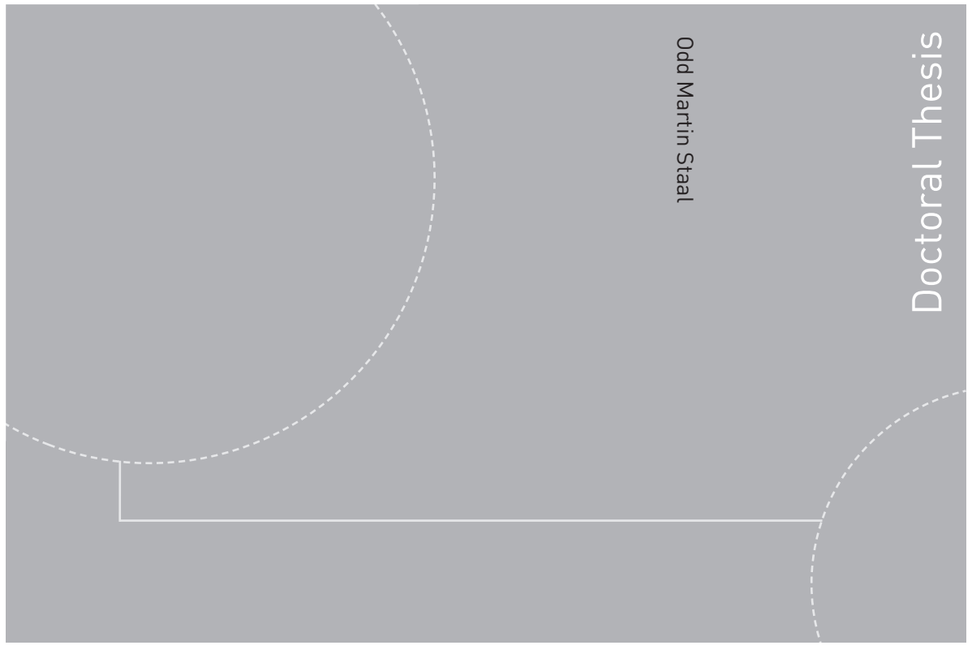
ISBN 978-82-326-3863-8 (printed version)
ISBN 978-82-326-3863-5 (electronic version)
ISSN 1503-8181. ITK-report 2019-4-W




Doctoral theses at NTNU, 2019:132

Odd Martin Staal

# Blood glucose dynamics:

## Identification, smoothing and real time estimation in free-living settings

NTNU
Norwegian University of
Science and Technology

Doctoral theses at NTNU, 2019:132

NTNU

**NTNU**
Norwegian University of
Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of
Science and Technology

Odd Martin Staal

# Blood glucose dynamics:

## Identification, smoothing and real time estimation in free-living settings

Thesis for the degree of Philosophiae Doctor

Trondheim, December 2018

Norwegian University of Science and Technology
Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of
Science and Technology

**NTNU**
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Information Technology
and Electrical Engineering
Department of Engineering Cybernetics



ISBN 978-82-326-3863-8 (printed version)
ISBN 978-82-326-3863-5 (electronic version)
ISSN 1503-8181.

ITK-report 2019-4-W

Doctoral theses at NTNU, 2019:132

Printed by Skipnes Kommunikasjon as

# Abstract

Diabetes currently affects about 9% of the world's population, and causes loss or impairment of the body's ability to control blood glucose (BG). Frequent BG measurements and medications, including insulin injections, are needed to keep BG as close as possible to the normal range.

Dynamic models that describe BG originated in clinical research settings and have helped towards understanding the physiology of diabetes and developing external systems that control BG. Glucose dynamics is highly person-dependent, so individualisation of the BG dynamics model to patient data is desirable. Many diabetes patients are currently using Continuous Glucose Monitors and insulin pumps, gathering large amounts of data from free-living settings, which can potentially be used to individualise models. One challenge with this approach is that the only available measurement from the internal state of the body in these data is the glucose concentration, whereas many models require more measurements, e.g. of plasma insulin, to avoid observability and identifiability problems. Another challenge is that the quality of the data from free-living settings is frequently poor, with missing data and other errors occurring both in the glucose data and in the logged meal data.

The aim of this PhD project has been to use BG dynamics models for real-time estimation and prediction based on continuous glucose data, using individualised models of BG dynamics, and achieving the individualisation based on data from free-living settings. Publications from the PhD project relate to glucose data cleaning, glucose sensor characterisation, identifiability analysis in blood glucose dynamics models, and meal detection in glucose data, and are included as appendices to the thesis. The main text of the thesis contributes a review of physiology of blood glucose dynamics, diabetes, diabetes related technology, blood glucose modelling and identification of dynamic models. The PhD work contributes to Prediktor Medical's development related to products for blood glucose measurement and prediction.

# Preface

This thesis is submitted in partial fulfilment of the requirements for the degree of philosophiae doctor (PhD) at the Norwegian University of Science and Technology (NTNU), Department of Engineering Cybernetics.

This PhD project is funded as an industrial PhD (NæringsPhD), grant number 242167, and has been carried out in cooperation with Prediktor Medical, Fredrikstad, NTNU, and Norges Forskningsråd (Research Council of Norway).

## Acknowledgements

I thank my great supervisors, Steinar Sælid, Øyvind Stavdahl, Anders Lyngvi Fougner and Sven Magnus Carlsen for all the academic, medical and technological support, the interesting conversations and discussions during these 4 years, and their help with the writing of papers and thesis. I thank my co-authors for the many helpful discussions and for all their work and cooperation during the writing of the papers included in this thesis.

I thank my company, Prediktor Medical, for the opportunity to work in this field, their long-term thinking when creating this Industrial PhD position, and their patience in giving me the time needed to finish it. I also thank the employees at Prediktor Medical that I have worked with. I hope to soon be of more use than I have been lately and look forward to being a full-time member of the team.

I thank Norges Forskningsråd for the grant to do this PhD project.

I thank Karl Christian Faye for the interesting conversations and sharing of experiences with using the Freestyle Libre.

I thank my ex-band for being good friends and making great music. In particular, I am grateful for the newly released EP, which helped keep my motivation up during the dark months of November and December when I was writing my thesis together. Best motivational song ever: https://spoti.fi/2S3kVCh

Most of all I thank my family.

Takk for at dere har holdt ut med mine sene kvelder med jobbing og skriving, og en tidvis mentalt fraværende mann/far/sønn/bror.

Malin: Takk for at du ga meg rom og tid til å gjennomføre dette arbeidet. Det hadde vært helt umulig uten deg. Takk for alt du har gjort og gjør for å holde familien og husholdningen sammen.

Mamma og pappa: Takk for at dere alltid stiller opp, og for all hjelp med ungepass og alt anna bra dere alltid tar med dere når dere kommer på besøk.

Magnus: Ja, nå kan vi bygge legoroboter.

Maria: Pappa virker igjen!

# Contents

# List of Abbreviations

| | |
|---|---|
| AP | Artificial Pancreas |
| BC | Bolus calculator |
| BG | Blood Glucose |
| BGL | Blood Glucose Level |
| CEG | Consensus or Clarke Error Grid |
| CF | Correction Factor (in the context of insulin bolus calculators) |
| CGM | Continuous Glucose Monitor |
| CHO | Carbohydrate |
| CSII | Continuous Subcutaneous Insulin Infusion |
| DM | Diabetes Mellitus |
| DM1 | Type 1 Diabetes Mellitus |
| DM2 | Type 2 Diabetes Mellitus |
| DMMS.R | Diabetes Mellitus Metabolic Simulator for Research |
| EGP | Endogeneous Glucose Production |
| EKF | Extended Kalman Filter |
| FFA | Free Fatty Acids |
| FGM | Flash Glucose Monitor |
| FL | Freestyle Libre |

| | |
|---|---|
| GGE | Glycemic Glucose Equivalents |
| GI | Glycemic index |
| GL | Glycemic Load |
| GV | Glycemic Variability / Glucose Variability |
| HCL | Hybrid Closed Loop |
| HWITL | Hardware in the Loop |
| IFG | Impaired Fasting Glycaemia |
| IG | Interstitial Glucose |
| IGT | Impaired Glucose Tolerance |
| IOB | Insulin on Board |
| KF | Kalman Filter |
| MARD | Mean Absolute Relative Difference |
| MC | Monte Carlo |
| MDI | Multiple Daily Injections |
| MPC | Model Predictive Control |
| NEFA | Non-Esterified Fatty Acids |
| NFC | Near Field Communication |
| NICS | Non-Identifiable Correlated Set |
| NIR | Near Infrared |
| OGTT | Oral Glucose Toleance Test |
| PEG | Parkes Error Grid |
| PF | Particle Filter |
| PI | Practical identifiability |
| PISA | Pressure Induced Sensor Attenuation |
| PK/PD | Pharmacokinetic/Pharmacodynamic |

| | |
|---|---|
| PPGR | Post-Prandial Glucose Response; The glucose response following a meal |
| PPGR | Post-prandial Glucose Response |
| RMSE | Root Mean Square Error |
| SA | Sensitivity Analysis |
| SI | Structural identifiability |
| SMBG | Self Monitoring of Blood Glucose |
| SVD | Singular Value Decomposition |
| T1DMS | Type 1 Diabetes Mellitus Simulator |
| UI | User interface |
| UKF | Unscented Kalman Filter |
| UVa-Padova | University of Virginia and Padova |
| VP | Virtual Patient |
| YSI | Yellow Springs Instrument (see Sec. 3.1.6) |

# Chapter 1

# Introduction

This thesis is about blood glucose dynamics models with the intended use of real time estimation and filtering. Individualization of such models, based on glucose and insulin data gathered in free-living settings is the main topic. Related topics are glucose data cleaning, in the form of smoothing, interpolation, outlier detection and uncertainty estimation. Unreported meal detection is also investigated for this purpose. Another topic is the characterization of glucose data from different commercially available glucose measurement devices and how these relate to each other. Also investigated are methods for determining the practical identifiability of model parameters, with the intention of finding a proper set of parameters to estimate in a model individualization setting.

## 1.1 Context: The BioMKR device

Prediktor Medical AS is developing the BioMKR device, aiming to make it the world's first usable non-invasive glucose monitor. An image of a prototype of the device is shown in Fig. 1.1.

The device is intended to be worn on the upper arm. It measures the near-infrared (NIR) diffuse reflectance of the skin in the wavelength region 900-1800$\mathrm{nm}$, along with bio-impedance measurements in the frequency range 1-25 $\mathrm{MHz}$. The NIR measurements are performed every second while the device is in contact with the skin, whereas the bio-impedance measurements are performed every 30 $\mathrm{s}$. The device also measures temperature and acceleration.

Non-invasive glucose measurement has been researched by many, focusing either on NIR [1, 2, 3, 4] or on bio-impedance [5, 6, 7]. There are several review papers

on the subject [8, 9, 10], most of which conclude that the problem is not solved, and that the main problems to be solved are related to the low specificity of the measurements and an unstable measurement situation on skin, including motion artifacts and a high number of confounding factors like sweating, temperature changes, fluctuations in blood flow and alterations of pressure.

Prediktor Medical's idea with BioMKR is to combine two promising non-invasive and miniaturizable glucose measurement principles from the literature, NIR and bio-impedance, with a model of glucose dynamics. This model can provide sensor fusion and noise reduction of the two noisy non-invasive measurements. This provides a path to overcoming some of the motion artefacts seen in the signals, and could enable sufficiently accurate blood glucose tracking to yield a useful monitoring device for people with diabetes.

Blood glucose (BG) dynamics models have applications in such a setting. Many of the models formulated have been used and identified based on clinical research setting data, where more measurements of internal states are available, e.g. insulin, glucagon, and where the inputs introduced (meals, insulin injections) are accurately known. To be of use in a product, the models must be usable based on data obtained in a free-living setting (as opposed to controlled hospital study settings). This implies use of only glucose data and information about meals and insulin injections.

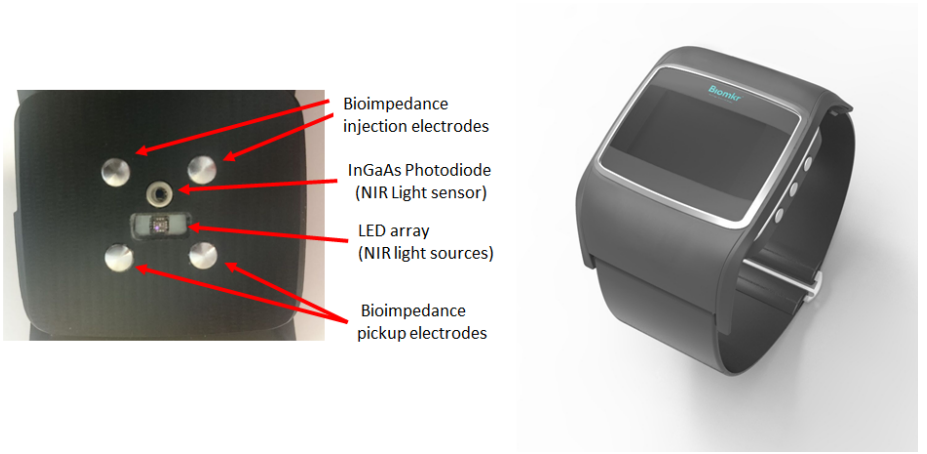



**Figure 1.1:** *Left*: Skin-facing side of the BioMKR device, showing bioimpedance and NIR sensors. *Right*: 3D model of device with band.

Some questions in this respect are:

1. Is it possible to predict glucose using only data from free-living settings?
2. Which inputs (meals, insulin, exercise etc) and what quality of data are needed to predict blood glucose?
3. What are the most relevant BG dynamic model models to use for this purpose?
4. How should the individualisation of such models be performed?

The PhD project has worked towards answering these questions, with the intention of using individualised BG dynamics models in the BioMKR product.

## 1.2 Publications

The following is a list of publications produced in this PhD project. The letter and number in boldcase corresponds to the chapter of the appendix that includes the full text of the article.

### 1.2.1 Journal articles

Paper **A.1** O. M. Staal, S. Sælid, A. Fougner, and Ø. Stavdahl, "Kalman smoothing for objective and automatic preprocessing of glucose data," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 218–226, Jan 2019

Paper **A.2** O. M. Staal, H. M. U. Hansen, S. C. Christiansen, A. L. Fougner, S. M. Carlsen, and Ø. Stavdahl, "Differences between flash glucose monitor and fingerprick measurements," *Biosensors*, vol. 8, no. 4, 2018

### 1.2.2 Conference articles

Paper **B.1** O. M. Staal, A. L. Fougner, S. Sælid, and Ø. Stavdahl, "Glucose-insulin metabolism model reduction and parameter selection using sensitivity analysis," *ACC*, Submitted to American Control Conference (ACC) 2019

### 1.2.3 Draft manuscripts

Paper **B.2** O. M. Staal, S. Sælid, A. L. Fougner and Ø. Stavdahl "Meal input estimation from Continuous Glucose Monitor data using Kalman filtering and hypothesis testing". *Draft manuscript*

### 1.2.4 Other works

The following works have also been made during the PhD project, but are not formally part of the thesis, and are not included in the appendix.

**Journal article**: (Accepted) C. Tronstad, O. Elvebakk, O. M. Staal, H. Kalvøy, J. O. Høgetveit, T. G. Jenssen, K. I. Birkeland, and Ø. G. Martinsen, "Non-invasive prediction of blood glucose trends during hypoglycemia," *Accepted for publication in Analytica Chimica Acta*, 2018

**Conference article**: (Published) C. Tronstad, O. M. Staal, S. Sælid, and Ø. G. Martinsen, "Model-based filtering for artifact and noise suppression with state estimation for electrodermal activity measurements in real time," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 2015, pp. 2750–2753

**Poster presentation**: (Presented) O. M. Staal, S. Sælid, T. V. Karstang, and Ø. Stavdahl, "Kalman smoothing of glucose data applied to Partial Least Squares modeling of non-invasive near-infrared measurements," in *International Conference on Advanced Technologies & Treatments for Diabetes (ATTD)*, 2017, Poster. Available: http://bit.ly/2kw8u5H

**Poster presentation**: (Presented) S. M. Carlsen, A. L. Fougner, T. V. Karstang, O. M. Staal, and S. C. Christiansen, "Continuous non-invasive glucose monitoring by sensor fusion of near infrared light and bioimpedance measurements: Results of a proof of concept study," in *International Conference on Advanced Technologies & Treatments for Diabetes (ATTD)*, 2016, Poster. Available: https://bit.ly/2OP2eVq

**Oral presentation**: (Presented) O. M. Staal, S. C. Christiansen, H. M. U. Hansen, A. L. Fougner, S. M. Carlsen, and Ø. Stavdahl, "Characterization of the error between SMBG and a flash glucose monitor in the first day of use," in *International Conference on Advanced Technologies & Treatments for Diabetes (ATTD)*, 2018, Oral presentation. Available: https://doi.org/10.13140/RG.2.2.34374.37447

## 1.3 Contributions

The contributions of this PhD thesis are as follows:

- A method to smooth, interpolate, detect outliers, and estimate uncertainty in glucose data, Paper A.1.
- A characterisation of errors in a Flash Glucose Monitor, Freestyle Libre from Abbott, Paper A.2.
- Use of sensitivity analysis and singular value decomposition to evaluate parameter identifiability in glucose-insulin metabolism models. The method was described and partly applied to the UVa-Padova model in Paper B.1.
- Use of Kalman filtering and a hypothesis testing algorithm to detect meals in CGM data, Paper B.2.
- A review of physiology, technology, models and identifiability in the context of blood glucose dynamics is provided in the main text of this thesis, chapters 2-5.
- The contribution to my company, Prediktor Medical, is the use of the theory and methods presented here in the development of the BioMKR device.

The contributions are further discussed in Chapter 8.

## 1.4 Structure of the thesis

The thesis is structured as follows:

Chapter 2 reviews background information that is of a medical and physiological nature, including the physiology of glucose metabolism, diabetes, and diabetes treatment.

Chapter 3 provides information about medical devices currently being used to monitor, manage and treat diabetes. It also reviews the concept of virtual patients, and how these can be used to improve processes and systems related to diabetes management.

Chapter 4 reviews state space modelling and models of blood glucose dynamics. Some models are described in full, e.g. a model in use in Prediktor Medical.

Chapter 5 looks into identification of such models and describes problems related to using data from free-living settings for identification.

Chapter 6 describes methods for real-time state and parameter estimation in BG dynamic models.

Chapter 7 provides some practical experiences in applying the methods described to data from free-living settings gathered in different experiments during the PhD project.

Chapter 8 discuss the contributions of this thesis with suggestions for future work.

Chapter 9 concludes the thesis.

The main papers of the thesis are reproduced in the appendices A.1, A.2, B.1 and B.2.

# Chapter 2

# Physiological and medical background

This chapter is a short summary of relevant physiology in glucose dynamics and diabetes. A common reference for the entire chapter is Hall [19]. Some of the content in this section is from the self-study performed as part of the PhD coursework.

## 2.1 Physiology of glucose metabolism

This section describes the mechanisms involved in keeping glucose in a narrow, normal range in healthy individuals. This control process is called glucose homeostasis. It is summarised in Fig. 2.1.

### 2.1.1 Overview

Humans get the energy they need from ingesting food. Food generally contains carbohydrates, proteins and lipids (fats).

Digestible carbohydrates in the diet (starches and sugars) are broken down by the digestive system into di- and monosaccharides before being absorbed from the gastrointestinal tract into the blood stream and transported to cells for usage or storage. Most of the saccharides in the human diet are converted into glucose before entering into the metabolic breakdown reactions. Therefore, glucose is an ubiquitous monosaccharide that is a vital energy source in human beings, and most other organisms from mammals to bacteria. It is used throughout the human body as fuel. Brain cells depend on a constant supply of glucose, and a too low glucose level will lead to loss of consciousness.

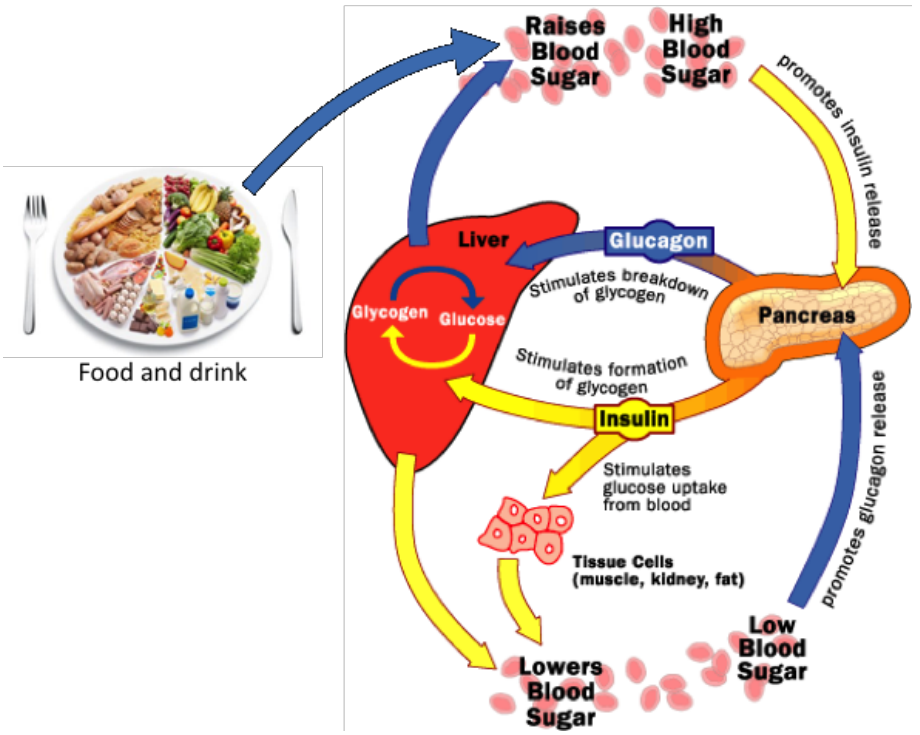



**Figure 2.1:** An overview of glucose homeostasis in healthy individuals.
Source: Adapted from HowStuffWorks.com

The other main source of fuel for the body is non-esterified fatty acids (NEFA), also called free fatty acids (FFA). FFAs are derived from dietary lipids, and is the main source of energy in skeletal muscle during rest or light exercise [20]. FFA is released into the blood stream by adipose tissues when needed, or stored as triglycerides when abundant, mainly in the adipose tissues.

Proteins are not considered a major source of energy in the human body in the fed state. Proteins from food are split into amino acids that are mainly used for building other proteins needed by the body. During starvation proteins come into play as an emergency energy source, as they can be converted for synthesising glucose.

### Relevant anatomical structures

The main anatomical structures involved in metabolism of nutrients are shown in Figure 2.2 and described in the following.

*The gastrointestinal (GI) system, or GI tract*, is used for the ingestion of food,

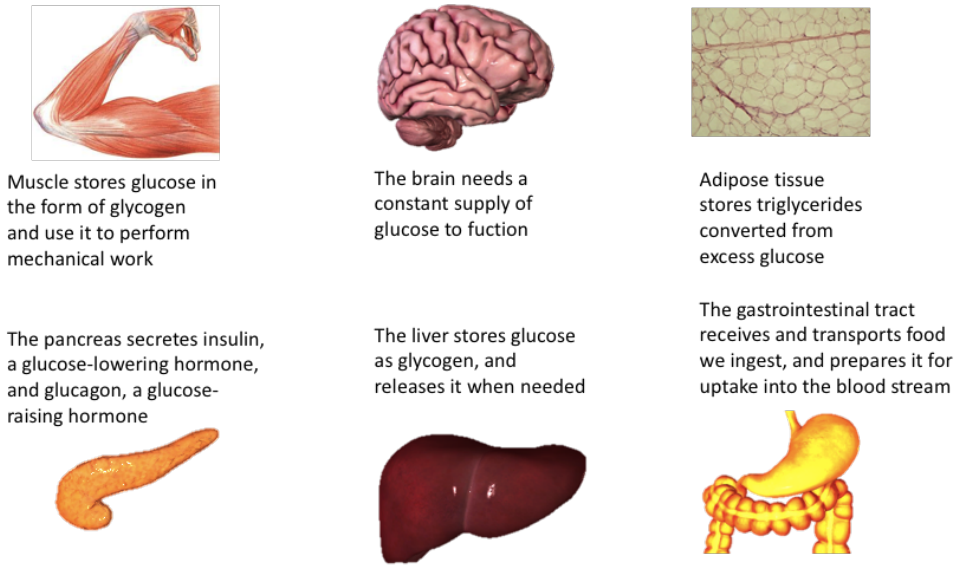



**Figure 2.2:** The main anatomical structures relevant to glucose metabolism

digestion of food into nutrients, transport of nutrients to blood and excretion of non-nutritional food content (faeces).

*The liver*, which filters blood drained from the GI system, stores glucose as glycogen (glucogenesis), converts stored glycogen to glucose (glycogenolysis), performs conversions of other nutrients to glucose (gluconeogenesis), and more.

*The pancreas* is a gland with two main functions. The *exocrine pancreas* produces pancreatic juice, digestive enzymes that are emptied into the duodenum and used in the GI tract to digest food. The endocrine pancreas produces hormones, mainly insulin, glucagon and amylin, which are secreted into the blood at the portal vein. These hormones control production, storage and consumption of glucose in the liver and in tissues throughout the body. In Type 1 diabetes, the beta cells that secrete insulin are destroyed, and the patient is dependent on externally supplied insulin to survive. In Type 2 diabetes the pancreas still secretes insulin, but the insulin sensitivity is diminished. More on this in Section 2.2.

Glucose is consumed and stored in the cells of the different tissues. The main tissues of interest are described in the following.

*Muscle cells* convert nutritional energy into mechanical work and heat. Muscle cells have substantial glycogen stores for internal glucose buffering, amounting

to roughly 400 grams of glycogen in an adult. During rest, the muscle membrane is only slightly permeable to glucose, and the muscle cells use 60% FFA and 40% glucose for fuel. During moderate or heavy exercise the muscles absorb and use larger amounts of glucose from the blood, as well as from internal glycogen stores. Muscle absorbs much glucose after meals if glucose and insulin is present. Depending on the activity level, this absorbed glucose is either used directly or used to replenish the glycogen stores of the muscle cells.

*Adipose tissue cells* store energy in the form of fatty acids as triglycerides. Adipose tissue release fatty acids when other energy sources are not available. Adipose tissue cells use glucose to synthesise the glycerol portion of the fat molecules stored.

*Brain cells* consume glucose. The brain metabolism accounts for 15-20% of the total body glucose consumption although the brain is only about 2% of the body by weight. Brain cells have a small glycogen store that only last for up to two minutes, and can use other energy substrates than glucose only with difficulty, thus the brain cells are highly dependent on a constant supply of glucose from the blood. Transport of glucose into brain cells is not insulin dependent.

*Red blood cells* metabolise glucose anaerobically, producing lactate that is converted back to glucose by the liver or oxidised by muscle cells. Red blood cells typically account for about 2% of the glucose metabolism.

*Liver cells* (hepatocytes) play one of the main roles in the control of blood glucose, as they are capable of buffering glucose in their glycogen stores, and can convert back and forth between glucose and other molecules. The liver's placement in the circulatory system is ideal for this purpose. The nutrient-rich blood drained from the gut ends up in the portal vein, which goes into the liver [19]. This means that the liver is first in line to receive any glucose ingested. Also, since the pancreas empties its secreted hormones insulin and glucagon into the portal vein, the liver is able to respond quickly to changes in all these signals. An illustration of the liver's placement relative to the GI tract and the pancreas is shown in Figure 2.3. The placement also naturally makes the concentration of pancreatic hormones and postprandial glucose higher in the liver than in the peripheral tissues [21]. When insulin is injected subcutaneously, e.g. in people with diabetes, the concentration of insulin seen by the liver does not reach the same levels as that which occurs naturally in healthy individuals where insulin secretion is functional.

The liver is capable of storing about 100 grams of glycogen. Normally roughly 60% of the glucose ingested in a meal is stored in the liver as glycogen and released when blood glucose is low again after the meal. The liver cells will start to convert glucose into fatty acids when their glycogen stores have been filled and glucose is

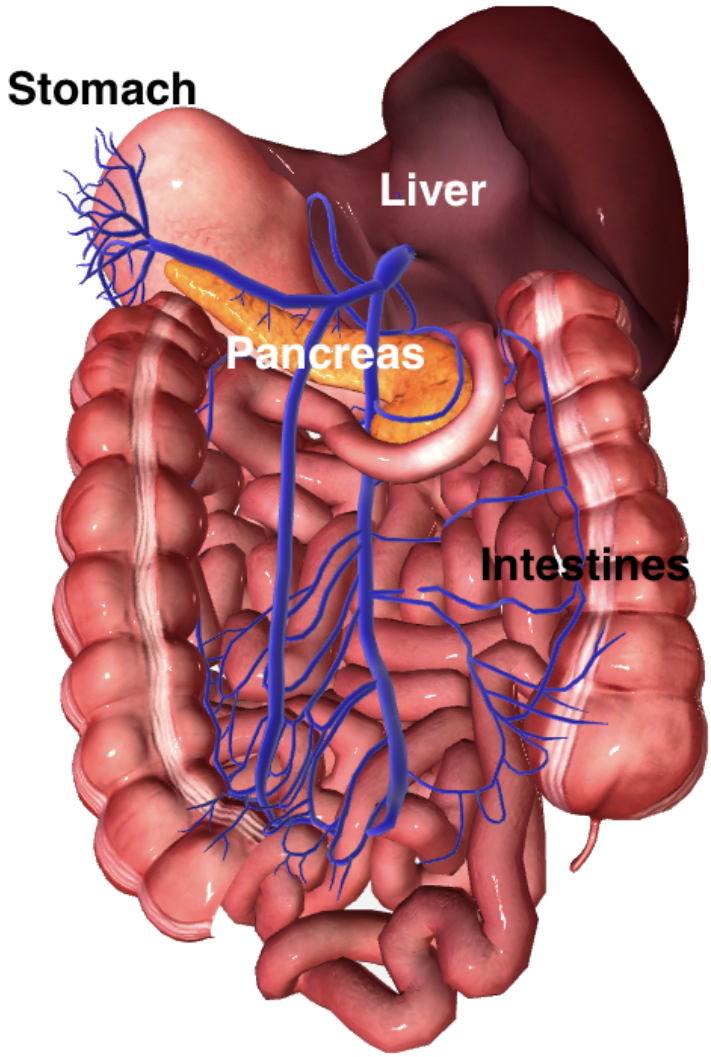



**Figure 2.3:** Posterior view of the pancreas (orange), gastrointestinal tract (pink), liver (dark red/brown) and hepatic portal venous system (blue). The large blue vein entering the liver is the portal vein. This vein receives blood that has been drained from the gastrointestinal tract, which is rich in glucose after a meal. Glucagon and insulin produced by the pancreas is released into the portal vein. Image is generated from human.biodigital.com.

still in excess.

### 2.1.2 Carbohydrate metabolism

Digestable carbohydrates are broken down into monosaccharides, mainly glucose, fructose and galactose. This digestion happens mainly in the small intestine, and the monosaccharides are transported through the gut wall and via the blood stream to the cell where it is used or stored. Other monosaccharides than glucose are converted into glucose intermediates as one of the first steps in the conversion processes, so glucose can be considered a sort of universal monosaccharide. Through a sequence of biochemical processes, glucose is converted via pyruvate to carbon dioxide and water, and in the process generates adenosine triphosphate (ATP) as a by-product. ATP is a molecule that is regarded as the universal energy currency in the body. ATP delivers energy to countless physiological processes in the body by losing one of its phosphate bonds and becoming adenosine diphosphate (ADP). One mol of glucose generates about 36 moles of ATP.

Depending on the availability of oxygen in the cell, the glucose is metabolised aerobically or anaerobically. *Aerobe* metabolism of glucose happens by oxidative phosphorylation of pyruvate, creating $CO_2$ and water as described above. $CO_2$ is eliminated from the body through respiration. In *anaerobe* metabolism of glucose, oxygen is not available, and glucose is converted via pyruvate to lactate, which is transported to the liver where it is turned back to pyruvate and glucose.

#### Transport, delivery and phosphorylation

It is common to talk of the process of glucose delivery from blood to cell in three steps [22]:

1. Delivery of glucose from blood to the interstitial fluid through the capillary wall. This is considered to be a diffusion process, i.e. driven mostly by concentration gradients.
2. Transport: The movement of glucose into the cell through the cell membrane. This is performed by a combination of active and passive transport mechanisms, see below.
3. Phosphorylation: When the glucose molecule gets inside the cell, it is phosphorylated. This is an irreversible process which has the effect of locking the molecule inside the cell.

*Glucose transport* Getting glucose through cell membranes is performed by glucose transporter membrane proteins. Either actively or passively. In active transport,

energy from ATP is spent to pull the molecule through the membrane. Active transport of glucose mainly happens in the kidneys by proteins called Sodium-Glucose Linked Transporter (SGLT). SGLT proteins co-transport sodium and glucose.

In passive transport, the molecules are transported through the cell membrane by facilitated diffusion. A membrane protein called a carrier protein provides a channel through which the molecule can move, or it binds to the molecule on the outside of the cell, then changes its shape so that the molecule can move into the cell. There are 12 different glucose transporter (GLUT) membrane proteins encoded in the human genome. The first four are most well-known:

*GLUT1* is found in high levels in the erythrocytes and in the blood-brain barrier, but is also responsible for a low level of basal glucose uptake in all cells.

*GLUT2* is found in renal tubular cells, liver cells and pancreatic beta cells. Allows bidirectional transport of glucose.

*GLUT3* is found mostly in neurons, and provides a basal uptake of glucose.

*GLUT4* is mainly found in adipose tissue and striated muscle cells, and is regulated by insulin. When insulin is not present, most GLUT4 proteins are stored in vesicles inside the cells and little or no glucose is transported through the cell membrane by GLUT4. The cell membrane has transmembrane insulin receptors. When the insulin receptors are triggered by an insulin molecule on the outside of the membrane, they trigger a signalling pathway inside the cell that causes the GLUT4 storage vesicles to release GLUT4 proteins, which binds to the cell membrane and starts transporting glucose into the cell.

The transporter proteins that are not insulin sensitive have differing affinity for glucose [23]. In GLUT1 and GLUT3 the $K_M$ constant of the Michaelis-Menten kinetics is close to 1 mmol/L, which results in a constant uptake of glucose from the blood stream by these transporters. GLUT2 however, has a higher $K_M$ of 15-20 mmol/L, meaning that the uptake is proportional to the glucose concentration for most of the achievable glucose range. This is useful for glucose-sensing cells, like pancreatic beta cells. This difference in $K_M$ is illustrated in Fig. 2.4. Michaelis-Menten kinetics is a nonlinear relationship found in many enzymatic reactions, and it can also describe glucose transport:

$$v_G = \frac{K_{max} G}{K_M + G} \tag{2.1}$$

here $v_G$ is the reaction/transport rate, $K_{max}$ is the maximum transport rate, G is concentration of glucose, and $K_M$ is the rate constant.

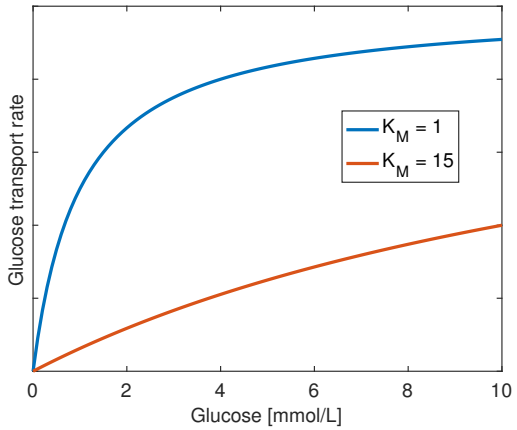



**Figure 2.4:** Michaelis-Menten kinetics for $K_M = 1$ and $K_M$=15. GLUT1 and GLUT3 has $K_M$ close to 1 (blue curve), and provides a near constant glucose transport when glucose is above 2 mmol/L. GLUT2, on the other hand, has a $K_M$ closer to 15 (red curve), and provides a glucose transport that is near linear with glucose concentration.

