure was due to the fact that the Fréchet distance was able to account for the sequential ordering and structure of time series. After defining the similarity measure for each virtual curve, the new subset of virtual subjects is generated by sequentially adding to the selected pool the virtual subject providing the best similarity score, i.e. the lowest Fréchet distance. Then, the number of subjects included in the selection depends on two indexes, FIT and S, defined as:

$$\mathrm{FIT} = 1 - \sqrt{\frac{\sum_{k=1}^{N}[\mathrm{y_{meas}}(\mathrm{t_k}) - \mathrm{y_{sim}}(\mathrm{t_k})]^2}{\sum_{k=1}^{N}[\mathrm{y_{meas}}(\mathrm{t_k}) - \mathrm{y_{mean}}]^2}} \tag{4.13}$$

where N is the number of samples, $\mathrm{y_{meas}}$ and $\mathrm{y_{sim}}$ are the mean or mean$\pm$SD experimental data and simulated curves respectively, $\mathrm{y_{mean}}$ the average $\mathrm{y_{meas}}$. The FIT index was used for each available measurement (i.e. Glucose, insulin, C-peptide and lactate average with/without SD), defining the score index S as:

$$\mathrm{S} = \frac{\sum_{M=1}^{4}\mathrm{FIT_M}}{4} \tag{4.14}$$

where M indicates the glucose, insulin, C-peptide, lactate profiles. If the standard deviations are also available in addition to the average curves, then the formula becomes:

$$\mathrm{S} = \frac{\sum_{M=1}^{4}\sum_{i=1}^{3}\mathrm{FIT_{Mi}}}{12} \tag{4.15}$$

where i indicates the i-th curve (i.e. the mean/mean-SD/mean+SD curves). The procedure stops when the S index reaches a value of 0.7, i.e. the mean or mean$\pm$SD profiles of the virtual subjects reaches on average a similarity of 70% with the original data. Such threshold was lower than the one used in the former study [71] since, here, glucose, insulin, C-peptide and lactate curves were used to extract the *in-silico* population, while in [71] only the first three mea-

surements were used, thus, complicating the extraction procedure.
While the procedure described above is used to exclude subjects that differed significantly from the target population in terms of substrate and/or hormone postprandial time courses, the second step consists in the extraction of the most representative 100 subjects from the subset previously created. Firstly, the simple classification metric (P) was used to determine if the subject's curves lie below or above the reference curves:

$$P_{j,M} = \sum_{k=1}^{N} \frac{y(t_k) - y_{meas}(t_k)}{y(t_k)} \tag{4.16}$$

where j indicates the j-th subject, k the k-th sample of the y measured (M: glucose/insulin/C-peptide/lactate) of N samples, while $y_{meas}(t_k)$ is the reference curve (i.e. the mean of experimental data). Each subject is classified as above the M-th reference curve if $P_{j,M}$ was ≥ 0, below otherwise. By doing so, subjects were clustered into groups, defined as:

$$groups = \begin{cases} P_G \geq 0; P_I \geq 0; P_{CP} \geq 0; P_L \geq 0 \\ P_G < 0; P_I \geq 0; P_{CP} \geq 0; P_L \geq 0 \\ P_G \geq 0; P_I < 0; P_{CP} \geq 0; P_L \geq 0 \\ \vdots \\ P_G < 0; P_I < 0; P_{CP} < 0; P_L < 0 \end{cases} \tag{4.17}$$

i.e. one group contains only the subjects for which the 4 conditions are verified (G: glucose, I: insulin, CP: C-peptide, L: Lactate).
Each group was randomly sampled to reach the number of 100 subjects. This step was done to overcome the problem of sampling with high frequency the more abundant groups, trying to maintain the proportion between them. In the simplest case, in which two groups were defined i.e. a single curve needs

to be matched e.g. the glucose average curve, 50 subjects were sampled above and 50 below that curve. While, if four groups were defined, i.e. two curves need to be matched e.g. the glucose and insulin average curves, 25 subjects per group were sampled, and so on. This procedure was repeated 1,000,000 times, mainly to contain the computational time needed (≈30min to extract the final population). The final population of *in-silico* subjects, which maximized the S index across all tested combinations (i.e. 1,000,000 x 100), is then selected. In Figure 4.4 a schematic representation of the above steps is shown.

At variance with the original article, where only mean and SD curves from a clinical trial were available, in our study real curves were available at single individual level. This enabled us first to use the S score to match the average data as it was done in [71], and, second, to conduct a more comprehensive analysis. In fact, beyond matching the average subject curves only, comparing the AUC and model parameter distributions between real and *in-silico* subjects, we were able to effectively assess the method's ability to capture the inter-subject variability.

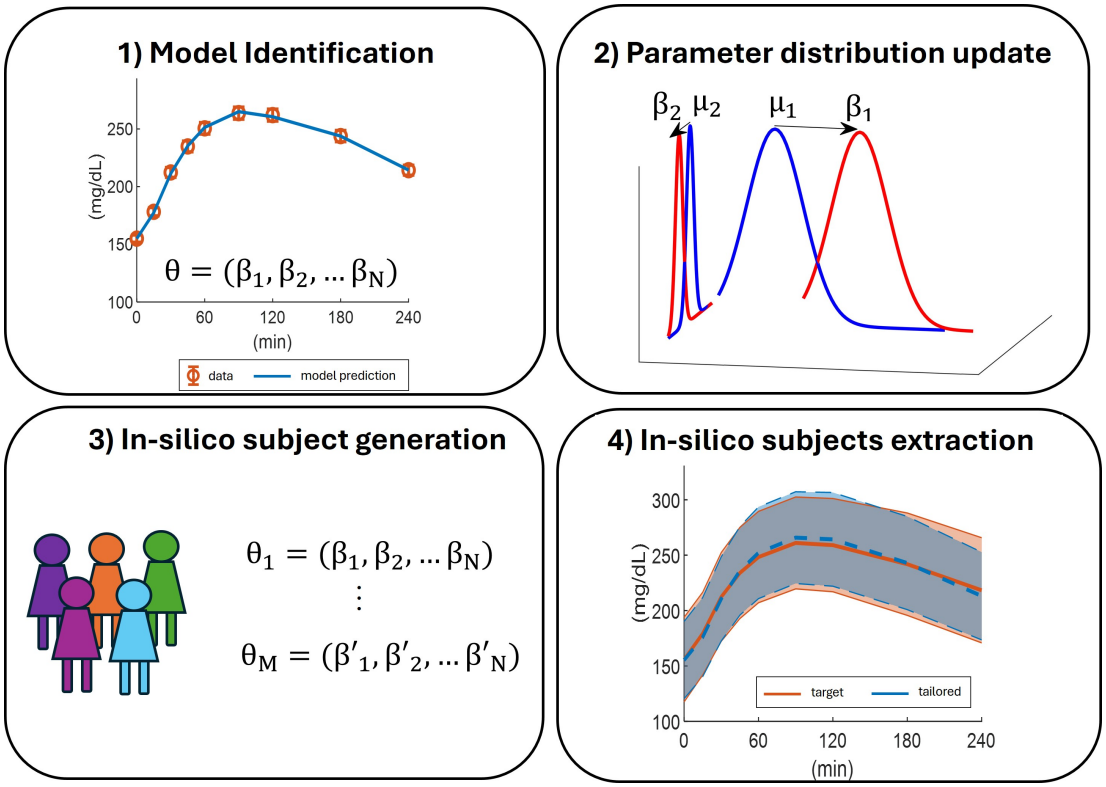

Figure 4.4: Schematic representation of the four steps to generate an *in-silico* population. Initially, the model is identified using clinical data to estimate the parameter vector (θ). Then, these estimates (β) are employed to refine the mean of the parameter distribution (μ). Finally, the *in-silico* population is generated by sampling from the updated parameter distribution, producing a set of M subjects each characterized by a unique set of parameters θ_i with i=1:M. After that, the *in-silico* subjects are used to simulate the same experimental protocol used to identify the model, and the set of 100 most representative *in-silico* subjects is extracted.

4.3 Results

This section describes the results of the simulation model application, i.e. the generation and extraction of the *in-silico* subjects. Furthermore, the model parameters of adolescents with and without obesity, examined in this study, are compared to those reported in the literature for healthy adults and individuals with type 2 diabetes.

In 4.3.1 the use of the simulation model is shown. In particular, the techniques previously described in Paragraph 4.2 were used to generate an *in-silico* population able to match the data measured during a real study. Finally, in 4.3.2 model parameter differences between lean subjects and those with obesity are shown. Also, comparisons with adults' parameters are described.

Further details on model development and selection are described in Appendix A.

4.3.1 Validation of the *in-silico* population

As already discussed, random sampling from the updated parameter distribution does not automatically imply a good match between *in-silico* curves and the experimental data. This was confirmed by the value of S score (Equation 4.15), calculated in the initially generated *in-silico* population, which was equal 0.3. Moreover, by comparing the AUC of glucose and lactate in the real and *in-silico* populations, statistically significant differences emerged (both P<0.001, Mann-Whitney test).