### 2.1.3 Hormones affecting metabolism

There is one main glucose-decreasing hormone, insulin, and 4 major glucose-increasing hormones, glucagon, epinephrine, cortisol and growth hormone.

#### Insulin

Insulin is a hormone that affects both carbohydrate, lipid and protein metabolism. Insulin is generally secreted by the beta cells in the islets of Langerhans in the endocrine pancreas as a response to increased glucose levels in the blood. Insulin triggers several mechanisms with blood glucose lowering effects:

- It enables glucose to enter cells in 80% of the body's tissues (e.g. not brain cells)
- It causes excess glucose to be stored as glycogen in muscles and the liver
- It causes excess glucose to be converted into fats and stored in adipose tissues

Insulin has a half-life in plasma of 5-6 minutes, meaning that most insulin that does not combine with receptors on the target cells will be cleared from the blood in 10-15 minutes.

Insulin combines with insulin receptors on the cell membranes of the target tissues

(mainly hepatic, muscle and adipose tissue cells), causing the membrane to become more permeable to glucose within a few seconds.

Insulin promotes fat synthesis and storage, and inhibits lipolysis in the adipose tissues. When insulin is not present, e.g. in Type 1 diabetes patients, the concentration of free fatty acids, triglycerides and cholesterol in the blood rises. The lack of insulin cause cells to use these fats as fuel in a way that leads to ketoacidosis, which is life threatening. Insulin also promotes protein synthesis and storage, and inhibits protein degradation. Therefore a lack of insulin inhibits synthesis of proteins and leads to degradation of proteins. This leads to protein wasting, weakness and organ damage.

Insulin is released by the beta cells in a two-phase fashion. When a blood glucose elevation is first sensed by the beta cells, they release an initial dose of insulin from a store of pre-formed insulin, raising the plasma insulin levels within 3-5 minutes after the blood glucose rise. During the next 5-10 minutes there is little secretion, leading the plasma insulin level to fall again. Then after about 15 minutes, the secondary phase begins, and newly synthesised insulin and more pre-formed insulin is secreted. Insulin secretion shutdown is fast, occurring within 3-5 minutes of the return to fasting glucose levels. Insulin secretion is pulsatile, with an intrapulse period of approximately 5 minutes [21].

Insulin secretion is not only a function of blood glucose concentration. Amino acids, for instance, potentiate the glucose stimulus for insulin secretion, meaning that when high levels of amino acids and glucose is simultaneously sensed in the blood by the beta cells, the insulin secretion can be nearly doubled compared to when only the same level of glucose is sensed [24]. The same is true for gastrointestinal hormones secreted by the GI tract when a person eats a meal. These hormones can also stimulate insulin secretion before any rise in blood glucose is seen. This anticipatory secretion of insulin may also be partly mediated by the nervous system, both as a response to stretch receptors of the stomach, or from the sympathetic nervous system, as in some individuals, merely thinking of food can lead to an insulin release [25].

When insulin is synthesised by the body, it is created as a long peptide chain called proinsulin. The middle part of the chain is called the C-peptide, and this part is cleaved off. The two end parts are connected to each other with disulphide bonds, forming insulin. So when insulin is secreted, an eqimolar amount of C-peptide is also secreted. Since C-peptide is not taken up in the liver like insulin is, and it has a half-life of 30-33 minutes, a measurement of the C-peptide levels in plasma reflects how much insulin has been secreted lately. This is a useful measurement in the diagnosis and treatment of diabetes, and can be used to distinguish between types

of diabetes (see Section 2.2), and to estimate viable beta cell mass. It can also be used to determine if a hypoglycemia was due to an overdose of externally supplied insulin or if it is due to a disease, since injected insulin contains no C-peptide.

### Glucagon

Glucagon is secreted by the alpha cells in the pancreas, and normally has a potent blood glucose raising effect. A microgram of glucagon can raise the blood glucose level by 1 mmol/L within 20 minutes. It works primarily by triggering glycogenolysis in the liver, freeing glucose. It also increases gluconeogenesis in the liver. A consequence of this is that if liver glycogen stores are depleted, glucagon has less or no effect on blood glucose levels.

Glucagon is secreted when blood glucose falls to hypoglycemic levels, and is reduced to normal level when normal blood glucose levels are restored.

Glucagon release is also inhibited by insulin release. The proximity of the alpha and beta cells in the pancreas make the alpha cells exposed to a high concentration of insulin when the beta cells are secreting, which is sensed by the alpha cells. Consequently, when insulin production is destroyed or impaired in diabetes, glucagon secretion is less inhibited. Subcutaneous injections of insulin does not restore this glucagon suppression, since the insulin concentrations seen by the alpha cells in this case are not nearly as high as they would be if neighbouring beta cells were secreting insulin.

In DM2 it has also been observed that in the absence of insulin, when blood glucose rises to hyperglycemic levels, a counter-intuitive increase of glucagon secretion occurs. This leads to glucagon hypersecretion and thus still further rising blood glucose levels. It has been hypothesised that in diabetic patients, hyperglycemia results not mainly from a lack of insulin, but from the glucagon excess resulting from the insulin lack [26]. Glucagon suppressant treatment in addition to insulin treatment in diabetes has been proposed [27], e.g by using leptin [28].

In DM1, however, the ability to synthesise glucagon become diminished, and according to [29], almost all DM1 individuals fail to produce an adequate glucagon response to hypoglycemia within five years of disease diagnosis. This is unfortunate since it means that one of the mechanisms protecting against hypoglycemia is lost.

Exercise has an increasing effect on glucagon levels in healthy individuals, to 4-5 times the normal level. This effect is seen also when blood glucose levels do not fall during exercise, so the mechanisms behind this exercise triggered glucagon secretion are not completely understood.

Glucagon for injection is available, and is used by some people with diabetes for hypoglycemia emergency treatment. It is also included in so-called dual-hormone AP concepts in addition to insulin.

### Amylin

Amylin is co-secreted with insulin by the pancreatic beta cell, and gives a sense of satiety. It also slows gastric emptying. Individuals with diabetes type 1 lose the ability to secrete amylin along with insulin.

### Other glucose-increasing hormones

*Epinephrine*, or adrenaline, is released by the adrenal gland and some neurons. It is involved in the fight-or-flight response, meaning that fear or nervousness will give higher levels of epinephrine. Exercise also leads to epinephrine release. Epinephrine raises the glucose level by stimulating glycogenolysis and gluconeogenesis [30]. Noradrenaline/norepinephrine is a slightly different hormone that has quite similar effects as epinephrine.

*Cortisol* is a glucocorticoid released by the adrenal gland in response to stress and low blood glucose. It raises the glucose level by increasing gluconeogenesis and inhibiting peripheral glucose uptake. Cortisol has a clear circadian rhythm, with highest secretion before waking up in the morning, and lowest around midnight [31].

*Growth hormone* is released by the pituitary gland, and increases protein synthesis and fatty acid mobilisation and decreases glucose utilisation.

### Thyroid hormones

The thyroid gland secretes two hormones that both increase the metabolic rate, thyroxine and triiodothyronine. Thyroxine is a prehormone that is converted to triiodothyronine within the cells, and it is this molecule that exerts the hormone effect. Lack of thyroid hormones gives a basal metabolic rate at about 50% below normal, while extreme excess of thyroid hormone secretion can lead to a basal metabolic rate of 100% above normal.

The dynamics of these hormones is rather slow, and the circulating levels are normally kept within strict limits. The hormones are bound to and stored by plasma and intercellular proteins before being used slowly over days and weeks. An injection of thyroxine takes effect after 2-3 days, reaching a maximum in 10-12 days. The effect lasts for 6-8 weeks. Triiodothyronine is substantially more rapid,

taking effect after 1 day and reaching a maximum after 2-3 days.

The thyroid hormones increase transcription of a large number of genes, thus synthesising more of the enzymes and transport proteins involved in nutrient transport and metabolism, and in general increase the metabolic activity throughout the body. Most individuals have normal thyroid hormone levels that can fluctuate somewhat in order to adapt to external changes, e.g. to colder climate. When thyroid hormone levels are higher than normal it is called hyperthyroidism. Adverse effects of prolonged hyperthyroidism include muscle (including heart) weakness induced by increased protein catabolism, muscle tremor, fatigue, sleeplessness, reduced bone mineralisation (i.e. osteoporosis) and mental issues like nervousness, anxiety and paranoia. Oppositely, hypothyroidism results in tiredness and weight gain.

### 2.1.4 Diurnal/circadian rhythms

Many processes in the body have a diurnal/circadian rhythm. This means that there are patterns that repeat in a roughly 24-hour cycle. Cortisol is a hormone that clearly behaves cyclically with the time of day, as described in Section 2.1.3. The experiments described in Hinshaw et al. [32] and Saad et al. [33] investigate the diurnal variability of insulin sensitivity. This has also been included in some simulations, e.g. Visentin et al. [34] claim that people can be classified according to 7 different classes of diurnal variation of their insulin sensitivity. The most frequent pattern was low sensitivity at breakfast and lunch, and high sensitivity at dinner, which was the case in 7 of the 20 healthy participants of the study, although the low number of participants and meals per participant make these classifications dubious.

Absorption of glucose from the gut also shows diurnal variations. It was found in [35] that the peak of the meal absorption of glucose occurs after $47 \pm 5$ minutes for the breakfast meal, and after $78 \pm 17$ minutes for the evening meal, with in-between meals nearly linearly interpolated. This has also been included in some simulations, see [36].

### 2.1.5 Other influences on blood glucose

There are many other factors that affect blood glucose, some of which are shown in Fig. 2.5.

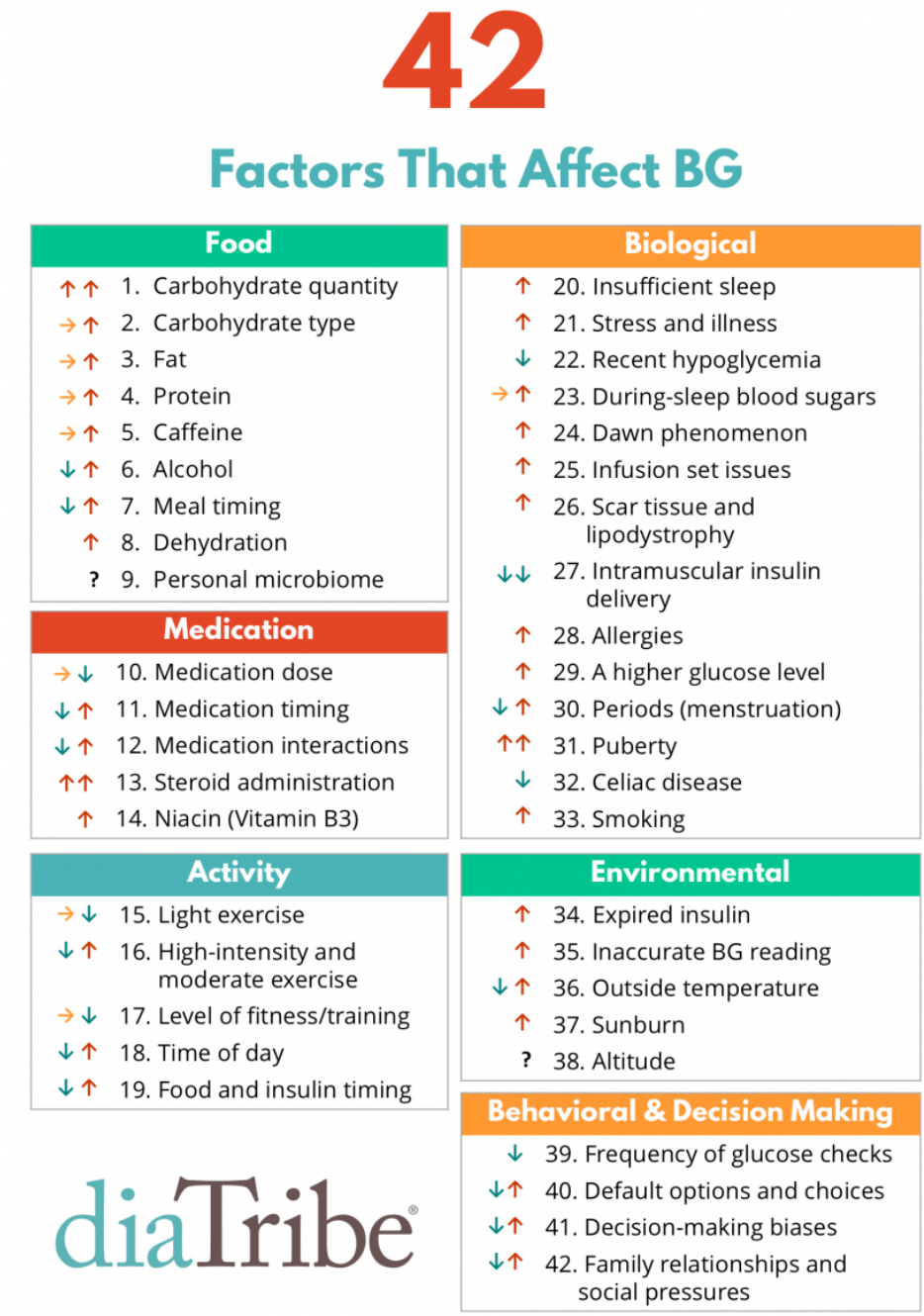



**Figure 2.5:** Factors influencing blood glucose. Image from https://bit.ly/2PVJnYZ

## 2.2 Diabetes

Diabetes mellitus is a condition of impaired carbohydrate, lipid and protein metabolism due to lack of insulin or decreased insulin sensitivity. The symptoms are high blood glucose levels, glucose in the urine, thirst and polyuria (excessive production of urine).

The World Health Organization (WHO) estimated that in 2014, 9% of the adult population suffered from diabetes. Worldwide, 1.5 million deaths in 2012 were directly related to diabetes, where 80% of these are in low- and middle-income countries [37].

### 2.2.1 Diabetes categories

Diabetes is usually categorised as follows, although the borders between the categories are not always clear cut, and some individuals may have a mixture of conditions.

#### Diabetes Mellitus Type 1 - DM1

In diabetes mellitus type 1 the body becomes unable to produce insulin, through autoimmune destruction of the beta cells in the pancreas. The reason for the autoimmune destruction is often unknown, but virus infections may be involved in a significant number of cases. The beta cell destruction is as of yet not preventable. The DM1 onset may be at any age, but with one summit in children and another in adolescents.

#### Diabetes Mellitus Type 2 - DM2

About 90% of diabetes cases are categorised as diabetes mellitus type 2, where the body becomes unable to utilize insulin effectively, i.e. the tissues of the body do not react normally to insulin. This is called increased insulin resistance, or decreased insulin sensitivity, and means that a normal amount of insulin does not reduce glucose as much as usual, or that an increased amount of insulin is needed to keep glucose at a normal level. The plasma insulin levels in DM2 patients are often higher than normal, to compensate for the low insulin sensitivity. However, many DM2 patients gradually develop impaired insulin secretion after some years due to beta cell exhaustion. DM2 and reduced insulin sensitivity is linked to obesity and lack of physical activity, but not all DM2 patients are overweight or obese, and not all overweight or obese individuals develop DM2 [38]. There are also other causes of increased insulin resistance.

#### Other types or subtypes of diabetes

*Gestational diabetes* This type of diabetes only occurs during pregnancy, and normally vanishes after delivery. Gestational diabetes associate with increased risk of pregnancy and birth complications, and increased maternal risk of developing DM2 in the future.

*MODY - Maturity Onset Diabetes of the Young* are a number of rare monogenetic forms of diabetes that usually are inherited.

*Secondary diabetes* In cases where the pancreas is destroyed, removed or damaged, e.g. due to pancreatitis, cancer or trauma, diabetes will ensue due to the loss of the beta cells and their insulin production.

#### Prediabetes

DM1 is of a binary character, you either have lost the ability to produce insulin, or you have not. DM2 is of a more continuous character, with more or less reduced insulin sensitivity, and/or more or less impaired insulin secretion. The term prediabetes describes individuals that are at risk of developing DM2, having elevated glycemic parameters compared to normal, but not yet exceeding thresholds set for diagnosing diabetes [39]. Interventions like weight loss, improved diet, exercise and quitting smoking can delay or prevent prediabetes from developing into DM2 [40]. Prediabetes and diabetes have traditionally been diagnosed based on an oral glucose tolerance test (OGTT). The HbA1c assay has also been accepted as a diagnostic tool for diabetes [41], where an HbA1c above 6.5% signifies diabetes. In a study by Acciaroli et al. [42] the states of healthy, prediabetes and type 2 diabetes is successfully predicted from CGM data. In another interesting paper [43] it is shown that signs of developing DM2 is visible in the glucose trajectories 3-6 years ahead of DM2 diagnosis, indicating that glucose curves as obtainable from CGMs may serve as a biomarker for DM2 in the prediabetic and possibly also in the healthy population.

### 2.2.2 Consequences and complications of diabetes

Treating diabetes and its complications is estimated to have cost 850 billion USD in 2017. An estimated 4 million people died from diabetes complications in 2017 [40].

The consequences of diabetes can be divided into acute and chronic. Among acute consequences there is diabetic ketoacedosis (DKA) and hypoglycemia. Among the chronic complications there is a host of micro- and macrovascular complications

arising from prolonged hyperglycemia.

#### Diabetic ketoacidosis

If insulin is absent, as in untreated DM1, glucose in the blood stream is not transported into the cells for use or storage. The high blood glucose causes the kidneys to excrete glucose, which leads to polyuria and dehydration. Since glucose is not available as fuel in the cells, they are forced to metabolise fatty acids instead. This produces ketones, which has a pH-lowering effect. The combination of dehydration and ketone production results in DKA, which can be fatal. According to Daneman [44], DKA is the main cause of early deaths in diabetes.

#### Hypoglycemia

When a diabetes patient starts administering insulin, the risk of hypoglycemia arises. An overdose of insulin will make cells transport all available glucose in the blood stream into the cells, and may bring the blood glucose below the hypoglycemia limit at 3.3 mmol/L. This has an effect on the brain, which needs a constant supply of glucose to function properly. Mild hypoglycemia gives cognitive symptoms like confusion or difficulties in speaking; intermediate hypoglycemia results in fainting and seizures, whereas coma and death may be the ultimate consequence of severe hypoglycemia.

#### Chronic complications

Higher than normal glucose concentration causes loss of function of proteins and glucose induced oxidative damage to tissues [45], which can cause long-term micro- and macrovascular damage resulting in neuropathy, retinopathy and nephropathy.

## 2.3 Treatment of diabetes

The target of diabetes treatment is to avoid DKA, hypoglycemia, and the long-term consequences of hyperglycemia, by controlling the blood glucose levels to be within the normal range. The normal range for glucose in healthy individuals lie between 4.0-5.5 mmol/L in the fasting state and may peak to 9 mmol/L in the postprandial state.

### 2.3.1 Treatment of type 1 diabetes

All DM1 patients have to be treated by injecting insulin. Modern insulin treatment of DM1 can be divided in two subcategories, those using multiple daily subcutaneous


**Figure 2.6:** An overview of diabetes treatment. Type 1 diabetes is treated with insulin therapy, while type 2 diabetes usually starts with lifestyle and diet intervention, before progressing to oral medication and for some, insulin therapy. Both groups of patients use glucose monitoring tools.

injections (MDI) and those using an insulin pump, called Continuous Subcutaneous Insulin Infusion (CSII). In both cases insulin is injected into the subcutaneous tissue, usually on the abdomen.

Patients that use MDI use insulin pens, which is an integrated device containing a syringe pre-filled with insulin, and a dosage mechanism, enabling the user to easily set the dose and inject insulin. Such users usually use two different types of insulin; rapid-acting insulin used in conjunction with meals, and slow-acting insulin, which is injected once or twice daily and that provides a basal insulin level. The basal insulin level prevents hyperglycemia in the fasting and semi-fasting state by inhibiting gluconeogenesis and glycogenolysis in the liver, and by providing a basal level of insulin, reduces the risk of developing DKA.

CSII users use an insulin pump, to achieve the same with only rapid-acting insulin. Insulin pumps are described in section 3.2.

Insulin cannot be given orally, since insulin is destroyed by the gastric enzymes, and for a long time the only option has been to inject the insulin. A recent development is inhaled insulin, which can provide an alternative to injected meal time insulin. Inhaled insulins provide more rapid absorption into the blood stream [46]. Inhaled insulins have not yet seen wide-spread use at the time of writing. Research into other routes of insulin delivery is ongoing, see Shah et al. [47] for a overview.

The development of new insulin analogues have led to a multitude of different insulins on the market, where the rapid-acting types strive to become ever more rapid-acting, and the slow-acting types strive for giving as stable basal insulin as possible with as few injections as possible. The current fast-acting insulins have onset times of 5-20 minutes, peak activity after 30-60 minutes, and a duration of action of 3-5 hours [48]. Current slow-acting insulins provide 24-hour duration of action and low peaking, meaning that slow insulin injections once daily can be sufficient to provide a basal insulin level. The relationship between insulin dose and insulin appearance in the blood stream is usually described using pharmacokinetic/pharmacodynamic (PK/PD) models [49]. A figure showing the PK/PD curves of some insulin variants is given in Fig. 2.7.

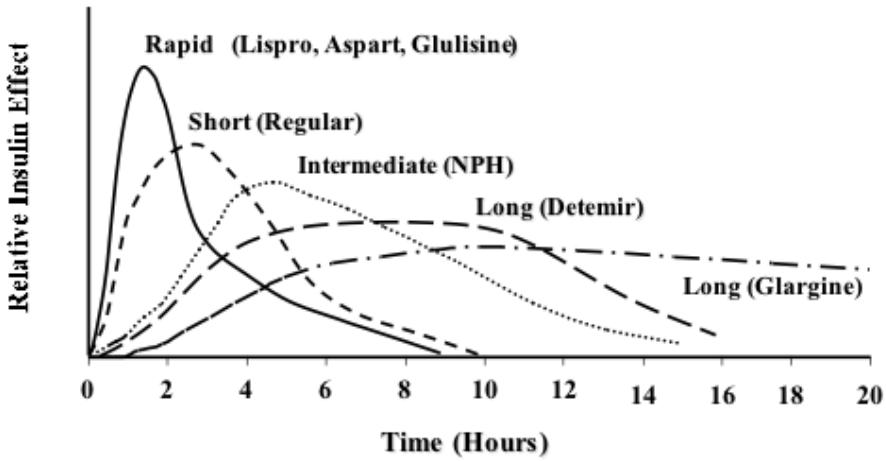



**Figure 2.7:** PK/PD curves for different insulin types. In addition to the insulin types shown here there is Fiasp, a ultra-short acting insulin.
Image by Anne Peters (CC BY 3.0) https://commons.wikimedia.org/w/index.php?curid=5856950

### 2.3.2 Treatment of type 2 diabetes

DM2 treatments usually start with lifestyle and diet interventions, then progress to oral antidiabetic medication (see below). Some DM2 patients are also treated with injected insulin, and insulin treatment gets more and more likely for each year after diagnosis of DM2 [48].

#### Oral antidiabetic medication

There are several oral medications for the treatment of DM2. The main ones are summarised below [50]:

*Sulfonylureas* increase the release of insulin from the beta cells in the pancreas, and only work when there is a residual endogenous production of insulin. They stimulate insulin release independent of glucose and may thus induce hypoglycemia.

*Metformin* decreases the blood glucose level in three ways. Firstly they may reduce glucose release from the liver by inhibiting gluconeogenesis and glycogenolysis. Secondly, they may increase insulin sensitivity and glucose uptake and utilisation in the peripheral tissues. Thirdly, they may delay intestinal glucose absorption.

*Thiazolidinediones, or glitazones* work by stimulating adipose tissue to increase the storage of fatty acids, which decreases the amount of FFA in the circulation, which in turn force cells to metabolise glucose to a greater extent.

*Alpha-glucosidase inhibitors* delay the digestion of carbohydrates in the gut, and thereby the uptake of glucose to blood.

*SLGT2 inhibitors* decreases the blood glucose level by inhibiting SGLT2, thereby preventing the kidneys from reabsorbing glucose from the pre-urine back into the blood, leading to more glucose being excreted in the urine.

*Incretins* are a group of hormones that stimulate a decrease in blood glucose levels by decreasing the rate of absorption of nutrients, or by causing an increase in insulin secretion. Medications in this category are Glucagon-like peptide (GLP)-1-analogues and Gastic inhibitory polypeptide (GIP), these need to be injected. Di-peptidyl peptidase (DPP)-4 inhibitor is another type of incretin that can be taken orally, to reduce inhibition of incretins already present in the body, thus prolonging the incretin effect.

### 2.3.3 Other diabetes treatments

#### Transplants

Beta cell transplants consists of transplanting donor beta cell islets into the pancreas or the liver. There, these cells will start to produce insulin. A limitation to this approach is the availability of donor beta cells, the high cost of preparing the cells from donor tissues, low survival of the transplanted beta cells and the need for immunosuppressive medication.

It is also possible to transplant the whole pancreas. This is usually done together with the kidneys, and is a high risk procedure that is usually only performed on patients with highly uncontrolled diabetes or diabetes complications.

In both above treatments the patient will have to use immunosuppressive drugs to prevent rejection of the transplanted organs or cells. Such treatment have major adverse effects and is a factor limiting the use of transplantation in treatment of DM1. Another factor is low availability of donors.

Stem cell research has the potential to restore the beta cell function in DM1 patients without or at least with reduced need for immuno-suppressants. The method consists of growing insulin-producing cells in vitro from the patient's own cells see e.g. [51, 52].

#### Glucose-sensitive insulin depot medication

Several research groups are investigating methods for creating molecules or nano-particles that are able to encapsulate insulin and release it in response to elevated glucose levels. If successful, this would enable 'depot insulin' that could be injected with intervals of days or weeks, and be able to automatically and passively avoid high levels of blood glucose [53, 54, 55].

### 2.3.4 The rule of halves

There is a rule of halves that applies to several medical areas, including diabetes. It states that among the currently estimated 425 million people with diabetes, only half are diagnosed. Of these, only half receive care for their condition. Of these, only half manage to achieve the treatment targets. Of these, only half get the desired outcome, which is to have no diabetes related complications, indicating that only 26 million, or 6% of the original 425 million people with diabetes, can lead complication-free lives . This is illustrated in Fig. 2.8

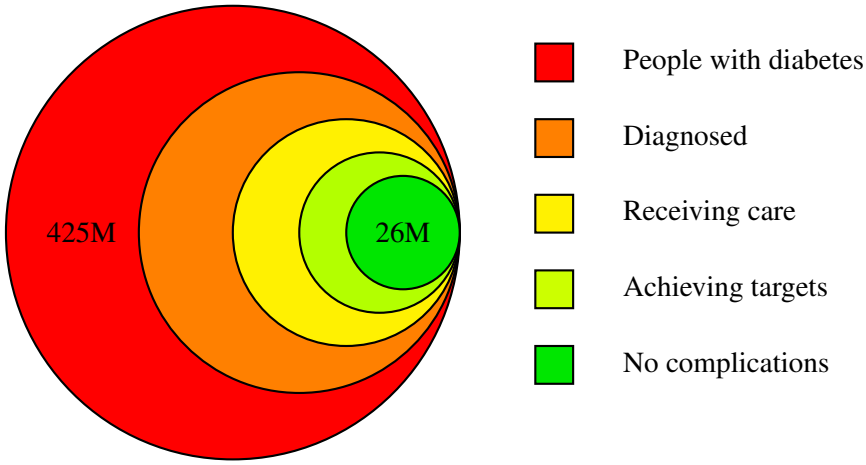



**Figure 2.8:** The rule of halves

## 2.4 Measurements relevant to diabetes

### 2.4.1 Blood glucose

Blood glucose measurement is used routinely in diagnosis, treatment and self-monitoring of diabetes, but is more complex and nuanced than one might initially assume. The following factors influence the level and variance of the glucose measurement:

- The specimen used (whole blood, plasma, serum)
- The site of blood collection (arterial, venous, capillary)
- For capillary blood, the blood flow in the capillary bed can affect the level
- The sampling procedure used
- The treatment of the blood sample between sampling and analysis

The glucose concentration of blood varies with where it is sampled. For instance, blood sampled from an artery typically has higher glucose concentration than blood sampled from the corresponding vein, due to uptake of glucose in the tissues between the artery and the vein. This difference is higher in the postprandial state [56]. There is also a difference between plasma and whole blood glucose, reported to be 15% lower in plasma, and this difference is hematocrit dependent. In addition, whole blood contains living cells, which keeps on consuming glucose also outside

the body, thus if the blood sample is not immediately analysed after being drawn a fall in glucose is observed [57].

Capillary blood glucose concentration varies with blood flow or perfusion through the capillary bed being sampled. In general, the less blood flow, the lower the measured blood glucose will be when sampling blood from this capillary bed, since the blood has longer time to lose its glucose to surrounding tissues during low blood flow. Capillary beds also contain mechanisms that control the amount of blood sent through the capillaries, these mechanisms are used for temperature regulation and to ensure that the tissue supplied by the capillary bed has sufficient nutrients and oxygen. Capillary blood flow is also pulsatile rather than continuous, this is called vasomotion, and the pulse frequency of vasomotion is independent of heart beat and respiration. Some places on the body have more constant capillary blood flow. However, capillary blood drawn from the finger, i.e. fingerpricks, has been found to be predominantly arterial blood, and thus closely approximates arterial blood glucose [58].

Systems for monitoring glucose are presented in Section 3.1

The SI unit of blood glucose concentration is mmol/L, sometimes written mM. Some countries, notably the US, use mg/dL. The conversion factor between these units is 18.018, i.e. 1 mmol/L glucose = 18.018 mg/dL.

### 2.4.2 Insulin, C-peptide and glucagon measurements

Plasma insulin, C-peptide and glucagon can be measured in blood samples in a laboratory using immunoassay methods. Such measurements are not available outside clinical or research settings [59].

Insulin is measured in international units, which is abbreviated IU, or often simply U which will be used in the following. In Norwegian the abbreviation IE is often used. The conversion factor from U to mg is 0.0347 for insulin, i.e. 1 U insulin = 0.0347 mg.

## 2.5 Variability

Modeling the glucose-insulin dynamics is a task that has occupied researchers for a long time, starting with Bolie in 1961 [60] through Bergman's minimal model [61] to today's maximal models [62]. Looking at data sets from clinical studies investigating the glucose rise or fall resulting from meals or insulin, the responses fit remarkably well to simple second order models, possibly with some delay [63]. This gives the impression that the system is not difficult to model, simulate and

predict. On the other hand, the amount of literature on the subject of glucose prediction indicates that this problem is far from solved. The unpredictable effects of a multitude of other events that happens to a person in his/her daily life make the dynamics in free-living settings only sporadically similar to the dynamics observed in clinical research settings. The nature of some of the variabilities causing low predictive value in blood glucose dynamic models are discussed in the following.

### 2.5.1 Meal variability

In most models found in the literature, meals are reduced to their carbohydrate content in grams, often abbreviated CHO, which is typically applied as an impulse input to the model. This is an oversimplification. It is known that two meals with the same amount of carbohydrates can behave quite differently in terms of what glucose excursions they will produce, and this has given rise to the concepts glycemic index (GI) [64] and glycemic load (GL) [65]. Meals with high glycemic load results in larger and more short-lived glucose fluctuations, see Fig. 2.9. Tables of glycemic index and load exist [66], and are used by dietitians and nutritional experts to give diet advises. It seems obvious that information about the glycemic index or glycemic load of meals need to be included in addition to just the carbohydrate amount, in order to give usable predictions [67].

GI/GL has some usability issues, since the GI index is difficult to measure accurately, and it seems to vary a lot with natural variation in different foods. As an example of the latter, an unripe banana may have a GI of 30, and an overripe banana may have a GI of 60. Another variability that reduces usability is that different sub-types of a food may have different properties. For instance potatoes, considered by most people to be a single, homogeneous type of food, consists of different breeds of potatoes. Looking at the table [66] we see a difference in GI for cooked potato ranging from 56 (Pontiac breed) to 118 (Sava breed). Even two different studies of the same breed (Pontiac) seems to have measured the GI quite differently, resulting in a GI of 56 in one study and $88 \pm 9$ in another. There are also studies showing that the preparation of a potato has much influence on the post prandial glucose response (PPGR), as raw potato, cooked potato, cooled cooked potato and reheated cooked potato [68, 69]. Other studies also report large between- and within-person variations [70, 71]. Such inconsistencies contribute to making GI less useful for glucose prediction purposes. Another measure is the Glycemic Glucose Equivalents (GGE) [72], which has seen less use.

The reproducibility issues seen in GI studies and tables may be explained through the work by Eran Segal and his research group. In an important paper [73] they explain how glycemic responses to food are person-dependent to a large degree. In

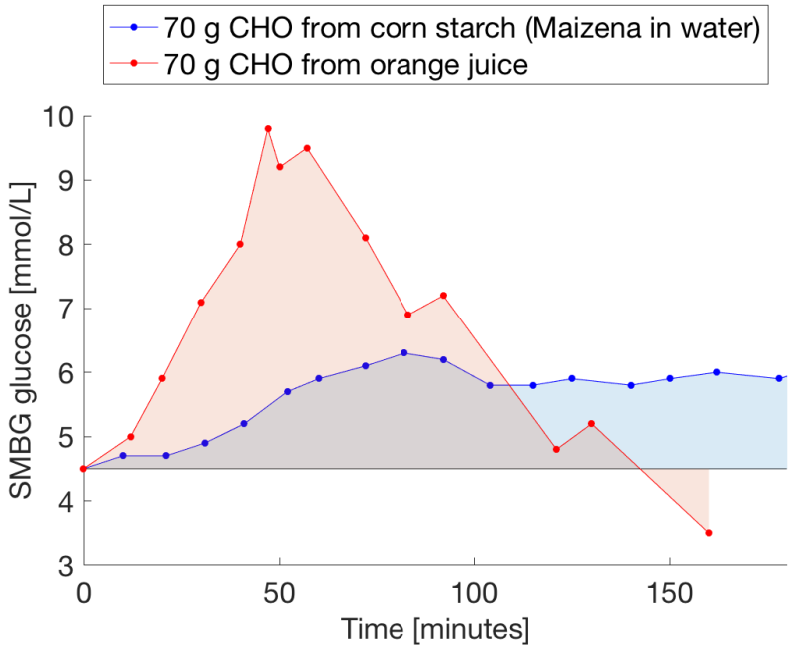



**Figure 2.9:** Differences between equal-carbohydrate meals with high GI vs low GI in a healthy individual. Meal is taken at time 0. The RMSE between these curves is 2.5 mmol/L, and the MARD between them is 36%. It is clear that predicting the post-prandial glucose from CHO counting alone is not reliable. The figure also illustrates a problem in the calculation of GI, which is based on the area under these curves. In this case, since the corn starch gives a prolonged glucose response above baseline the GI computed would be similar or even higher than the GI of the orange juice. This goes against the common interpretation of GI, which is that a high-GI meal will produce a high, spiky glucose response. Data from the author using a Freestyle Freedom Lite.

other words, a meal that gives a spiky and high glucose response in one individual may not produce response at all in another. This is illustrated in Fig. 2.10. The implication is that GI is not a property of the food only, but rather is determined by a combination of the food and the person consuming the food. In other words, it seems each individual needs an individualised GI table.

Even if these issues with GI had not been present, most people rarely eat only one type of food in a meal. The act of computing the composite GI of a meal consisting of many different types of food is non-trivial, and would only be performed reliably and consistently by the most diligent of users.

Cescon et al. [74] use the glycemic index of a meal to adapt four of the parameters

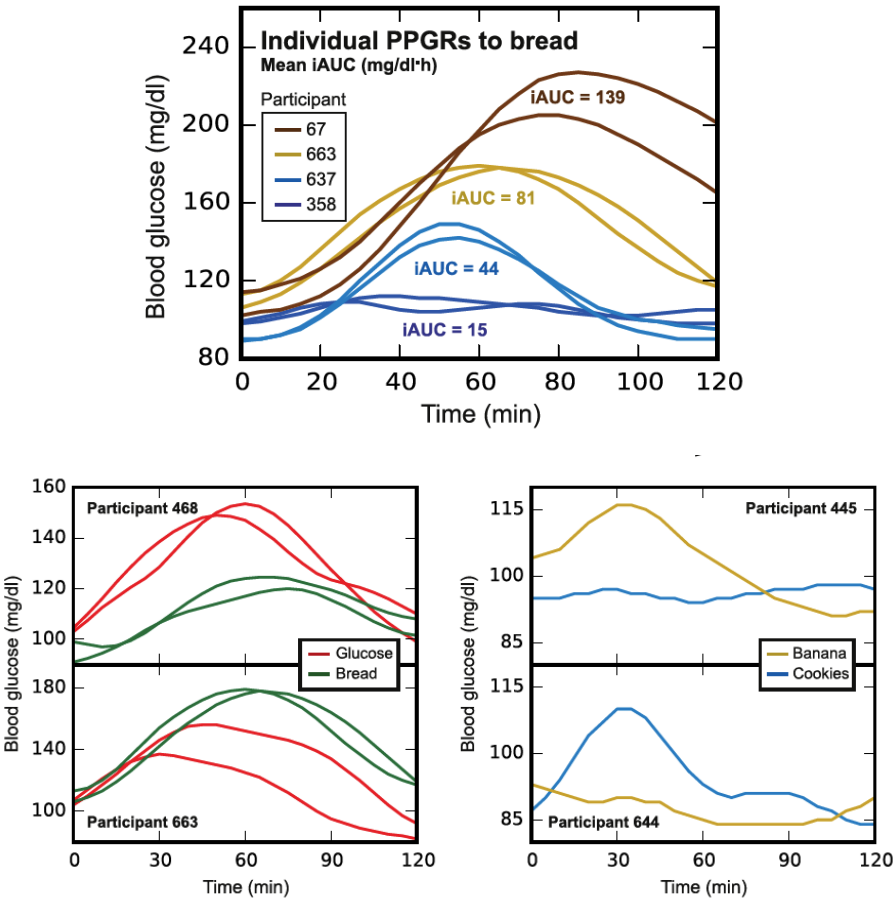



**Figure 2.10:** Upper panel shows the post-prandial glycemic response (PPGR) of four different participants to the same meal (bread). Note the high inter-person variability, but low intra-person variability. Lower left shows how one participant had a higher PPGR from bread than from pure glucose, contrary to what would be expected from GI tables. Lower right show how two different participants have opposite reactions to bananas and cookies. Reproduced from From Fig. 2 in [73], with permission.

in the UVa-Padova meal model. Bao et al. [75] found that using GL in addition to carbohydrate content improved the predictive ability, increasing the variance explained from 59% to 85% for single-food meals. In an article by Rozendaal et al. [76], both the use of a single CHO number and CHO + GI/GL is dismissed as not providing enough predictive ability of the glucose response following a meal, and a

more complex model of the meal absorption subsystem is presented.

### 2.5.2 Insulin input variability

When persons with diabetes administer insulin, the same dose will not have exactly the same effect from injection to injection, even if the dose remains constant. This has been investigated by several researchers [77, 78]. There are several factors that affect the insulin response, e.g. injection site, skin and ambient temperature, movement, exercise, injection into muscle vs subcutaneous tissue, age, smoking, lipohypertrophies and the thickness of subcutaneous tissue [79]. For pump users the possibility of partial or complete obstruction of the infusion set is a common cause of insulin input variability [80].

### 2.5.3 Insulin sensitivity variation

Insulin sensitivity is a parameter in most models, implying that it is near constant. However, there is much evidence that this quantity is more accurately represented as a slowly varying state. The insulin sensitivity varies throughout the day, and is reported to be affected by a great many factors [79]. Factors that are believed to decrease insulin sensitivity are: High blood glucose for over 12 hours, recent hypoglycemia, meal bolus set delayed in comparison to the meal, end of night (dawn effect), stress, puberty, pregnancy, overweight or weight increase, high blood pressure, smoking infections, fever, inactivity, surgery, ketoacidosis, some medications, some diseases and menstruation (for some). Factors that are believed to decrease insulin sensitivity are low blood glucose, weight loss, physical activity, early night and breast feeding. These factors and their influence on insulin sensitivity are likely person dependent.

# Chapter 3

# Technical background

This chapter provides background information related to technical devices involved in glucose measurement, insulin delivery and blood glucose/diabetes related data logging. The concept of virtual patients is also discussed.

## 3.1 Glucose monitoring systems

The monitoring of glucose is important in order to control blood glucose to stay within the normal range. Different systems for measuring blood and interstitial glucose exists, and are described in the following.

### 3.1.1 Self Monitoring of Blood Glucose (SMBG)

Self Monitoring of Blood Glucose is performed by people with diabetes in everyday life using a hand-held blood glucose meter. The sample is taken from capillary blood from the finger, placed on a strip and inserted into a hand-held meter. The electrochemical analysis of the blood takes about 5 seconds in recent systems, producing a glucose concentration that is often considered to be a direct measurement of plasma glucose. There are subtle differences between capillary blood and plasma glucose, but meters are usually calibrated to provide plasma glucose readings [81].

The main disadvantages of SMBG is the need for pricking the finger to draw blood, its discontinuous nature and a moderate risk of sampling error, e.g. insufficient washing of the finger before sampling, effects of hemolysis due to pressing the finger too hard to squeeze out blood, etc.


**Figure 3.1:** SMBG meter. The strip is an electrochemical device that receives a small drop of blood. The blood is transported into the strip by lateral flow, and comes into contact with enzymes inside the strip (hexokinase or glucose oxidase) that generate a current proportional to the glucose concentration in the blood. The current is read by the meter, and converted to a glucose concentration.

### 3.1.2 Continuous glucose monitor (CGM)

CGMs work by inserting a electrochemical sensor into the subcutaneous tissue, usually on the abdomen, that stays there during the sensor wear. The sensor is placed subcutaneously using a needle which retracts once the sensor is in place. The sensor is small and is by most users considered painless during normal use. In most systems the sensor needs to be replaced every 7 days. The sensor is connected to a transmitter that sits on the outside of the body. The transmitter sends the data to a receiver that displays the glucose value to the user. Data processing on the raw amperometric signal from the subcutaneous sensor is performed either in the transmitter or on the receiver to filter out noise and artefacts. This processing varies between CGM brands and models, but invariably results in some delay. In recently marketed systems the receiver can be a cellphone. CGMs measure interstitial glucose, which adds physiological delay compared to blood glucose. This means that when blood glucose is stable, the CGM measures glucose values quite similar to fingerpricks, but when the blood glucose is changing, and in particular when it is changing rapidly, the CGM lags behind the blood glucose measurements, and typically does not get to the same peak level. The delay reported in the literature varies from 5 to 40 minutes. Schmelzeisen-Redeker et al. [82] suggests that a large portion of the physiological delay variability is patient dependent. It has been observed that the lag is greater when the glucose level increases compared to when

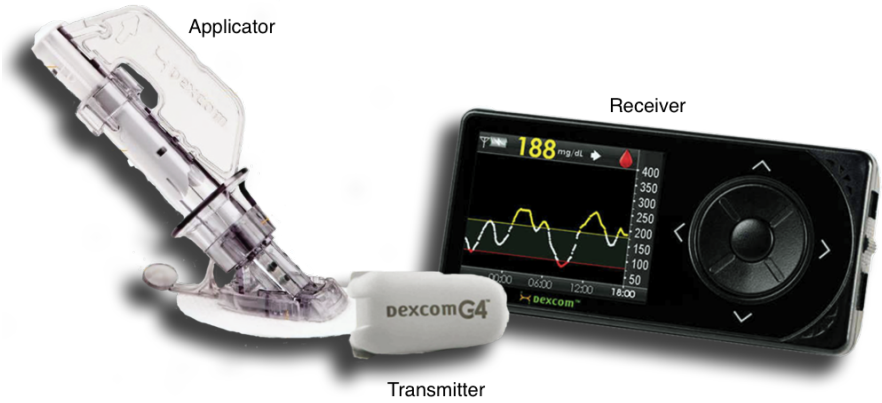



**Figure 3.2:** The DexCom G4 system. The item shown on the left is the applicator, containing the electrochemical sensor and a needle-based insertion mechanism that brings the electrochemical sensor into the subcutaneous tissue. Once the sensor is in place, the transmitter is connected to the sensor, and starts sending glucose data to the receiver.

it decreases [83]. Due to differences in sensor design and algorithms there will be differences between the CGM devices with respect to the delay dynamics of the CGM signal compared to blood glucose. This was shown to be the case in [84]. This is also discussed in [85], where it is argued that these device differences have confounded the issue of what the physiological delay really is, and that delays longer than 10 minutes are due to CGM device delay, not physiology. This is supported by the findings in other studies [86, 87], where a lag less than 10 minutes is found using microdialysis and radioactive tracer techniques. An argument for CGMs is that the brain glucose level, which is the important one with respect to hypoglycemia, is more similar to the CGM glucose measurement than blood glucose measurement [88].

A confounding factor in this is the sensor-to-sensor variability, which could also have elements of between-patient variability. As investigated as part of this PhD project (see paper A.2), the FreeStyle Libre FGM system has individual biases and lags that may follow the sensor or the patient, or a combination of both. This has implications with regards to how such data should be used for research purposes, especially data that spans several sensors (over one week). It is likely that different CGM systems (Medtronic, DexCom) have different characteristics with regards to this. Depending on what the CGM data is to be used for, the lag is more or

less important. For retrospective analysis of the data by a clinician or the patient, the variability and mean level of the signal is of much higher importance than the exact timeliness. For the purpose of real-time alerts of hypoglycemia or for artificial pancreas applications, any delay larger than zero is detrimental to system performance. In research use of CGM data, for instance in estimating parameters of BG dynamic models, it is important to be aware of these issues with the data to avoid them affecting parameter estimates.

Retrospective analysis of CGM data also have some other errors that occur: Sometimes the signal is missing, either because the sensor is temporarily malfunctioning, or because the receiver is not in the range of the sensor. Another commonly seen artefact in CGM data is Pressure Induced Sensor Attenuation (PISA), where the CGM signal is seen to fall and then recover while blood glucose in reality is stable. This effect comes from pressing on the sensor, e.g. lying on the sensor during sleep [89].

### 3.1.3 Flash Glucose Monitor (FGM)

At the time of writing, there is only one Flash Glucose Monitor on the market, Abbott's FreeStyle Libre (FL). On the sensor side, the FGM is similar to conventional CGMs, using electrochemical sensing of glucose in the interstitial fluid using a sensor inserted subcutaneously on the upper arm. The main difference between a FGM and a CGM is that the FGM only produces a result when a reader is used to scan it. This implies that the FGM has no alarm capability; it can only notify the user of hypo- or hyperglycemia if the user happens to scan the sensor at the times when these events are about to happen or are happening.

The FL is the first CGM-like system that is factory calibrated, meaning that it does not require a calibration SMBG measurement. It has also been approved by the FDA as a non-adjunctive device, meaning that no confirmatory SMBG measurement is required in order to take treatment decisions based on the FL measurement. DexCom G6 is another recently launched CGM system requiring no calibration.

One of the papers produced as part of this PhD project describes the errors of an FGM compared to SMBG measurements, focusing on the biases and lags between the sensors. See Appendix A.2.

### 3.1.4 Glycated haemoglobin (HbA1c)

HbA1c, or sometimes just A1c, is a measure of the binding of glucose to haemoglobin molecules of red blood cells. Glucose binds to haemoglobin irreversibly, so in a sense the haemoglobin molecules of red blood cells are computing the integral of


**Figure 3.3:** The FreeStyle Libre from Abbott. Shown here is the sensor (white) situated on the arm and the receiver (black) used to scan the sensor. The scanning can also be done using a Near Field Communication (NFC) enabled smartphone. Image from creative commons, license: CC BY-SA 4.0

the glucose levels the blood cells are subjected to. Since red blood cells live for about 3 months the fraction of glycated to un-glycated haemoglobin correlates with the level of glucose over the last 6-8 weeks, and this ratio is what HbA1c measures and reports in percent. The %HbA1c value can be transformed into an estimated average glucose (eAG) in mg/dL through a linear formula [90].

$$\text{eAG} = \%\text{HbA1c} \cdot 28.7 - 46.7 \tag{3.1}$$

HbA1c can be utilised in models of glucose dynamics. For instance, many models have a person-specific $G_b$ parameter describing the basal glucose level of the patient in $\mathrm{mg/dL}$. One approach to set this parameter is to use the last known HbA1c measurement of the individual and set $G_b = \text{eAG}$ computed with the above formula [91, 92].

HbA1c is used for monitoring and diagnosing diabetes. A HbA1c measurement greater than 48 mmol/mol (6.5%) is diagnostic of diabetes, and a often used treatment target is to get HbA1c below 53 mmol/mol (7%). An HbA1c above 75 mmol/mol (9%) is associated with significant risk of long-term complications. There are point-of-care analysers for measuring HbA1c [93], as well as laboratory analysers.

### 3.1.5 Glucose variability

While HbA1c tells some of the story of a diabetes patient's level of control, there are indications that the glucose variability (GV) matters as much or more than the mean level captured by HbA1c. Which method to use to measure GV and whether or not GV should be included as a treatment target is being debated [94, 95, 96].

### 3.1.6 Laboratory methods

There is no internationally recognised reference method for determining blood glucose, however the de-facto standard method is to sample venous blood and analyse it using a desktop analyser called YSI 2300 Stat, which is used as the reference method by more than 80% of glucometer manufacturers, according to YSI [97]. Glucose measurements with the YSI Stat 2300 are often just called YSI measurements.

### 3.1.7 Non-invasive measurement of blood glucose

There have been a lot of attempts at creating a blood glucose measurement system that does not require drawing blood or inserting sensors through the skin. There are multiple reviews of non-invasive technologies [98, 99, 100]. Although some devices have been marketed, so far no non-invasive glucose measurement device has been successful in being accurate enough and pain-free enough to see extensive use.

### 3.1.8 Interferents

CGM measurements can be affected by other chemical constituents in the blood in a way that results in a false reading of glucose [101, 102, 103]. Basu et al. [101] investigated the influence of different drugs on CGM readings in two different CGM systems (Dexcom G4 and the Medtronic Guardian Sof-Sensor). Among the drugs tested were paracetamol (acetaminophen) and ethanol (wine). They concluded that there was evidence that the electrochemistry of the CGM systems is perturbed by these compounds, and the strongest effect was seen with paracetamol.

Interferents can also apply to SMBG measurements [104, 81], e.g. high levels of uric acid or triglycerides in the blood. Externally taken medication like paracetamol and ascorbic acid [105] may also influence the reading.

## 3.2 Insulin pumps

Traditionally, insulin has been administered to patients using syringes and insulin pens, which are described in Section 2.3. Insulin pumps provide a convenient alternative for insulin administration.

Insulin pumps are small systems containing a reservoir of insulin, a pump and an infusion set. An image of an insulin pump in use is shown in Fig. 3.4. Insulin pumps work by injecting rapid-acting insulin via the infusion set into subcutaneous tissue, usually on the abdomen. The user selects basal and bolus dosages to be injected. The basal dosage is injected automatically and semi-continuously throughout the day and night as small and frequent injections, while the bolus dosages are injected when the user chooses to, usually in conjunction with meals. In most pumps the basal dose can be programmed on a 24-hour basis, and the user can typically define several different basal programs and select the most suitable one for the day depending on the user's predicted level of activity, stress or other factors that may affect the glucose dynamics. For instance, some insulin pump users choose to have a different basal insulin programs during weekends than during workdays.

The use of rapid-acting insulin also for providing the basal insulin level is an advantage over the MDI treatment regime and slow-acting insulin; if a change to the basal insulin dosage is needed it can be in effect within 30-120 minutes, while a slow-acting insulin injection cannot be aborted or adjusted once it has been set.

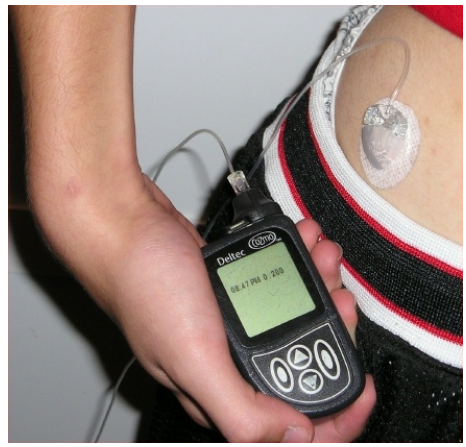



**Figure 3.4:** An insulin pump in use. Image from creative commons. License: public domain.

Insulin pumps take a lot of the inconvenience out of insulin administration, but adds

some new risks, e.g. blocked infusion set [80].

## 3.3 Logging systems

There are many factors other than meal and insulin that affect blood glucose. In order to understand and predict these effects they need to be logged together with the CGM and insulin pump data.

Many of the systems for glucose monitoring and insulin administration described above contain logging systems. For the glucose meters it is commonly the case that only the measurements are logged, but such data can be useful for retrospective analysis. CGM/FGM systems, in addition to storing logs of the actual glucose measurements, commonly has functionality to log other data from the patient. For instance, the FreeStyle Libre has the possibility to log rapid-acting insulin doses, long-acting insulin doses and meals, among other data. The data can be input with a value (units of insulin, and grams CHO in the case of meals) or as non-numeric, for those cases where the user cannot for some reason input the value (e.g. if the meal CHO content is hard to estimate). Insulin pumps have similar functionality, with a log of the insulin it has been administering. Insulin pumps that have bolus calculators (BC) typically log the estimated CHO contents of meals that the user input when using the BC. and also in some cases possibility to log other data.

Although these individual system logs provide useful functionality, they are not in much use by users, and the use of such extra functionality stops after having used the system for a while. Some of the reason for this may be the rather limited user interface (UI) of the CGM/FGM/insulin pump devices. As more systems move the primary user interface from a custom device with limited UI capabilities to a smart phone this becomes much less of a problem, and more logging of auxiliary data may result.

A noteworthy system for logging of diabetes related data is Tidepool, which seeks to bring together data from glucose monitors and insulin pumps from different vendors in a cloud-based system [106, 107]. Tidepool also has an app for logging events that may be relevant for the retrospective interpretation of the glucose and pump data. Such systems are useful for homogenising diabetes related data and adding meta-information that is necessary for correct interpretation of the data.

## 3.4 Virtual patients

A virtual patient (VP) is an interactive computer representation that is able to emulate the real response or behaviour of a real patient in some respect. There are

many different areas of medicine where the concept of a VP is useful, cost-effective and risk reducing [108], including diabetes. Some current and potential future applications of virtual patients in diabetes are:

1. Simulating the effect of treatments, see Section 3.4.1
2. Personalised model-based control in artificial pancreas, see Section 3.4.2
3. Diabetes advisors in the form of apps, see Section 3.4.4
4. Algorithms for improved CGM measurements, see Section 3.4.3

### 3.4.1 Diabetes patient simulator

VPs can be used in simulators to simulate how certain treatments will affect the patient. A treatment that is possible to simulate could be a bolus calculator, which is a formula for computing the best insulin dosage based on information about a meal about to be taken. Another treatment is a medical device, like a sensor-augmented pump, or an artificial pancreas system. The workings of such systems can be simulated by letting the system receive simulated glucose measurements from the virtual patient, and output insulin bolus commands to the virtual patient, thus simulating the response the system would have on a patient. It is common to only use part of the device software in such a test, e.g. the main algorithm, but it would also be possible to do hardware-in-the-loop (HWITL) simulations, where simulated glucose sensors provide a glucose signal that the external system uses to compute an insulin input to the model. This is illustrated in Fig. 3.5.

Performing simulations on a virtual patient population to determine the performance and safety of an algorithm or a technical system is called performing *in silico* experiments. *In silico* experiments reduces the cost and risk of clinical trials, since several algorithms can be tested quickly and reproducibly. Another benefit of *in silico* experiments is the possibility to include elements that would be unethical and/or dangerous to subject real patients to, for instance overdoses of insulin, severe hypoglycemia, long periods in hyperglycemia and starvation periods. The system's response to such situations is of course important to assess, but may be impossible to do safely in *in vivo* clinical trials.

An often cited simulator is the UVa-Padova T1DM simulator (T1DMS), which as the name implies simulates T1DM patients [62, 109]. T1DMS contains a set of VPs of 100 adults, 100 adolescents and 100 children. The educational version of the software has 10 VPs from each category. The model running in the UVa-Padova T1DMS is a model with 18 states and more than 40 parameters. The virtual patients consist of this model and a parameter set drawn from a joint distribution of

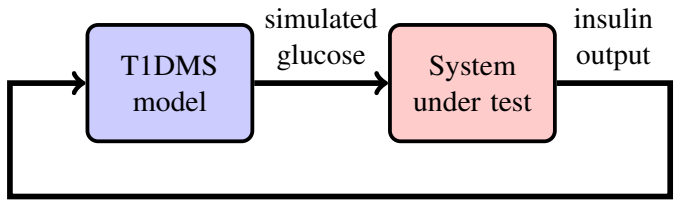



**Figure 3.5:** Use of the T1DM simulator in an *in silico* experiment to determine the properties of a system under test. The system under test may either be a software representation of the real system, or the real system itself (HWITL), with adapters for interfacing with the simulator, e.g. inputing simulated glucose measurements and outputing insulin pump commands to the simulator

parameters. A recent development by the same company is the Diabetes Mellitus Metabolic Simulator for Research (DMMS.R), which expands the simulator to also include DM2 diabetes and prediabetes virtual patients [110].

### 3.4.2 Artificial pancreas / closed loop glucose control

The artificial pancreas aims to reduce the burden on the person with diabetes by closing the loop of the glucose measurement and insulin delivery, i.e. by interconnecting CGM and insulin pump and letting the insulin pump automatically decide to inject insulin based on the CGM signal.

The only commercial artificial pancreas system currently available is Medtronic 670G, a so-called Hybrid Closed Loop (HCL) system. This system only adjusts the basal glucose rate, and does not automatically compensate for meals.

Model Predictive Control (MPC) is one of the algorithm families that can be used in an AP to control the glucose. In MPC, a BG dynamic model is used to predict the future glucose response, and the next control input (or sequence of inputs) is determined by solving an optimisation problem. The cost function of the optimisation problem is typically related to some treatment target, e.g. minimising the distance from a target glucose, or time outside the normal glucose zone (Zone-MPC, [111]).

The artificial pancreas system is potentially lethal to its user if it malfunctions [80], most notable is the risk of insulin overdosing. This makes device development harder, since more stringent regulatory requirements needs to be fulfilled for such systems. Diabetes is also a very variable disease, making it hard to design a control system that fits all. Consequently, it is difficult to get AP systems to market.

Some users have grown impatient with waiting for commercial actors, and have

created their own AP systems based on commercially available insulin pumps and CGM systems. The OpenAPS group [112, 113] provides recipes for interconnecting CGM systems and insulin pumps and putting in a control algorithm, resulting in a Do-It-Yourself (DIY) artificial pancreas. Since each system is built by the person that is going to use it, all risks are taken by the users. It also means that each system is unique, and the group has adopted a notation for this, $N = 1$, implying that each system is tailor made to its user, by its user. This of course requires technical competency and know-how, some of which is provided by the OpenAPS website, but this solution is clearly not for all. According to the website, as of December 3 2018 there were more than $(n = 1) * 1,015$ users of various types of DIY closed loop implementations around the world.

### 3.4.3 Personalised CGMs

A BG dynamic model may be used to filter signals from a continuous glucose monitoring system, by providing estimates in times of signal fallout, and filtering out non-physiological high-frequency components of the signal. In the context of Prediktor Medical's development of BioMKR this is an interesting use case.

Another possibility is to use models to filter signals in order to try to remove some of the lag in the data compared to SMBG measurements, see e.g. the work by Facchinetti et al. [114]. There are signs in the data from CGMs and FGMs that such models are in place, see an illustration of this in Fig. 3.6.

Another interesting possible future (and perhaps far-fetched) use of BG dynamic models in relation to CGM could be for sensor calibration. In theory, a BG dynamic model could be individualised to a such degree that it could detect that a newly inserted CGM sensor has a different bias and lag than the previous one, and correct for this.

### 3.4.4 Diabetes advisor apps

Several use cases have been envisioned for BG dynamic models used in advisor apps [115, 116]. For instance, a properly calibrated BG dynamic model could provide future glucose predictions based on the CGM measurements and input data (meals, insulin) until present, and could give advice about meals and insulin boluses, suggest changes to the basal insulin dosing, or warn about predicted future hyper- or hypoglycemias.

In a longer time perspective, an advisor app might track the parameter values of a patient over time, looking for patterns that could be relevant for patients and/or caretakers, for instance decreasing insulin sensitivity.

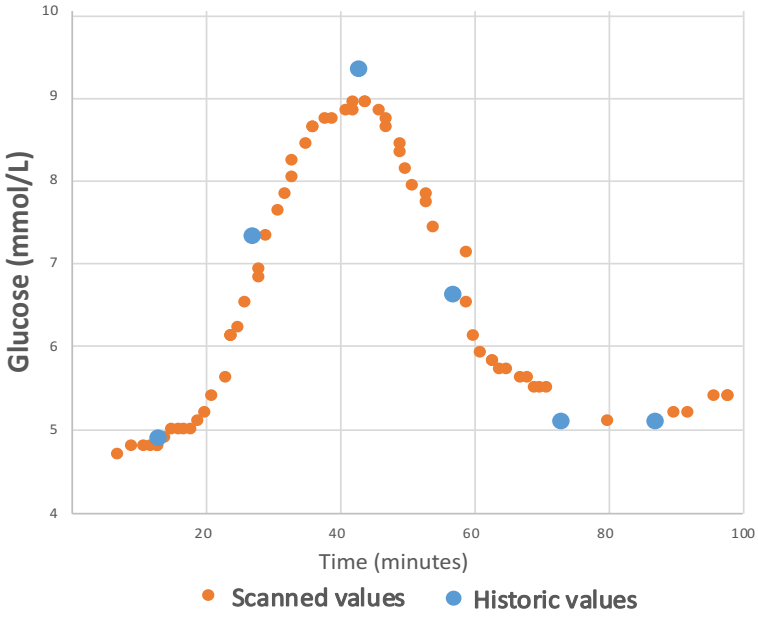



**Figure 3.6:** Logged data from a FreeStyle Libre(FL) FGM. The sensor was scanned every minute, producing the orange points. The FL generates the blue points, called historic glucose in the log file. Some kind of filtering is evidently present since the historic value peak is larger than any of the scanned ones. Based on time stamps and the ordering in the log file, the blue points are computed at the end of every 15-minute window but are given the time stamp of the beginning of the window. Data from the author.

# Chapter 4

# Models of glucose dynamics

## 4.1 Overview

A blood glucose (BG) dynamic model is any model that describes the dynamics of blood glucose, given inputs such as meals, insulin dosing and potentially other measured quantities that affect blood glucose. A virtual patient (VP) for diabetes may consist of a BG dynamic model with or without patient adapted parameters. A BG dynamic model may be intended to describe a certain subset of the population, e.g. only T1DM patients, or it may be able to describe all groups by using different parameters, including healthy individuals. As an example, a model may be developed to fit healthy and DM2 subjects. By setting the parameters involved with endogenous insulin production to zero, a suitable model for DM1 may result.

BG dynamic models have been a topic of research since the 1960s, and there are several reviews of the different models that have been investigated, see e.g. [117, 118, 119, 120].

There are many different uses of a BG dynamic model. Some are used purely for simulation, where all inputs and parameters are assumed fully known, and the purpose is to generate glucose trajectories that could be representative of real glucose data, e.g. as in the virtual patient concept and *in silico* testing. Such models can be very complex, with a large number of states and parameters, and have the ability to describe many of the effects that can be seen in glucose data. The drawback of these models is that they require measurement of many different states in order to be observable and/or to identify parameters of the model, which is a necessary property in models that shall be used for other purposes than simulation. Other models are constructed for the purpose of estimation, prediction and/or

control. These models are typically considerably less complex than the complex simulation models, which implies lower fidelity and ability to model the more complex glucose phenomena. There is a trade-off that needs to be made in selecting the correct model for the intended application.

In this PhD project we are interested in models that provide a way to predict glucose from measurements and input information that can realistically be obtained in a free-living setting. This means measurements like SMBG and CGM, and input information like meals, insulin boluses (fast and slow-acting). Some of this information is available from external systems that the person with diabetes may be using. A CGM provides glucose measurements and an insulin pump provides insulin information. Given the cost of CGMs and insulin pumps, such technology is currently only available to a quite small percentage of the people suffering from diabetes globally.

## 4.2 State space models

State space models are a family of models where a vector $\mathbf{x}$ of states is propagated according to a state transition function

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, \mathbf{u}, \mathbf{v}) \tag{4.1}$$

where $\mathbf{p}$ is a vector of parameters, $\mathbf{u}$ is a vector of inputs, and $\mathbf{v}$ is a vector of disturbances, also called process noises.

The measurements available from the system are given by

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{u}, \mathbf{w}) \tag{4.2}$$

where $\mathbf{y}$ is a vector of measurements and $\mathbf{w}$ is a vector of measurement noises.

How much of the glucoregulatory system that is modelled depends on the model and its intended use. An illustration of the different submodels that may be included in a BG dynamic model is given in Fig 4.1.

The glucose/insulin dynamics central model is present in most BG dynamics models, but is only applicable in situations where glucose and insulin is applied intravenously, e.g. in clamping studies. In most cases a submodel for the appearance of glucose is needed, i.e. a meal dynamics model. To model the effect of injecting insulin, an exogenous insulin submodel is needed. The measurement dynamics may be of interest to model if the measurement data are from CGM systems that may be considerably influenced by the plasma-interstitial diffusion delay that such submodels describe. Exercise influence is interesting to add for some usage scenarios,

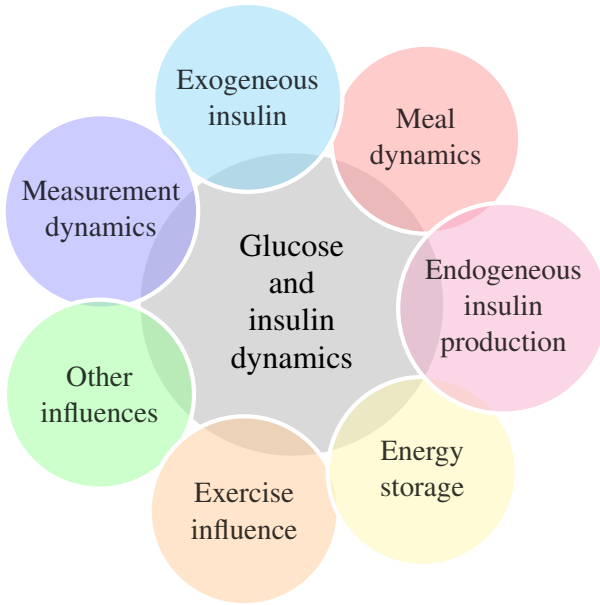



**Figure 4.1:** Illustration of the different modules BG dynamic models may consist of

and in free-living settings the exercise influence is a major cause of prediction error [121], and is considered one of the more difficult situations to handle for artificial pancreas applications.

#### States

In the BG dynamic models of the literature, the state definition and naming usually differ from model to model, but there are some commonalities. It is commonly the case that the state vector has at least the following states:

- $G_p$: Plasma glucose
- $I_p$: Plasma insulin
- $I_r$: Remote insulin

Remote insulin is sometimes given the symbol $X$ and may be modelled as an insulin action, which is not an insulin concentration, but a time constant with units (1/min).

Other states of interest and included in some models are:

- $G_i$: Interstitial fluid glucose

- $I_s$: Subcutaneous insulin (often more than one compartment is modelled, i.e. $I_s$ is split into $I_{s1}$ and $I_{s2}$ )
- $G_s$: Glucose in the stomach
- $G_g$: Glucose in the gut
- Insulin in other compartments, e.g. in liver tissue
- Glycogen storage in liver
- Glycogen storage in muscle
- Glucagon concentration in the blood
- Exercise related states
- Alcohol related states

#### Inputs and disturbances

The inputs most often applied to BG dynamic models are

- $u_{CHO}$: Amount of carbohydrates in meal
- $u_{Ir}$: Rapid-acting insulin

Other inputs that may be supported are

- $u_{Is}$: Slow-acting insulin
- $u_g$: Glucagon injection (currently glucagon is only injected as an emergency procedure in case of hypoglycemia, but in the future it may be part of a dual-hormone AP concept)
- $u_{Ii}$: Inhaled insulin (currently not in widespread use)
- $u_{hr}$: Heart rate (as a measure of exercise intensity)
- $u_a$: Activity (for exercise modelling, e.g. from a classifier based on accelerometer data)

Disturbances or process noise in BG dynamic models are all the effects not modelled or not measurable. Effects on BG caused by stress, disease, circadian rhythms and menstrual cycle are usually not explicitly modelled, and are therefore considered disturbances. Note that in models where e.g. exercise is not explicitly modelled, exercise effects would also be considered a disturbance, or process noise.

The distinction between input and disturbance is not always clear cut. Is it sensible to model something as a control input if the time or magnitude of the input is

uncertain, difficult to define or measure, as is the case with meals? Conversely, is it sensible to model something as a disturbance if it can be accurately measured? In the context of BG dynamic models, meals and insulin are usually used as if they are known, deterministic inputs. While this is true to some extent for insulin when the insulin is administered by an insulin pump, it is far from true for meals, see Section 2.5.1.

#### Parameters

The parameters of BG dynamic models range in number from 4 for the most minimalistic models, to more than 40 for the complex models. The distinction between a state and parameter is often that a parameter is constant, while a state is dynamic. This distinction can get blurred, for instance by slowly varying parameters. A prime example is the insulin sensitivity, which is often modelled as a fixed parameter, but as discussed in Section 2.5.3 should perhaps be considered a state, or a disturbance with a known pattern of variation.

#### Measurements

In a free-living setting the only measurements from a person with diabetes is glucose data. Measurements of insulin and glucagon is currently only available in lab settings. For insulin pump users the amount of insulin that the pump has injected (or more accurately: what it has tried to inject) is available. This could potentially also become true of glucagon, however glucagon pump systems are currently only used for research in dual-hormone artificial pancreas, and are presently not available to the general public.

The glucose data is provided by CGM and SMBG systems. CGM data is dense, typically one sample every five minutes, but has the disadvantage that it measures the glucose in a compartment remote from the one of primary interest, i.e. it measures interstitial fluid glucose instead of BG. On the other hand, SMBG data is a measure of the BG directly, but it is sparse; most users perform only a few SMBG measurements per day.

## 4.3 Submodels

Some of the submodels that are relevant to add to a BG dynamic model is reviewed in the following subsections.

### 4.3.1 Measurement dynamics

CGM systems measure interstitial fluid glucose, which lags behind plasma glucose as measured with SMBG [122, 123, 85]. The processes involved in the [124]. A commonly used model of the plasma to interstitial glucose dynamics is

$$\dot{G}_i = \frac{1}{\tau_i}(G_p - G_i) \tag{4.3}$$

Here, $G_p$ is plasma glucose, $G_i$ is interstitial fluid glucose, and $\tau_i$ is a time constant governing the diffusion between these compartments. This model was used in one of the articles published as part of this PhD project (paper A.2), to evaluate the lags present in FGM measurements.

Nine different submodels of the plasma-to-interstitial glucose kinetics have been investigated by Wilinska et al. [125], and they found the model that best explained their data had a term for glucose disappearance in the interstitial tissue, as well as a insulin-regulated diffusion from plasma to interstitial fluid. Another study into these kinetics was performed by Schiavon et al. [126], and they found that a model similar to the one in Eq. 4.3 but with more parameters, is appropriate to model these dynamics. Although these more complex models exist, the simpler model in Eq. 4.3 is widely used in modelling efforts based solely on CGM data, due to identifiability reasons.