On the other hands, by using the method described in Chapter 4.2 and published in [71], the population of selected 100 *in-silico* subjects reached a much better similarity score of S=0.70 (Figure 4.5, second row). By comparing AUCs and the parameter distributions (Table 4.2), no statistically significant differences emerged between the real population and the extracted one. In particular, the substrate/hormone AUCs of the real and *in-silico* populations

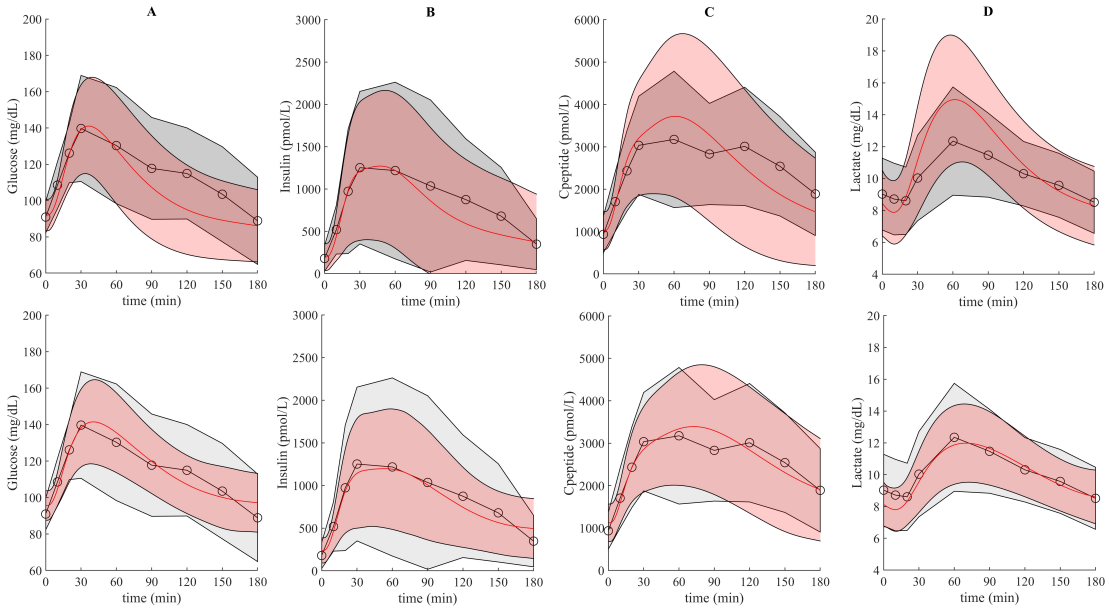

Figure 4.5: Plasma glucose (column A), insulin (column B), C-peptide (column C) and lactate (column D) data (Mean±SD circle and grey area, N=44) vs simulated curves (red line and pink area) in the initially generated 3000 (first row) and in the optimal 100 subjects (second row).

resulted to be indistinguishable after the selection (P=0.6 for glucose, insulin and lactate; P=0.9 for C-peptide). Additionally, the equivalence test rejected the null hypothesis with P= 0.0281, 0.0172, 0.0063, and 0.011, for glucose, insulin, C-peptide and lactate respectively. Moreover, the parameters directly involved into lactate kinetic description, i.e. fr, k_L and V_{mx}, were not statistically different in the two populations, with P=0.5, 0.5 and 0.2, respectively. Again, the equivalence test yielded significant results, with P<0.002 for each variable. It is worth noting that, also for k_{p3}, Φs and Φd the equivalence resulted to be significant, with P<0.01. We concluded that the observed differences were small enough to be considered practically irrelevant within these specific equivalence margins (see Paragraph 3.2.4). As previously reported, the match index (S, Equation 4.15) in the former study [71] was higher (i.e. 0.93) compared to the results of this thesis but it was probably due to the higher molecules variability of our real population compared to the former one. In addition, the requirement of matching also the lactate curve complicated the process of population extraction.

Table 4.2: Real and *in-silico* populations comparsion

	Real population	***in-silico* population**	**P**
AUC			
Glucose (mg/dL min x10^4)	2.1 [1.8-2.3]	2.1 [1.9-2.2]	0.6
Insulin (pmol/L min x10^4)	13 [8-20]	14 [10-18]	0.6
C-peptide (pmol/L min x10^4)	46 [34-58]	45 [36-57]	0.9
Lactate (mg/dL min x10^4)	0.18 [0.16-0.20]	0.19 [0.17-0.21]	0.6
Model key parameters			
fr (%)	19 [12-29]	16 [10-29]	0.5
k_L (min^{-1})	0.075 [0.053-0.115]	0.074 [0.043-0.117]	0.5
k_{p3} (mg/kg/min/pmol/L)	0.009 [0.005-0.014]	0.007 [0.004-0.012]	0.1
V_{mx} (mg/kg/min/pmol/L)	0.036 [0.015-0.048]	0.025 [0.010-0.043]	0.2
Φs (10^{-9} min^{-1})	64 [42-113]	77 [57-106]	0.1
Φd (10^{-9})	1301 [963-1885]	1289 [904-1725]	0.6

Median and [25th-75th] percentile of AUC and model parameter distributions of the real population and selected 100 *in-silico* subjects. P indicate the p-value of the comparison between real and *in-silico* values provided by the Mann-Whitney test. Note that, for all the variables, the equivalence test yielded to significant results.

4.3.2 Impact of obesity degree and age

In this paragraph, the impact of obesity degree on key simulation model parameters is discussed. Furthermore, similarly to [58], in which adolescents and adults overweight/with obesity were compared, adolescent model parameters were compared with those estimated in adult with/without T2D of [44], underlying the existing differences related to age.

First, model parameters estimated (see Appendix A.2.1 for details) in lean adolescents were compared with those obtained in individuals with obesity (Table 4.3). Despite failed to achieve statistical significance, probably due to the limited sample size, lean adolescents have an apparent increased fractional lactate clearance rate (k_L) compared to adolescents with obesity, similarly to what already observed with the minimal model (0.126 [0.076-0.394] vs 0.065 [0.045-0.095] min^{-1}, P=0.11). Probably, the impaired lactate utilization is the cause of higher lactate concentration in adolescents with obesity compared to

leans. Also, a trend appeared in the glucose disposal insulin sensitivity (V_{mx}), underlying again the increased insulin resistance state with the worsening of obesity.

Table 4.3: Comparison between subjects with/without obesity

Parameter	Without obesity	With obesity	P
$EGP_b (mg/kg/min)$	2.90 [2.63-3.09]	2.40 [2.21-2.74]	0.13
$k_{p3} (mg/kg/min/pmol/L)$	0.013 [0.011-0.015]	0.008 [0.004-0.015]	0.22
V_{mx} (mg/kg/min/pmol/L)	0.047 [0.022-0.056]	0.032 [0.008-0.043]	0.08
Φs (10^{-9} min^{-1})	101.2 [62.05-126.8]	57.3 [40.9-93.1]	0.07
Φd (10^{-9})	1505.8 [929.6-1991.5]	1168.7 [903.5-1719.4]	0.73
fr (%)	21.7 [15.2-52.8]	17.5 [11.0-25.9]	0.56
k_L (min^{-1})	0.126 [0.076-0.394]	0.065 [0.045-0.095]	0.11

Comparisons between adolescents with/without obesity. P indicate the p-value of the comparison between the two groups by using Mann-Whitney test. Note that, for all the variables, the equivalence test failed to reject the null hypothesis, while the non-inferiority was always significant, except for EGP_b, Φd and fr (lean subjects were used as reference group).

Despite that, for both V_{mx} and k_L, the equivalence test did not rejected the null hypothesis (P= 0.87, 0.75 respectively), while the non-inferiority was significant for both (P= 0.001, 0.004, using lean subjects as reference group). The results, particularly for k_L, highlight that the data only support that the fractional lactate clearance rate of lean group was not lower than that of subjects with obesity, aligning with the findings of the minimal model (see Paragraph 3.3.3).

Second, similarly to [58], adolescents with obesity/overweight were compared with healthy/T2D adults with obesity or overweight (Table 4.4). EGP_b was significantly higher in adolescents (P<10^{-6}) compared with adults, as previously shown in [69][70]. Similarly, the disposal insulin sensitivity index was significantly lower in adolescents compared to healthy adults (P=0.004), but not different from that characterizing subjects with T2D, in agreement with [57][58].

Note that, despite the equivalence test failed to reject the null hypothesis, the non-inferiority was significant (P<10^{-4}, using adolescent as reference group). Differently, the hepatic insulin sensitivity (k_{p3}) was not statistically different in adolescents and adults. However, only the non-inferiority test between adolescents and healthy adults was significant (P<10^{-6}, using adolescent as reference group), while the equivalence was significant between adolescents and adults with T2D (P=0.018, using adolescent as reference group). Insulin secretion indexes were both significantly higher in adolescence compared to adults (P<10^{-6} for all the comparisons), in agreement with [58][77]. As previously pointed out, lactate related parameters cannot be compared due to the absence of lactate data in adults. These results demonstrated that, even though the model identification procedure utilized an *a priori* knowledge derived from adult subjects (see Appendix A.2.1 for details), it successfully accounted for existing age-related differences. Moreover, when the joint parameter distribution was updated with adolescent parameters, significant relationships emerged that were not present when using adult data only. For instance, an inverse relationship between EGPb and age was identified (Figure 4.6), consistent with findings from [69].

Table 4.4: Comparison between adolescents and adults with/without T2D

Parameter	Adolescents	Adults without diabetes	Adults with T2D
BMI (kg/m^2)	33.8 [30.8-40.4]	28.1 [26.8-30.1]	33.8 [29.2-37.2]
$EGP_b(mg/kg/min)$	2.38 [2.21-2.71]	1.79[1.66-1.92]*	1.57 [1.41-1.91]*
$k_{p3}(mg/kg/min/pmol/L)$	0.008 [0.004-0.015]	0.009 [0.007-0.015]	0.006[0.004-0.008]
V_{mx} (mg/kg/min/pmol/L)	0.033 [0.012-0.044]	0.047 [0.039-0.062]*	0.034 [0.019-0.051]
$\Phi_s(10^{-9}\ \mathrm{min}^{-1})$	57.5 [41.2-112.1]	34.6 [27.9-42.6]*	21.0 [12.78-29.3]*
$\Phi_d(10^{-9})$	1220.7 [963.1-1885.3]	548.8 [409.7-724.7]*	295.7 [172.1-415.6]*

Comparisons between adolescents and adults with/without T2D. The subjects were all with obesity/overweight to be consistent with [58] * indicates statistically significant difference compared to adolescents.