To be able to determine the parameters of the measurement dynamics (including $\tau_i$) it is necessary to have frequent SMBG measurements in addition to CGM measurement, which is seldom encountered in data sets from free-living settings. This is likely the reason that many models either use the model Eq. 4.3 with $\tau_i$ fixed to 10 minutes, or consider this dynamics negligible, as in the model described in Section 4.4.1.

### 4.3.2 Meal dynamics

A often used model of meal dynamics is the one proposed by [127], which constitute the meal submodel in the UVa-Padova T1DM simulator model. The submodel has three states and 12 parameters, of which one is the body weight (BW) and another is the glucose content in the last meal. The model is given by

$$\mathbf{x} = \begin{bmatrix} Q_{sto1} \\ Q_{sto2} \\ Q_{gut} \end{bmatrix} \begin{matrix} \text{- Solid phase glucose in stomach} \\ \text{- Liquid phase glucose in stomach} \\ \text{- Glucose mass in intestine} \end{matrix}$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \dot{Q}_{sto1} \\ \dot{Q}_{sto2} \\ \dot{Q}_{gut} \end{bmatrix} = \begin{bmatrix} -k_{gri} Q_{sto1} + u_{CHO} \\ -k_{empt}(Q_{sto}) Q_{sto2} + k_{gri} Q_{sto1} \\ -k_{abs} Q_{gut} + k_{empt}(Q_{sto}) \end{bmatrix}$$

Here the $Q_{sto}$ is the sum of $Q_{sto1}$ and $Q_{sto2}$, and the $k_{empt}(Q_{sto})$ term is a non-linear function with 7 parameters [128].

The rate of appearance of glucose into blood plasma from this model is given by

$$R_a = \frac{f k_{abs} Q_{gut}}{BW} \tag{4.4}$$

The relative complexity of this model in terms of states and parameters makes is suitable for simulation, but not for real-time estimation and prediction purposes unless nearly all parameters are fixed. As discussed in Section 2.5.1 the meal variability is such that one could wonder if it makes sense in meal modelling to have a set of parameters for a person, or if the parameters should be different for each meal. The answer is likely some hybrid, where some aspects of the meal dynamics are considered person related, and some that are considered to be meal-related. The answers to these questions determine how good a prediction one can expect to make with a meal model that only accepts the CHO content of the meal as input.

In the papers by Rozendaal et al. [76], Maas et al. [129] a different model is presented. This model uses fewer states but more parameters, and the meal is modelled by its carbohydrate content and four of these parameters. Tables are provided with identified parameter values per meal for different foods. This model provides an alternative to the model described above, but was not implemented or tested as part of this PhD project.

### 4.3.3 Subcutaneous insulin dynamics

The timing of insulin's effect on blood glucose after an insulin injection is an important parameter in the control of blood glucose, both in artificial pancreas applications and in the manual control performed daily by people with diabetes. For insulin boluses in relation to meals with rapid-acting insulin, the effect should

ideally have an onset as quickly as possible, and the insulin should be cleared from the blood within an hour. For slow-acting insulin the ideal curve is a constant level. The insulin types on the market approximate these ideals to varying degrees, see Fig. 2.7.

The study Nucci and Cobelli [130] provides an overview of different insulin subsystem models, and provides a critical review of their properties. Wilinska et al. [131] test a set of SC models and their ability to describe data from insulin lispro injections. They find that a model with one slow and one fast absorption channel and local insulin degradation fit their data the best.

It seems there are no similar studies as those mentioned above related to slow/long-acting insulin types. However, a similar model as the one used for rapid-acting insulin, but with longer time constant, can be used. See Section 4.4.3 for an example of this.

### 4.3.4 Exercise

The influence of exercise and physical activity is considered a significant difficulty in terms of modelling glucose dynamics, glucose control in artificial pancreas systems and for the person with diabetes [121, 132]. To illustrate the difficulty of predicting the effect of exercise, it is commonly found that in aerobic exercise of moderate intensity, blood glucose generally decreases, while in anaerobic exercise the blood glucose is often found to increase, whereas mixed exercise like ball games can produce both effects [133].

A model by [134] has been formulated to explain the influence of exercise on the glucose dynamics. In this model there is a short-term and a long-term influence of exercise on meal absorption and glucose utilisation. For an example of how this exercise model can be included in a BG dynamic model model, see See Section 4.4.3, states $x_Y$ and $x_Z$.

Another recently published model of physical exercise is Palumbo et al. [135] which includes modelling of many other hormones, and concentrations not normally found in BG dynamic models, including epinephrine, lactate, alanine, glycerol and free fatty acids (FFA) .

## 4.4 Blood Glucose Dynamics models

There are many models available in the literature for describing blood glucose dynamics, spanning from linear, minimal models to highly non-linear models with large numbers of states and parameters. Chee and Fernando [118] provides a good

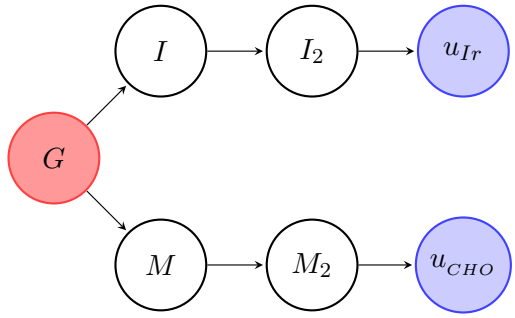



**Figure 4.2:** Inference graph of Magdelaine's 5-state linear model. White and red circles are states. Red states can be measured in a free-living setting. Blue circles are inputs. An edge in the graph from state A to state B signifies that the differential equation governing A has terms containing B, implying that information about B can be inferred by monitoring A [141]. All states are self-referencing, but the loop edges usually signifying this have been removed from the graph to reduce clutter.

overview of some models. Since the focus of this thesis is on models that can be used in a free-living setting, the more complex models are not given much attention here, due to the restricted identifiability of such models given only data from free-living settings. The models considered in this thesis are given in the table below, and are described in more detail in the following subsections.

| Model Name | States | Parameters | Inputs | Reference(s) |
|---|---|---|---|---|
| Magdelaine | 5 | 5 | 2 | [136, 137] |
| IVP | 7 | 10 | 2 | [138] |
| GlucoPred | 14 | 41 | 4 | This thesis |
| Uva-Padova | 18 | 48 | 3 | [62, 139, 34] |

**Table 4.1:** Table of models investigated as part of this PhD work

### 4.4.1 Magdelaine's models

Magdelaine et al. [136] postulated a simple linear model describing long-term dynamics of the glucose-insulin system. The model had 5 states with 9 parameters, of which 3 are set from knowledge of body weight. The model can be further reduced to improve identifiability [137], resulting in 5 parameters. An inference graph of the model is shown in Fig. 4.2. Use of inference graphs in dynamic systems are described in Liu et al. [140].

The reduced 5-state model is:

$$\mathbf{x} = \begin{bmatrix} G \\ I \\ I_2 \\ M \\ M_2 \end{bmatrix} \begin{array}{l} \text{- CGM glucose} \\ \text{- Plasma insulin} \\ \text{- Input compartment insulin} \\ \text{- Meal rate of appearance} \\ \text{- Input compartment meal} \end{array}$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \dot{G} \\ \dot{I} \\ \dot{I}_2 \\ \dot{M} \\ \dot{M}_2 \end{bmatrix} = \begin{bmatrix} \theta_1 - \theta_2 I + \theta_4 M \\ \frac{1}{\theta_3}(I_2 - I) \\ -\frac{1}{\theta_3} I_2 + u_{Ir} \\ \frac{1}{\theta_5}(M_2 - M) \\ -\frac{1}{\theta_5} M_2 + u_{CHO} \end{bmatrix}$$

The vector of unknown parameters is $\boldsymbol{\theta} = [\theta_1 \dots \theta_5]^\mathsf{T}$. Since the parameters have been lumped together they have lost a clear connection to physiology, but $\theta_1$ describes the upward or downward drift in the data in the absence of any inputs, $\theta_2$ describes the effect of insulin on glucose, $\theta_3$ is a time constant for insulin transport from subcutaneous to plasma, $\theta_4$ is the effect of a meal on glucose, and $\theta_5$ is a time constant for meal transport from ingestion to appearance in plasma.

### 4.4.2 Identifiable Virtual Patient (IVP) model

The IVP model as originally formulated [35, 138] has five states ($G_p$,$G_{isf}$ $I_p$, $I_{sc}$, and $I_{eff}$) and an analytic equation describing $R_a$, the rate of appearance of glucose from meals [138]:

$$R_a = \frac{u_{meal}}{V_G \tau_m^2} t e^{-\frac{t}{\tau_m}} \tag{4.5}$$

The $t$ here is a time variable that starts at the time of meal intake. Such a formulation is not practical for simulations containing several meals. Recognising Eq. 4.5 as the solution of a second order linear system, we can use the following two compartment model as an equivalent representation:

$$\dot{Q}_{gut} = \frac{1}{\tau_m}(Q_{sto} - Q_{gut}) \tag{4.6}$$

$$\dot{Q}_{sto} = -\frac{1}{\tau_m} Q_{sto} + u_{meal} \tag{4.7}$$

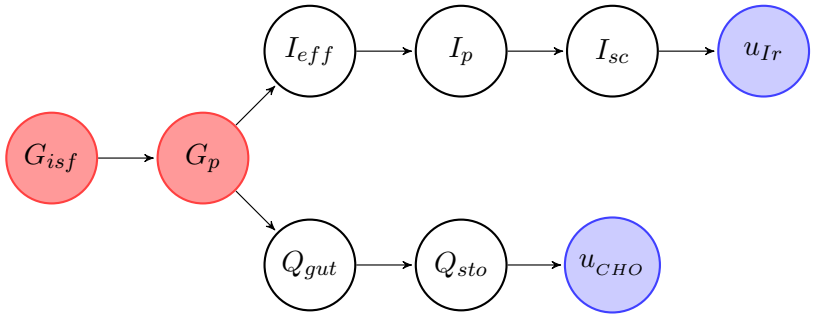



**Figure 4.3:** Inference graph of the IVP model. Legend as in Fig. 4.2

where two new states $Q_{gut}$ and $Q_{sto}$ have been introduced, and

$$R_a = \frac{Q_{gut}}{V_G \tau_m} \tag{4.8}$$

The IVP model is shown as an inference graph in Fig. 4.3 and is given by:

$$\mathbf{x} = \begin{bmatrix} G_p \\ G_{isf} \\ I_p \\ I_{sc} \\ I_{eff} \\ Q_{sto} \\ Q_{gut} \end{bmatrix} \begin{array}{l} \text{- Plasma glucose concentration} \\ \text{- Interstitial (CGM) glucose} \\ \text{- Plasma insulin} \\ \text{- Subcutaneous insulin} \\ \text{- Effective insulin (called X in many texts)} \\ \text{- Stomach compartment glucose content} \\ \text{- Gut compartment glucose content} \end{array}$$

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} \dot{G}_p \\ \dot{G}_{isf} \\ \dot{I}_p \\ \dot{I}_{sc} \\ \dot{I}_{eff} \\ \dot{Q}_{sto} \\ \dot{Q}_{gut} \end{bmatrix} = \begin{bmatrix} (GEZI + I_{eff}) * G_p + EGP + \frac{Q_{gut}}{V_G \tau_m} \\ \frac{1}{\tau_{isf}}(G_p - G_{isf}) \\ \frac{1}{\tau_2}(I_{sc} - I_p) \\ \frac{1}{\tau_1}\left(I_{sc} + 10^6 \cdot \frac{u_{Ir}}{C_I}\right) \\ p_2(S_I I_p - I_{eff}) \\ -\frac{1}{\tau_m} Q_{sto} + 10^3 \cdot u_{CHO} \\ \frac{1}{\tau_m}(Q_{sto} - Q_{gut}) \end{bmatrix}$$

The vector of unknown parameters is

$$\boldsymbol{\theta} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ C_I \\ S_I \\ p_2 \\ V_g \\ \tau_m \\ GEZI \\ EGP \\ \tau_{isf} \end{bmatrix} \begin{array}{l} \text{- Insulin PK time constant 1} \\ \text{- Insulin PK time constant 2} \\ \text{- Insulin clearance} \\ \text{- Insulin sensitivity} \\ \text{- Time constant insulin action} \\ \text{- Distribution volume of glucose} \\ \text{- Time constant of meal carbohydrate transfer to plasma} \\ \text{- Glucose efficiency at zero insulin} \\ \text{- Endogenous glucose production} \\ \text{- Time constant of glucose diffusion from plasma to ISF} \end{array}$$

The IVP is based on the minimal model by Bergman, [61].

Another model based on Bergmans minimal model is the Subcutaneous oral glucose minimal model (SOGMM) [92], described and analysed for identifiability in Garcia-Tirado et al. [91]. Some differences between SOGMM and IVP is that in SOGMM the insulin passes through two compartments before entering plasma, but there is only one parameter governing this transfer, versus the three in IVP ($C_I$, $\tau_1$ and $\tau_2$). The SOGMM model as originally formulated by Patek et al. [92] includes subcutaneous glucose sensing dynamics, while the identifiability analysis done by Garcia-Tirado et al. [91] neglects this extra state and parameter and therefore assumes that the glucose signal from a CGM is a valid proxy for the plasma glucose. The SOGMM model with CGM measurement dynamics has 8 states and 14 parameters, of which 5 are free. The other 9 are fixed, one of them ($G_b$) through knowledge of the patient's HbA1c level. At first glance, the SOGMM seems to be a good alternative to the IVP model for use in the situations that Prediktor Medical are interested in. The model was discovered close to the end of the PhD project, but was analysed to some degree and found to be comparable to the IVP, however with some improved identifiability properties.

### 4.4.3 GlucoPred model

This model was originally formulated by Steinar Sælid in Prediktor Medical, and has been the main focus in this PhD project. The original model also included muscle glycogen, free fatty acids and triglyceride stores. These more obviously non-observable states were removed from the model in a first model reduction step. The reduced version of the model is described in the following. The model includes glucose, insulin, meal, subcutaneous insulin (rapid and slow), exercise, glucagon dynamics, and liver glycogen stores. The model has elements from both the minimal model and the UVa-Padova model, the exercise submodel from [134], and other extensions. Since it contains an endogenous insulin production term it can be used to type 2 diabetes and healthy individuals in addition to type 1 diabetes. An inference graph of the model is shown in Fig. 4.4.

In this model a slightly different notation is adopted. All states start with $x$, all parameters with $p$, constants start with $k$, and symbols starting with other letters are functions. The notation $X^+$ should be taken to mean that the term X is made non-negative, i.e. if $X < 0$, $X^+ = 0$, otherwise $X^+ = X$ . Another notational convention is $X^{\in[a,b]}$ which should be taken to mean that X is limited to be in the range $[a, b]$

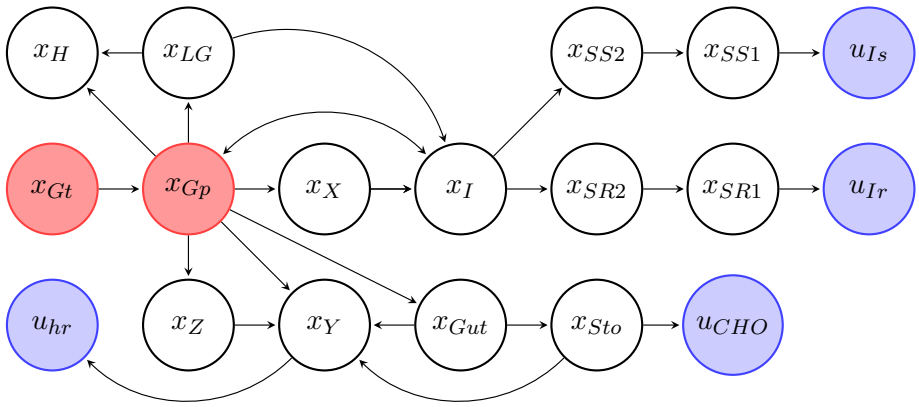



**Figure 4.4:** Inference graph of the GlucoPred model. Legend as in Fig. 4.2

The states are

$$\mathbf{x} = \begin{bmatrix} x_{Gp} \\ x_{Gt} \\ x_I \\ x_X \\ x_{SR1} \\ x_{SR2} \\ x_{SS1} \\ x_{SS2} \\ x_H \\ x_{LG} \\ x_{Sto} \\ x_{Gut} \\ x_Y \\ x_Z \end{bmatrix} \begin{array}{l} \text{- Glucose concentration in plasma [mg/dL]} \\ \text{- Glucose concentration at measurement site [mg/dL]} \\ \text{- Insulin concentration in central compartment [mU/L]} \\ \text{- Insulin concentration in remote compartment [mU/L]} \\ \text{- Rapid insulin concentration in SC compartiment 1 [mU/L]} \\ \text{- Rapid insulin concentration in SC compartment 2 [mU/L]} \\ \text{- Slow insulin concentration in SC compartment [mU/L]} \\ \text{- Slow insulin concentration in SC compartment 2 [mU/L]} \\ \text{- Glucagon concentration in plasma [pg/ml]} \\ \text{- Liver glycogen store [g]} \\ \text{- Glucose in stomach [g]} \\ \text{- Glucose in gut [g]} \\ \text{- Exercise short term effect [unitless]} \\ \text{- Exercise long term effect [unitless]} \end{array} \tag{4.9}$$

The inputs are

$$\mathbf{u} = \begin{bmatrix} u_{Ir} \\ u_{Is} \\ u_{hr} \\ u_{CHO} \end{bmatrix} \begin{array}{l} \text{- Injected rapid insulin [U]} \\ \text{- Injected slow insulin [U]} \\ \text{- Heartrate [beats/min]} \\ \text{- Meal intake, carbohydrates [g]} \end{array} \tag{4.10}$$

The state transition equations are:

$$\dot{x}_{Gp} = G_{prod}(x_H, x_{LG}, x_I) - G_{use}(x_{Gp}, x_X, x_Z, x_Y) + R_a(x_{Gut}) \tag{4.11}$$
$$\dot{x}_{Gt} = p_{kDel}(x_{Gp} - x_{Gt}) \tag{4.12}$$
$$\dot{x}_I = -p_n x_I + I_{endo}(x_{Gp}, \dot{x}_{Gp}) + I_{exo}(x_{SR2}, x_{SS2}) \tag{4.13}$$
$$\dot{x}_X = -p_2(x_X - x_I)^+ \tag{4.14}$$
$$\dot{x}_{SR1} = -\frac{x_{SR1}}{T_d} + u_{Ir} \tag{4.15}$$
$$\dot{x}_{SR2} = \frac{x_{SR1} - x_{SR2}}{T_d} \tag{4.16}$$
$$\dot{x}_{SS1} = -\frac{x_{SS1}}{T_{ds}} + u_{Is} \tag{4.17}$$
$$\dot{x}_{SS2} = \frac{x_{SS1} - x_{SS2}}{T_{ds}} \tag{4.18}$$
$$\dot{x}_H = p_n(S_H - x_H) \tag{4.19}$$
$$\dot{x}_{LG} = -10^{-3} Q_{liver}(x_H, x_{LG}, x_I) \tag{4.20}$$
$$\dot{x}_{Sto} = -f_{abs}(x_Y) x_{Sto} + u_{CHO} \tag{4.21}$$
$$\dot{x}_{Gut} = -f_{abs}(x_Y)(x_{Gut} - x_{Sto}) \tag{4.22}$$
$$\dot{x}_Y = \frac{1}{p_{TY}}(f_{hr}(u_{hr}) - x_Y) \tag{4.23}$$
$$\dot{x}_Z = -\frac{1}{p_{Tmax}} x_Z + f_Y(x_Y)(1 - x_Z) \tag{4.24}$$

The functions embedded in the above equations are given in the following.

The glucose production is given by

$$G_{prod} = \frac{1}{p_{Vg}} \left(p_{Eneo} + Q_{liver}(x_H, x_{LG}, x_I) + p_{Gneo} p_{kHL} x_H\right) \tag{4.25}$$

where the liver net production is given by

$$Q_{liver} = (1 - p_{Gneo}) p_{kHL} x_H \frac{x_{LG}}{p_{MlMx}} - p_{kIL} f_{kM}(x_{LG}) \frac{x_I}{x_I + p_{Ihalf}} \tag{4.26}$$

$f_{kM}$ is a reduction factor when liver glucagon stores are full, given by

$$f_{kM} = \left(1 - p_{sComp}\left(\frac{x_{LG}}{p_{MlMx}} - 1\right)\right)^+ \tag{4.27}$$

The glucose usage is given by

$$G_{use} = p_{Uii} + p_{kgb}x_{Gp} + f_{idep}(x_X, x_Y, x_Z)\frac{x_{Gp}}{p_{Kmg} + x_{Gp}} \tag{4.28}$$

where the insulin dependent, exercise modified utilization is given by

$$f_{idep} = p_{Si}x_X + p_{kgm} + p_{kglg} + p_\alpha x_X x_Z + p_\beta x_Y \tag{4.29}$$

The rate of appearance of a meal is given by

$$R_a = \frac{10^3 * p_f * f_{abs}(x_Y) * x_{Gut}}{p_{Vg}} \tag{4.30}$$

The absorption function in the equation for $R_a$, $x_{Sto}$ and $x_{Gut}$ is given by

$$f_{abs} = (p_{mealGI}p_{kabs}(1 - p_r x_Y))^{\in[0,1]} \tag{4.31}$$

and contain a parameter $p_{mealGI}$ between 0 and 1 that is related to the glycemic index of the last meal (set to 1 if unknown). Note that it also depends on recent exercise.

The endogenous insulin production is given by

$$I_{endo} = p_n \left( p_{Rm}\dot{x}^+_{Gp} + \frac{p_{Ri}}{1 + \left(\frac{p_{GI}}{x_{Gp}}\right)^{p_{ni}}} \right) \tag{4.32}$$

The contribution to plasma insulin from exogenous insulin input is

$$I_{exo} = \frac{10^3}{p_{Vi}} \left( \frac{x_{SR2}}{p_{Td}} + \frac{x_{SS2}}{p_{Tds}} \right) \tag{4.33}$$

The production of glucagon is given by

$$S_H = S_{H1} + S_{H2} \tag{4.34}$$

$$S_{H1} = \left( p_{Rhbas} + (p_{Rhmax} - p_{Rhbas}) \left( 1 - \frac{1}{1 + \left(\frac{p_{GH}}{x_{Gp}}\right)^{p_{nh}}} \right) \right)^+ \tag{4.35}$$

$$S_{H2} = \left( \frac{p_{kDia}}{1 - (p_{aP}(p_{I0} - x_I))^{\in[0,k_{IHmax}]}} \left( \frac{x_{Gp}}{p_{GH}} \right)^2 \right)^{\in[0,k_{SHmax}]} \tag{4.36}$$

In the exercise submodel the expressions are given by the following:

$$f_{hr} = \left( \frac{u_{hr} - p_{HRB}}{p_{HRM} - p_{HRB}} \right)^{\in[0,1]} \tag{4.37}$$

$$f_Y = \frac{(x_Y k_a)^{k_{nY}}}{1 + (x_Y k_a)^{k_{nY}}} \tag{4.38}$$

The parameters for the Prediktor model are based on the models it has been derived from, namely the minimal model [61], the UVa-Padova model [62] and the exercise model from [134].

### 4.4.4 UVa-Padova model

The UVa-Padova model is perhaps the most well known model in the area of glucose-insulin dynamics modelling. The model equations have been described in many articles [62, 134, 139, 142, 143] and books [144] and are not repeated here. The model accepts meals in the form of a CHO value, units of insulin, and glucagon injections as inputs. The model was originally postulated based on data from 204 healthy individuals, and later modified to be used for type 1 diabetes [120]. An inference graph of the UVa-Padova model is shown in Fig. 4.5.

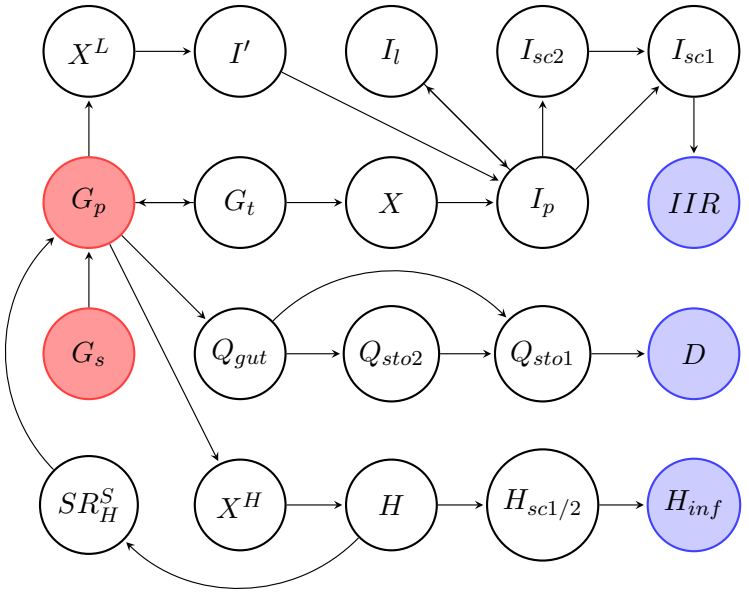



**Figure 4.5:** Inference graph of the UVa-Padova model. Legend as in Fig. 4.2

## 4.5 Other blood glucose modelling and prediction strategies

While this thesis has focused on state space models of glucose-insulin dynamics, there are many other approaches to glucose prediction to be found in the literature. Since this thesis has been focused on state-space models I do not describe these other types of models in more detail. The interested reader can find a good overview of BG prediction strategies in the review article by Oviedo et al. [145]. A book on the subject is also available [146].

## 4.6 Model selection

When selecting a model to be used as an individualisable blood glucose dynamics model, the choice of model structure is important. In addition, the model needs to be parameterised, meaning that we need to select a subset of parameters that are to be considered free, and individualisable through parameter estimation, and some that shall be fixed and not estimated. This parameterisation determines which dynamics the model will be able to fit. As such, model selection can be said to encompass the choice of a model structure *and* a parameterisation. I.e. the model selection should specify which parameters that should be considered free and which should be made constant, sub-population dependent or computed from other patient information (HbA1c level, body mass, BMI etc). This selection of parameters should be guided by observability and identifiability analyses, which is treated in detail in the next chapter.

Model selection often consists of finding the right trade off between ability to fit data and model parsimony, and of course predictive ability. This is a general problem, and different criteria have been developed to rank models on how they perform. Akaike's Information Criterion (AIC) is given by

$$AIC = 2N_p - 2ln(L) \tag{4.39}$$

here $N_p$ is the number of (free) parameters in the model, and $L$ is a measure of the fit. Originally $L$ is the likelihood of the measurements given the model, which is often replaced by the sum of squared prediction errors (see Section 5.3.2). The Bayesian Information Criterium (BIC) is very similar:

$$BIC = ln(N_y)N_p - 2ln(L) \tag{4.40}$$

here the number of data points $N_y$ is also taken into account. The lower AIC and BIC a model has, the better .

# Chapter 5

# Identification

Blood glucose dynamics is variable on different levels: between different populations (healthy, prediabetes, type 2 diabetes, type 1 diabetes), between individuals in each population, and intra-individual variation over time is also present.

Thus, most blood glucose dynamics models need to be individualised, by finding the model structure and parameterisation that is most capable of describing or predicting glucose data from a specific patient. This task is also known as system identification. If the model structure is known and only the parameter set is unknown it is referred to as parameter identification or parameter estimation.

The system identification task is traditionally broken down into the following separate tasks [147]:

1. Gather data containing inputs and outputs from the system, and if possible, measurements of initial states
2. Select a suitable model structure
3. Estimate the parameters of the model

   - This frequently also entails selecting which parameters from the full set to estimate, and fixing the rest.
4. Validate the identified model on an independent data set that was not used for model selection/identification

The following sections discuss these steps in the context of BG dynamic model identification for individuals based on data obtainable in a free-living context.

## 5.1 Data collection

How glucose and meal/insulin data are collected is crucial to the success of BG dynamic model parameter estimation. The experiment design is important in this respect, to ensure that the information gathered provides the most information possible about the model parameters.

There are many aspects to consider when conducting a clinical study to look at BG dynamics:

1. Duration of the identification experiment
2. Timing and magnitude of each controllable input (meals, insulin)
3. Timing, frequency and type of measurements:
    - Glucose (CGM, SMBG, YSI)
    - Insulin
    - Glucagon
4. Amount of restriction of the patient during the experiment (lie down / sit down / allowed to walk around / free-living / doing exercise)
5. Measurement of any measurable disturbances
6. Logging of unintended events and un-measurable, but detectable disturbances
7. Recording validation data sets in addition to the training / identification data

### 5.1.1 Experiment design

In system identification theory there is the concept of *persistently exciting inputs*, which are a necessity for successful identification of parameters. Ljung [147] describes how such inputs can be characterised and constructed to be matched to the model that we wish to identify parameters in. A rule of thumb is that the input signal must contain at least as many frequencies as there are parameters to identify.

While such engineered perturbations are often possible to apply without risk in technical systems, in a BG dynamic model the real system being "perturbed" is a human being and it could be dangerous and unethical to apply the inputs in terms of meals and insulin that would be ideal for parameter estimation. In other words, the experiment design for BG dynamic model parameter estimation has constraints given by the risk the patient is subjected to during the identification experiment.

Another common constraint in research settings is the length of the experiment in time, since the research clinic visit in which the identification experiment is

performed is of limited duration. The glucose-insulin dynamics is relatively slow, and it typically takes several hours to record the glucose dynamics triggered by one meal or insulin bolus. These conditions put limits on how much information that can be gathered about a person with regards to BG dynamic model parameters in clinical settings.

Data from free-living settings from people with diabetes that are using a CGM and that reliably log meals and insulin injections (or insulin input data from an insulin pump, if that is used) can be used for identification [148, 136, 149], and provides a path to using longer-term data for identification of the model parameters. For motivated individuals it is not unrealistic to achieve a month of logging of such data. However, although the time span of identification data sets can be dramatically improved in data from free-living settings compared to clinical research data, the accuracy in the data is generally lower, mostly due to incomplete or inaccurate meal and insulin logging. Another point is that in clinical research we have more freedom to alter the inputs in order to improve the information content in the data, while in data from free-living settings we have to make do with what we get. A third difference is that in data from free-living settings adverse events like hypoglycemia or prolonged periods of hyperglycemia occur naturally, while clinical research experiment protocols often go to great lengths to avoid such situations.

### 5.1.2 Glucose Tolerance Tests

There are several standard protocols in use to determine glucose dynamics, some of which are mentioned here.

The Oral Glucose Tolerance Test (OGTT) consists of ingesting a 75g dose of pure glucose, after fasting for the last 8 hours. In the most common variant of the test, BG is measured before the dose is taken (fasting BG) and 2 hours after glucose intake. The purpose is to determine the diabetic state of the individual. Fasting BGL should be less than 6.1 mmol/L, and BG at the two-hour point should be less than 7.8 mmol/L to be considered normal. Repeated measurements of fasting BG above 7.0, or BG above 11.1 at the 2-hour point are diagnostic of diabetes.

The reproducibility of the OGTT has been assessed by several investigators [150, 151], and found to be somewhat lacking, and some argue that its use should be discontinued [152]. Munang et al. [153] find that 1 in 4 do not get the same gestational diabetes diagnosis on a repeat OGTT as in the first, when these are repeated within a week of each other. [154] tested the OGTT on 60 overweight children twice, with less than a month between sessions, and found that from the 60 OGTTs performed in each session, 10 showed Impaired Glucose Tolerance (IGT) in the first session, and 12 had IGT in the second session, but only 3 subjects got

the same IGT result in both tests, indicating that the reproducibility of the test is questionable.

The Mixed-meal tolerance test (MTT) is also in use, where a more realistic meal is consumed instead of using pure glucose. The meal composition does not seem to be standardised, but it may consist of eating a mixed meal of 1 g glucose per kg body weight, 2 scrambled eggs and pieces of bacon as reported in Dalla Man et al. [155]. Shankar et al. [156] also use the MTT.

The Intravenous glucose tolerance test (IVGTT) is more invasive since the glucose is administered directly into the blood stream. The glucose and insulin curves resulting from an IVGTT is similar to the impulse response from a first order system, see e.g. Cobelli et al. [120].

### 5.1.3 Data from free-living use

Compared to the clinical research settings described above, data from free-living settings are considerably less structured and complete, since they rely on the user to remember to estimate meals and log them. On the plus side, data from a CGM and insulin pump user can span months and years. Common problems in CGM data from free-living settings is shown in Fig. 5.1, where some of the data from the OhioT1DM data set is shown [157].

An important task when using data from free-living settings for model estimation is to clean the data. This data cleaning should ideally be automatic and objective. Some cleaning tasks in free-living glucose data sets are:

- Remove glucose measurement outliers
- Correct measurement errors (PISAs, missing data)
- Handle jumps related to calibration
- Detect missing meals
- Evaluate correctness of reported meal value or time

One of the articles published as part of this PhD project deals with glucose data smoothing, interpolation and uncertainty estimation, for this purpose. See A.1. The article describes a method that uses a Kalman smoother to preprocess glucose data, providing interpolation, smoothing and outlier detection. The method also estimates the uncertainty in the data. An illustration of the method is shown in Fig. 5.2 and its use on a problematic data set is shown in Fig. 5.3.

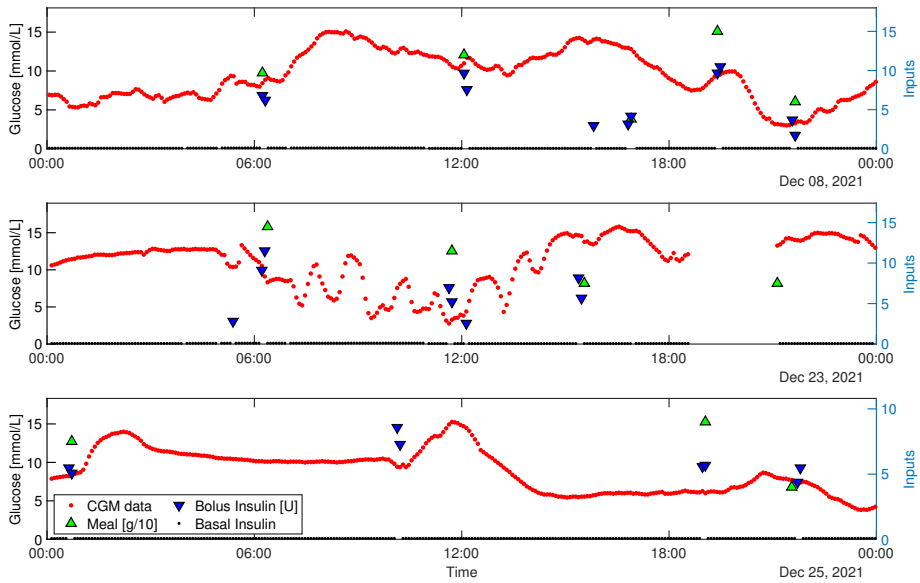



**Figure 5.1:** Examples of commonly encountered issues in CGM data. From the OhioT1DM data set, participant 570. *Upper panel:* OK data set with seemingly all inputs registered. *Middle panel:* Oscillatory data from 6 to 16 o'clock, could be real glucose fluctuations due to some unlogged events or could be PISAs. This data set also exhibits a period of missing data after 18 o'clock that could be a sensor change. *Lower panel:* A probable case of missed meal registration around lunchtime. Only insulin boluses are logged, but the glucose increases.