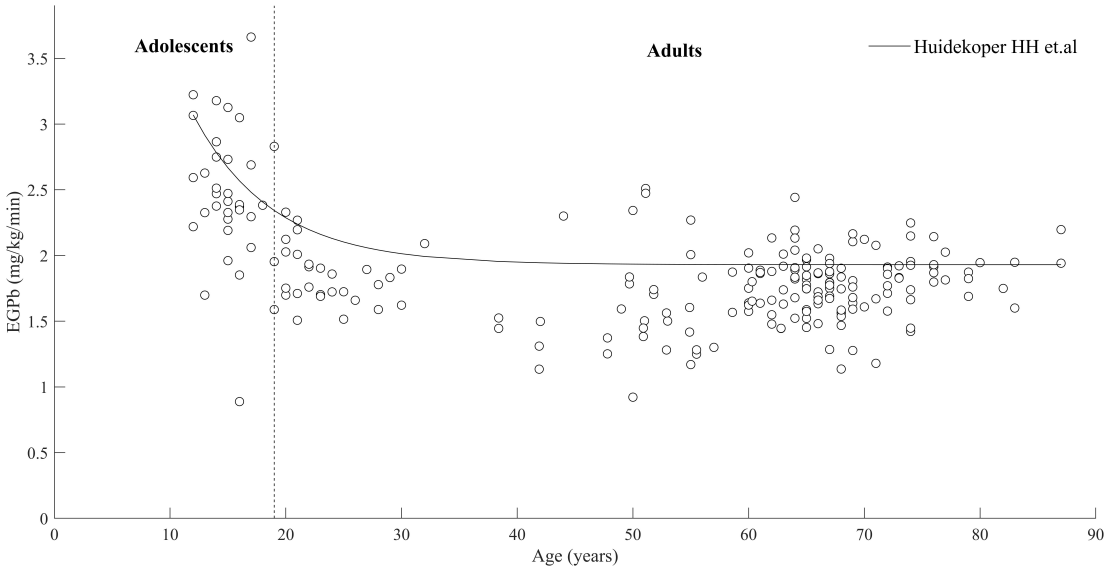

Figure 4.6: Relationship between age and basal endogenous glucose production (EGPb). Circle reported the estimated values of EGPb in adolescents and in healthy/with T2D adults (all with obesity/overweight). Continuous black line reported the relationship estimated in [69].

Chapter 5

Discussion

Lactate is no longer considered a glycolytic waste product since it is involved in several metabolic processes [1]-[4],[7][8] and related to pathophysiological states, such as obesity, diabetes [17][18][28] and MASLD [13]-[16]. In addition, it showed also a role as a signaling molecule [11][12] , underlying the key role of this molecule in human metabolism. The possibility to assess lactate kinetic through a minimal model and simulate its dynamics using computer simulation would be important to assess its impact in insulin resistance states, and to possibly provide a tool to test drugs affecting its metabolism. Some minimal models have been proposed in the literature to assess lactate kinetics, but these were developed and tested during non-physiological tests such as IVGTT [36][38] or tracer lactate infusion [39]. Here, the first aim of this thesis was to extend the domain of validity of such models to more physiological tests, like an OGTT. Moreover, to the best of our knowledge, none of the existing glucose-insulin simulators account for lactate kinetics [42]-[46], which would be a useful tool for testing drugs and treatments focusing on glucose-lactate relationship.

5.1 Oral Glucose-Lactate Minimal model

In the first part of this thesis, we tested several minimal models trying to describe both glucose and lactate kinetics during OGTT, showing that the integration of a single compartment description of lactate kinetics in the Single Tracer Oral Minimal Model, STOMM [41], was the best choice in this type of protocol [54]. In particular, we found that simple models, like Model 1, in which no insulin effect was assumed on lactate production rate, combined with fully conversion of glucose to lactate, or Models 2 and 3, which assumed that lactate production rate is controlled by the net insulin action, with or without fully conversion of glucose to lactate, provided unsatisfactory results in terms of model ability to fit the data. This is likely due to the fact that, during postprandial condition, lactate is primarily determined by the rate of glucose disposal, which strongly depends on insulin effect. This is why Models 4-6 were based on STOMM [41], which considered the insulin effects on endogenous glucose production and glucose disposal separately. Despite that, both Model 5 and 6 showed poor *a posteriori identifiability* properties. In case of Model 5, this probably occurred because lactate control on its own production cannot be clearly observed during a test in which lactate concentrations did not rise as much as during exercise or lactate infusion, while, in case of Model 6, the model structure was likely too complicated to achieved a good *a posteriori identifiability* [36].

All the literature models [36][38][39] assumed a two compartment description of lactate kinetic, while our minimal model did not. We thought that, since IVGTT produced a rapid perturbation of the glucose-lactate system, the two compartment description was necessary to account for the delay between glucose and lactate while, during an OGTT, the perturbation was more gentle making a single compartment model enough to fit the data. The model we

selected as best is based on the STOMM, which was originally identified using glucose tracer data which allowed to segregate glucose production and disposal. Here, tracers have not been employed but the model shown to be identifiable thanks to the inclusion of lactate. Indeed, the model demonstrated strong predictive accuracy and *a posteriori identifiability* properties. Therefore, we concluded that it provided the most effective description of glucose and lactate kinetics during the OGTT.

It is worth noting that, during the identification procedure, we assumed a Gaussian *a priori* distribution for certain model parameters (see Paragraph 3.2.2). The mean values for these parameters were set based on the estimates obtained from previous studies [40][41], which involved adult populations, while the standard deviations were set to the 20% of the respective mean values [59]. Furthermore, no correlation among parameters was assumed. This choice reflects, on one side, a reasonable level of uncertainty in the parameters while guaranteeing to estimate realistic and meaningful values and, on the other side, the paucity of information available from the literature which precluded the possibility to use nongaussian distribution or to assume correlations among model parameters. The decision to use these assumptions was guided by two main considerations. In particular, an initial guess for the model parameters was necessary, despite our prior knowledge was derived from adult's parameters of [40][41] while the model was being applied to an adolescent population. In fact, adolescents may exhibit different physiological responses compared to adults [57][58], making the use of a prior derived from adult parameter estimates ineffective if not potentially deleterious (see Paragraph 3.3.3). As a consequence, adopting a standard deviation of 20% for the Gaussian distribution served to constrain the parameters within realistic bounds balancing the incorporation of *a priori* knowledge with the flexibility to

accommodate the specificities of the adolescent population. For other parameters, such as SI_L and SI_D and parameters of lactate kinetics, a non-informative prior was adopted to avoid introducing potential bias in the estimates of these key variables. Indeed, although the model was not directly used in adults and individual with diabetes due to the lack of data of these populations, the comparisons of the insulin sensitivities (i.e. SI_L and SI_D) estimated in the adults of [41] and our adolescents were concordant with literature studies [57].

Finally, the model was also able to capture physiological differences between subjects with/without obesity. SI_L and SI_D estimated by the model were significantly lower in adolescents with obesity compared to leans. A pattern also emerged for *fr* and k_L, with the first higher and the second lower in subjects with obesity, but only k_L was proved to be non-inferior in lean compared to subjects with obesity with the specific margin adopted. Probably, the small number of lean subjects available limited this kind of analysis and further investigations are needed to confirm those possible differences. It is important to emphasize that our study was not originally designed to assess differences between subjects with and without obesity. Despite that, the over-basal AUC of the estimated LPR was significantly higher in lean subjects, underlying that, in subjects with obesity, the capability of increasing lactate in post-prandial condition is limited by the degree of insulin resistance, which is strongly related to the obesity state [55], as already observed in [68].

5.2 Oral Glucose-Lactate Simulation model

In the second part of the thesis, a simulator model of glucose, insulin, C-peptide and lactate was developed [72]. To do so, the lactate kinetic minimal model was adapted to be directly integrated into the Padova T2D simulator [44]. This was necessary since, in the simulator, the glucose subsystem was described by a two compartment model, while, in the minimal model, a single compartment description was adopted. That was also why different models of lactate kinetics were developed and tested (see Appendix A) to prove that the adapted minimal model would not be outperformed by others. Briefly, despite all the tested models provided good performance in predicting lactate data (see Appendix A.3 for details) again the minimal model-based structure was selected as the best one.

It is important to point out that the original glucose-insulin-C-peptide simulator [44] was developed by using tracer data of non-diabetic and T2D adults. As it was for the minimal model, the simulator was identified using *a priori* knowledge derived from adults (see Appendix A.2.1 for details). In fact, the identification of the simulation model was more difficult by using a Fisher approach, as shown previously in [71][78], thus we used the covariance matrix of model parameters determined in [44], which included a correlation among model parameters and subject anthropometric characteristics, such as age and obesity degree. Although using more informative priors, the simulation model was able to account for the physiological differences related to the age, as previously shown in the literature [57][58][77], making it suitable to describe also such population (see Paragraph 4.3.2). In fact, the *in-silico* population matched the real one both in terms of substrate/hormone AUCs and model parameter distributions, highlighting that the method was able to reproduce the inter subjects' variability of a real population. These results were also supported by

the equivalence test, which yielded always significant results.
Finally, by comparing lean adolescents to those with obesity, no statistical differences were identified between the two groups for the key model parameters. Furthermore, the results of the equivalence and non-inferiority tests were consistent with those obtained with the minimal model, particularly for lactate-related parameters. This highlights that, despite the structural differences and distinct purposes of the two models (one designed for parameter estimation and the other for simulation), their outcomes were practically the same.

5.3 Limitations of the present work

Despite the good results achieved by both the minimal and the simulation models proposed in this work, some limitations must be acknowledged.
One of the most important is that the models were developed and validated exclusively using lactate data from adolescents. This inherently limits the ability to generalize the findings to adults, whose metabolic responses may differ due to various age-related factors, such as changes in insulin sensitivity, insulin secretion, and overall metabolism, as already shown in [57][58][77]. Prior studies have demonstrated significant differences in insulin sensitivity between adults and adolescents, and the models were able to account for that (see Paragraphs 3.3.3 and 4.3.2), but the role of lactate in adult populations remains unexplored in this thesis due to the absence of adult's lactate data. Moreover, the adolescents studied did not have conditions such as diabetes or other metabolic diseases. While obesity itself impacted both glucose and lactate kinetics (see Paragraphs 3.3.3 and 4.3.2), it is expected that the effects of conditions like T2D or MASLD would be more pronounced. Thus, the absence of subjects with

more advanced metabolic disorders limits the model's ability to describe the entire range of potential lactate and glucose metabolic dysfunctions. While the current study identified some important differences between adolescents with and without obesity, these findings could provide an incomplete picture of the impact of obesity on this metabolite due to the small number of lean subjects analyzed. In fact, although the estimated LPR was significantly different between adolescents with/without obesity, lactate-related parameters were not and the impact of obesity on LPR was mainly attributed to disparities in insulin sensitivity, which also play a major role in determining lactate synthesis. To possibly confirm these observations and to robustly assess if and how much obesity influence lactate clearance and conversion rate, a larger sample size and a broader range of metabolic conditions are necessary.