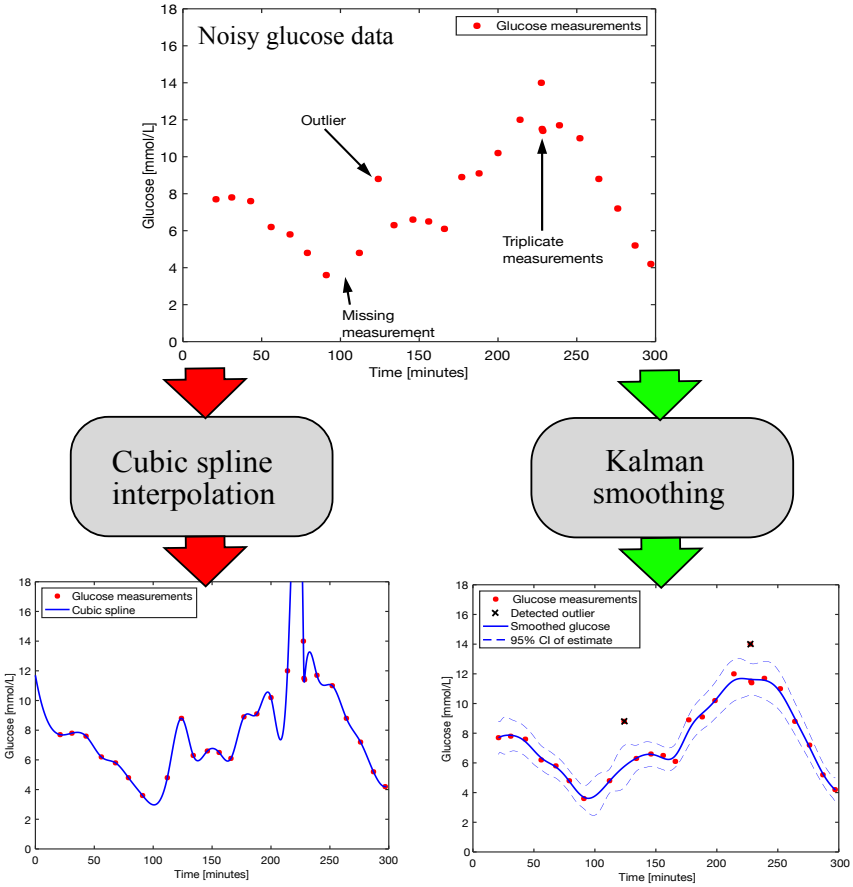



**Figure 5.2:** An illustration of the method for cleaning of glucose data described in article A.1. Compared to other methods for interpolation (e.g. cubic spline), a glucose specific model running in a Kalman smoother enables outlier detection and uncertainty estimation from glucose data.

## 5.2 Identify model structure

The question of choosing the correct model among those mentioned in Chap. 4 depends heavily on the intended use of the model, available measurements, and available data about inputs and disturbances.

Many of the models that have been developed for describing glucose dynamics assume that insulin measurements are available, such measurements are not available in data from free-living settings. To ensure observability of states and identifiability of parameters when the number of measurements are reduced, the model complexity

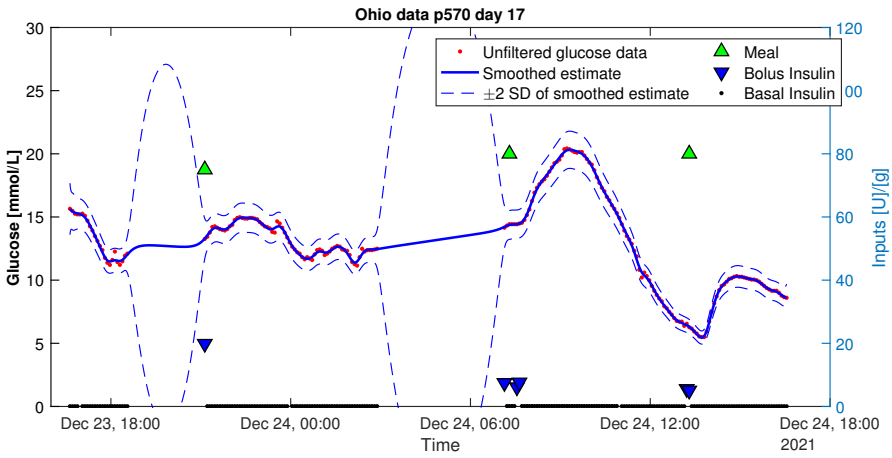



**Figure 5.3:** The smoother used on a 24-hour portion of data from the Ohio T1DM data set, where there are some periods of missing data. The uncertainty band shown by the dashed line increases when there are missing data, and envelope the likely paths glucose can have taken through these regions. This uncertainty band can be used to evaluate the usability of the data set as a whole, or used as the measurement variance in downstream processing.

must likely be reduced. Paper B.1 discusses this in more detail, discussing how sensitivity analysis and singular value decomposition can be used to investigate parameter identifiability and guide model reduction.

## 5.3 Parameter estimation

In essence, the parameter estimation task is to select a parameter vector $\widehat{\boldsymbol{\theta}}$ that makes our model best describe a set of data $\boldsymbol{Y}$.

This is achieved by minimisation of some function $L$ that describes the goodness of fit to the data

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, L(\boldsymbol{\theta}, \boldsymbol{Y}) \tag{5.1}$$

The task of estimating parameters in a BG dynamic model can be broken down as follows:

- Prerequisite: data from the patient is available in the form of glucose measurements, insulin injection values and meal information, and possibly other inputs.
- Prerequisite: A BG dynamic model model has been chosen
- Determine which parameters of the model to estimate
- Decide on objective function $L$ to optimize, here there are many options, e.g.

  - should $L$ be based on the glucose level or the glucose rate or a combination?

  - should measurements far from the basal level be weighted higher?

  - should we make allowances for slightly missed dynamics (i.e. "approving" model predictions that reach the right levels, but not exactly at the correct time)?
- Determine which algorithm to use for the numerical optimisation
- Perform optimisation
- Validate parameters

Initial states are often unknown, and if so, also need to be estimated. The estimates of initial states will depend on the model parameters. To illustrate this, imagine that a downward trend is observed in the beginning of a glucose data set. This can be taken to mean that the patient had a certain initial insulin level. However, if the insulin sensitivity parameter is increased, the initial insulin level needs to be decreased to keep the fit good. This indicates that the initial insulin level and the insulin sensitivity is not individually identifiable from only the initial data. When more data is considered it can provide a more accurate estimate of the insulin

sensitivity, which in turns enables more accurate estimation of the insulin initial state.

A similar case happens for meal estimation in a data cleaning setting, where we would like to estimate the most likely time and value of a suspected missing meal registration. This can be based on a BG dynamic model model, but the parameters of the model will determine the magnitude of any detected meal in CHO.

Conversely, during model parameter estimation the achieved parameter values depend on the meal (and other input) values being accurate. If a meal has been incorrectly estimated or forgotten altogether, this can have an impact on what values the parameters get from the parameter estimation.

Thus the problem bites its own tail: The initial states and meal estimation depends on the model parameters, and the model parameters depend on the inputs and the initial states. One approach to solving this issue is to apply an iterative process of initial state, meal and parameter estimation, that starts out with a nominal parameter set, identifies initial states, then uses this model to find missing meals meals, then estimated model parameters. From this refined model another iteration can be done, iterating until convergence is achieved. This suggested iterative procedure is illustrated in Fig. 5.4. Adding meals and initial states to be estimated is equivalent to adding new parameters to the model, and it decreases overall identifiability.

### 5.3.1 Parameter space reduction

Most BG dynamic models are not identifiable from CGM data alone, at least not the full parameter set. The identifiability generally decreases as the number of parameters to be estimated increases, and selecting a parsimonious model is a necessary first step to improve identifiability.

Fixation of parameters is often needed. This consists of setting the most known, and/or the least identifiable parameter(s) and/or the least variable parameter to constant values. The value can be based on population averages, and sometimes it makes sense to use different parameter values for different sub-populations of virtual patients (DM1, DM2, Healthy).

Some parameters can be set to a fixed value based on other patient data. For instance, the parameter $V_g$, which is often present in BG dynamic model meal submodels, describing the glucose distribution volume, has a correlation with the body mass, so setting the value of this parameter based on the body mass is sensible. Other parameter values may be predictable from e.g. BMI, HbA1c level, age or duration of diabetes. However, such relationships are often not known *a priori* and must be established after data collection.

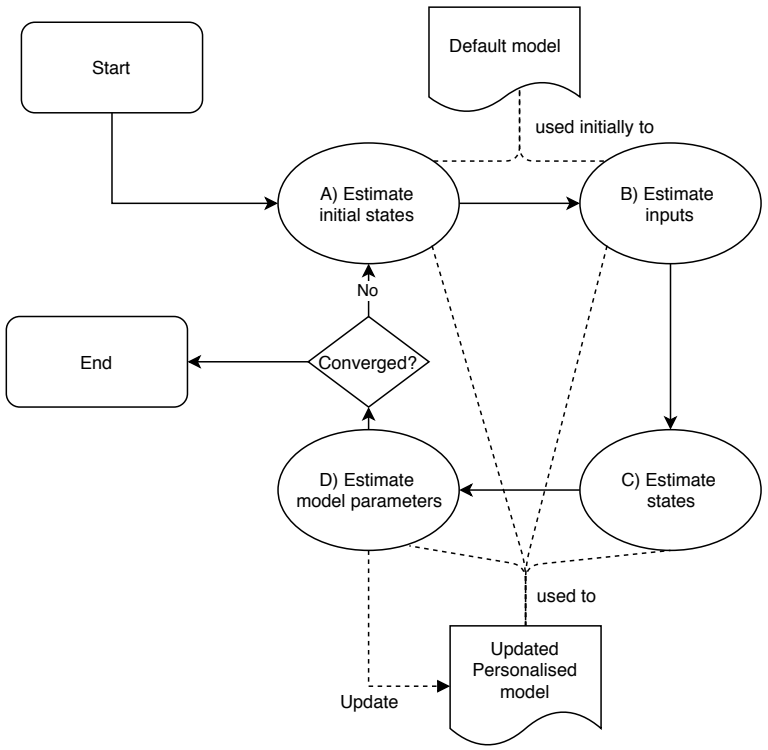



**Figure 5.4:** Suggested iterative approach to estimating initial states, missing meals and parameters in BG dynamic model models

### 5.3.2 Simulation based cost function

The parameter estimation problem is an optimisation problem where the cost function is some function of the sum of the squared errors between the measurements and the predicted values from the model. If the prediction errors are defined to be the differences between the measurements and a simulation starting at some initial state $\mathbf{x}_0$, we can use the Root Mean Squared Error (RMSE) as a cost function:

$$\textbf{RMSE} = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(y_i - \widehat{y}_i)^2} \tag{5.2}$$

where $y_i$ is the measurement at time step $i$ and $\widehat{y}_i$ is the corresponding prediction from the simulation at this time. Another commonly seen cost function weigh each squared error by the uncertainty of the measurement:

$$\textbf{Weighted RMSE} = \sqrt{\frac{1}{N}\sum_{i=0}^{N}\frac{(y_i - \widehat{y}_i)^2}{\sigma_i^2}} \tag{5.3}$$

here $\sigma_i^2$ is the variance of the measurement at time $i$. Minimisation of the RMSE is the same as minimisation of the sum of squared errors (SSE)

$$\textbf{Weighted SSE} = \sum_{i=0}^{N}\frac{(y_i - \widehat{y}_i)^2}{\sigma_i^2} \tag{5.4}$$

Minimisation of this cost function gives the same as the maximum likelihood estimate (ML) when the prediction error $y_i - \widehat{y}_i$ is assumed to be Gaussian.

Del Favero et al. [158] introduced the Glucose Specific Root Mean Square Error (gRMSE). This measure is based on regular RMSE, but has an added penalty function to penalize errors that are more dangerous in a clinical sense, inspired by the Clarke Error Grid [159]. The penalty function gives a high penalty for overestimating glucose when the true glucose is low, moderate penalty for underestimating when true glucose is high, and no penalty for the other errors, since these have less clinical importance. It is implemented with sigmoid functions parameterised to activate at 85 mg/dL and 155 mg/dL for the low and high glucose range, respectively.

Other cost functions that can be used are FIT [147]:

$$FIT = 100 \times \left(1 - \frac{|y - \widehat{y}|}{|y - \bar{y}|}\right) \tag{5.5}$$

which is closely related to percentage of variance accounted for:

$$VAF = 100\% \times \left(1 - \frac{||y - \widehat{y}||}{||y - \bar{y}||}\right) \tag{5.6}$$

A simulation based cost function is fine when looking at synthetic data generated from the same model as the one we are trying to fit, with no noise. This is of course not realistic, but as described in section 5.6.3 using such data can be a useful exercise to investigate identifiability and to test the parameter estimation algorithm. When faced with real data that is noisy, have missing elements, or contain other effects than what is modelled, this approach is not so robust. To see why, consider a parameter estimation based on 24 hours of free living, and assume that the data fits well to a model for the last 23 hours, but in the beginning of the data set some effect happens that is not modelled, that gives an increase in the glucose that the model cannot explain. If this effect is large enough to give large residuals between measurements and model predictions for a significant part of the data set, it is likely that the parameter estimation for other and unrelated parameters will be adjusted to try to compensate for the un-modelled effect. Often such effects can cause parameter estimation fails completely, resulting in parameter values that make the predicted glucose lie at the mean of the data set used for identification. In glucose data the un-modelled effect that we speak of here may be caused by sudden exertions, stressful experiences, etc. Another commonly encountered practical problem is that the data set does not contain information about all the meals that were really ingested. In such cases, a cost function that is based on the innovations from a Kalman filter can give better results. This is illustrated in Fig. 5.5. Kalman filters are described in more detail in Section 6.1, and filtering based cost functions are described in the next section.

### 5.3.3 Filtering based cost function

In sequential estimation the cost function is based on the one-step-ahead prediction error, i.e. the error of our prediction of measurement $y_{k+1}$ given all the measurements up to $y_k$. This one-measurement-ahead prediction is a function of $\theta$ and is denoted $\bar{y}_{k+1|k}(\theta)$ in the following. When the estimation is done in a Kalman filter, the cost function that we want to minimise becomes the sum of squares of the innovations in the Kalman filter:

$$L(\theta) = \sum_{k=0}^{N} (y_i - \bar{y}_{k+1|k}(\theta))^2 = \sum_{k=0}^{N} (y_i - h(\bar{x}_{k+1|k}(\theta)))^2 = \sum_{k=0}^{N} \varepsilon(\theta)^2 \tag{5.7}$$

where $\varepsilon(\theta)$ is the innovation of a Kalman filter running a model with parameter set $\theta$.

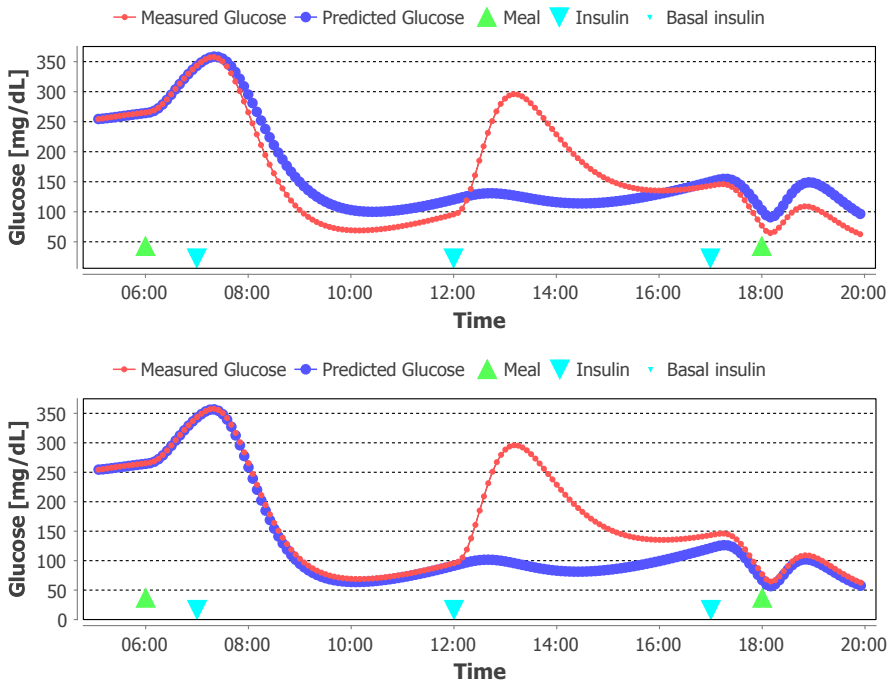



**Figure 5.5:** An illustration of the difference of using simulation based and filtering based cost functions. The data set is generated from the IVP model, using a nominal parameter set, three meals and three insulin boluses, and simulated measurements are taken at 5 minute sampling interval, without added noise, shown as the red points in the graphs. Then the information about the lunch meal at noon is removed from the identification data in order to emulate that the user forgot to log the meal. The glucose data and the manipulated input data are used for parameter estimation. Only the insulin sensitivity parameter $S_I$ is estimated here, other parameters and initial states are assumed to be known. The upper graph shows the results of using a simulation based cost function, while the lower graph shows the result of a filtering based cost function. The blue points are simulations with the resulting $S_I$ parameter. We see that in the first case the parameter estimation has decided on a trajectory that tries to avoid some of the penalty in the region of the missing meal by selecting a parameter value that makes the graph move towards the mean level of the data. This results in a value for $S_I$ that is 66% of the true value. The filtering based cost function produces a better result, as the predicted curve fits the data except where the meal information is missing. The estimated value for $S_I$ is within 10% of the true value with the filtering based cost function.

The gradient of this cost function with respect to the parameters is:

$$\frac{dL}{d\theta} = \sum_{k=0}^{N} \frac{d}{d\theta}\varepsilon(\theta)^2 = \sum_{k=0}^{N} 2\varepsilon(\theta)\frac{d}{d\theta}\varepsilon(\theta) = -2\sum_{k=0}^{N} \varepsilon(\theta)\frac{d}{d\theta}\hat{y}_{k+1|k}(\theta) \tag{5.8}$$

The last term $\frac{d}{d\theta}\widehat{y}_{k+1|k}(\theta)$ is the derivative of the predicted output with respect to the parameters, i.e. the parameter-to-output sensitivity discussed in 5.6. Thus, a Kalman filter that computes the sensitivity matrix during the filtering process can be used to compute the gradient, which can be used in the search for the optimum in the parameter estimation problem.

### 5.3.4 Optimisation problem

Once the parameter estimation problem has been cast as an optimisation problem it needs to be optimised using some numerical optimisation algorithm.

If parameters are highly correlated, the parameter estimation can fail. Such correlations can be detected by the sensitivity analysis with singular value decomposition as discussed in 5.6.1. Thus, any highly correlated pairs or sets of parameters must be eliminated from the model through reduction or parameter fixation before estimation is attempted.

To improve the convergence rate and ability of the parameter estimation method to find the true minimum, we can use Newton-Raphson's algorithm. This method alters the direction of the search by multiplying by the inverse of the Hessian matrix. The Hessian matrix can be approximated by the Fisher Information Matrix (FIM), which will be described in Eq. 5.11. The FIM can also be computed as part of a Kalman filter pass through the data. This method has similarities with natural gradient estimation [160, 161].

Another way of estimating parameters in a sequential way that is also suitable for real-time use is to use a Kalman filter where the state vector has been augmented with the vector of parameters to be estimated [162]. More on this in Section 6.5.

## 5.4 Input estimation for data cleaning purposes

It is normally the case in dynamic systems that the input $\mathbf{u}$ is a *control* input, implying it is something we have full control over and that can be measured with a high degree of accuracy. As discussed previously, in a BG dynamic model this is not the case, at least not in data from free-living settings. Meals are more like disturbances to the system than control inputs, even though in most of the literature about BG dynamic models they are treated as deterministic and perfectly known both in time and amount.

In Section 2.5.1 we discussed how simplifying a meal to only its CHO amount is an oversimplification. This fact, and the fact that most people with diabetes have difficulties to correctly estimate this amount, and often forget it entirely, is a strong

indication that meal information may be the most relevant to identify.

Meal times and values can be estimated analogous to parameter estimation but with meal times and values as the variables to be optimised. In the case of cleaning data for BG dynamic model parameter estimation, the parameter vector to be optimised can be expanded to include elements of the scenario. This of course increases the number of free parameters and could be detrimental to identifiability.

The above approach assumes that we have a batch of data to estimate the meal from, which is the case in offline parameter estimation settings. The problem of meal detection and estimation is also of interest in online settings, most notably to detect unannounced meals in an artificial pancreas context [163, 164], since being able to detect that a meal has been taken is important to achieve good post-prandial glucose control.

As part of this PhD project a meal estimator that is based on Kalman filtering and hypothesis testing for input estimation based on a Chan-Hu-Plant algorithm [165] was investigated. The method is still in a preliminary stage, but has the potential to estimate the most likely size and time of meal input. The method is further described in paper B.2.

### 5.4.1 Model validation

The last item on the list of system identification tasks is model validation. Model validation is typically done by evaluating the model's performance on a separate data set that it should be able to predict, and that does not contain any data used to identify the model. For BG dynamic models, this often means data for the same person, but from a different time span than that used to identify parameters, e.g. data from the day after. The validation may consist of achieving a good enough prediction of the validation data, e.g. as measured by MARD or by 15 or 30 minute prediction errors.

## 5.5 Identifiability

A definition of identifiability of a parameter is given by Raue et al. [166], and is repeated here: A parameter $\theta_i$ is *structurally identifiable* if its estimate is a unique minimiser of a cost function that describes the goodness of fit to data, often a likelihood function. It is *practically identifiable* if the confidence interval of its estimate has finite size.

Judging by the amount of literature on the topic of identifiability it is far from straightforward, neither does it seem to have reached a consensus on which methods

are most appropriate to use [167, 168, 169, 166, 170, 171, 172, 91, 173, 174].

The consequences of non-identifiability are several:

- Identified parameter values are non-unique, implying that small perturbations of the input data or measurements can lead to significantly altered parameters
- Tracking parameters over time, i.e. longitudinal studies, does not make sense
- Low predictive value in the identified model
- Problematic parameter estimation optimisation problem, will more easily get stuck
- Wrong conclusions drawn from experiments, see e.g. Alarid-Escudero et al. [175].

### 5.5.1 Observability vs identifiability

Observability is related to the internal states and the outputs of a system. If you can determine the value of the initial state vector $\mathbf{x}_0$ by observing the inputs and the outputs of a system from a start time $t_0$ to an end time $t_N$, the system is observable. A common observability test in linear systems is to calculate the observability matrix and compute its rank, this is called the Kalman Rank Criterion. In nonlinear systems a similar test exists, called the nonlinear Observability Rank Criterion, that relies on computing a matrix of Lie derivatives.

Identifiability is a property relating to the parameters of the system. Since any system with constant parameters can be transformed into an equivalent system where the state vector has been augmented with all the uncertain parameters identifiability can be considered a special case of observability [167]. Conversely, the initial states are often included as part of the parameter set to be estimated, so the distinction between observability and identifiability is not a clear one.

### 5.5.2 Structural and practical identifiability

We divide identifiability into structural identifiability (SI) and practical identifiability (PI).

Structural non-identifiability occurs when there are structural problems with the model that makes identification impossible. Examples are the use of products, ratios and sums of parameters in such a way that the parameters are not individually identifiable. A trivial example to illustrate this is the system $\dot{x}_1 = -(a+b)x_1, y1 = x_1$. Here, $a$ and $b$ cannot be independently determined, only their sum.

Structural non-identifiability cannot be remedied by more frequent or more accurate sampling, it can only be eliminated through model restructuring. This may consist of adding one or more measurements, or reducing the number of parameters.

Chis et al. [176] provide a good overview of different methods for analysing structural identifiability. A common complaint about these methods is that they become intractable when the dimension of the problem becomes high. Another is that the results are hard to interpret.

Structural identifiability is a necessary condition for being able to identify parameters of a model, but it is not sufficient. There are many ways that the parameter estimation may fail even for structurally identifiable models: too infrequent sampling, too inaccurate sampling, insufficient or improper activation of the system by inputs; all of these and others can lead to little or no information in the output data about the parameters of interest.

Network-based analysis of dynamic systems has recently been introduced and can provide useful results in large nonlinear dynamic systems [177, 140] however it has been criticised for not finding all non-observability issues with models [178]. A component of this method is to construct the inference graph for the model from the differential equations. This is a useful step to get a first overview of a model, and to see if there are states that cannot be reached from the measured states. Some inference graphs are shown in the presented models in Section 4.4.

### 5.5.3 Practical identifiability methods

In the problem of identification of BG dynamic models based on data from free-living settings, we have CGM data, collected at 5 minute sampling with a certain measurement accuracy. Meal and insulin inputs are logged by the user to his or her insulin pump or CGM. we would like to know which parameters we can identify based on the data that we have. This question can be answered through practical identifiability (PI) methods. PI methods will also detect structural identifiability issues, and is arguably more straightforward to do. Therefore this thesis has focused more on practical identifiability than on structural identifiability.

One method for practical identifiability described by Raue et al. [169, 166] is capable of assessing both structural and practical identifiability. The method depends on the profile likelihood, which can be computed around for each parameter individually. The profile likelihood is the minimum of the objective function optimised over all parameters except the one of interest. The profile likelihood is explored around a parameter vector of interest. If the profile is flat, this indicates that the parameter is structurally non-identifiable. The profile likelihood can also be used to determine

the confidence interval of each parameter.

Another method is well explained in Stigter et al. [168]. The method consists of performing singular value decomposition of sensitivity profiles to determine which parameters there is any information about in a set of data. The method will be discussed in detail after an introduction to sensitivity analysis and the Fisher Information Matrix.

## 5.6 Sensitivity analysis

Sensitivity analysis is an important tool with regards to determining which parameters are possible to estimate based on a given input scenario and measurements. The parameter-to-output sensitivity can be computed for any dynamic system, a set of inputs and a set of measurements. The system is simulated from an initial state and using a nominal parameter vector, and the sensitivity trajectories are computed, describing how much the output would change at each point in time due to a change in a parameter. This is intuitively related to parameter identifiability; if the output does not change when one of the parameters changes, the parameter is not identifiable based on the data. Said differently, if we get sensitivity trajectories that are identical to zero, the data contained no information about the parameter. Such zero sensitivities can result if the model suffers from structural non-identifiability, e.g. that some parts of the system is not observable from the measurements we have. It may also result if the input scenario does not excite all parts of the system. An example of this could be a BG dynamic model that contains an exercise sub-model, but the input data contains no heart-rate data. Then all parameters related to the exercise model will show up as having zero sensitivity.

Similarly, if the output has a very slight parameter-to-output sensitivity this can be judged against the accuracy of the measurements. For example, if the glucose measurement device used to generate our data have a precision of $\pm 0.2\ \mathrm{mmol/L}$, then sensitivities that never reach this level point to parameters that are practically unidentifiable or difficult to identify.

A third case detectable by sensitivity analysis is correlated parameters, meaning that two or more parameters are highly correlated. When this happens, a change in output due to one parameter's change can be close to perfectly counteracted by a change in another parameter. The cause of such correlations could be structural identifiability issues, but this can also be caused by unfortunate correlations in the input data. A commonly encountered problem in BG dynamic model data sets that is of this kind is that the meal and insulin input happens simultaneously, since the insulin bolus is typically set in conjunction with a meal. This will lead

meal-related and insulin-related parameters to be correlated. A common remedy for this employed in clinical research that wishes to separate this dynamics is to delay the insulin bolus compared to the meal, as this improves the information content in the data by separating these effects.

Sensitivity analysis is particularly useful in nonlinear models, that can have parts that are inactive in parts of the state space. An example of this would be for instance renal excretion, which is typically modelled as a ramp function; once the blood glucose reaches a certain threshold, the kidneys start excreting glucose into the urine at a certain rate that is proportional with the glucose level, below this threshold, no renal secretion of glucose occurs. In the Uva-Padova model, this is modelled as a function $E(t)$ that assumes the value $k_{e1}[G(t) - k_{e2}]$ when $G(t) > k_{e2}$ and zero otherwise [62]. Here $G(t)$ is the glucose levels, and $k_{e1}, k_{e2}$ are parameters. If a data set only contains blood glucose measurements beneath the threshold level $k_{e2}$, the sensitivity of the output to changes in the parameters governing renal excretion will be zero, meaning that the data does not provide any information about these parameters, thus they are unidentifiable from the data. If instead an input scenario that triggers hyperglycemia is used, the threshold is exceeded, and the influence of these parameters are visible as non-zero sensitivities.

### 5.6.1 Singular value decomposition of sensitivity profiles

Stigter et al. [168] describes a straightforward method of analysing several aspects of practical identifiability in any dynamic system, using parameter to output sensitivities and singular value decomposition (SVD).

The method consists of first computing the $n_x \times n_p$ parameter-to-state sensitivity matrix $\mathbf{S}_x(t_k)$ for each time step $t_k$, based on a simulation of the system using a nominal parameter vector and spanning the measurements of interest.

The following system is solved through numerical integration over a time interval that spans the measurements, with initial value $\mathbf{S}_x(0) = 0$ :

$$\dot{\mathbf{S}}_x = \frac{\partial \mathbf{f}}{\partial \mathbf{x}^T}\mathbf{S}_x + \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}^T} \tag{5.9}$$

where (5.9) follows from differentiating the system equation of the model (Eq. 4.1) with respect to $\boldsymbol{\theta}$.

From the sequence of matrices $\mathbf{S}_x(t_k)$ we can generate the rows of an output sensitivity matrix, $\mathbf{S}_y$, where each row corresponds to each measurement time. For the linear measurement equation often encountered in BG dynamic models this becomes simply $H\mathbf{S}_x(t_k)$. $\mathbf{S}_y$ will then contain the influence each parameter would

have on the outputs at every measurement, and it is an $n_y \times n_p$ matrix, where $n_y$ is the number of measurements and $n_p$ is the number of parameters.

Note that in this method the actual values of the measurements are not used, only the time of each measurement. In some cases it is of interest to force the system to pass through the states that would produce a specific set of measurements, especially in non-linear systems where different state trajectories lead to different observability and identifiability. In this case a Kalman filter estimating the same model can be used to "pull" the system through states that explain the actual observed measurements.

The $\mathbf{S}_y$ matrix is immediately useful. Each column in the matrix is a sensitivity time series, and if any columns consists of all zeros, then there is no information in the gathered data about the parameter corresponding to that column. In addition, we can put limits on how close to zero we allow columns to be. Intuitively, if the maximum sensitivity observed in a column is much lower than the resolution of our measurement, then there is not much hope of identifying the corresponding parameter from real, noisy data.

The $\mathbf{S}_y$ matrix can be further analysed by decomposing it using Singular Value Decomposition (SVD). SVD decomposes $\mathbf{S}_y$ into one $n_y \times n_p$ matrix $\mathbf{\Sigma}$ which is zero except for the diagonal elements, the singular values (SV) of matrix $S_y$. The decomposition also produces an $n_y \times n_y$ matrix $\mathbf{U}$ of left singular vectors and an $n_p \times n_p$ matrix $\mathbf{V}$ of right singular vectors (RSV) such that

$$\mathbf{S}_y = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{5.10}$$

The RSVs found as column vectors in $\mathbf{V}$ contain information about the correlation between the parameter-to-output sensitivity curves in $S_y$, and the singular value for the corresponding RSV describe the strength of this combination in relation to the others. An interesting property of the SVD of the sensitivity curves is that if two or more parameters are not individually identifiable this will be seen as a drop in singular value magnitude [168], and the corresponding RSV has nonzero elements corresponding to parameters that are not individually identifiable from the data.

The method is local, in the sense that it looks at sensitivities along a specific trajectory given by a specific initial state and parameter set, indicating that the results obtained with the method about observability/identifiability are valid at this point and a small area around it. [168] claims that concatenating the sensitivity curves from several runs starting from different places in parameter and initial state space greatly improves the robustness of the method, making it less local.

### 5.6.2 Fisher Information Matrix analysis

Fisher's Information Matrix (FIM) can be computed based on the sensitivity analysis matrices described in Section 5.6.1 and can be used to determine the practical identifiability of parameters [179, 180, 147]. To construct an approximation of Fisher's Information Matrix we use the sensitivity matrices described in Section 5.6.1 and knowledge about the measurement noise:

$$\text{FIM} = \sum_{k=1}^{N} \mathbf{S}_x(k)^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{S}_x(k) \tag{5.11}$$

where $S_x(k)$ is the $n_x \times n_p$ parameter-to-state sensitivity matrix described in Section 5.6, **R** is the measurement covariance matrix and **H** is the measurement matrix. The index k is here over all available measurements.

The inverse of the FIM is the Cramer-Rao Lower Bound (CRLB), which provides a lower bound on the variance of the estimate of the parameters that have been considered in the sensitivity analysis.

The derivation of this form of the FIM is given in the following, the measurement noise is assumed to be white and Gaussian with covariance **R**.

Fisher's Information Matrix (FIM) is given by

$$\begin{aligned} \text{FIM} &= E\left\{ \left( \frac{\partial \ln p(z|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left( \frac{\partial \ln p(z|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \right\} \\ &= -E\left\{ \frac{\partial}{\partial \boldsymbol{\theta}} \left( \frac{\partial \ln p(z|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right\} \end{aligned} \tag{5.12}$$

where z is the measurement, $\boldsymbol{\theta}$ is the parameter vector and $p(z|\theta)$ is the likelihood function.

To arrive at the equation in Eq. 5.11 we must assume a Gaussian distribution for the likelihood of measurement $z_k$ at time step k:

$p(z|\boldsymbol{\theta}) = \mathcal{N}(z_k; \mathbf{H}\mathbf{x}_k(\boldsymbol{\theta}), \mathbf{R})$

where $\mathbf{x}_k(\boldsymbol{\theta})$ is the state at time k. The value of $\mathbf{x}_k$ is provided through simulation or Kalman filtering of the system, and it depends on the parameter vector $\boldsymbol{\theta}$. We also want to switch to differentiation with respect to $\boldsymbol{\theta}$ since we here are interested in the parameter identifiability. We also assume only scalar measurement (plasma glucose) in the following for notational simplicity.

The Gaussian assumption simplifies the log likelihood function in Eq. (5.12) to:

$$\ln p(z|\boldsymbol{\theta}) = c - \frac{1}{2}(z_k - \mathbf{H}\mathbf{x}_k(\boldsymbol{\theta}))\mathbf{R}^{-1}(z_k - \mathbf{H}\mathbf{x}_k(\boldsymbol{\theta})) \tag{5.13}$$

where c is a constant.

After differentiation w.r.t. $\boldsymbol{\theta}$, multiplying with the transpose, and taking the expectation, we arrive at Eq. 5.11

### 5.6.3 Simulation studies

A straight-forward and practical test of identifiability of a model under a specific measurement regime is to do simulation studies [147] . We run our selected parameter estimation on a set of synthetic measurements created by simulating the model using some scenario of inputs, fixed initial conditions and a known parameter set $\boldsymbol{\theta}_{true}$ that we select, preferably one that makes the model fit some real CGM data set of interest reasonably well.

This simulation produces a set of no noise measurements at a measurement interval that we choose (e.g. 5 minutes to emulate normal CGM data). If the identification procedure fails to recover the true parameter set, then something is amiss, either with the optimisation routine, the observability/ identifiability of the model, or the data. The nature of the failure may tell us something; If the optimisation finds an optimum $\boldsymbol{\theta}_{opt}$ that has cost equal to the cost of $\boldsymbol{\theta}_{true}$ (usually zero in non-noisy data), this means that the solution to the problem is not unique, and there is an obvious non-identifiability problem. When the optimisation finds an optimum that has a higher cost than the cost of $\boldsymbol{\theta}_{true}$, then there is a problem with the optimisation strategy. It has converged to a local minimum, and efforts to make the optimisation better able to find the global minimum could be tried, e.g. choosing a different cost function or search algorithm. But in this case it may also be a good idea to investigate if there are highly correlated sets of parameters, as described in Section 5.6.1, since this can also make the parameter estimation search fail.