Other than this, the models presented in this thesis were based on data collected during a 3- to 4-hour OGTT. The short duration of the OGTT precluded the possibility of capturing more prolonged or dynamic lactate metabolic fluctuations, which could be observed over longer periods, such as an entire day, during fasting or exercise conditions.

In addition, the lack of EGP data in this study precluded the possibility to model lactate's gluconeogenic contribution. Nonetheless, the model's estimates of EGP were consistent with findings of previous studies [69][70] that used tracer techniques, suggesting that the omission of gluconeogenesis may not have drastically affected the results under the current protocol. However, this limitation does highlight the need for future models to incorporate additional metabolic pathways, such as gluconeogenesis, to capture lactate's full role in metabolic regulation.

The models were designed and validated exclusively in the context of an OGTT. For instance, lactate dynamics may differ significantly in response to different

types of nutrient challenges, such as mixed meal tolerance test (MMTT), which contains fats and proteins in addition to carbohydrates. Similarly, exercise is another condition under which lactate metabolism would likely follow a distinct pattern, yet the current models do not account for such variability. As a result, the models' application is limited to these specific experimental conditions.

5.4 Future Development

A first future development of this work, given the significant correlations observed between model parameters and subjects' characteristics, is to resort to a mixed effects approach. In fact, using this methodology, subject-specific covariates, such as anthropometric variables, can be integrated into the model structure to enhance its applicability and accuracy. For example, Denti et al. discussed the use of nonlinear mixed-effects modeling for the IVGTT glucose minimal model [37], which first improved parameter estimate precision, but also accounted for between-subject variability by including demographic characteristics as covariates [79]. Also, the population-based approach could be useful in the case of sparse sampling, for which accurate parameter estimation cannot be determined. In fact this method allow borrowing information across all subjects simultaneously, quantifying population features directly, and subsequently, deriving individual parameter estimates [80]. However, as usually done in the field, before approaching such methodology a physiological model, as done in this work, must be developed first.

Secondly, the present work investigated only the lactate kinetic so far, without taking into consideration its signaling role and relationship with molecules other than glucose. Future studies may consider to determine lactate effect or relationship with other substrates, such as FFA, to improve our ability to quantitatively assess the intricate metabolic interactions in humans. This would be particularly important to study obesity, diabetes and MASLD, in which such molecules represent predictors of the disease state. In [81] we have developed a minimal model of the FFA kinetic in adults with/without T2D during three-5h MMTT (i.e. breakfast, lunch and dinner), showing the existence of a daily pattern in some key model parameters and fluxes, along with differences between the two groups. The possibility to analyze glucose, lactate, and FFA data

simultaneously and throughout an entire day could significantly enhance the models capabilities, e.g. to conduct comprehensive simulations that span a full day. This advancement would enable the inclusion of circadian fluctuations, like it was done for parameters representing insulin sensitivity in subjects with T1D [82].

Additionally, measuring lactate levels during various daily-life activities, such as physical exercise, could enhance the accuracy and predictive capabilities of current glucose models under these specific conditions [83]. This approach would not only improve the existing models but also allow for the description and integration of other metabolic pathways, like the role of lactate in gluconeogenesis as observed during exercise [84]. From this point of view, the use of minimally invasive lactate sensors, like the continuous glucose monitoring (CGM) systems currently employed for glucose tracking, represents a promising future development to monitor both glucose and lactate during everyday life conditions. This would enable, first the application of the glucose-lactate minimal model in real-world, free-living conditions without the need of resorting to OGTT data, and, therefore, significantly reducing the burden for the patients. This has recently done, for instance, by designing the Minimally Invasive Oral Glucose Minimal Model (MI-OMM) [85], which successfully substituted plasma glucose data with interstitial glucose data from wearable sensors, demonstrating strong agreement between parameters estimated using plasma measurements and CGM. Extending this approach to minimally invasive lactate measurements could enhance the usability of the glucose-lactate minimal model, broadening its application to everyday life. Finally, incorporating minimally invasive lactate measurements into simulation models, could enable testing a new generation of control algorithms/decision support system for managing Type 1 [48] and Type 2 diabetes [51], accounting for this important

metabolite.

Appendices

Appendix A

Simulation Model Development

A.1 Lactate subsystem models

[1]As previously described, the glucose subsystem was formerly developed using tracer experiment and it was able to predict real glucose data, along with fluxes [42][43][44], before been used for simulation purposes. As described in paragraph 4.1.3, the lactate kinetic minimal model was adapted to facilitate its integration into the simulator (its adaptation is reported in Eq. 4.10 and is here referred as Model A1). To ensure that Model A1 would not be outperformed by other ones, once it is coupled with a two compartment model of glucose kinetics, two additional lactate models (Models A2 and A3, described below) based on a very recent publication [39], were tested.

Model A2

Model A2 has the same structure of Model A1, with the difference that lactate plasma volume, V_L, was no more assumed to be equal to V_G, but, based on

[1]This appendix contains material published in [72]

[39], it was estimated during the model identification (see Paragraph A.2.1).

Model A3

Model A3 differs from the previous two since it assumes a two compartments description of lactate kinetics, as recently proposed by Romeres et.al. in [39]. In the original work, the model was identified on lactate tracer data, during exercise under euglycemic and hyperglycemic clamp with a primed infusion of [1-^{13}C] lactate tracer, and used to assess the impact of exercise on lactate clearance. The model equations are:

$$\begin{cases} \dot{Q}_{LP}(t) = -k_{21L} \cdot Q_{LP}(t) + 2fr[U_{ii}(t) + U_{id}(t)] + k_{12L} \cdot Q_{LT}(t) & Q_{LP}(0) = L_b \cdot V_L \\ \dot{Q}_{LT}(t) = -(k_L + k_{12L}) \cdot Q_{LT}(t) + k_{21L} \cdot Q_{LP}(t) & Q_{LT}(0) = \frac{k_{21L} \cdot L_b \cdot V_L}{k_L + k_{12L}} \\ L(t) = \frac{Q_{LP}(t)}{V_L} \end{cases} \tag{A.1}$$

with Q_{LP} and Q_{LT} (mg/kg) being the lactate plasma and tissue masses. The model assumes that exchanges between plasma and tissue compartments occur, with rate parameters k_{21L} and k_{12L} (min^{-1}), while only tissue lactate undergoes an irreversible utilization with rate parameter k_L (min^{-1}). A schematic representation of the three new tested models is reported in Figure A.1.

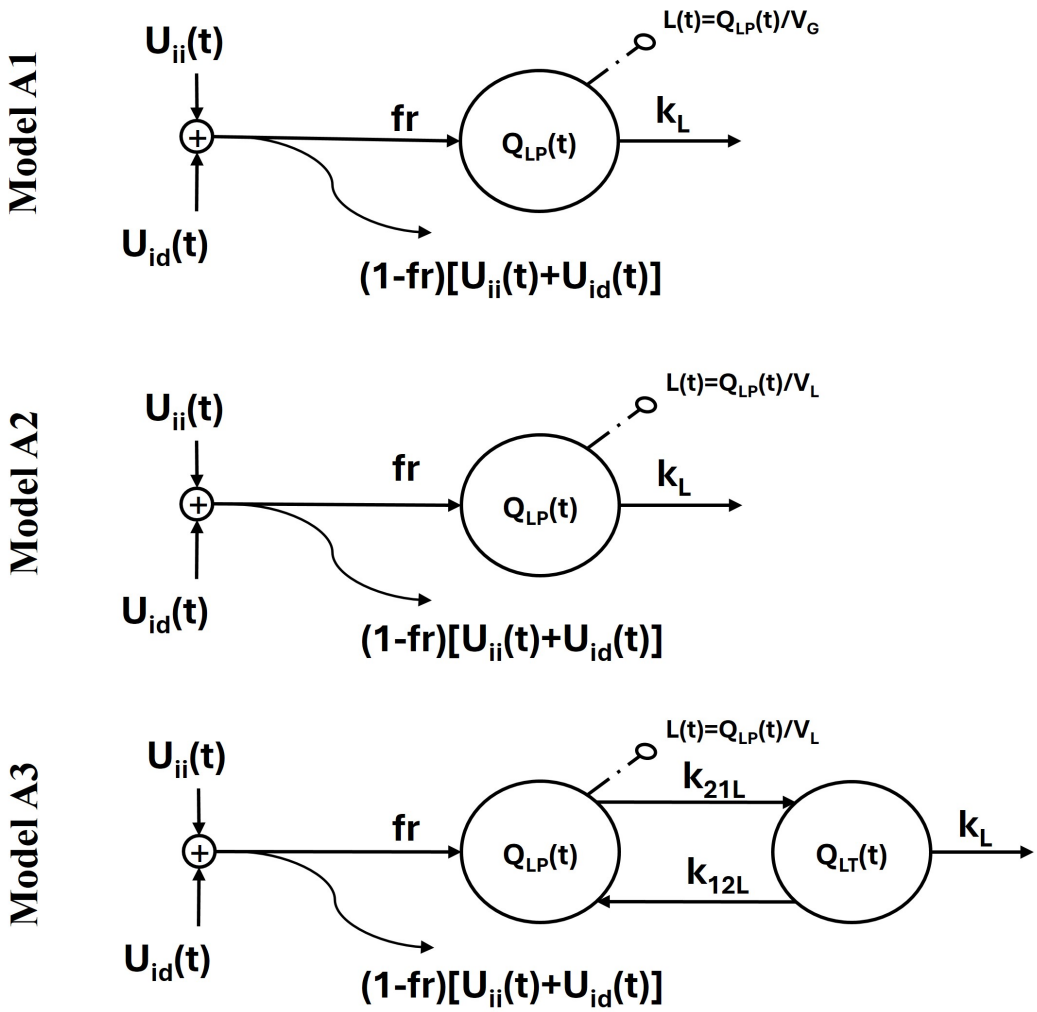

Figure A.1: Schematic representation of the three new tested lactate kinetic models. Continuous lines represent fluxes, while dashed-dotted lines with circle indicate the measured plasma concentrations. U_{ii} and U_{id} are the glucose insulin independent and insulin dependent utilizations respectively (Equation 4.8); *fr* is the fraction of the total glucose utilization that generates lactate; Q_{LP} and Q_{LT} are the lactate mass in the plasma and tissue respectively; L is the lactate plasma concentration; k_{12L} and k_{21L} are the rate constant of the exchanges between plasma and tissue (Model A3); k_L is the rate constant of plasma lactate utilization (Model A1-A2) and tissue lactate utilization (Model A3); V_G and V_L are the glucose and lactate plasma volume respectively.

A.2 Model identification and assessment

In the *Parameter estimation* section, the main assumptions and method of parameters estimation are presented. In the *Model assessment* section, the methods used to evaluate model performance along with the criteria used to select the best performing one are reported. Finally, in the *Statistical Analysis*, the statistical tests used to compare the models are described.

A.2.1 Parameter estimation

Originally, the T2Ds was developed by using system decomposition and forcing function strategy since EGP, glucose rate of appearance from the meal (R_a), and glucose rate of disappearance (R_d) were determined thanks to a triple-tracer experiment [44]. In our study, these measurements were not available, therefore we fitted the model on all the available experimental measurements. It is worth noting that the lactate availability allowed to segregate glucose production from disposal, since lactate production is a fraction of glucose utilization only. In addition, similarly to [71][78], the Bayesian MAP estimator was employed, exploiting the *a priori* knowledge derived from a set of 255 adults subjects (204 healthy and 51 T2D)[44], since the model was too complex to be identified with a Fisher approach. In particular, the *a priori* probability density function of model parameters was assumed to be a lognormal distribution with mean μ_p and covariance Σ_p, determined in [44]. It is also worth noting that parameters representing lactate kinetics were assumed to have infinite variance and zero correlation with the other model parameters in Model A1, while, V_L in Model A2 and V_L , k_{21L} , and k_{12L} in Model A3, were assumed to have *a priori* information with mean derived from the one estimated in [39] (i.e., V_L = 47 mL/kg, k_{12L} = 0.07 min^{-1}, and k_{21L} = 0.81 min^{-1}) and

standard deviation equal to 20% of the mean. Each model was identified on glucose, lactate, insulin and C-peptide data, simultaneously. The assumptions on insulin and C-peptide measurement error were the same adopted in Toffolo et.al. [86], while, for glucose and lactate, measurement errors were assumed to be independent, Gaussian, with zero mean and standard deviation assumed to be known (coefficient of variation CV = 2%), as previously done. Thus, the covariance matrices of the measurement error (Σ_G, Σ_L, Σ_I, Σ_{CP}) were:

$$\begin{aligned}
\Sigma_G &= \begin{bmatrix} (0.02G_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (0.02G_{NG})^2 \end{bmatrix} \\
\Sigma_L &= \begin{bmatrix} (0.02L_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (0.02L_{NL})^2 \end{bmatrix} \\
\Sigma_I &= \begin{bmatrix} 6 + 0.0055 \cdot {I_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 6 + 0.0055 \cdot {I_{NI}}^2 \end{bmatrix} \\
\Sigma_{CP} &= \begin{bmatrix} 2000 + 0.001 \cdot {CP_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 2000 + 0.001 \cdot {CP_{NC}}^2 \end{bmatrix}
\end{aligned} \tag{A.2}$$

where G, L, I and CP indicate glucose, lactate, insulin and C-peptide plasma concentrations respectively, while N_G, N_L, N_I, N_{CP} are the glucose, lactate, insulin and C-peptide number of samples, respectively.

The objective function to be minimized was thus:

$$\begin{aligned} \text{OBJ}(\hat{p}) = &(Y_G - \hat{Y}_G)^T \cdot \Sigma_G^{-1} \cdot (Y_G - \hat{Y}_G) + (Y_L - \hat{Y}_L)^T \cdot \Sigma_L^{-1} \cdot (Y_L - \hat{Y}_L) \\ &+(Y_I - \hat{Y}_I)^T \cdot \Sigma_I^{-1} \cdot (Y_I - \hat{Y}_I) + (Y_{CP} - \hat{Y}_{CP})^T \cdot \Sigma_{CP}^{-1} \cdot (Y_{CP} - \hat{Y}_{CP}) \\ &+(\hat{p} - \mu_p)^T \cdot \Sigma_p^{-1} \cdot (\hat{p} - \mu_p) \end{aligned} \tag{A.3}$$

where Y and $\hat{Y}$ were the experimental data and the model prediction vectors, respectively, with subscripts G, L, I, and CP indicating the glucose, lactate, insulin, and C-peptide plasma concentrations; Σ_G, Σ_L, Σ_I, Σ_{CP} were the covariance matrices of the measurement errors; $\hat{p}$ was the log-transformed parameter vector to be estimated, while μ_p was its expected value and Σ_p was its covariance matrix.

A.2.2 Model assessment

As previously done (section 3.2.3), model assessment was performed by checking if the following criteria were met: the normality [62] and randomness [63] of the weighed residuals; the ability of the model to fit the data, assessed by means of the WRSS, as well as visual inspection; the ability to provide precisely estimated parameter estimates (*a posteriori identifiability*) assessed by the coefficient of variation (CV%) and the percentage of subjects with at least one parameter estimated with CV%>100% (N_{CV}). The models which passed the previous tests were compared by means of a parsimony criterion, i.e. the BIC, defined as in Equation 3.19.

A.2.3 Statistical analysis

Data and results are reported as median [25th -75th], if not normally distributed, or mean±SD, if normally distributed. Paired and unpaired compar-

isons were done by using Student-T test/ Wilcoxon test, or unpaired T test/ Mann Whitney test respectively, depending on distributions (assessed using Shapiro-Wilk test [64]). The non-significant comparisons were also analyzed by performing equivalence and non-inferiority tests, as previously described in 3.2.4. Significance level was assumed as α=0.05 for two-sided comparisons tests, α=0.025 otherwise. The analysis were conducted in Matlab R2022b [61] and in R (version 4.3.2) [67].

A.3 Model selection

Three models were tested and compared as described in paragraphs A.1 and A.2.2. All of them showed good performance in terms of randomness and normality of the residuals. The ability of the models to predict the data were satisfactory and no statistical differences appeared on WRSS distributions (P=0.75). However, Model A3 was discarded since it provided imprecise parameter estimates, with 20% of the subjects providing at least one parameter estimated with CV%>100%. In fact, the model considered exchanges between the two compartments representing the plasma and tissue lactate and these were probably not observable during postprandial conditions, as they were in the former study. Models A1 and A2, which were comparable in terms of *a posteriori identifiability* were thus compared in terms of BIC, showing that Model A1 performed better than A2 (P=0.02). In addition, V_L estimated in Model A2 was not statistically different (Wilcoxon test P=0.85, equivalence test P<10^{-10}) and strongly correlated with V_G (ρ=0.97, P<10^{-4}), confirming that, in this experimental condition and with this model structure, assuming equal lactate and glucose plasma volume was the best choice. To conclude, Model A1 was selected as the best one and integrated into the simulator. Table A.1 summa-

rizes these results.

Table A.1: Simulation Model selection

Model	Run Test	Normality Test	WRSS	CV%	N_{CV}	BIC
A1	**95**%	**98**%	**400 [300-523]**	**19 [9-29]**%	**9**%	**114 [112-124]**
A2	97%	100%	370 [279-503]	17 [8-28]%	11%	118 [116-128]*
A3	97%	100%	406 [268-518]	21 [11-33]%	20%	124 [122-135]**

Data are reported as median and [25th-75th] percentile or percentages. The bold values highlight the best selected model. WRSS: weighted residual sum of squares; N_{CV}: percentage of subjects with at least one parameters estimated with CV>100%; BIC: Bayesian information criterion. *, ** indicate p<0.05 and <0.01, with respect to Model A1.

In Figure A.2 the weighted residuals obtained by using each tested model for glucose, lactate, insulin and C-peptide are reported. In Tables A.2 and A.3 , estimated parameters, along with the precision of the estimates, are reported.