This method is also usable for investigating state observability, only then we would fix the parameter set, and let the initial states be considered free variables. If the identification procedure does not find the true initial states this means that the system is not observable. Observability is a prerequisite for identifiability, so this analysis should be done first. Or, one can do as Stigter et al. [168] and do both parameters and initial states in the same analysis.

A case study of this method on the Identifiable Virtual Patient model is given in the following section, and a similar case study is given in Paper B.1.

## 5.7 Case study : The Identifiable Virtual Patient model

The IVP model as described in Sec. 4.4.2 has 7 states and 10 unknown parameters. Compared to the original formulation, our modifications have added parameters to be estimated, possibly reducing the identifiability of the overall model. Note also that the "Identifiable" in the name of the IVP model is in a situation where insulin is also measured, so we expect to run into some identifiability issues when these inputs are not available, as we assume in the following.

The question we would like to answer is:

> Which parameters are we able to estimate from only CGM measurements in the IVP model?

We construct a scenario with some meal inputs and insulin inputs (listed in Table 5.1), then simulate it with the IVP model using parameter set $\boldsymbol{\theta}_0$, given below. The simulation produces synthetic measurements $\{y^i_{synth}\}$, which we use as input to the parameter estimation procedure. The identification procedure produces a parameter estimate $\widehat{\boldsymbol{\theta}}$ through an optimisation search using Newton-Raphson optimisation in a Kalman filter as described in Sec. 5.3.4. The optimisation is initialised with initial guess parameter sets generated by randomly drawing parameter vectors. Each individual parameter is drawn based on its defined range based on parameter values found in published studies of the IVP model [35, 138]. If $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$, we conclude that the model parameters are identifiable by the given data.

| Time | Type | Value | Time | Type | Value |
|---|---|---|---|---|---|
| 06:00 | Meal | 30 g | 12:00 | Insulin | 5 U |
| 07:00 | Insulin | 10 U | 17:00 | Insulin | 25 U |
| 12:00 | Meal | 100 g | 18:00 | Meal | 100 g |

**Table 5.1:** Details of scenario 1

$$\boldsymbol{\theta}_{0_{IVP}} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ C_I \\ S_I \\ p_2 \\ V_g \\ \tau_m \\ GEZI \\ EGP \\ \tau_{isf} \end{bmatrix} = \begin{bmatrix} 50 \\ 25 \\ 1000 \\ 5 \cdot 10^{-4} \\ 0.01 \\ 160 \\ 25 \\ 5 \cdot 10^{-3} \\ 1.0 \\ 10 \end{bmatrix}$$

A sensitivity analysis with SVD was performed around this nominal parameter value. The initial value for $G_p$ was set to 100 $\mathrm{mg/dL}$, all other initial states were zero. The resulting trajectory and analysis is shown in Fig. 5.6. Here we see that the last singular value (SV) is clearly close to zero, indicating a non-identifiable correlated set of parameters [168]. This is also evidenced from the result of the parameter estimation, where several parameters are estimated to values far from the true values, and the fit is not good (not shown). The right singular vector (RSV) corresponding to the smallest SV shows that the correlated parameters are $S_I$ and $C_I$, the insulin sensitivity and insulin clearance, respectively. The plot of the sensitivities for these parameters shows that they are anti-correlated. The sensitivity curves are shown in the bottom of 5.6.

To try to understand why $C_I$ and $S_I$ are not identifiable here, we look at the state transition equations for the IVP model, Eq. 4.9 in Sec. 4.4.2. We see that $C_I$ acts as a scaling parameter on the injected insulin, deciding what insulin concentration will result in $I_{sc}$ and consequently also in $I_p$ from a certain insulin injection. In other words, there is a curve in the $C_I$ – $S_I$ parameter space producing the same outputs. This also implies that the states $I_p$ and $I_{sc}$ will reach different values depending on how $C_I$ and $S_I$ are set. Remembering that the IVP model was intended for use in situations where insulin measurements are available, we can understand that this problem appears when we try to use this model with only CGM measurements. It seems that insulin concentration in the IVP model is free to be at any level (given by $C_I$ and $S_I$) , when it is no longer forced to be at a specific level by insulin measurements.

In the intended use with data from free-living settings, insulin is not available, so we need to reduce the model, and one way to do that is to fix $C_I$. We fix $C_I$ to its true value and move on with the analysis with $C_I$ excluded to see if there are more parameters that are non-identifiable.

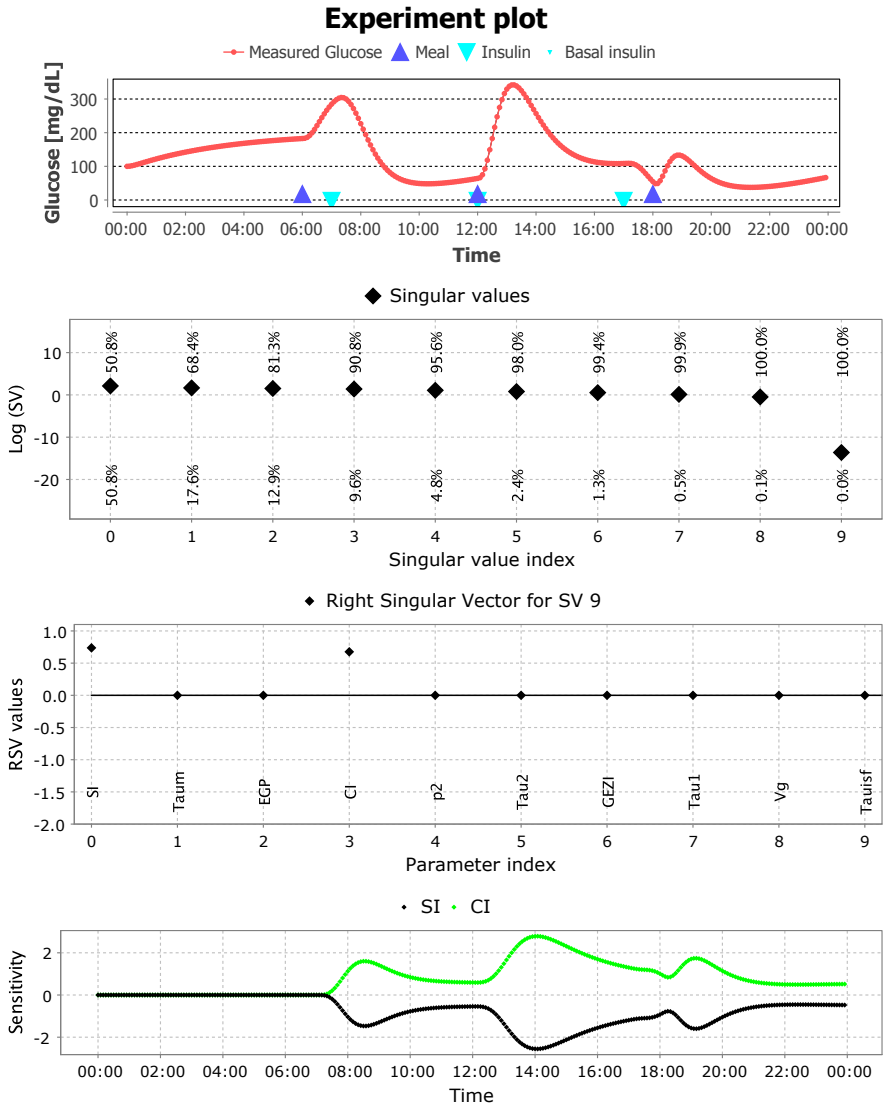



**Figure 5.6:** Sensitivity analysis SVD result for the IVP model with full parameter set. *Uppermost panel:* Trajectory of the simulation. *Upper middle panel:* Singular values (SV), with SV 8 having a clear drop compared to the others. The percentage numbers are the magnitudes of each singular value compared to the sum of singular values, with individual percentages at the bottom and accumulated percentages at the top. *Lower middle panel:* Right Singular Vector (RSV) corresponding to SV 8. This vector has nonzero elements corresponding to parameters $S_I$ and $C_I$. *Lower panel:* Sensitivities of the $S_I$ and $C_I$ parameters against time.

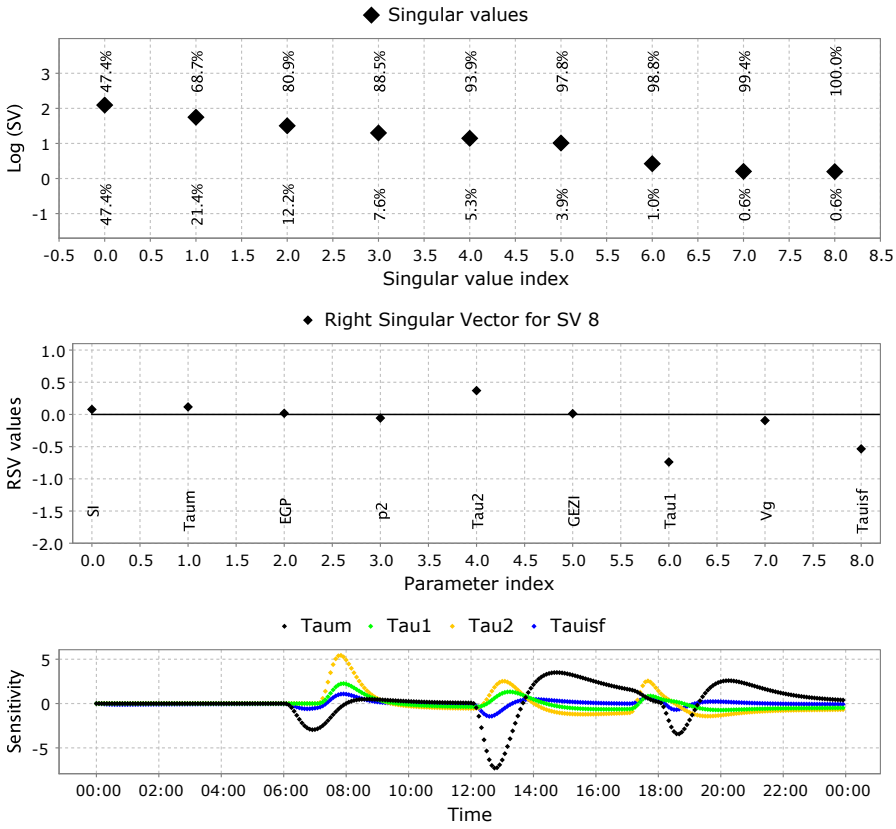



**Figure 5.7:** Run 2 of the analysis, with $C_I$ fixed. *Upper panel:* Singular values (SV). *Middle panel:* Right Singular Vector (RSV) corresponding to SV 7. This vector has nonzero elements corresponding to parameters $\tau_{isf}$,$\tau_2$, $\tau_1$ and $\tau_m$. *Lower panel:* Sensitivities of the implicated parameters against time.

A re-run of the analysis with $C_I$ fixed produces a SV plot where there is no obvious gap like in the first run, but it appears that the three last SVs are distinctly lower than the rest. The result is shown in Fig. 5.7. The parameter estimation is better, but some of the parameter values are still not close to their true value. The last RSV from the lowest SV show that $\tau_1$,$\tau_2$,$\tau_m$ and $\tau_{isf}$ are correlated. From the sensitivity plot of these parameters, we see that although the sensitivities are no longer perfectly correlated as we saw in the first run, the signals still have enough correlation with each other to give identifiability problems. $\tau_1$,$\tau_2$ are related to insulin transport, $\tau_m$ is related to digestion of the meal, and $\tau_{isf}$ is related to the CGM measurement dynamics. From this we can gather that much of the correlation comes from the correlation of the inputs themselves. Making the scenario more

realistic by adding basal insulin injections could alleviate this, but scenarios like the one used is commonly found in real data from free-living use, since it is natural to bolus close to when a meal is taken. The fact that $\tau_{isf}$ is also implicated tells us that the use of CGM measurements instead of direct blood glucose makes it harder to discern the different time constants involved in the dynamics.

It is possible to continue using this method, investigating the lowest SV and its corresponding RSV, fixing the implicated parameter(s) and continuing until the lowest remaining SV is above some limit, e.g. 10% of the sum of SVs, or until parameter estimation produces the correct parameter vector.

A simulation based analysis with sensitivity analysis and singular value decomposition as illustrated here is a useful tool to investigate many other questions of interest, e.g. how sensitive the parameter identifiability is to measurement noise or increased sampling interval, or how typical CGM sensor problems like biases, lags, calibration errors and PISAs affect our ability to identify parameters, and the influence of other errors commonly seen in data from free-living settings like missing or incorrectly reported meals. This type of analysis is important for being able to evaluate the feasibility of parameter estimation in different BG models given a specific set of data from free-living settings.

# Chapter 6

# Real-time estimation in Blood Glucose Dynamics Models

Once a model has been identified for a patient, it can be used for real-time estimation and prediction. Commonly only state estimation is performed, but in some cases it is also desirable to track slowly varying parameter values. In a state estimation configuration for use in blood glucose estimation, the estimator is provided CGM measurements, SMBG measurements (if any) and input information, e.g. meals and insulin that are supplied by the user. The output is a state estimate with mean and covariance, which can be used for prediction. Such predictions are useful in several usage scenarios ranging from diabetes advisor apps, for providing an estimate during CGM sensor dropouts, or in model predictive control (MPC).

For linear systems with Gaussian noise processes, the Kalman filter (KF) is applicable and provides optimal state estimates. The systems we encounter in glucose-insulin dynamics are usually not linear, and the noise processes we encounter are not Gaussian. In this case it is more appropriate to use extensions of the Kalman filter, e.g. Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF), or sequential Monte Carlo filtering, also known as Particle Filter (PF), [181]. The Kalman filter and the different approaches are described briefly in the following.

## 6.1 State estimation methods

The estimators described here all construct an estimate $\hat{\mathbf{x}}_k$ of the state vector $\mathbf{x}$ from a set of measurements $\mathbf{y}$, i.e. they describe the probability $p(\mathbf{x}_k|\mathbf{y}_1, \mathbf{y}_2 \cdots, \mathbf{y}_k)$. In the case of Kalman filtering, this is done recursively, meaning that the state estimate at $k+1$ is based only on the state estimate at $k$ and the measurement at $k+1$. In

other words, the state estimate $\hat{\mathbf{x}}_k$ contains all the information gathered from the measurements $\mathbf{y}_1, \mathbf{y}_2 \cdots, \mathbf{y}_k$.

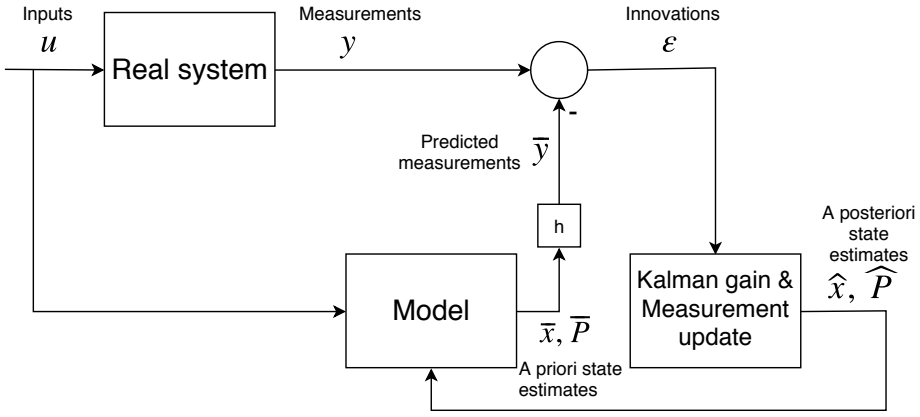



**Figure 6.1:** Block diagram of a Kalman filter. The Kalman filter works by running a model of the process to be estimated/filtered through one or more time updates, producing an *a priori* estimate of the state, $\bar{\mathbf{x}}$. When a measurement is available, a measurement update is performed, generating an *a posteriori* state estimate $\hat{\mathbf{x}}$. The process continues with new time updates until the time of the next measurement. The uncertainty of the estimate is also propagated in the form of covariance matrices $\bar{\mathbf{P}}$ and $\hat{\mathbf{P}}$.

### 6.1.1 Kalman filtering

A Kalman filter is shown in Fig. 6.1. The Kalman filter equations are [182]:

Time updates:

$$\bar{\mathbf{x}}_k = \Phi_{k-1}\hat{\mathbf{x}}_{k-1} \tag{6.1}$$

$$\bar{P}_k = \Phi\hat{P}_{k-1}\Phi^{\mathsf{T}} + Q \tag{6.2}$$

Measurement updates:

$$K_k = \bar{P}_k H^{\mathsf{T}}(H\bar{P}_k H^{\mathsf{T}} + R_k)^{-1} \tag{6.3}$$

$$\hat{\mathbf{x}}_k = K_k(\mathbf{y}_k - H_k\bar{\mathbf{x}}_k) \tag{6.4}$$

$$\hat{P}_k = (I - K_k H)\bar{P}_k \tag{6.5}$$

here $\bar{\mathbf{x}}_k$ is the a priori state estimate at time $k$ with covariance matrix $\bar{P}_k$, $\hat{\mathbf{x}}_k$ is the a posteriori state estimate, having covariance matrix $\hat{P}_k$. $\Phi$ is the state transition matrix, $H$ is the measurement matrix, $Q$ is the process noise covariance matrix, and $R$ is the measurement covariance matrix.

When the system is linear and all uncertainties follow Gaussian distributions, the Kalman filter is an optimal estimator of x.

### 6.1.2 Extended Kalman filtering

The Extended Kalman filter (EKF) extends the application of Kalman filters to nonlinear systems. The EKF uses the same equations as the regular Kalman filter, except the transition matrix $\Phi$ and the measurement matrix $H$ are now time-varying, and are found for each time step as Jacobian matrices evaluated at the last estimate, i.e.

$$\Phi_k = \left.\frac{d\mathbf{f}}{d\mathbf{x}^\mathsf{T}}\right|_{\hat{\mathbf{x}}_k} \tag{6.6}$$

$$H_k = \left.\frac{d\mathbf{g}}{d\mathbf{x}^\mathsf{T}}\right|_{\hat{\mathbf{x}}_k} \tag{6.7}$$

### 6.1.3 Unscented Kalman filtering

The linearization performed by the EKF at each time step approximates the nonlinear transition and/or measurement functions in order to be able to keep propagating Gaussian distributions as per the normal Kalman filter, and for some systems this works well. Sometimes the nonlinear system's transition function and/or measurement function are accurately known. In this case it makes more sense to approximate the probability distribution, which in most cases is inaccurately known anyway. The Unscented Kalman Filter (UKF) [183] achieves this by representing the probability distribution of the state by its so-called sigma points, which are cleverly selected points that ensures that the set of points and the distribution they encode has the same mean and variance. The sigma points are transformed by the transition function $f(\mathbf{x}, \mathbf{u})$ a,d the measurement function $g(\mathbf{x}, \mathbf{u})$ and each point is weighted by the likelihood of the observed measurement y. A Gaussian is constructed from the mean and variance of the transformed and weighted sigma points. From here, a new set of sigma points are computed and the process starts again with the next time step.

### 6.1.4 Particle filters

Although the UKF deals with non-linearity in a better way than the EKF, it still assumes that the Gaussian is a good approximation for the probability distribution of the state estimate. In some cases the distribution can be far from Gaussian. In a particle filter the distribution is represented by a a swarm of particles drawn from the distribution. The PF concept has similarities with UKF in how the particles are

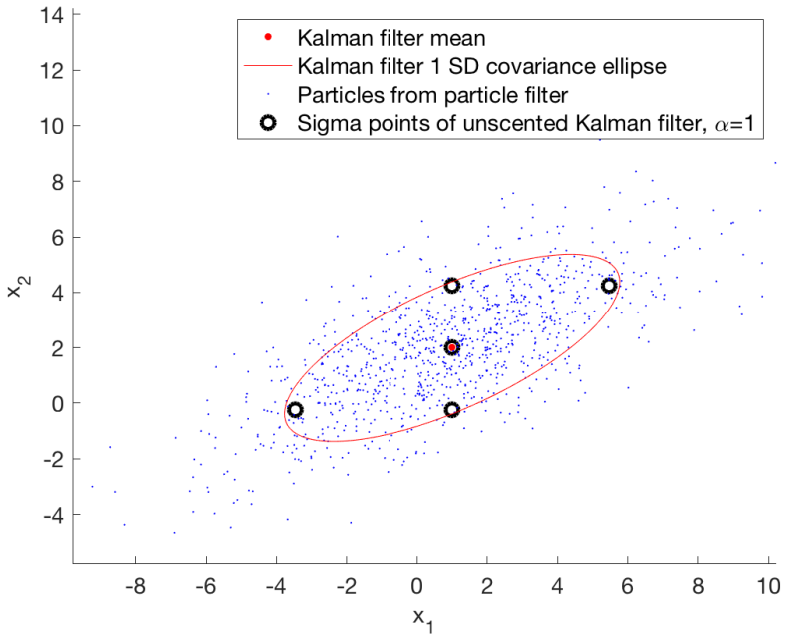



**Figure 6.2:** How uncertainty is represented by different estimators. The Kalman filter represents the estimate by the mean and covariance matrix of a Gaussian distribution, which can be drawn as an error ellipsoid (red). An UKF represents the estimate as a set of sigma points (shown as black circles). The particle filter represents the distribution by a set of samples drawn from the distribution (blue points). The sigma points and the particles shown in this figure have the same mean and covariance as that represented by the error ellipsoid.

propagated and weighted by measurements, but in PF the particles are never turned back into a distribution. Removal of the most unlikely particles and resampling is needed to keep the swarm from becoming degenerate.

An illustration of the different ways the uncertainty of a pair of states is represented by a KF, UKF and PF is shown in Fig. 6.2.

## 6.2 State estimation in blood glucose dynamic models

State estimation in BG dynamic model is done for many purposes. Some researchers are interested in the state values themselves, e.g. estimating the plasma insulin dynamics from CGM data [184]. For others, the state estimation is performed in order to get the best possible starting point for a prediction into the future, either for MPC [111] or advisor app purposes. The latter usage is also the case that is

most relevant for Prediktor Medical. For a metabolism model to be useful in a product like the BioMKR, it needs to be able to perform state estimation while the measurements from the device are valid, and be able to provide predictions of the glucose while measurements are invalid or absent, e.g. in periods of poor signal quality due to movements, or in short periods of the sensor being off skin.

The state estimation methods presented are all suitable for use in a real-time settings, and even particle filters would be possible to run on an embedded system like the BioMKR, provided that the number of particles are not excessive. There are however some model related questions, that are independent of the choice of estimator algorithm, that needs to be answered:

1. Does the estimation recover the true internal state in ideal conditions?
2. Does the estimation recover the true internal state in realistic sampling and noise conditions?
3. What is the best model with regards to being able to predict glucose in periods of absent measurements?

The latter question is perhaps the most important, at least for Prediktor Medical applications.

## 6.3 Noise modelling

When using Kalman filters, the noise processes we choose determine the performance of the filtering to a large degree. In the Kalman filtering theory a process noise is added to the transition function, and a measurement noise to the measurement function. The balancing of these noise processes determine how the output from the model is believed relative to the measurements from the system.

### 6.3.1 Measurement noise

The measurement noise to be used by the Kalman filter should be based on the actual measurement system in use. The regular Kalman filter works best if the measurement error can be somehow represented as a zero-mean Gaussian distribution with a certain variance. Another approach is to model the noise as a coloured noise process, by adding interconnected noise states where one of them is driven by Gaussian white noise.

There are noise models for both SMBG [185, 186] and CGM measurements [187, 188, 189, 190] available, and they are discussed in the following.

### SMBG measurement noise

In the study by Vettoretti et al. [185] the performance of two SMBG meters (One Touch Ultra 2 (OTU) and Bayer Contour Next (BCN)) is compared against YSI measurements. The distributions of the measurement errors are gaussian-like, but skewed towards positive errors. The authors find that the error is best represented by a constant variance in the lower range and a relative variance in the higher range, but find different thresholds for dividing lower and higher range (75 mg/dl in OTU and 115 mg/dL in BCN). While the distribution found is neither zero-mean or Gaussian, it seems a reasonable approximation to use zero-mean normal distribution for these errors in Kalman filtering that uses SMBG data. The relative standard deviations for the measurements in the high range found in this study was 5% for OTU and 10% for BCN, implying that different variances for SMBG meters is appropriate.

Often the type and model of the SMBG meter that produced the data is not included in the data to analyse, and if it is, the exact model to use to describe the error is not known for the data. In such cases an option is to use a worst-case variance on the measurements. The ISO 15197 standard for blood glucose meters [186] provides this. An SMBG device that conforms to the ISO 15197:2015 standard should have an error within $\pm 0.83$ mmol/L (15 mg/dL) when the real blood glucose level is less than or equal to 5.55 mmol/L (100 mg/dL), and less than $\pm 15\%$ error for higher glucose levels. The standard specifies that 95% of all measurements shall fall within this limit. Interpreting this as a 95% confidence interval the limits correspond to roughly 2 SDs in a normal distribution. We see that also here we operate with different variances in different glucose ranges.

The noise models described above with the use of two zones, one low and one high using an absolute variance in the low glucose range and a relative variance in the high glucose range seems to be common in the description of SMBG meter errors. This is an example of heteroscedastic noise, where the noise level is not constant, but depends on the magnitude of the measurement. This error model can be included in a Kalman filter by setting the value of the measurement noise matrix $R$ for each measurement based on the actual value of the measurement, e.g. in the case of ISO 15197, use measurement variance $\sigma_R^2 = 0.172$ [mmol$^2$/L$^2$] for measurements below 5.55 mmol/L, and $\sigma_R^2 = 0.0056y^2$ [mmol$^2$/L$^2$] for measurements above 5.55 mmol/L, where $y$ is the measured glucose value. This approach is used in the Kalman smoother described in Appendix A.1.

#### CGM noises

The noise on CGM measurements is not well represented by Gaussian white noise. In the study by Facchinetti et al. [188], four CGMs of type Dexcom Seven Plus (D7P) were worn by each of the 19 patients participating in the study. The authors assume that all four sensors measure the same ISF glucose concentration, which follows the plasma glucose concentration through the model described in Eq. 4.3 but each sensor is affected by different body reactions to the insertion and sensor drifts.

The D7P is an older CGM device, launched in 2009, having a reported MARD of 16%. It thus has significantly poorer performance than newer systems. For comparison, the most recent CGM from Dexcom, G6, has a MARD rating of 9%. In Facchinetti et al. [191] a similar analysis is performed for Dexcom sensor, the G4 Platinum, which is more recent than the D7P but a predecessor of G6, having an intermediate MARD of 9 – 13% (depending on software version). Biagi et al. [189] use the same method to estimate the error model for the Medtronic Paradigm Veo Enlite CGM sensor. The models are all on the same form, but have different parameter values. The model is $CGM(k) = G_{isf}(k) + a(k) + b(k)$, where $CGM_k$ is the k-th CGM measurement, $G_{isf}$ is the interstitial glucose concentration at this time, considered to follow the dynamics described in Eq. 4.3. $a(k)$ and $b(k)$ are CGM-system specific noises , and for the Medtronic Paradigm Veo Enlite CGM sensor they are given by the following equations [163]:

$$a(k) = 1.584 * a(k-1) - 0.8842 * a(k-1) + 0.1798 * a(k) + \omega_a(k) \quad (6.8)$$

$$b(k) = 1.367 * b(k-1) - 0.4816 * b(k-2) + \omega_b(k) \quad (6.9)$$

This model does not take into include some of the larger and less random artefacts that may affect CGM measurements, e.g. PISAs.

### 6.3.2 Process noise

The process noises used in the Kalman filter and described by the matrix $Q$ in Eq. 6.2 are more difficult to find correct values for, and this often turns into a tuning exercise. The process noise makes the filter more robust against disturbances entering the system and it also allows modelling errors, i.e. discrepancies between the real system and the model. In BG dynamic models both are present.

The modelling errors are always present since there is no model that explains everything that is going on in the glucose dynamics. Models that shall be estimated based on CGM and insulin data need to be quite simplistic in order to be observable. Thus the modelling error increases, and the process noise needs to be set higher.

The disturbances to the system are several, and depending on which inputs the model accepts, different influences are considered inputs or disturbances. This was discussed also in Section 4.2.

## 6.4 Hidden nonlinearities

A common issue with Kalman filtering in BG dynamic models is that the models contain nonlinearities that are often not explicitly stated. For instance, states that describe concentrations shall be strictly positive. Such nonlinearities are present in many models, even those that are considered linear, e.g. Magdelaine's model described in Section 4.4.1. This needs to be handled in two ways. First, when the system equations are implemented and used in the time update of the Kalman filter, checks must be put in place that makes sure concentrations do not go negative, unless the transition function enforces this. Also we must make sure that the measurement update of the Kalman filter does not produce states that are negative.

This can be handled differently in different types of Kalman filters. In an EKF, states that should be non-negative that still go negative during measurement updates must be checked and can be forced positive after the measurement update, however this should also result in a change in the state covariance estimate of the filter. In a UKF the sigma points that reach invalid states can be moved to the closest valid state [192]. In a particle filter a properly implemented transition function $\mathbf{f}$ will ensure that particles are not propagated to invalid states, but the process noise added to each particle could push it across the border(s). Such particles can be removed as part of the resampling step of the PF.

## 6.5 Online parameter estimation

For parameters that vary over time it is possible to do real-time/online parameter estimation as described in Gelb [162]. A Kalman filter is used, and the state vector is augmented with the parameters to be estimated:

$$\boldsymbol{x}_a = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\theta} \end{bmatrix} \tag{6.10}$$

The parameters are usually modelled as having no dynamics, $\dot{\boldsymbol{\theta}} = 0$. For more volatile parameters a noise model that allows the parameter to drift slowly can be used, e.g. a white noise process.

Since the state augmentation makes the resulting system nonlinear, an EKF (or UKF or PF) must be used. The linearised, discretised transition matrix for the augmented

system for use in an EKF is

$$\boldsymbol{\Phi}_a = \begin{bmatrix} \boldsymbol{\Phi} & \frac{d\boldsymbol{f}}{d\boldsymbol{\theta}^T}\Delta t \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}_{\hat{\mathbf{x}}_k, \hat{\boldsymbol{\theta}}_k} \tag{6.11}$$

Running a Kalman filter on this augmented state vector will estimate the parameter value together with the states. The initial variance on the parameters and the process noise tuning becomes important here, since the balance between state noises and parameter noises determines the results to a large degree. If the initial parameter uncertainty is small, the process noise affecting parameters is zero, and the covariance of the states is relatively larger, the parameter will stay close to the initial value. If its initial variance is made larger or it is considered to be dynamic by adding process noise to it, it will move more throughout the estimation session.

## 6.6 Online input estimation

The shortcomings of data from free-living settings, namely missing and inaccurate data, especially meal data, is as much a problem in real-time estimation as in other settings, e.g identification, as discussed in Chap. 5.

The input estimator presented in paper B.2 is intended to detect missing meals during real-time state estimation. The method is still in a preliminary stage of development, and has not been yet been tested in combination with real-time estimation of glucose dynamics.

## 6.7 Estimating fewer states in the GlucoPred model

Since the Prediktor model has low observability, an attempt was made to run a hybrid estimation, where the least observable states were run in "ballistic mode", i.e. they were not estimated by the Kalman filter, only simulated. The states from the GlucoPred model that were set to "ballistic" were $x_{SR1}$, $x_{SR2}$, $x_{SS1}$, $x_{SS2}$, $x_{LG}$, $x_{Sto}$, $x_Y$ and $x_Z$. While this helped to some degree on the real-time estimation of the remaining states, and keeping the filter from diverging, it was not found to achieve the necessary performance, and was not further investigated.

## 6.8 Tracking-oriented glucose models

An alternative to starting from complex, physiology-based models that are quite far from observable and identifiable using data from free-living settings, we can instead start in the other end, with small, empirical models that are observable

and identifiable, but lack the connection to physiology. Such models were also attempted during the testing of the BioMKR.

# Chapter 7

# Practical experiences

This chapter presents some of my experiences of putting the theory mentioned in previous chapters into practical use for use in real-time estimation of glucose dynamics in the BioMKR device.

## 7.1 Studies

As part of the development and testing of the BioMKR device (see Section 1.1) Prediktor Medical conducted several clinical studies. In these studies the main focus was on evaluating the near infrared and bioimpedance measurement system's ability to measure glucose, and a secondary goal was to record data that could be used for identification and testing blood glucose dynamics models. The BioMKR device and other reference measurement equipment used in these experiments were capable of logging meals and insulin doses and the BioMKR device also contains an accelerometer that can be used to monitor the activity level of the person wearing the device.

The first study of BioMKR was conducted in 2015, shortly after this PhD project was started, and consisted of measuring with 6 BioMKRs on each patient while the patient was lying still in a bead in a research ward undergoing a clamping procedure designed to achieve a specific glucose response. Only glucose was measured, with an YSI 2300 Stat. 12 participants with Type 1 diabetes participated. Since this was a clamping study, the data collected were not considered representative for free-living use. The short duration of the session and only one clinic visit for each participant that these data were not suitable for metabolism modelling.

The second study was conducted in 2016, where a clinical research ward session

was followed by 7 days of free-living use for 17 patients. The BioMKR device was used, and participants logged SMBG measurements, meals and insulin with it. However, data was mostly incomplete, and the metabolism model identification that could be performed based on it was not reliable and there was not enough data to validate models.

The third study was conducted in 2017, again with a clinical research ward visit on the first day, followed by 7 nights of use, using a Freestyle Libre device as a reference glucose measurement, together with frequent SMBG measurements in the first day visit. This provided an opportunity to investigate the differences between FGM and SMBG measurements that is reported in paper A.2. At this point the meal and insulin logging mechanism of the device had been turned off, since the device had been moved to be on the upper arm, making data input via the buttons on the device difficult. The results from the second study also showed that the completeness and quality of data collected in this manner was poor. Participants were equipped with Freestyle Libre devices that were to provide glucose reference measurements during the non-invasive experiments. The participants were encouraged to use the logging functionality in the Freestyle Libre, and a short instruction for how to log data of interest was provided to the participants. However, 27 out of 39 participants entered nothing, and for the rest there we no more than 2 entries of any kind. Thus, the data collected through Freestyle Libre logs in the BioMKR III study was considered not usable for metabolism model identification for any of the participants.

These experiences led to the realisation that in most cases we could not base a metabolism model parameter estimation on the data that is self-reported by the participants, not even for highly motivated study participants. A way to remedy to this this could lie in methods to preprocess the data to clean and reconstruct missing data where possible. This led to investigations into glucose data cleaning, uncertainty estimation and automatic missing meal detection, as described in papers A.1 and B.2.

## 7.2 Real-time estimation of glucose in periods of missing data

The BioMKR device (see Section 1.1) is strapped to the upper arm, and it produces glucose estimates from NIR and bioimpedance data. The quality of the glucose estimates depends heavily on the stability of the skin-device interface. Factors affecting this interface are pressure, skin moisture, skin temperature etc. The device is also intended to be taken off during short periods, e.g. while showering. In such transient periods of high signal noise, or short spans of missing measurements, a BG dynamic model running in a Kalman filter is useful for suppressing noise and

for doing predictions.

Such filtering was attempted on NIR and bioimpedance data using various BG dynamic model, and worked well as a smoothing method. For bridging gaps in data, however, it was too often found to be lacking in performance. One reason for this was improper state estimation during the OK periods of the signal, both due to improper noise modelling and a too complex and unobservable model in use in the Kalman filter. This led to investigations into observability and identifiability, and reduction of the model in use in the BioMKR device, as described in paper B.1. Another reason why such use is problematic has been discussed earlier: if an input is introduced into the system during a period of missing data, but not announced to the system, the predictions will be wrong. This led to investigations into meal detection and estimation, as described in paper B.1.

# Chapter 8

# Discussion

The papers included in this thesis are about different topics related to glucose data and blood glucose dynamics modelling. They are related to a common theme; Identification of blood glucose dynamics models based on data from free-living settings, as illustrated in Fig. 8.1.