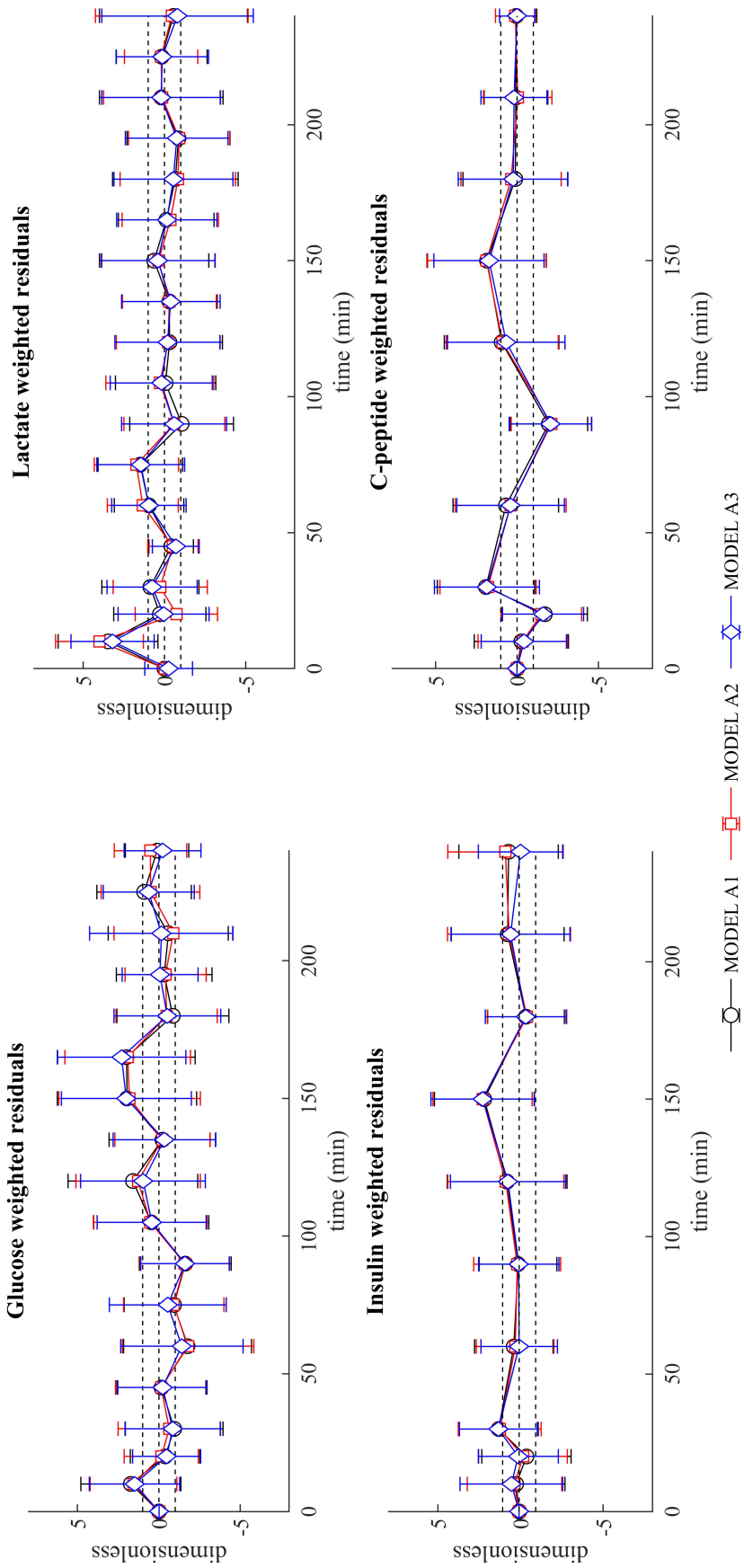

Figure A.2: Mean±SD glucose and lactate (first row), insulin and C-peptide (second row) weighted residuals obtained with Model A1 (black line and circles), Model A2 (red line and squares), and Model A3 (blue line and diamonds). Horizontal lines indicate [-1: 1] range.

Table A.2: Glucose subsystem parameters

Parameter	Description and unit	Estimate (median [25th-75th])	Precision (CV%) (median [25th-75th])
Gastro-intestinal tract model Eq. 4.5			
k_{abs}	Rate constant of intestine absorption (min^{-1})	0.321 [0.185-0.401]	34 [27-42]
k_{max}	Maximum gastric emptying rate (min^{-1})	0.045 [0.035-0.053]	11 [7-15]
k_{min}	Minimum gastric emptying rate (min^{-1})	0.007 [0.005-0.010]	16 [9-23]
b	Inflection point of gastric emptying curve (dimensionless)	0.811 [0.727-0.867]	3 [2-4]
d	Inflection point of gastric emptying curve (dimensionless)	0.163 [0.139-0.209]	14 [10-27]
Endogenous glucose production Eq. 4.7			
EGP_b	Basal endogenous glucose production (mg(kg/min)	2.47 [2.31-2.84]	8 [7-10]
k_{p2}	Hepatic glucose effectiveness (min^{-1})	0.008 [0.005-0.011]	23 [15-37]
k_{p3}	Hepatic insulin sensitivity (mg/kg/min per pmol/L)	0.009 [0.005-0.014]	22 [16-38]
k_{p4}	Portal insulin sensitivity (mg/kg/min per pmol/kg)	0.039 [0.029-0.054]	33 [24-48]
k_i	Rate of delayed insulin action (min^{-1})	0.012 [0.006-0.016]	24 [16-35]
Glucose utilization Eq. 4.8			
V_{mx}	Insulin sensitivity on glucose utilization (mg/kg/min per pmol/L)	0.035 [0.015-0.048]	21 [17-35]
p_{2U}	Rate of insulin action on EGP (min^{-1})	0.042 [0.029-0.065]	18 [13-32]
Km0	Glucose mass in Michealis Menten relation (mg/kg)	248.13 [231.38-268.77]	7 [6-8]
Glucose kinetics Eq. 4.4			
V_G	Distribution volume (dL/kg)	1.75 [1.62-1.83]	6 [5-6]
k_1	Rate parameter (min^{-1})	0.039 [0.019-0.066]	19 [14-24]
k_2	Rate parameter (min^{-1})	0.136 [0.078-0.162]	26 [24-34]

Median and [25th -75th] percentile estimated model parameters with CV%. The remaining not reported parameters were calculated during the identification process to guarantee the steady state condition at time 0 min or fixed using literature values.

Table A.3: Lactate, Insulin and C-peptide subsystem parameters

Parameter	Description and unit	Estimate (median [25th-75th])	Precision (CV%) (median [25th-75th])
Lactate kinetics Eq. 4.10			
fr	Fraction of the glucose utilization that generates lactate (dimensionless)	18 [11-29]	23 [19-28]
k_L	Fractional lactate disposal rate (min^{-1})	0.075 [0.054-0.110]	19 [15-25]
Insulin and C-peptide secretion Eq. 4.1			
Φd	β-cell responsivity to glucose rate of change (10^{-9})	1301 [963-1885]	6 [5-8]
α	Delay between glucose and insulin secretion (min^{-1})	0.053 [0.027-0.163]	15 [10-26]
Φs	β-cell responsivity to glucose (10^{-9} min^{-1})	63 [42-113]	4 [4-5]
Insulin kinetics Eq. 4.3			
V_i	Distribution volume (L/kg)	0.036 [0.023-0.050]	25 [20-30]
m_1	Rate parameter (min^{-1})	0.71 [0.34-1.17]	35 [29-44]
m_4	Rate parameter (min^{-1})	0.27 [0.18-0.41]	29 [23-39]
m_5	Rate parameter (min^{-1})	0.13 [0.08-0.22]	32 [24-51]
m_6	Rate parameter (min^{-1})	0.017 [0.012-0.022]	25 [20-35]
a_G	Glucose control on HE (dL/mg)	0.010 [0.007-0.014]	20 [15-24]

Median and [25th -75th] percentile estimated model parameters with CV%. The remaining not reported parameters were calculated during the identification process to guarantee the steady state condition at time 0 min or fixed using literature values.

Bibliography

[1] G. A. Brooks, H. Dubouchaud, M. Brown, J. P. Sicurello, and C. E. Butz, "Role of mitochondrial lactate dehydrogenase and lactate oxidation in the intracellular lactate shuttle," *Proceedings of the National Academy of Sciences*, vol. 96, pp. 1129–1134, Feb. 1999. Publisher: Proceedings of the National Academy of Sciences.

[2] G. A. Brooks, "Lactate shuttles in Nature," *Biochemical Society Transactions*, vol. 30, pp. 258–264, Apr. 2002.

[3] G. A. Brooks, "The Science and Translation of Lactate Shuttle Theory," *Cell Metabolism*, vol. 27, pp. 757–785, Apr. 2018. Publisher: Elsevier.

[4] G. A. Brooks, J. A. Arevalo, A. D. Osmond, R. G. Leija, C. C. Curl, and A. P. Tovar, "Lactate in contemporary biology: a phoenix risen," *The Journal of Physiology*, vol. 600, pp. 1229–1251, Mar. 2022.

[5] C. F. Cori and G. T. Cori, "GLYCOGEN FORMATION IN THE LIVER FROM *d*- AND *l*-LACTIC ACID," *Journal of Biological Chemistry*, vol. 81, pp. 389–403, Feb. 1929.

[6] J. Katz and J. A. Tayek, "Gluconeogenesis and the Cori cycle in 12-, 20-, and 40-h-fasted humans," *The American Journal of Physiology*, vol. 275, pp. E537–542, Sept. 1998.

[7] M. J. Rogatzki, B. S. Ferguson, M. L. Goodwin, and L. B. Gladden, "Lactate is always the end product of glycolysis," *Frontiers in Neuroscience*, vol. 9, p. 22, Feb. 2015.

[8] L. B. Gladden, "Lactate metabolism: a new paradigm for the third millennium," *The Journal of Physiology*, vol. 558, pp. 5–30, July 2004.

[9] A. Bonen, "The expression of lactate transporters (MCT1 and MCT4) in heart and muscle," *European Journal of Applied Physiology*, vol. 86, pp. 6–11, Nov. 2001.

[10] P. J. Magistretti and L. Pellerin, "Astrocytes Couple Synaptic Activity to Glucose Utilization in the Brain," *Physiology*, vol. 14, pp. 177–182, Oct. 1999.

[11] M. Nalbandian and M. Takeda, "Lactate as a Signaling Molecule That Regulates Exercise-Induced Adaptations," *Biology*, vol. 5, p. 38, Dec. 2016. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[12] V. Mosienko, A. G. Teschemacher, and S. Kasparov, "Is L-Lactate a Novel Signaling Molecule in the Brain?," *Journal of Cerebral Blood Flow & Metabolism*, vol. 35, pp. 1069–1075, July 2015. Publisher: SAGE Publications Ltd STM.

[13] N. Santoro, S. Caprio, B. Pierpont, M. Van Name, M. Savoye, and E. J. Parks, "Hepatic De Novo Lipogenesis in Obese Youth Is Modulated by a Common Variant in the GCKR Gene," *The Journal of Clinical Endocrinology and Metabolism*, vol. 100, pp. E1125–E1132, Aug. 2015.

[14] Y. Takahashi and T. Fukusato, "Pediatric nonalcoholic fatty liver disease: Overview with emphasis on histology," *World Journal of Gastroenterol-*

ogy, vol. 16, pp. 5280–5285, Nov. 2010. Publisher: Baishideng Publishing Group Inc.

[15] Q. Lu, X. Tian, H. Wu, J. Huang, M. Li, Z. Mei, L. Zhou, H. Xie, and S. Zheng, "Metabolic Changes of Hepatocytes in NAFLD," *Frontiers in Physiology*, vol. 12, Aug. 2021. Publisher: Frontiers.

[16] T. Wang, K. Chen, W. Yao, R. Zheng, Q. He, J. Xia, J. Li, Y. Shao, L. Zhang, L. Huang, L. Qin, M. Xu, Z. Zhang, D. Pan, Z. Li, and F. Huang, "Acetylation of lactate dehydrogenase B drives NAFLD progression by impairing lactate clearance," *Journal of Hepatology*, vol. 74, pp. 1038–1052, May 2021.

[17] M. Adeva-Andany, M. López-Ojén, R. Funcasta-Calderón, E. Ameneiros-Rodríguez, C. Donapetry-García, M. Vila-Altesor, and J. Rodríguez-Seijas, "Comprehensive review on lactate metabolism in human health," *Mitochondrion*, vol. 17, pp. 76–100, July 2014.

[18] M.-T. van der Merwe, G. P. Schlaphoff, N. J. Crowther, I. H. Boyd, I. P. Gray, B. I. Joffe, and P. N. Lönnroth, "Lactate and Glycerol Release from Adipose Tissue in Lean, Obese, and Diabetic Women from South Africa1," *The Journal of Clinical Endocrinology & Metabolism*, vol. 86, pp. 3296–3303, July 2001.

[19] R. Weiss, S. Dufour, S. E. Taksali, W. V. Tamborlane, K. F. Petersen, R. C. Bonadonna, L. Boselli, G. Barbetta, K. Allen, F. Rife, M. Savoye, J. Dziura, R. Sherwin, G. I. Shulman, and S. Caprio, "Prediabetes in obese youth: a syndrome of impaired glucose tolerance, severe insulin resistance, and altered myocellular and abdominal fat partitioning," *The Lancet*, vol. 362, pp. 951–957, Sept. 2003. Publisher: Elsevier.

[20] K. F. Petersen, S. Dufour, D. Befroy, R. Garcia, and G. I. Shulman, "Impaired Mitochondrial Activity in the Insulin-Resistant Offspring of Patients with Type 2 Diabetes," *New England Journal of Medicine*, vol. 350, pp. 664–671, Feb. 2004. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa031314.

[21] K. L. Donnelly, C. I. Smith, S. J. Schwarzenberg, J. Jessurun, M. D. Boldt, and E. J. Parks, "Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease," *Journal of Clinical Investigation*, vol. 115, pp. 1343–1351, May 2005.

[22] E. J. Parks and M. K. Hellerstein, "Carbohydrate-induced hypertriacylglycerolemia: historical perspective and review of biological mechanisms123," *The American Journal of Clinical Nutrition*, vol. 71, pp. 412–433, Feb. 2000.

[23] Y. Ogawa, K. Imajo, Y. Honda, T. Kessoku, W. Tomeno, S. Kato, K. Fujita, M. Yoneda, S. Saito, Y. Saigusa, H. Hyogo, Y. Sumida, Y. Itoh, K. Eguchi, T. Yamanaka, K. Wada, and A. Nakajima, "Palmitate-induced lipotoxicity is crucial for the pathogenesis of nonalcoholic fatty liver disease in cooperation with gut-derived endotoxin," *Scientific Reports*, vol. 8, p. 11365, July 2018.

[24] E. D'Adamo, A. M. G. Cali, R. Weiss, N. Santoro, B. Pierpont, V. Northrup, and S. Caprio, "Central role of fatty liver in the pathogenesis of insulin resistance in obese adolescents," *Diabetes Care*, vol. 33, pp. 1817–1822, Aug. 2010.

[25] E. Castillo-Leon, C. E. Cioffi, and M. B. Vos, "Perspectives on youth-onset nonalcoholic fatty liver disease," *Endocrinology, Diabetes & Metabolism*,

vol. 3, p. e00184, Oct. 2020.

[26] E. Barbieri, N. Santoro, and G. R. Umano, "Clinical features and metabolic complications for non-alcoholic fatty liver disease (NAFLD) in youth with obesity," *Frontiers in Endocrinology*, vol. 14, p. 1062341, Jan. 2023.

[27] J. B. Schwimmer, R. Deutsch, J. B. Rauch, C. Behling, R. Newbury, and J. E. Lavine, "Obesity, insulin resistance, and other clinicopathological correlates of pediatric nonalcoholic fatty liver disease," *The Journal of Pediatrics*, vol. 143, pp. 500–505, Oct. 2003.

[28] Y. Lin, M. Bai, S. Wang, L. Chen, Z. Li, C. Li, P. Cao, and Y. Chen, "Lactate Is a Key Mediator That Links Obesity to Insulin Resistance via Modulating Cytokine Production From Adipose Tissue," *Diabetes*, vol. 71, pp. 637–652, Apr. 2022.

[29] R. Durand, M. Galli, M. Chenavard, D. Bandiera, H. Freund, and L. A. Messonnier, "Modelling of Blood Lactate Time-Courses During Exercise and/or the Subsequent Recovery: Limitations and Few Perspectives," *Frontiers in Physiology*, vol. 12, Oct. 2021. Publisher: Frontiers.

[30] H. Freund, S. Oyono-Enguelle, A. Heitz, J. Marbach, C. Ott, and M. Gartner, "Effect of exercise duration on lactate kinetics after short muscular exercise," *European Journal of Applied Physiology and Occupational Physiology*, vol. 58, pp. 534–542, Mar. 1989.

[31] R. Beneke, M. Hütler, M. Jung, and R. M. Leithäuser, "Modeling the blood lactate kinetics at maximal short-term exercise conditions in children, adolescents, and adults," *Journal of Applied Physiology*, vol. 99, pp. 499–504, Aug. 2005. Publisher: American Physiological Society.

[32] O. J. Quittmann, T. Abel, S. Zeller, T. Foitschik, and H. K. Strüder, "Lactate kinetics in handcycling under various exercise modalities and their relationship to performance measures in able-bodied participants," *European Journal of Applied Physiology*, vol. 118, pp. 1493–1505, July 2018.

[33] P. Wahl, Z. Yue, C. Zinner, W. Bloch, and J. Mester, "A mathematical model for lactate transport to red blood cells," *The Journal of Physiological Sciences : JPS*, vol. 61, pp. 93–102, Dec. 2010.

[34] J. Kim, G. M. Saidel, and M. E. Cabrera, "Multi-Scale Computational Model of Fuel Homeostasis During Exercise: Effect of Hormonal Control," *Annals of Biomedical Engineering*, vol. 35, pp. 69–90, Jan. 2007.

[35] M. C. Palumbo, M. Morettini, P. Tieri, F. Diele, M. Sacchetti, and F. Castiglione, "Personalizing physical exercise in a computational model of fuel homeostasis," *PLOS Computational Biology*, vol. 14, p. e1006073, Apr. 2018. Publisher: Public Library of Science.

[36] R. M. Watanabe, J. Lovejoy, G. M. Steil, M. DiGirolamo, and R. N. Bergman, "Insulin sensitivity accounts for glucose and lactate kinetics after intravenous glucose injection," *Diabetes*, vol. 44, pp. 954–962, Aug. 1995.

[37] R. N. Bergman, Y. Z. Ider, C. R. Bowden, and C. Cobelli, "Quantitative estimation of insulin sensitivity.," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 236, p. E667, Jun 1979.

[38] D. Stefanovski, J. H. Youn, M. Rees, R. M. Watanabe, M. Ader, V. Ionut, A. U. Jackson, M. Boehnke, F. S. Collins, and R. N. Bergman, "Estimating hepatic glucokinase activity using a simple model of lactate kinetics," *Diabetes Care*, vol. 35, pp. 1015–1020, May 2012.

[39] D. Romeres, Y. Yadav, F. N. U. Ruchi, R. Carter, C. Cobelli, R. Basu, and A. Basu, "Hyperglycemia suppresses Lactate Clearance during Exercise in Type 1 Diabetes," *The Journal of Clinical Endocrinology and Metabolism*, p. dgae005, Jan. 2024.

[40] C. Dalla Man, A. Caumo, and C. Cobelli, "The oral glucose minimal model: Estimation of insulin sensitivity from a meal test," *IEEE Transactions on Biomedical Engineering*, vol. 49, pp. 419–429, May 2002.

[41] R. Visentin, C. Dalla Man, R. Basu, A. Basu, R. A. Rizza, and C. Cobelli, "Hepatic insulin sensitivity in healthy and prediabetic subjects: from a dual- to a single-tracer oral minimal model," *American Journal of Physiology - Endocrinology and Metabolism*, vol. 309, pp. E161–E167, July 2015.

[42] C. Dalla Man, R. A. Rizza, and C. Cobelli, "Meal simulation model of the glucose-insulin system," *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 1740–1749, Oct 2007.

[43] C. Dalla Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The UVA/PADOVA type 1 diabetes simulator," *Journal of Diabetes Science and Technology*, vol. 8, pp. 26–34, Jan 2014.

[44] R. Visentin, C. Cobelli, and C. Dalla Man, "The Padova Type 2 Diabetes Simulator from Triple-Tracer Single-Meal Studies: In Silico Trials Also Possible in Rare but Not-So-Rare Individuals," *Diabetes Technology & Therapeutics*, vol. 22, pp. 892–903, Dec. 2020.

[45] R. Hovorka, F. Shojaee-Moradie, P. V. Carroll, L. J. Chassin, I. J. Gowrie, N. C. Jackson, R. S. Tudor, A. M. Umpleby, and R. H. Jones, "Partitioning glucose distribution/transport, disposal, and endogenous production

during IVGTT," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 282, pp. E992–E1007, May 2002. Publisher: American Physiological Society.

[46] J. T. Sorensen, *A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes*. PhD thesis, Massachusetts Institute of Technology, 1985.

[47] J. Chertoff, M. Chisum, B. Garcia, and J. Lascano, "Lactate kinetics in sepsis and septic shock: a review of the literature and rationale for further research," *Journal of Intensive Care*, vol. 3, p. 39, 2015.

[48] C. Toffanin, R. Visentin, M. Messori, F. D. Palma, L. Magni, and C. Cobelli, "Toward a Run-to-Run Adaptive Artificial Pancreas: In Silico Results," *IEEE transactions on bio-medical engineering*, vol. 65, pp. 479–488, Mar. 2018.

[49] M. Schiavon, R. Visentin, C. Giegerich, J. Sieber, C. Dalla Man, C. Cobelli, and T. Klabunde, "In Silico Head-to-Head Comparison of Insulin Glargine 300U/mL and Insulin Degludec 100U/mL in Type 1 Diabetes," *Diabetes Technology & Therapeutics*, vol. 22, pp. 553–561, Aug. 2020.

[50] R. Visentin, M. Schiavon, C. Giegerich, T. Klabunde, C. D. Man, and C. Cobelli, "Incorporating Long-Acting Insulin Glargine Into the UVA/Padova Type 1 Diabetes Simulator for In Silico Testing of MDI Therapies," *IEEE Transactions on Biomedical Engineering*, vol. 66, pp. 2889–2896, Oct. 2019. Conference Name: IEEE Transactions on Biomedical Engineering.