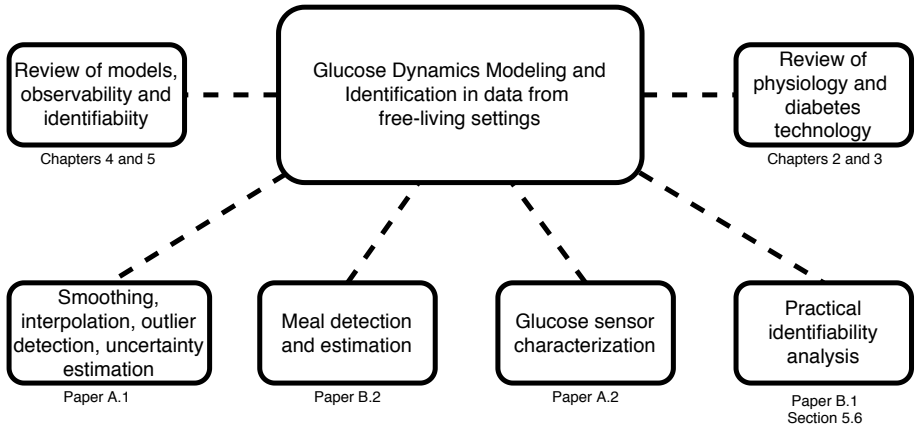



**Figure 8.1:** The main theme and related contributions of the work underlying this thesis

## 8.1 Related research

There are several other efforts in this direction:

- Messori et al. [149] try two different approaches, one black-box non-parametric method and one grey-box identification technique that uses a linearised

version of the UVa-Padova model, and perform model identification and validation on in silico data from the full UVa-Padova T1DMS model. Toffanin et al. [148] builds on this approach and use a two-step method estimate parameters based on data from free-living settings. In the first step they use simulation data from an average virtual patient in the UVa/Padova T1DM simulator to estimate parts of their target model, which is a linear model of significant reduced complexity compared to the non-linear, 18-state UVa-Padova T1DM simulator model. In this step they can design the inputs as they wish in order to give the most informative data for model estimation, thereby bypassing the limitations discussed in Section 5.1.1 related to patient risk. In step 2 of the procedure a small subset of the target model's person-specific parameters are estimated using CGM data from free-living settings. This provides a way to make the simpler model behave as close as possible to the more complex UVa-Padova model that the simple model is able to represent, and then personalise some parameters to make the model fit the real data. As such the simpler target model inherits some of the dynamics embedded in the physiologically well-founded UVa-Padova model. The authors comment that the collected data had some technical issues that necessitated careful data selection before model identification and validation.

- Cescon et al. [63] report on the identification of 7 parameters in a model describing meal and insulin transport using second-order transfer functions. They use administration of insulin 2 hours after meals to increase the information content in the data and improve identifiability of parameters. Their study focused on single meals (breakfast).
- Magdelaine et al. [136] adapt their model (also described in Section 4.4.1 of this thesis) to CGM data and achieve fits that seem to represent the glucose in the long term, but that seems to gloss over some of the glucose excursions, perhaps due to missing information about some meals.
- Boiroux et al. [193] investigate methods for identifying parameters of models based on CGM data and compare optimisation method with regards to finding the correct parameter values in simulated CGM data.
- Hovorka et al. [194] use Bayesian parameter estimation to estimate parameters in a non-linear model intended for MPC glucose control, based on intravenously drawn blood glucose.
- Visentin et al. [36] describe a method to individualise the UVa-Padova model based on Bayesian parameter estimation, however plasma glucose and insulin concentrations is needed in order to do the identification.

From the above short overview of related work it is obvious that many studies

are performed on in silico data and there are few reports of problem-free use of real data from free-living settings to accomplish model individualisation. To fully achieve model individualisation based on real data from free-living settings, additional methods and knowledge seem to be needed. Clean and complete data are important in order to do model identification. It is also of importance to know the characteristics of the glucose sensing equipment, like bias and lag, that if uncounted for can affect the quality of the model individualisation. Methods for deciding which parameters to identify based on the available data that is available are also needed. In a system for model individualisation based on data from free-living settings such operations should ideally be done without human intervention. The papers that are part of this thesis are contributions towards these ends, and are described individually in the next section.

## 8.2 Contributions of individual papers

**Paper** A.1 describes a method for processing glucose data to provide interpolation, smoothing, outlier detection and uncertainty estimation, all of which are useful in preprocessing of data for model identification.

Finegood and Bergman [195] noted the importance of smoothing measurement data before attempting state or parameter estimation, and propose an optimal segments method for doing so. Their method is comparable to ours but lacks the uncertainty estimation in the interpolated signal that our method provides through the estimated variance. This measure of uncertainty is useful in a state/parameter estimation setting. Del Favero et al. [196][197][198] use a method for retrofitting CGM data where the intention seems to be to retrospectively improve the accuracy of a CGM signal and bring it closer to YSI or SMBG values, given data sets that contain both CGM and SMBG measurements.

**Paper** B.2 is about meal estimation. We use a Chan-Hu-Plant (CHP) input estimator, originally designed to detect pilot manoeuvres in tracked aircraft, to detect meals in CGM data. In addition to detecting meals, the method also estimates the time and CHO content of the input. Meal detection from CGM data has been investigated by many in artificial pancreas research settings, and it is frequently criticised for being too late; detection times are often 40 minutes or more after the meal was taken, and in an artificial pancreas this is often too late to avoid large postprandial glucose excursions. The meal detector we proposed seems to give as late a detection as other methods, but its ability to detect the time and the value of the meal means that it could be a useful tool for data cleaning purposes in an offline setting, e.g. as a preprocessing step before model identification to find meals that were taken, but not reported, or meals that are reported too late or too early. Further research is

needed to establish if this is possible.

While our suggested method for meal estimation has similarities with some of the recently proposed methods for meal detection from CGM data [163, 164], we believe it is novel in its use of the CHP estimator for glucose data. A method that also estimates the meal size is described by Mahmoudi et al. [163]. They use a Kalman filter to estimate the states in an individualised second-order linear model for insulin and meals, and use tests on the KF innovations and the estimated states to detect the onset of change due to a meal. Once the meal is detected, a smoothing method is used to identify the value of the meal in g CHO after it has been detected. They test the method on data generated from the UVa-Padova T1DMS simulator. Kölle et al. [164] use a Moving Horizon Estimator (MHE) together with Linear Discriminant Analysis (LDA) to detect meals, which also is model-based, however this method reports the detection time only, not the estimated meal time or value. The method is tested on simulated, single-meal data and on real CGM data [199].

**Paper** A.2 investigates the characteristics of the data from FreeStyle Libre (FL), and determines the bias and lag of data in day one of using this sensor. The performance of the FreeStyle Libre has been investigated by many [200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210] often focused on the total, overall MARD. In our study we look at individual MARDs and find large differences between participants. This may have implications for the use of data from Freestyle Libre for model identification purposes. It also is of interest to users of the Freestyle Libre system and their caretakers, since the biases observed in our study could be something that is not expected in a system that claims to be factory calibrated. This study contributes to the body of knowledge of glucose sensor systems, which has been extensively studied for conventional CGM systems [187, 188, 189, 85, 84, 123].

**Paper** B.1 makes use of the method described by Stigter et al. [168] to investigate identifiability properties of parameters in the UVa-Padova model, as a case study of the method applied to BG dynamic models. This method is useful to understand which parameters that can be estimated given a model and a certain data set. Garcia-Tirado et al. [91] use other methods for the same purpose, investigating structural and practical identifiability in control-oriented models, among them the Magdelaine model and the SOGMM model described in Sections 4.4.1 and 4.4.2 of this thesis.

Finally, the content of this thesis provides a review of relevant background related to diabetes, measurement equipment, BG dynamics models and model identification, providing a short introduction to many of the topics needed in this field, and with pointers to more thorough treatments of the topics, that should prove useful to researchers new to the field.

## 8.3 Suggestions for future work

There are several unsolved problems in the field of getting BG dynamics models individualised and into practical use. Some suggestions for future work are:

- Compare the method presented in Paper B.2 with other methods for meal detection and estimation, e.g. Mahmoudi et al. [163], to see if the methods are comparable, both in on-line, artificial pancreas oriented settings and in off-line, data cleaning oriented settings.
- Investigate the feasibility of using meal estimation methods as a data cleaning method in data from free-living settings to improve data yield and quality of BG dynamics model individualisation.
- Creating a (preferably open-source) software framework for defining, simulating and estimating BG dynamics models. Methods for observability testing and identifiability testing should be included in such a software framework, and the data cleaning methods described in this thesis could also be included.
- Investigations into using other ways than only the CHO count to represent meals in BG dynamics models, e.g. as proposed in Rozendaal et al. [76], and investigate if more complex meal modelling improves predictive ability.

# Chapter 9

# Concluding remarks

This thesis has reviewed several concepts in the field of blood glucose dynamics related to diabetes, including glucose monitoring, modelling and identification. Publications from the PhD project are given in the appendix and are related to glucose data cleaning, glucose sensor characterisation, identifiability analysis in blood glucose dynamics models, and meal detection in glucose data.

The implications of this PhD project for the scientific community are:

- The Kalman smoother described in paper A.1 should be useful in many different systems for automatic glucose data processing. The method has been made publicly available and is therefore easy to utilise.
- The comparison of the Freestyle Libre to SMBG measurements and the characterisation of the errors in terms of biases and lags described in paper A.2 should be of interest to researchers wanting to use FGM data in their research. The data from this study has been made publicly available.
- The method for structural and practical identifiability analysis of blood glucose dynamics models described in paper B.1 can be of use to other researchers needing to reduce their models or check identifiability properties.
- The meal detection algorithm described in B.2 describes a new approach to meal estimation that performs estimation of the meal in both time and value that could prove useful in artificial pancreas applications, but also for glucose data cleaning purposes. This method is at a preliminary stage as it is not fully developed and tested.
- The background theory described in this thesis provides an introduction and a review of physiology, diabetes, glucose measurement equipment, models

and model identification.

The investigation into differences between FGM and SMBG measurements (paper A.2) is relevant for diabetes patients that use or intend to use the Freestyle Libre, as well as to health care personnel working with diabetes patients, since it enables them to better understand the nature of FGM inaccuracies compared to SMBG measurement, and to interpret FGM measurements.

Implications of this work with respect to industry are foremost the contributions made in my company, Prediktor Medical. This has consisted of insights into various aspects of metabolism modelling and identification. The work has resulted in several prototype implementations of real-time estimators of blood glucose that has been used in different prototypes of the BioMKR product, and contributions to their software for identifying metabolism parameters. This includes the methods for glucose data cleaning. The work also supported the clinical studies performed by Prediktor Medical by evaluating the Freestyle Libre FGM as a substitute for SMBG measurements.

# Appendix A

# Published work

## A.1 Kalman smoothing for objective and automatic preprocessing of glucose data

Journal article published in *IEEE Journal of Biomedical Health Informatics*

O. M. Staal, S. Sælid, A. Fougner, and Ø. Stavdahl, "Kalman smoothing for objective and automatic preprocessing of glucose data," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 218–226, Jan 2019.

Full text preprint version is included on the next page.

# Kalman Smoothing for Objective and Automatic Preprocessing of Glucose Data

Odd Martin Staal, *Member, IEEE*, Steinar Sælid, Anders Fougner, *Member, IEEE* and Øyvind Stavdahl, *Member, IEEE*

***Abstract*— A method for preprocessing a time series of glucose measurements based on Kalman smoothing is presented. Given a glucose data time series that may be irregularly sampled, the method outputs an interpolated time series of glucose estimates with mean and variance. The method can provide homogenization of glucose data collected from different devices by using separate measurement noise parameters for differing glucose measurement equipment. We establish a link between the ISO 15197 standard and the measurement noise variance used by the Kalman smoother for Self Monitoring of Blood Glucose (SMBG) measurements. The method provides phaseless smoothing, and it can automatically correct errors in the original datasets like small fallouts and erroneous readings when surrounding data allows. The estimated variance can be used for deciding at which times the data are trustworthy. The method can be used as a preprocessing step in many kinds of glucose data processing and analysis tasks, such as computing the Mean Absolute Relative Deviation (MARD) between measurement systems, or estimating the plasma-to-interstital fluid glucose dynamics of continuous glucose monitor (CGM) or Flash Glucose Monitor (FGM) signals. The method is demonstrated on SMBG and FGM glucose data from a clinical study. A Matlab implementation of the method is publicly available.**

## I. Introduction

DIABETES is a disease suffered by close to 10% of the world's adult population [1].

To avoid the acute and chronic consequences of diabetes, the blood glucose level should be controlled to a level as close as possible to the normal range. To achieve such control, people with diabetes need to measure their blood glucose level frequently.

Blood glucose is most often measured using Self Monitoring Blood Glucose (SMBG) meters. These devices provide discontinuous blood glucose readings, by analyzing a drop of blood that the user applies to a test strip. Advances in sensor technology has resulted in Continuous Glucose Monitor (CGM) systems, which measure the interstitial fluid (ISF) glucose continuously with a thin electrochemical sensor inserted under the skin, and reporting a filtered measurement value usually every 5 minutes. A recent addition to the family of glucose devices is the Flash Glucose Monitor (FGM) [2], that provides data only when the sensor is scanned, but then it provides a current glucose estimate and 8 hours of historic 15-minute interval data, thus providing a kind of hybrid between the two data types.

Tremendous amounts of glucose data are generated with such devices, both in research projects, commercial product development and by far the most data are generated in normal use of the devices by individual users.

Errors are present in most glucose data sets, and include sensor system errors and user errors. For SMBG systems a common source of error is incorrect sampling procedure [3], e.g. forgetting to clean the finger before sampling. There is also a baseline variation caused by strip manufacturing variability. For CGM systems, pressure induced sensor attenuation (PISA) errors are common [4], [5], usually resulting from the patient lying on the sensor. Fallouts, bias and latencies are other occurences in CGM data [6], and some or all of these errors also apply to FGM.

The primary purpose of SMBGs and CGMs is to provide real-time information about glucose levels, allowing the user to take informed decisions about insulin dosing, meals and exercise, thus helping to avoid hyper- and hypoglycemia. There are many possible secondary uses of the glucose data in offline settings. One possibility is to estimate the parameters of glucose-insulin metabolism models, which has been of interest to many, see e.g. [7]–[10], thereby obtaining personalized models of the glucose-insulin system. Such models have many uses, for instance in model predictive control (MPC) of an artificial pancreas (AP), see e.g. [11]. Since estimates of model parameters will suffer from noisy or erroneous measurements [12], data should be cleaned and smoothed before parameter estimation commences.

Mean Absolute Relative (MARD) analysis is commonly used to characterize and compare sensor systems. This analysis is also sensitive to errors in the data sets, and to some researchers it is useful to be able to detect outliers in the reference signal, typically SMBG or laboratory measurements. MARD analysis can also require interpolation in order to be able to properly align data points between the different sensor systems.

This paper proposes a preprocessing method for interpolation of glucose data and suppression/removal of outliers in an automatic and objective manner. The methods based on

Manuscript received October 7, 2017; revised January 26, 2018; accepted February 25, 2018. *(Corresponding author: Odd Martin Staal)*
This work was supported by the Norwegian Research Council
The authors are with the Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway.
Odd Martin Staal and Steinar Sælid are also with Prediktor Medical, Fredrikstad, Norway.

Kalman smoothing, and converts a time series of possibly irregularly spaced blood glucose measurements to a continuous time series of interpolated estimates with mean and variance (uncertainty). The method is able to cater for different measurement devices by using device-specific noise models. The method can also cater for other glucose dynamics models than those presented here, including models with person-specific parameters.

The presented method is suitable for homogenizing glucose data sets, which is valuable in an offline automated data processing setting to increase data yield. The method also provides a consistent interpolation between glucose values without introducing interpolation artefacts like those that may result from other methods, e.g. cubic spline interpolation. Since the smoothing is applied in a forward-backward manner, a phaseless smoothing and interpolation of the noise data is achieved. Finally, the method takes into account the uncertainty of the measurements, which can be used to determine which parts of the measurement series are trustworthy and which are not.The method was originally intended for processing SMBG data, but is extendable to include CGM/FGM data as well, which is demonstrated. The method is tested on real SMBG and FGM glucose data from a clinical study. This paper expands on ideas previously presented in [13].

The key contributions of this paper are:

- Use of Kalman smoothing applied to glucose data, resulting in phaseless smoothing and interpolation.
- Glucose data interpolation with realistic uncertainty estimates
- Derivation of a measurement noise model from ISO 15197:2015 [14] for SMBG measurements.
- Possibility of combining glucose data sources with different noise characteristics

## II. Method

In this section Kalman filtering and fixed-interval Kalman smoothing is revisited, and their application to glucose data is described.

### A. Kalman filtering

The Kalman filter theory assumes that the signal $\mathbf{y}_k$ to be filtered is generated by a system on the form

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{v}_k \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \tag{2}$$

Here $\mathbf{x}$ is the system internal state, $\mathbf{u}$ is the input to the system, and $\mathbf{y}$ is the measured output, $\mathbf{v}$ is the process noise, and $\mathbf{w}$ is the measurement noise. These noise processes are assumed to be white, zero mean Gaussian, with covariance given by matrices $Q$ and $R$, respectively:

$$\mathbf{v}_k \sim \mathcal{N}(0, Q) \qquad \mathbf{w}_k \sim \mathcal{N}(0, R) \tag{3}$$

The Kalman filter computes a state estimate $\hat{\mathbf{x}}$ and a state covariance matrix $\hat{P}$ for each time step. The system model in Eq. (1) is used to predict the state one step ahead in time. This is called the *time update*, and it results in an *a priori estimate* denoted by $\bar{\mathbf{x}}$ and $\bar{P}$. Then the *measurement update* is performed. This updates the a priori estimate with a measurement with known measurement noise, to produce an *a posteriori estimate* $\hat{\mathbf{x}}$ and $\hat{P}$ . In glucose data sets it is commonly the case that measurements are infrequent and/or taken with irregular intervals. The filter handles this by doing several time updates per measurement update. In time steps where no measurement is available, the a posteriori estimate is set equal to the a priori estimate. The Kalman filter equations are [15]:

$$\bar{\mathbf{x}}_k = \Phi_{k-1}\hat{\mathbf{x}}_{k-1} + B_{k-1}\mathbf{u}_{k-1} \tag{4}$$

$$\bar{P}_k = \Phi_{k-1}\hat{P}_{k-1}\Phi_{k-1}^T + Q_{k-1} \tag{5}$$

$$K_k = \bar{P}_k H_k^T (H_k \bar{P}_k H_k^T + R_k)^{-1} \tag{6}$$

$$\hat{\mathbf{x}}_k = K_k(\mathbf{y}_k - H_k\bar{\mathbf{x}}_k) \tag{7}$$

$$\hat{P}_k = (I - K_k H_k)\bar{P}_k \tag{8}$$

$\Phi$ is the state transition matrix and $H$ is the measurement matrix. If the system and/or measurement equations are non-linear, these matrices will in general be time-variant, and can be found from linearizing $\mathbf{f}$ and $\mathbf{h}$ in Eqs. (1) and (2) at each time step around the most recent estimate resulting in the Extended Kalman Filter (EKF). Two different sets of linear system equations have been tested in this work, and are described in section II-D. An augmentation of the state space for CGM-SMBG parameter estimation is described in Sec. II-D.3, and this augmentation makes the resulting system nonlinear, necessitating EKF.

### B. Kalman smoothing

The Kalman filter described above is suitable for real-time processing of data, as it only uses past data to produce its estimates. In an offline setting, where the whole data set is available, smoothing can be used to get further improvement of the estimates. Smoothing uses all the available data before and after time $k$ to produce the estimate at time $k$. The Rauch-Tung-Striebel (RTS) algorithm [16] accomplishes this. RTS first makes a forward pass through the data using the normal Kalman filter Eqs. (4-6), storing the sequences of a priori and a posteriori estimates $\bar{\mathbf{x}}_k$, $\hat{\mathbf{x}}_k$ and state covariance matrices $\bar{P}_k$ and $\hat{P}_k$. These are then used as input to a backward pass that computes the smoothed estimates $\hat{\mathbf{x}}_k^s$ and $\hat{P}_k^s$ as follows [15]:

$$C_k = \hat{P}_k \Phi_k \bar{P}_{k+1}^{-1} \tag{9}$$

$$\hat{\mathbf{x}}_k^s = \hat{\mathbf{x}}_k + C_k(\hat{\mathbf{x}}_{k+1}^s - \bar{\mathbf{x}}_{k+1}) \tag{10}$$

$$\hat{P}_k^s = \hat{P}_k + C_k(\hat{P}_{k+1}^s - \bar{P}_{k+1})C_k^T \tag{11}$$

The filtering and smoothing process is illustrated in Fig. 1. The error bands in this figure (and the rest of this paper) are based on 2 standard deviations (SD), approximating a 95% confidence interval under the Gaussian assumption. The SD is the square root of the diagonal element in the $\hat{P}_k^s$ or $\hat{P}_k$ covariance matrix that corresponds to the glucose state. We have used fixed-interval smoothing in all the results reported

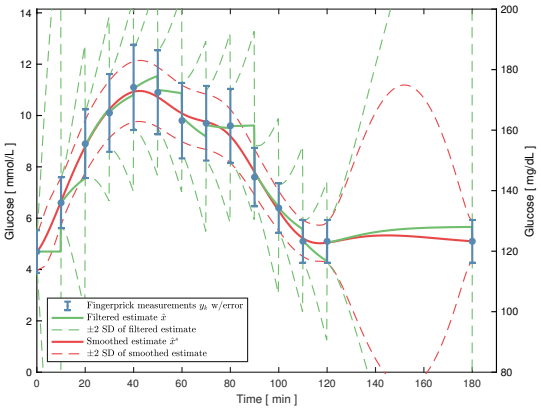



Fig. 1. Filtering vs smoothing in a 3-hr data set of glucose SMBG measurements (blue points, with ISO 15197:2015 error indicated with blue bars). The estimates are shown with solid line for the mean and dashed lines for the 2 SD error band approximating a 95% confidence interval. The forward pass filtering result is shown in green. This estimate is seen to jump every time a measurement arrives, setting a new rate estimate. The smoothed estimate is shown in red. Note how the error of the smoother estimate is smaller than the error bars in the original measurements. This is possible due to the proximity in time of the measurements. In the last hour of the recording the sampling rate is reduced to one sample per hour, and it is seen that the error band of the smoother estimate (red dashed line) grows with time, describing the uncertainty of the estimate during the time of no measurements.

here, utilizing all the available data, i.e. the interval is the entire data set. A fixed-lag smoother implementation would also be possible, where a fixed window of data is used to estimate the state at the start of the window, which is interesting in a near real-time setting.

The error band could be used to determine when it makes sense to use the interpolated values output by the smoother for further analysis, and when the estimates are too uncertain to be used. This could be done by applying a threshold for maximum allowed estimated error band, for example.

### C. Noise modeling

An important issue in use of Kalman filtering is the modeling of the noise processes $\mathbf{v}$ and $\mathbf{w}$, more specifically the values of the covariance matrices $Q$ and $R$.

*1) Measurement noise modeling:* The measurement noise process $\mathbf{w}$ needs to be set according to which measurement system has generated the data. In a highly accurate laboratory system for blood glucose analysis it is appropriate to use a low variance. If a more inaccurate blood gluose measurement system like an SMBG meter is used, a higher variance on the measurement noise should be used.

The ISO 15197:2015 standard [14] is applicable for SMBG meters, providing limits for measurement error that SMBG manufacturers must comply with. Let us recode these limits into normal distribution variances to be used in the Kalman filter. An SMBG device that conforms to the ISO 15197:2015 standard should have an error within $\pm 0.83$ mmol/L (15 mg/dL) when the real blood glucose level is less than or equal to 5.55 mmol/L (100 mg/dL), and less than $\pm 15\%$ error for higher glucose levels. The standard specifies that 95% of all measurements shall fall within this limit. Interpreting this as a 95% confidence interval the limits correspond to roughly 2 SDs in a normal distribution.

Therefore, to approximate the ISO 15197:2015 measurement variance $\sigma_R^2$ we have that $2\sigma_R = 0.83$, i.e. $\sigma_R^2 = 0.172$ [mmol$^2$/L$^2$] for measurements below 5.55 mmol/L, and $2\sigma_R = 0.15y$, i.e. $\sigma_R^2 = 0.0056y^2$ for measurement above the limit, where $y$ is the measured glucose value.

The above value(s) for $R$ may serve as conservative defaults if no other information is available. Another value for $R$ could and should be used if more details about the measurement variance is known. Variance information could be found from the following:

- The 2003 version of the ISO 15197 standard has wider limits, so to smooth data from older devices conforming to the old standard, those limits should be used.
- Several SMBG meter manufacturers currently promote their products as having a performance significantly better than the IS015197:2015 limits. For some SMBG devices a more suitable measurement variance may be available from the device documentation.
- In other cases there may be independent analyses of SMBG accuracy that provides information about the error model to use, e.g. as in [17].

*2) Process noise modeling:* The process noise $\mathbf{v}$ should account for all the noise originating from modeling errors, as well as unknown disturbances affecting the system.

One important point in this context is whether or not meals and insulin injections are considered as inputs or disturbances in the model. For the application considered here, inputless models are preferred. In other words, we treat meals and insulin injections as unknown disturbances to the system. This is done to make the method more generally applicable, since many glucose data sets lack meal and/or insulin information that is accurate enough to be of use in this context.

As a consequence, the variance of the process noise must be set large enough to accomodate glucose excursions originating from meals or insulin injections. Thus, the $Q$ matrix should be chosen such that the error band of the estimate grows quickly enough to envelop worst case glucose excursions, like right after a meal or an insulin injection. This method was used to tune the process noise covariance matrix $Q$ in this work, and is illustrated in Fig. 2. A dataset containing a large meal is used. Those measurements that are most informative about the meal onset have been held back from the smoother. The error band of the estimate (red dashed line in in Fig. 2) grows when measurements are not available, and the growth rate depends on the value of $Q$. The goal of the tuning is to make the error band conservative enough to encompass the held-back measurements, which represent a worst case glucose excursion from a meal. Since the glucose lowering effect of insulin is comparable to the glucose rising effect of meal digestion [18], we tuned the process noise based on meal cases only.

Values for the $Q$ matrix depend on the unit used to represent glucose by the model, and values given in the following

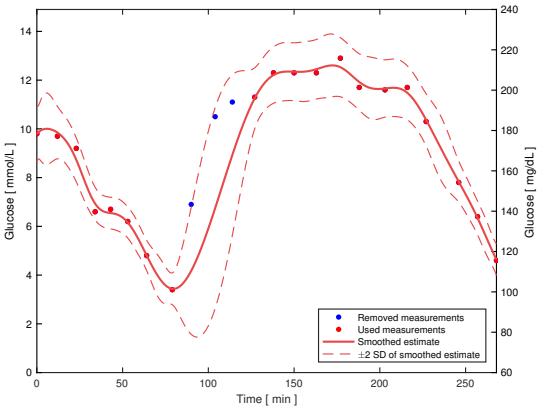



Fig. 2. Illustration of the process noise tuning. Red dots are SMBG measurements available to the smoother, while the blue measurements have been held back.

sections are for models using mmol/L as the unit for the internal state of glucose. The conversion factor from mmol/L to mg/dL is 18.02 and needs to be squared when converting variances.

### D. Glucose dynamics modeling

The Kalman filter uses a dynamic model that describes the assumed internal dynamics of the system generating the measurements that are to be filtered. This model is used to provide predictions of the state between each measurement.

Several glucose metabolism models exist in the literature, and research into such models is ongoing. The models range from minimal [19], [20], via intermediate complexity [21]–[24] to maximal models [25], [26]. A common trait in these models is that they require information of the amount of glucose and insulin entering the system, which is not always present in a glucose dataset. When such data are present, they are often erroneous and/or incomplete, unless they have been recorded in strictly controlled research settings.

Another issue with the more complex glucose metabolism models is that they are often non-observable based on glucose measurements alone. This means that the internal state of the model can not be computed based on the measurements. Observability is a requirement for Kalman filtering.

To summarize, the model used in the smoother should be observable and have no external inputs. The two models described below satisfy both these requirements. One is a simple rate-only model, and another is inspired by the model in [24], but further simplified for filtering usage (central-remote rate model). Both models operate without insulin and meal input knowledge.

*1) Model 1: Rate-only model:* This model is perhaps the simplest dynamic system that could be said to represent glucose. The state vector consists of plasma glucose and its rate, $\mathbf{x} = [G_p \quad \dot{G}_p]^T$, and the model is $\dot{\mathbf{x}} = A\mathbf{x}$ with $A = \left[\begin{smallmatrix} 0 & 1 \\ 0 & -a \end{smallmatrix}\right]$, where $a$ is a small number determining the decay of an observed rate of change. $a$ has been set to 0.05 in this work. Larger values of $a$ will make the rate of change decay faster towards zero. Setting $a$ to zero gives a system where an observed rate of change is allowed to continue indefinitely. This is clearly not physiological and can be detrimental for smoothing performance in data sets where there are long periods of missing measurements.

The system is discretized by setting $\Phi = e^{A\Delta t}$ where the time step $\Delta t$ has been set to 10 s in this work.

The process noise was set to $Q = \left[\begin{smallmatrix} 0 & 0 \\ 0 & q_{m1} \end{smallmatrix}\right]\Delta t$ for this model, where $q_{m1}$ is a tunable value, set to 0.005 mmol$^2$/L$^2$ in this work. It has been found using the worst case analysis described in Sec. II-C.2 and illustrated in Fig. 2. The presence of $\Delta t$ in the $Q$ matrix is a convenience to automatically adjust the process noise if the discretization time step interval changes.

*2) Model 2: Central-remote rate model:* In this model, the glucose rate from model 1 is divided into two states, $C_c$ and $C_r$, where $C_c$ occupies a central compartment, and $C_r$ occupies a remote compartment. In this model, any input (meal/insulin) first affects a central compartment and then with a first order delay diffuses over to the remote compartment where it takes effect on the blood glucose. Another interpretation of the $C_c$ and $C_r$ states is that they describe meal effects when positive, and insulin effects when negative. The effects of any other blood glucose increasing or decreasing phenomena are lumped into the same states.

The state vector of this model is $\mathbf{x} = [G_p \quad C_c \quad C_r]^T$. The state transition equations are

$$\dot{G}_p = C_r \tag{12}$$

$$\dot{C}_c = -\frac{1}{T_d}C_c \tag{13}$$

$$\dot{C}_r = \frac{1}{T_d}(C_c - C_r) \tag{14}$$

where $T_d$ is a parameter of the model, a time constant that describes the rapidness of flow between compartments.

The process noise is set to only directly influence the central compartment, i.e. $Q = \left[\begin{smallmatrix} 0 & 0 & 0 \\ 0 & q_{m2} & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]\Delta t$. Here $q_{m2}$ is again a tunable parameter, set to 0.02 mmol$^2$/L$^2$ in this work, using the same method for finding the process noise level as previously described. The parameter $T_d$ also affects the variance development; the lower the $T_d$, the faster the process noise on state $C_c$ propagates to $C_r$. Thus, $T_d$ needs to be tuned together with the process noise, and was set to 600 s in this work. Unless otherwise stated, figures and results in this paper are generated using model 2 as the glucose dynamics model.

*3) Plasma-ISF glucose dynamics:* An interesting extension to the method presented is possible when both SMBG and CGM measurements are present in the data set to be smoothed.

The smoother described above can be expanded to provide sensor fusion of the two measurement types. Some usage scenarios for such processing include:

- With sparse SMBG and dense CGM data, bias correction of the CGM data is possible
- With dense SMBG and dense CGM data, estimation of person-specific plasma-ISF dynamics is possible in addition to bias
- Improved outlier detection and removal by combining SMBG and CGM data

If SMBG and CGM measurements are both to be used by the smoother, a model describing the plasma-ISF dynamics is needed. Plasma-ISF dynamics has been investigated by several groups, see e.g. [10], [27]–[30]. Preference is here again given to a simple model that do not require other information than the blood glucose and the CGM measured glucose, for observability reasons and for the method to be more generally applicable.

The dynamics between the SMBG and CGM measurements can be modeled by augmenting the glucose dynamics systems described in above sections with a new state $G_{isf}$, having a first order dynamics relationship to the blood glucose, as in [10]:

$$\dot{G_{isf}}(t) = \frac{1}{T_{isf}}(G_p - G_{isf}) \tag{15}$$

where $G_{isf}$ is interstitial glucose, $G_p$ is plasma glucose, and $T_{isf}$ is the time constant governing the diffusion process. This differential equation can be added to any model that describes the dynamics of $G_p$ (e.g. the ones described in previous sections) to augment them with the ability to use CGM measurements in combination with blood glucose measurements (e.g. SMBG).

To be able to use the CGM measurements, the Kalman filter needs to be expanded to accomodate the extra measurement. One way to do this is to use a two-row measurement matrix $H$, where the second row measures $G_{isf}$. In the likely situation that both measurements are not available at the same instant, a better strategy is to only use scalar measurements in the Kalman filter, switching to the appropriate $H$ single-row matrix depending on which measurement is available in a given instant.

A commonly occuring error in CGM measurements are bias errors. This bias can be included in the Kalman filter and estimated as part of the smoothing procedure. The measurement equation becomes:

$$y_{cgm,k} = G_{isf,k} + b_{cgm} + w_{cgm,k} \tag{16}$$

where $b_{cgm}$ is an unknown bias to be estimated, by including $b_{cgm}$ in the state space of the Kalman filter, using zero derivative, zero process noise and a non-zero initial variance that is large enough to describe biases that may be encountered. The zero process noise indicates that the parameter is modeled as an unknown constant. Adding a small noise on the parameter allows it to have some drift; this may also be beneficial for mathematical/numerical reasons to avoid degenerate covariance matrices in the smoothing operation.

The presence of a Gaussian white noise process $w_{cgm,k}$ can be debated. Some researchers have claimed that the CGM measurement noise in CGMs is non-Gaussian, e.g. Breton et al. [28] found that a Johnson distibution was more appropriate. Others argue that inferring the error distribution is confounded by modelling inaccuracies in the plasma-ISF dynamics and/or calibration errors, [31]. As stated in [32], not all CGM systems are equal. Along the lines of the SMBG measurement noise modeling of Sec. II-C.1, we consider the model given by Eqs. (15) and (16) a useful default when no more information is available, and if more detailed information about the CGM errors are available for the data to be smoothed, e.g. as given in [28], [32], [33], these models can be used instead.

The time constant $T_{isf}$ affects the bias estimation, so simultaneous estimation of the bias and the time constant is in order. Thus $T_{isf}$ is also added to the state space in the same fashion as the bias, and estimated by the smoother.

### E. Outlier suppression and removal

The method described above works well as is to suppress outliers in the data, due to the smoothing introduced by nearby points.

In some applications outlier suppression is not enough, and it is more desirable to remove the outliers altogether. The described method lends itself to this task, too. The process noise tuning described in Sec.II-C.2 ensures that the error band of the estimate grows rapidly enough to encompass meals and insulin injections. Thus, the error band of the estimate can be used to determine which measurements are likely outliers. This criterium is then well-founded, as it will reject only those measurements that are unlikely based on surrounding data, and taking worst-case, but possible, glucose fluctuations into consideration. We can base such a removal algorithm both on the filter (forward pass) estimate and the smoother (backward pass) estimate.

Outlier detection based on the filter estimate error band is only capable of detecting gross outliers. As discussed above, the process noise used is set relatively high in order to give correct variance development under the assumption that meals/insulin injections may occur at any time. This makes the filter estimate develop a large variance quite quickly after a measurement, as seen in Fig. 1.

Outlier detection based on the smoother estimate error band is more promising. This is intuitively a more sound approach, since also data after the suspected outlier is used to determine if it is an outlier. An example of outlier suppression and removal in a dataset containing a likely outlier is shown in Fig. 5. A block diagram showing the process of filtering and smoothing used for outlier detection is shown in the left side of Fig. 3.