[51] J. Bonet, R. Visentin, and C. Dalla Man, "Smart Algorithms for Efficient Insulin Therapy Initiation in Individuals With Type 2 Diabetes: An in Silico Study," *Journal of Diabetes Science and Technology*, p. 19322968241245930, Apr. 2024.

[52] L. D. Hammer, H. C. Kraemer, D. M. Wilson, P. L. Ritter, and S. M. Dornbusch, "Standardized percentile curves of body-mass index for children and adolescents," *American Journal of Diseases of Children (1960)*, vol. 145, pp. 259–263, Mar. 1991.

[53] The Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, "Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus," *Diabetes Care*, vol. 26, pp. s5–s20, Jan. 2003.

[54] J. Bonet, B. Galuppo, N. Santoro, and C. D. Man, "A New Oral Model to Assess Postprandial Lactate Production Rate," *IEEE Transactions on Biomedical Engineering*, vol. 69, pp. 1533–1540, Apr. 2022. Conference Name: IEEE Transactions on Biomedical Engineering.

[55] J. Bonet, D. Fox, R. Nelson, M. B. Nelson, L. Nelson, C. Fernandez, E. Barbieri, C. Dalla Man, and N. Santoro, "Modelling and assessment of glucose-lactate kinetics in youth with overweight, obesity and metabolic dysfunction-associated steatotic liver disease: A pilot study," *Diabetes, Obesity & Metabolism*, vol. 26, pp. 3207–3212, Aug. 2024.

[56] G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò, "DAISY: a new software tool to test global identifiability of biological and physiological systems," *Computer Methods and Programs in Biomedicine*, vol. 88, pp. 52–61, Oct. 2007.

[57] S. Arslanian, J. Y. Kim, A. Nasr, F. Bacha, H. Tfayli, S. Lee, and F. G. S. Toledo, "Insulin sensitivity across the lifespan from obese adolescents to obese adults with impaired glucose tolerance: Who is worse off?," *Pediatric Diabetes*, vol. 19, pp. 205–211, Mar. 2018.

[58] M. E. Chen, A. G. Chandramouli, R. V. Considine, T. S. Hannon, and K. J. Mather, "Comparison of β-cell Function Between Overweight/Obese Adults and Adolescents Across the Spectrum of Glycemia," *Diabetes Care*, vol. 41, pp. 318–325, Feb. 2018.

[59] A. Basu, C. Dalla Man, R. Basu, G. Toffolo, C. Cobelli, and R. A. Rizza, "Effects of Type 2 Diabetes on Insulin Secretion, Insulin Action, Glucose Effectiveness, and Postprandial Glucose Metabolism," *Diabetes Care*, vol. 32, pp. 866–872, May 2009.

[60] C. Dalla Man, A. Caumo, R. Basu, R. Rizza, G. Toffolo, and C. Cobelli, "Minimal model estimation of glucose absorption and insulin sensitivity from oral test: validation with a tracer method," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 287, pp. E637–E643, Oct. 2004. Publisher: American Physiological Society.

[61] T. M. Inc., "Matlab version: 9.13.0 (r2022b)," 2022.

[62] T. W. Anderson and D. A. Darling, "A test of goodness of fit," *Journal of the American Statistical Association*, vol. 49, pp. 765–769, 1954. Place: US Publisher: American Statistical Association.

[63] A. Wald and J. Wolfowitz, "On a Test Whether Two Samples are from the Same Population," *The Annals of Mathematical Statistics*, vol. 11, pp. 147–162, June 1940. Publisher: Institute of Mathematical Statistics.

[64] S. S. SHAPIRO and M. B. WILK, "An analysis of variance test for normality (complete samples)†," *Biometrika*, vol. 52, pp. 591–611, Dec. 1965.

[65] Aaron, R., Caldwell, and aut, "Exploring equivalence testing with the updated toster r package," *PsyArXiv*, 2022.

[66] Daniel, Lakens, and aut, "Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses," *Social Psychological and Personality Science*, vol. 1, pp. 1–8, 2017.

[67] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

[68] J. Lovejoy, F. D. Newby, S. S. P. Gebhart, and M. DiGirolamo, "Insulin resistance in obesity is associated with elevated basal lactate levels and diminished lactate appearance following intravenous glucose and insulin," *Metabolism - Clinical and Experimental*, vol. 41, pp. 22–27, Jan. 1992. Publisher: Elsevier.

[69] H. H. Huidekoper, M. T. Ackermans, A. F. C. Ruiter, H. P. Sauerwein, and F. A. Wijburg, "Endogenous glucose production from infancy to adulthood: a non-linear regression model," *Archives of Disease in Childhood*, vol. 99, pp. 1098–1102, Dec. 2014. Publisher: BMJ Publishing Group Ltd Section: Original article.

[70] G. Toffolo, C. D. Man, C. Cobelli, and A. L. Sunehag, "Glucose Fluxes During OGTT in Adolescents Assessed by a Stable Isotope Triple Tracer Method," *Journal of Pediatric Endocrinology and Metabolism*, vol. 21, pp. 31–46, Jan. 2008. Publisher: De Gruyter.

[71] R. Visentin, M. Schiavon, J. Bonet, M. Riz, B. Wagenhuber, and C. D. Man, "Tailoring the Padova Type 2 Diabetes Simulator for Treatment Guidance in Target Populations," *IEEE Transactions on Biomedical Engineering*, pp. 1–9, 2024. Conference Name: IEEE Transactions on Biomedical Engineering.

[72] J. Bonet, E. Barbieri, N. Santoro, and C. Dalla Man, "Modeling Glucose, Insulin, C-Peptide, and Lactate Interplay in Adolescents During an

Oral Glucose Tolerance Test," *Journal of Diabetes Science and Technology*, p. 19322968241266825, July 2024. Publisher: SAGE Publications Inc.

[73] E. Van Cauter, F. Mestrez, J. Sturis, and K. S. Polonsky, "Estimation of insulin secretion rates from C-peptide levels. Comparison of individual and standard kinetic parameters for C-peptide clearance," *Diabetes*, vol. 41, pp. 368–377, Mar. 1992.

[74] R. S. Sherwin, K. J. Kramer, J. D. Tobin, P. A. Insel, J. E. Liljenquist, M. Berman, and R. Andres, "A Model of the Kinetics of Insulin in Man," *The Journal of Clinical Investigation*, vol. 53, pp. 1481–1492, May 1974. Publisher: American Society for Clinical Investigation.

[75] "Reprint of: Mahalanobis, P.C. (1936) "On the Generalised Distance in Statistics."," *Sankhya A*, vol. 80, pp. 1–7, Dec. 2018.

[76] T. Eiter and H. Mannila, "Computing Discrete Fréchet Distance ," 1994.

[77] K. M. Utzschneider, M. T. Tripputi, A. Kozedub, K. J. Mather, K. J. Nadeau, S. L. Edelstein, T. S. Hannon, S. A. Arslanian, M. Cree-Green, T. A. Buchanan, S. Caprio, A. Mari, and RISE Consortium, "β-cells in youth with impaired glucose tolerance or early type 2 diabetes secrete more insulin and are more responsive than in adults," *Pediatric Diabetes*, vol. 21, pp. 1421–1429, Dec. 2020.

[78] R. Visentin, C. D. Man, and C. Cobelli, "One-Day Bayesian Cloning of Type 1 Diabetes Subjects: Toward a Single-Day UVA/Padova Type 1 Diabetes Simulator," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 2416–2424, Nov. 2016. Conference Name: IEEE Transactions on Biomedical Engineering.

[79] P. Denti, A. Bertoldo, P. Vicini, and C. Cobelli, "IVGTT glucose minimal model covariate selection by nonlinear mixed-effects approach," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 298, pp. E950–E960, May 2010. Publisher: American Physiological Society.

[80] P. Denti, A. Bertoldo, P. Vicini, and C. Cobelli, "Nonlinear Mixed Effects to Improve Glucose Minimal Model Parameter Estimation: A Simulation Study in Intensive and Sparse Sampling," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 2156–2166, Sept. 2009. Conference Name: IEEE Transactions on Biomedical Engineering.

[81] J. Bonet, Y. Yadav, J. Miles, A. Basu, C. Cobelli, R. Basu, and C. Dalla Man, "A new oral model of free fatty acid kinetics to assess lipolysis in subjects with and without type 2 diabetes," *American Journal of Physiology. Endocrinology and Metabolism*, vol. 325, pp. E163–E170, Aug. 2023.

[82] R. Visentin, C. Dalla Man, Y. C. Kudva, A. Basu, and C. Cobelli, "Circadian Variability of Insulin Sensitivity: Physiological Input for In Silico Artificial Pancreas," *Diabetes Technology & Therapeutics*, vol. 17, pp. 1–7, Jan. 2015.

[83] D. Romeres, M. Schiavon, A. Basu, C. Cobelli, R. Basu, and C. Dalla Man, "Exercise effect on insulin-dependent and insulin-independent glucose utilization in healthy individuals and individuals with type 1 diabetes: a modeling study," *American Journal of Physiology - Endocrinology and Metabolism*, vol. 321, pp. E122–E129, July 2021.

[84] L. A. Messonnier, B. Chatel, C.-A. W. Emhoff, L. Blervaque, and S. Oyono-Enguéllé, "Delayed Rebound of Glycemia During Recovery Following Short-Duration High-Intensity Exercise: Are There Lactate and Glucose

Metabolism Interactions?," *Frontiers in Nutrition*, vol. 8, p. 734152, Nov. 2021.

[85] E. Faggionato, M. Schiavon, L. Ekhlaspour, B. A. Buckingham, and C. Dalla Man, "The minimally-invasive oral glucose minimal model: Estimation of gastric retention, glucose rate of appearance, and insulin sensitivity from type 1 diabetes data collected in real-life conditions," *IEEE Transactions on Biomedical Engineering*, vol. 71, pp. 977–986, Mar 2024.

[86] G. Toffolo, M. Campioni, R. Basu, R. A. Rizza, and C. Cobelli, "A minimal model of insulin secretion and kinetics to assess hepatic insulin extraction," *American Journal of Physiology-Endocrinology and Metabolism*, vol. 290, pp. E169–E176, Jan. 2006. Publisher: American Physiological Society.