### F. Replicate handling

Another case related to outlier removal is a situation that is common in SMBG data sets, where there are two or more measurements close together in time. This may be specified in the study protocol the data has been recorded under, or may be the result of the normal behaviour of some users when they get measurements they believe might be incorrect, repeating a measurement immediately. One such dataset is shown in Fig. 4, where the smoother output is compared to a cubic spline interpolation of the same data. Cubic spline is often used to provide interpolation between points, and the motivation for this comparison is to showcase how risky this can be in an automated setting.

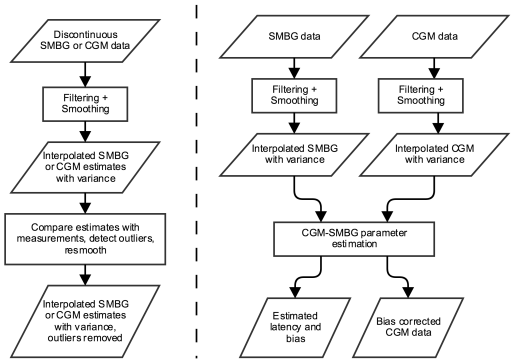



Fig. 3. A block diagram of the interpolation and outlier removal processing (left side) and the data fusion and CGM-SMBG parameter estimation processing (right side)

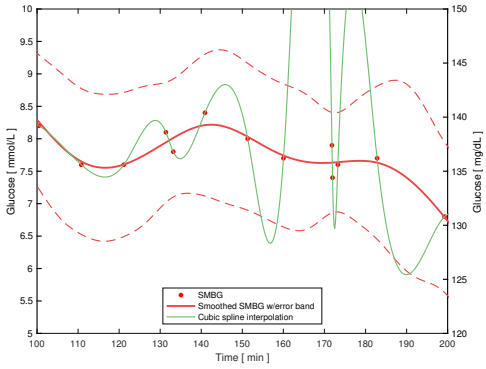



Fig. 4. Automatic handling of triplicates. The smoother estimate (red curve) goes through the mean of the triplicate measurements at t=172 min. A cubic spline interpolation of the same data is shown in green for comparison, where a bad reaction to the proximity in time of the triplicate measurements is seen.

### G. Implementation

A Kalman smoother for SMBG measurements and a CGM-SMBG dynamics parameter estimator have been implemented in Matlab, and is publicly available [34]. The estimator uses the smoother individually on the SMBG and CGM measurements, and then uses the resulting smoothed and outlier removed curves to estimate the CGM bias and time constant, as illustrated in the right side of Fig. 3.

### H. Testing

The smoother was tested on data sets recorded as reference measurements in a study of BioMKR®, a novel non-invasive glucose sensor being developed by Prediktor Medical, Fredrikstad, Norway. The study was approved by the regional ethical comittee, Study ID: REK Midt 2016/1127.

The study recruited 39 type I diabetes subjects. A calibration session was performed for each subject, where reference glucose was collected every ten minutes with an SMBG meter (Freestyle Freedom Lite, Abbot). A Flash Glucose Monitor (FGM; Freestyle Libre, Abbot), worn on the upper right or left arm, was scanned at the same time as the SMBG measurements were taken. Glucose increase was achieved by sugary drinks, and decreased by insulin injections. The calibration session data sets ranged from 2.6 to 5.8 hours, with a mean duration of 4.6 hours. The SMBG measurements glucose data ranged from 3 to 26 mmol/L, with a mean of 8.5 mmol/L. Only the SMBG and FGM data from the calibration sessions in this study were used as test sets in this paper. The SMBG data were used for testing the smoothing/interpolation and outlier removal.

The approach for combining SMBG measurements and CGM measurements with simultaneous estimation of $T_{isf}$ and $b_{cgm}$ as described in Sec. II-D.3 was tested with synthetic data sets with known time constant and bias. These were generated by simulating the system and measurement Eqs. (15) and (16), using the real SMBG measurements from the calibration sessions as the $G_p$ state and varying $T_{isf}$ (1, 5, 10 and 20 min) and $b_{cgm}$ (-2 to 2 in steps of 0.5). Thus, 36 simulated CGM curves were generated for each of the 39 calibration sessions, resulting in more than 1400 synthetic data sets for testing the parameter estimation, and comparing the estimate with the known true values for the parameters. After using these data sets to determine the smoother's ability to find the parameters, the SMBG-CGM estimator was also tested on the 39 real Freestyle Libre data sets.

## III. Results

### A. Outlier suppression and removal

Using model 2, the smoother automatically found three outliers in the 39 data sets. One is the dataset shown in Fig. 5, where the point at t=161 min is too high by about 4 mmol/L. The others were too high by about 3 and 2 mmol/L, respectively, and are shown in Fig. 6. Removal of these outliers resulted in a change of MARD between CGM and SMBG measurements of 0.3, 0.7 and 1.5 percentage points for the sets, respectively. This illustrates the impact such outliers in the reference signal can have for subsequent analyses like a MARD computation.

Manual inspection of the other data sets found measurements that could be regarded as more moderate outliers, these are suppressed by the smoother, but not removed.

If instead model 1 was used in the smoother, the result was the same except for one case: the outlier shown in the top panel of Fig. 6 was not removed. It is clear from the figure that this point is just barely outside the error band when using model 2. Model 1 has a slightly faster variance development compared to model 2, making this point fall inside the error band instead of outside, and that is the reason for this point not being labeled an outlier by the smoother when using model 1. This shows that the choice of glucose dynamics model and its process noise parameters can determine the outcome of the outlier removal to some extent.

### B. CGM measurement parameter estimation

The SMBG-CGM parameter estimation was run on the synthetic test sets. It estimated the bias to within 0.1 mmol/L

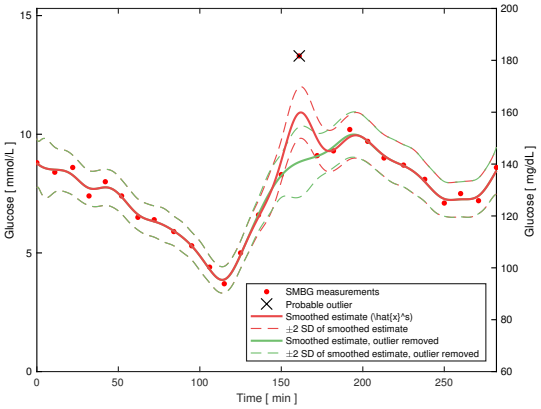



Fig. 5. Outlier suppression and removal. A likely SMBG outlier is present at t=161 min. The surrounding points make the first pass smoother estimate (solid red line) follow a curve that has suppressed the outlier, but the estimate is still drawn towards the outlying value. Removal of outliers can be based on the error band of the smoother (red dashed line). The smoother can be rerun with outliers removed for an improved estimate, this is shown in green. Note how the error band of the resmoothed estimate is now slightly larger where the outlier has been removed.

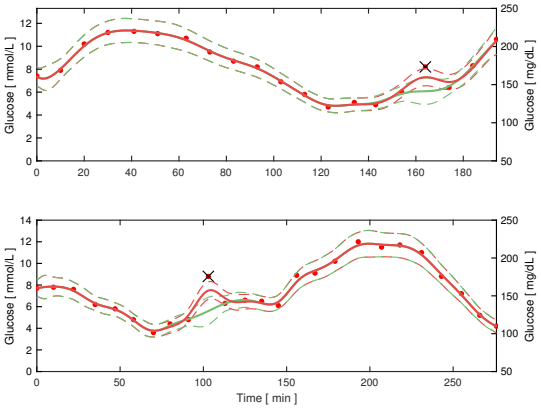



Fig. 6. Two other outliers detected by the method. Legend as in Fig. 5

of the true bias in 95% of the cases with a maximum error of 0.14 mmol/L, and the time constant to within 1 min of the true value in 81% of the cases with a maximum error of 1.8 min.

When run on the real Freestyle Libre data from the 39 data sets from the clinical study, the smoother estimated biases ranging from -1.8 to 1.5 mmol/L and time constants ranging from 1 to 24 min. Parameter estimations are shown for three selected runs in Fig. 7.

## IV. Discussion

The Kalman smoothing methods presented here are useful for various tasks in automatic processing of glucose data for research and commercial purposes. The smoother presented is model based, and two simplistic glucose dynamics models

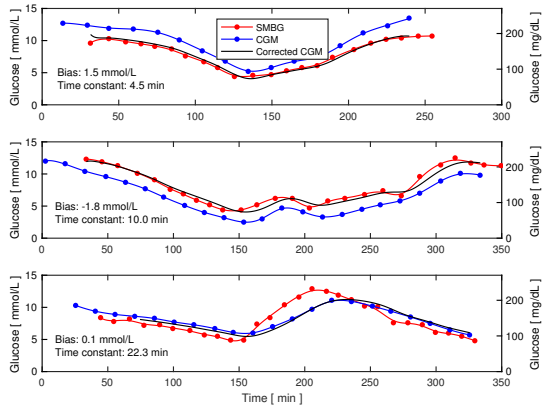



Fig. 7. Parameter estimation of the CGM-SMBG dynamics in three runs from the calibration sets. The three runs display different SMBG-CGM dynamics; the two first sets have clear bias effects in different directions, whereas the last set shows a clear latency effect. Biases and time constants found by the smoother are given in the plots. The black curves are bias corrected CGM signals, produced by retrieving the $\boldsymbol{G_{isf}}$ state after smoothing

were investigated for use in the smoother. The more plausibly the model describes real glucose dynamics, the more confidence we can have that the smoother will give realistic estimates and variances. On the other hand, the model must be observable with glucose as the sole measurement, which limits the complexity of the models that may be used. The choice of model and noise parameter settings gives room for some subjectivity, but it provides a way to be explicit about the assumptions made, which helps provide reproducability.

There are many examples of Kalman filtering applied to glucose in the literature, largely applied in online settings, e.g. for denoising CGM data in real time [35], or for state estimation in MPC settings [11]. Our approach differs in that it uses Kalman *smoothing*. This implies that the method is usable only in offline settings, where all the data is available. Such retrospective settings are common, at least in research, and especially in research related to glucose-insulin metabolism models. One commonly encountered task in metabolism model research is parameter estimation in the models. The importance of smoothing measurement data before attempting state or parameter estimation is acknowledged in [12], where different methods for smoothing glucose data are evaluated. Their optimal segments method is comparable to ours in the resulting curve of smoother estimate means, but lacks the information about the uncertainty in the interpolated signal that our method provides through the estimated variance. This measure of uncertainty is directly useful in a state/parameter estimation setting.

A noteable feature of the output from our smoother is how rapidly the variance grows in periods of no measurements. However, due to the tuning done to encompass worst case glucose fluctuations, this rapid uncertainty development *is realistic* when meal and insulin inputs are considered to be unknown. This feature of our method is useful in SMBG datasets

to get a measure of the uncertainty between measured points, as illustrated in Fig. 1. This enables an informed decision about which interpolated values to include in subsequent analyses.

The strengths of the method include a minimum of assumptions made and high customizability. No person-specific parameters or information about meal and insulin inputs are needed to use the smoother. The method does encode some information about glucose variability and glucose sensor error modeling that could be useful as defaults when no more information is available, and if more detailed information about sensor error, patient glucose dynamics or meal/insulin inputs are available, the method can include this. This includes the use of models with person-specific parameters.

One could also envision clinical use where methods as described here are used to clean and correct SMBG and/or CGM data before they are displayed to users and their caretakers. This could potentially reduce the burden on the users and caretakers in having to know about common errors in the data and trying to mentally correct for them.

The CGM-SMBG parameter estimation implemented as an extended test case in this work finds the correct bias and time constant in synthetic data sets. It finds parameter values in real FGM data sets that are plausible and comparable to those found by other groups for other CGM systems [32], [33] showing that simultaneous estimation of these parameters can be done as part of the smoothing. This is an important preprocessing step when using CGM data, as using biased data could influence downstream processing. For instance in metabolism model parameter estimation, where bias correction and knowledge of the plasma-interstitial fluid time constant is needed to prevent CGM device-specific dynamics affecting the estimation of person-specific parameters. Our SMBG and FGM data does not allow us to conclude whether the biases we observed originate from the SMBG or the FGM measurements, but we assume the latter, since bias/calibration error is commonly found to be the largest error in CGM systems [32], [33], whereas SMBG measurement errors have been found to be uncorrelated in time [17]. It should be noted that glucose data sets with both CGM and frequent SMBG measurements as those analysed here, rarely occur outside research settings.

This work has considered glucose data sets, but the method should be applicable for other biomedical measurements that behave similarly, e.g. lactate. Some of the parameters (e.g. measurement and process noises) will need altering, but the general method should be applicable.

## V. Conclusion

A Kalman smoother for automatic and objective preprocessing of glucose data has been presented, providing interpolation, outlier removal, replicate handling and uncertainty estimation in glucose data. Properties of the method have been discussed, and its performance on human glucose data sets containing SMBG and FGM measurements has been demonstrated. The method is recommended over some other methods that may be used for such tasks, e.g. linear or cubic spline interpolation, due to its noise suppression properties and its ability to estimate realistic variance (uncertainty) at each interpolated point. A Matlab implementation of the described method is publicly available [34].

## References


[1] International Diabetes Federation. (2015) IDF diabetes atlas, 7th edn. [Online]. Available: http://www.diabetesatlas.org

[2] T. Bailey, B. W. Bode, M. P. Christiansen, L. J. Klaff, and S. Alva, "The performance and usability of a factory-calibrated Flash Glucose Monitoring system," *Diabetes Technol. Ther.*, vol. 17, no. 11, pp. 787–794, Nov 2015.

[3] G. Freckmann, A. Baumstark, N. Jendrike, D. Rittmeyer, S. Pleus, and C. Haug, "Accuracy evaluation of four blood glucose monitoring systems in the hands of intended users and trained personnel based on ISO 15197 requirements," *Diabetes Technol. Ther.*, vol. 19, no. 4, pp. 246–254, Apr 2017.

[4] M. D. Breton, R. Hinzmann, E. Campos-Nanez, S. Riddle, M. Schoemaker, and G. Schmelzeisen-Redeker, "Analysis of the accuracy and performance of a Continuous Glucose Monitoring sensor prototype: an in-silico study using the UVa/Padova type 1 diabetes simulator," *J Diabetes Sci Technol*, vol. 11, no. 3, pp. 545–552, May 2017.

[5] A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli, "Modeling transient disconnections and compression artifacts of continuous glucose sensors," *Diabetes Technol. Ther.*, vol. 18, no. 4, pp. 264–272, Apr 2016.

[6] Z. Mahmoudi, D. Boiroux, M. Hagdrup, K. Nørgaard, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Application of the continuous-discrete extended Kalman filter for fault detection in continuous glucose monitors for type 1 diabetes," in *2016 European Control Conference (ECC)*, June 2016, pp. 714–719.

[7] E. Carson and C. Cobelli, *Modelling Methodology for Physiology and Medicine*. Elsevier, 2014.

[8] Q. Wang, P. Molenaar, S. Harsh, K. Freeman, J. Xie, C. Gold, M. Rovine, and J. Ulbrecht, "Personalized state-space modeling of glucose dynamics for type 1 diabetes using continuously monitored glucose, insulin dose, and meal intake: An extended Kalman filter approach," *Journal of diabetes science and technology*, vol. 8, no. 2, pp. 331–345, 2014.

[9] R. Visentin, C. Man, and C. Cobelli, "One-day Bayesian cloning of type 1 diabetes subjects: Towards a single-day UVa/Padova type 1 diabetes simulator," *IEEE Trans Biomed Eng*, Feb 2016.

[10] D. Boiroux, M. Hagdrup, Z. Mahmoudi, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Model identification using continuous glucose monitoring data for type 1 diabetes," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 759–764, 2016.

[11] R. Gondhalekar, E. Dassau, and F. J. Doyle, "Periodic zone-MPC with asymmetric costs for outpatient-ready safety of an artificial pancreas to treat type 1 diabetes," *Automatica*, vol. 71, no. Supplement C, pp. 237 – 246, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109816301388

[12] D. T. Finegood and R. N. Bergman, "Optimal segments: a method for smoothing tracer data to calculate metabolic fluxes," *American Journal of Physiology-Endocrinology And Metabolism*, vol. 244, no. 5, pp. E472–E479, 1983.

[13] O. M. Staal, S. Sælid, T. V. Karstang, and Ø. Stavdahl, "Kalman smoothing of glucose data applied to Partial Least Squares modeling of non-invasive near-infrared measurements," in *Advanced Technologies & Treatments for Diabetes*, 2017, Poster. Available: http://bit.ly/2kw8u5H.

[14] ISO, "In vitro diagnostic test systems – requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus," International Organization for Standardization, Geneva, CH, Standard 15197:2015, May 2015.

[15] F. Gustafsson, *Statistical Sensor Fusion*. Studentlitteratur AB, Lund, 2010.

[16] H. E. Rauch, C. Striebel, and F. Tung, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[17] M. Vettoretti, A. Facchinetti, G. Sparacino, and C. Cobelli, "A model of self-monitoring blood glucose measurement error," *J Diabetes Sci Technol*, vol. 11, no. 4, pp. 724–735, Jul 2017.

[18] B. P. Kovatchev, W. L. Clarke, M. Breton, K. Brayman, and A. McCall, "Quantifying temporal glucose variability in diabetes via continuous glucose monitoring: mathematical methods and clinical application," *Diabetes Technol. Ther.*, vol. 7, no. 6, pp. 849–862, Dec 2005.

[19] 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, no. 6, p. E667, 1979.

[20] C. Cobelli, C. Dalla Man, G. Toffolo, R. Basu, A. Vella, and R. Rizza, "The oral minimal model method," *Diabetes*, vol. 63, no. 4, pp. 1203–1213, Apr 2014.

[21] S. S. Kanderian, S. A. Weinzimer, and G. M. Steil, "The identifiable virtual patient model: comparison of simulation and clinical closed-loop study results," *J Diabetes Sci Technol*, vol. 6, no. 2, pp. 371–379, Mar 2012.

[22] R. Hovorka, V. Canonico, L. J. Chassin, U. Haueter, M. Massi-Benedetti, M. Orsini Federici, T. R. Pieber, H. C. Schaller, L. Schaupp, T. Vering, and M. E. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiol Meas*, vol. 25, no. 4, pp. 905–920, Aug 2004.

[23] Y. Ruan, M. E. Wilinska, H. Thabit, and R. Hovorka, "Modelling day-to-day variability of glucose-insulin regulation over 12-week home use of closed-loop insulin delivery," *IEEE Trans Biomed Eng*, Sep 2016.

[24] N. Magdelaine, L. Chaillous, I. Guilhem, J. Y. Poirier, M. Krempf, C. H. Moog, and E. L. Carpentier, "A long-term model of the glucose-insulin dynamics of type 1 diabetes," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1546–1552, June 2015.

[25] C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, and C. Cobelli, "The UVa/Padova type 1 diabetes simulator: new features," *J Diabetes Sci Technol*, vol. 8, no. 1, pp. 26–34, Jan 2014.

[26] J. T. Sorensen, "A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes," Ph.D. dissertation, Massachusetts Institute of Technology, 1985.

[27] M. E. Wilinska, M. Bodenlenz, L. J. Chassin, H. C. Schaller, L. A. Schaupp, T. R. Pieber, and R. Hovorka, "Interstitial glucose kinetics in subjects with type 1 diabetes under physiologic conditions," *Metab. Clin. Exp.*, vol. 53, no. 11, pp. 1484–1491, Nov 2004.

[28] M. Breton and B. Kovatchev, "Analysis, modeling, and simulation of the accuracy of continuous glucose sensors," *Journal of diabetes science and technology*, vol. 2, no. 5, pp. 853–862, 2008.

[29] K. Rebrin, N. F. Sheppard, and G. M. Steil, "Use of subcutaneous interstitial fluid glucose to estimate blood glucose: revisiting delay and sensor offset," *J Diabetes Sci Technol*, vol. 4, no. 5, pp. 1087–1098, Sep 2010.

[30] G. Schmelzeisen-Redeker, M. Schoemaker, H. Kirchsteiger, G. Freckmann, L. Heinemann, and L. Del Re, "Time delay of CGM sensors: Relevance, causes, and countermeasures," *J Diabetes Sci Technol*, vol. 9, no. 5, pp. 1006–1015, Sep 2015.

[31] A. Facchinetti, G. Sparacino, and C. Cobelli, "Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies," *J Diabetes Sci Technol*, vol. 4, no. 1, pp. 4–14, Jan 2010.

[32] L. Biagi, C. M. Ramkissoon, A. Facchinetti, Y. Leal, and J. Vehi, "Modeling the error of the Medtronic Paradigm Veo Enlite glucose sensor," *Sensors (Basel)*, vol. 17, no. 6, Jun 2017.

[33] A. Facchinetti, S. Del Favero, G. Sparacino, and C. Cobelli, "Model of glucose sensor error components: identification and assessment for new Dexcom G4 generation devices," *Med Biol Eng Comput*, vol. 53, no. 12, pp. 1259–1269, Dec 2015.

[34] O. M. Staal. (2017) kalman-smoothing-glucose, GitHub repository. [Online]. Available: https://github.com/omstaal/kalman-smoothing-glucose

[35] A. Facchinetti, G. Sparacino, and C. Cobelli, "Enhanced accuracy of continuous glucose monitoring by online extended Kalman filtering," *Diabetes Technol. Ther.*, vol. 12, no. 5, pp. 353–363, May 2010.

## A.2 Differences between Flash Glucose Monitor and fingerprick measurements

Journal article published in *MDPI Biosensors*

O. M. Staal, H. M. U. Hansen, S. C. Christiansen, A. L. Fougner, S. M. Carlsen, and Ø. Stavdahl, "Differences between flash glucose monitor and fingerprick measurements," *Biosensors*, vol. 8, no. 4, 2018.

Full text is included on the next page.

The paper is also available at

https://doi.org/10.3390/bios8040093

biosensors

MDPI

Article

# Differences Between Flash Glucose Monitor and Fingerprick Measurements

**Odd Martin Staal [1,2,*], Heidi Marie Umbach Hansen [2], Sverre Christian Christiansen [3,4], Anders Lyngvi Fougner [1], Sven Magnus Carlsen [3,4] and Øyvind Stavdahl [1]**

1 Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway; anders.fougner@ntnu.no (A.L.F.); oyvind.stavdahl@ntnu.no (Ø.S.)
2 Prediktor Medical, 1630 Gamle Fredrikstad, Norway; heidi@prediktor.no
3 Department of Clinical and Molecular Medicine, NTNU, 7491 Trondheim, Norway; sverre.christiansen@ntnu.no (S.C.C.); sven.carlsen@ntnu.no (S.M.C.)
4 Department of Endocrinology, St. Olavs University Hospital, 7491 Trondheim, Norway
* Correspondence: odd.m.staal@ntnu.no

Received: 27 August 2018; Accepted: 15 October 2018; Published: 17 October 2018

**Abstract:** Freestyle Libre (FL) is a factory calibrated Flash Glucose Monitor (FGM). We investigated Mean Absolute Relative Difference (MARD) between Self Monitoring of Blood Glucose (SMBG) and FL measurements in the first day of sensor wear in 39 subjects with Type 1 diabetes. The overall MARD was 12.3%, while the individual MARDs ranged from 4% to 25%. Five participants had a MARD $\geq$ 20%. We estimated bias and lag between the FL and SMBG measurements. The estimated biases range from $-1.8$ mmol/L to 1.4 mmol/L, and lags range from 2 min to 24 min. Bias is identified as a main cause of poor individual MARDs. The biases seem to persist in days 2–7 of sensor usage. All cases of MARD $\geq$ 20% in the first day are eliminated by bias correction, and overall MARD is reduced from 12.3% to 9.2%, indicating that adding support for voluntary user-supplied bias correction in the FL could improve its performance.

**Keywords:** blood glucose; measurement; error analysis; continuous glucose monitor; flash glucose monitor; self monitoring of blood glucose

## 1. Introduction

People with diabetes need to control their blood glucose level to be as close as possible to the normal range, in order to avoid acute and chronic consequences of the disease. Continuous glucose monitoring (CGM) is an important tool for people with diabetes, primarily to detect potentially dangerous blood glucose levels and to assist in insulin bolusing. Secondarily, CGMs help patients with diabetes to understand the dynamics of their blood glucose levels, e.g. learning which foods give what glucose responses, which activities or situations trigger glucose fluctuation, or how their glucose level varies overnight. CGMs are so-called minimally invasive glucose measurement devices, meaning that they have a small electrochemical sensor inserted under the skin for the duration of the sensor wear, which at present is between seven and 14 days [1].

Due to the diffusion time of glucose from the capillaries to the subcutaneous interstitial fluid and diffusion across sensor membranes, CGM measurements are delayed compared to glucose measured in blood [2–5]. Consequently, CGM systems have a disadvantage compared to direct blood glucose measurements, and lag is a known issue with present CGM systems that users need to be aware of. However, recent CGM systems provide accurate results despite this lag, achieving Mean Absolute Relative Difference (MARD) of below 10% when comparing against blood glucose measurements [6].

A recent addition to the family of glucose monitoring devices is the Flash Glucose Monitor (FGM) [7], of which there is currently only one system on the market—Abbott's Freestyle Libre (FL) [8].

The FL uses the same minimally invasive electrochemical sensing principle as conventional CGMs. A main difference between FGM and conventional CGM is in its usage; FL only provides a reading when the user scans the sensor using a hand held scanner device or a Near Field Communication (NFC) enabled smart phone. The current glucose level along with historic glucose data for the last 8 h is displayed on the scanner/smart phone. A consequence of this user-initiated data transfer between sensor and display unit is that the FL cannot provide alarm functionality, like CGMs do. Although the common usage pattern of FGM is different from CGM, in parts of this paper we treat the FGM data as if it is from a CGM, since the measurement principle is common between the systems, and our FGM data are frequently sampled.

CGMs have until recently required calibration against finger capillary blood measurements provided by Self Monitoring of Blood Glucose (SMBG) meters. Calibration is usually performed twice daily, and is intended to combat drift and minimize bias. However, the user supplied calibration values may also impair the overall accuracy of the system, e.g., if the user does not input the SMBG values correctly, the calibration is performed in periods of high glucose variation, or if the SMBG measurements are not correct. Calibration of CGMs is discussed in detail by Acciaroli et al. [9].

The FL is factory calibrated, meaning that the user does not need to calibrate it with SMBG measurements during the sensor wear. This is marketed as a profound improvement in the world of CGMs, and is likely responsible for a great deal of the popularity of the FL system, since the twice-daily SMBG calibration of most of its current CGM competitors is a burden to the users. However, the FL does not provide a means to bias-correct the measurements even if the user wants to. The FL has a built-in SMBG meter in the scanning device that would easily enable bias correction, but according to Bailey et al. [7], the built-in SMBG reader of the FL has no influence on the FGM readings. DexCom G6 is another calibration-free system that has recently been launched.

The introduction of calibration-free systems like FL eliminates the risk of failed calibration due to user error, but it comes with a price: it eliminates a mechanism that can reduce or ideally remove an inherent sensor bias due to either inaccurate factory calibration or some sensor-person interaction effect.

CGM/FGM data are useful for research purposes. For some uses it is important to identify the a model of the error in the datax. This has been done for conventional CGM systems by Facchinetti et al. [10]. The introduction of factory calibration means that a different model needs to be applied for systems like the FL. Characterization experiments like the one reported in this paper are needed to accomplish this. Being familiar with the characteristics of FGM signals and any limitations or challenges related to factory calibration is also important to patients and health care professionals.

In this paper we investigate the accuracy of FL measurements compared to SMBG measurements, and we investigate in detail the characteristics of the errors in the FL data, focusing on biases and lags in the FL glucose estimates. This has previously been requested by other researchers [11].

The current paper expands upon work presented orally at the Advanced Technologies & Treatments for Diabetes (ATTD) conference in Vienna, February 2018.

## 2. Method

### *2.1. Data Collection*

In a study of 39 individuals with Type 1 diabetes, simultaneous SMBG (Freestyle Freedom Lite, Abbott) and FGM (Freestyle Libre, Abbott) measurements were taken every 10 min. These sessions lasted from 2.5–6 h and were performed in a research ward within 24 h of FL sensor insertion and activation.

In these sessions the participants were non-fasting and used their regular insulin regime and sugary drinks or meals to manipulate their glucose level. The collected glucose responses typically had three flanks (i.e., up/down/up, or down/up/down) with approx. 5 mmol/L as the lower turning point and approx. 10 mmol/L as the upper turning point, respectively. The overall range of glucose in

our dataset as measured by SMBG was 3 mmol/L to 26 mmol/L. Four representative examples of the SMBG and FGM data resulting from these sessions are shown in Figure 1.

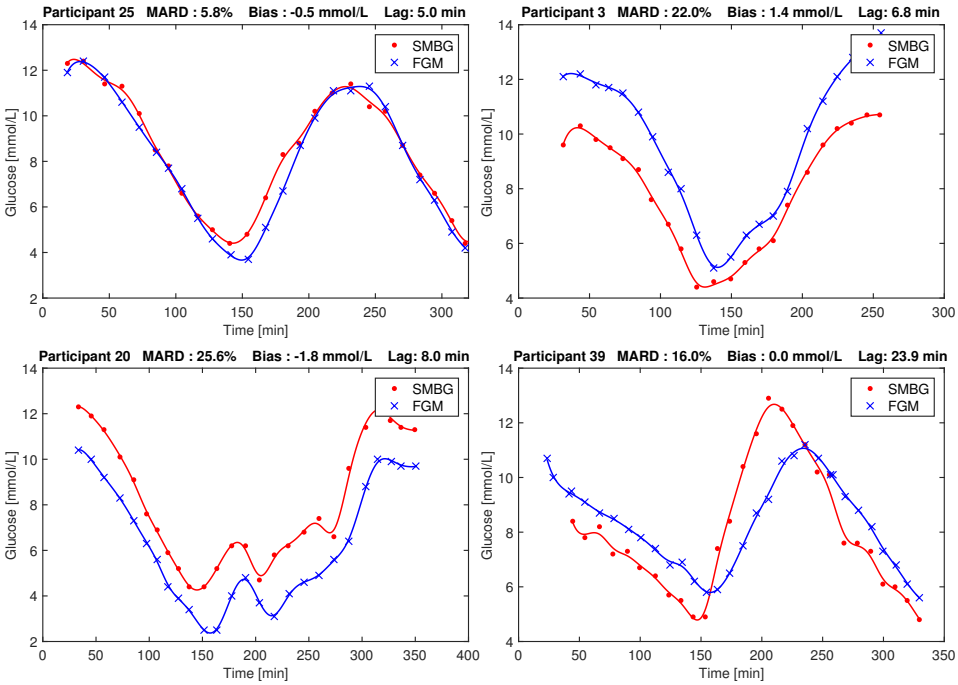



**Figure 1.** Examples of Flash Glucose Monitor (FGM) and Self Monitoring of Blood Glucose (SMBG) recordings in our study. **Top left:** Low Mean Absolute Relative Difference (MARD). **Top right:** Large positive bias. **Bottom left:** Large negative bias. **Bottom right:** Large lag. Points are measurements, and the line between points is the result of smoothing, as described by Staal et al. [12].

After the frequent-sampling session, the participants went on to use the sensor in their everyday environment for days 2–7 after insertion, performing sporadic SMBG measurements and scanning the FL at least every 8 hours. Of the 39 sensors, 11 were dislodged before the 7 days had passed. Data from these sensors were included in the analysis, implying that for these participants, the data from days 2–7 are incomplete.

All participants signed consent forms prior to the study. The study was approved by the Regional Ethics Committee (REK Midt 2016/1172). The participant demographics is given in Table 1.

**Table 1.** Participant demographics, mean (range).

| | |
|---|---|
| Sex | 14 female, 25 male |
| Age | 42 (18–72) years |
| Duration of diabetes | 23 (3–45) years |
| Body Mass Index (BMI) | 27 (21–38) kg/m$^2$ |

The paired FL and SMBG measurements are included in the supplementary data. The demographics data are not made publicly available, since they may enable identification of the study participants.

### 2.2. Data Analysis

Looking at the initial session (day 1) data, the MARD between corresponding FL scanned values ($y_{FGM}$) and SMBG data points ($y_{SMBG}$) was computed, both on an overall level and on a per individual level ($MARD_p$).

$$MARD = \frac{1}{N_{all}} \sum_{i=1}^{N_{all}} \frac{|y_{i_{FGM}} - y_{i_{SMBG}}|}{y_{i_{SMBG}}} \quad (1)$$

$$MARD_p = \frac{1}{N_p} \sum_{i=i_p}^{i_p+N_p} \frac{|y_{i_{FGM}} - y_{i_{SMBG}}|}{y_{i_{SMBG}}} \quad (2)$$

here $N_{all}$ is the overall number of paired points in our study. $i_p$ is the first measurement from participant $p$ and $N_p$ are the number of paired points for participant $p$.

A Parkes/Consensus error grid (PEG) analysis [13] was performed on an overall level, using all paired points.

The data from all runs were plotted to investigate the characteristics and reason for high MARDs, and it was seen that bias and lag effects were present in the data. Therefore, an estimator based on a Rauch-Tung-Striebel Kalman smoother [12,14,15] was implemented to estimate the bias and lag from the data, using the following dynamic model of the FGM measurements:

$$\dot{G}_i = \frac{1}{\tau_i}(G_p - G_i) + v(t) \quad (3)$$

$$y_{FGM,k} = G_{i,k} + b_{FGM} + w_k \quad (4)$$

Here, $G_p$ is plasma glucose, $G_i$ is interstitial fluid glucose, and $\tau_i$ is a time constant governing the diffusion between these compartments. This time constant also models any diffusion dynamics across the FGM sensor membranes. The process noise is modeled by $v(t)$. Furthermore, $y_{FGM,k}$ is the FGM measurement at time step k, $b_{FGM}$ is a bias constant, and $w_k$ is the measurement noise process. The parameters $b_{FGM}$ and $\tau_i$ were estimated per individual data set using the Kalman smoother and the above model, using the same noise modeling as Staal et al. [12]. The word "lag" is sometimes interpreted as a pure time delay, or a combination of a pure time delay and a time constant [16], but in this paper we use lag as a synonym for the time constant.

We investigated the effect on MARD of correcting the day 1 data for only bias, only lag, and both. The bias correction was done simply by subtracting the estimated bias $b_{FGM}$ per participant from all measurements $y_{i_{FGM}}$ from that participant, producing $y_{FGM,Bcorr}$. The lag correction is more complicated, and can be done in different ways [17,18]. We used Equations (3) and (4) in a Kalman smoother that use the participant lag, bias and FGM data to produce lag- and bias-corrected FGM measurements, $y_{FGM,BLcorr}$. By not providing the bias to the smoother we can produce an only lag-corrected signal, $y_{FGM,Lcorr}$, to investigate the effect on MARD of only correcting for lag. By providing neither bias nor lag information, we can produce an uncorrected signal $y_{FGM,smoothed}$ that has been subjected to all the processing of our method but has not really corrected for anything. This latter signal was used to get an idea of how much MARD is influenced by the smoothing introduced by our method. Both Kalman smoothers described above for parameter estimation and state estimation are working in fixed interval mode, i.e., they use all data in the data set, thus this is an offline method. A detailed description of the state estimation Kalman filter is provided in Appendix A.

The bias and lag estimates computed by the Kalman smoother are accurate because of the many points used. Basing a real-time bias estimation on this method is not practical due to the many SMBG measurements required. In a practical bias calibration, only one or two data points per day should be used, as in normal CGM calibration regimes. We therefore also computed a 1-point bias correction, finding a bias $b_{FGM,1p}$ based on the error in the first paired data point from each day 1 session. Correcting for this bias produces a signal $y_{FGM,Bcorr1p}$. We also computed a 2-point bias

$b_{FGM,2p}$ based on the mean of the errors in the first and last paired point from each session, producing the signal $y_{FGM,Bcorr2p}$. The different biases were used to correct data from day 1 and days 2–7. A block diagram providing an overview of the method we employed is shown in Figure 2.

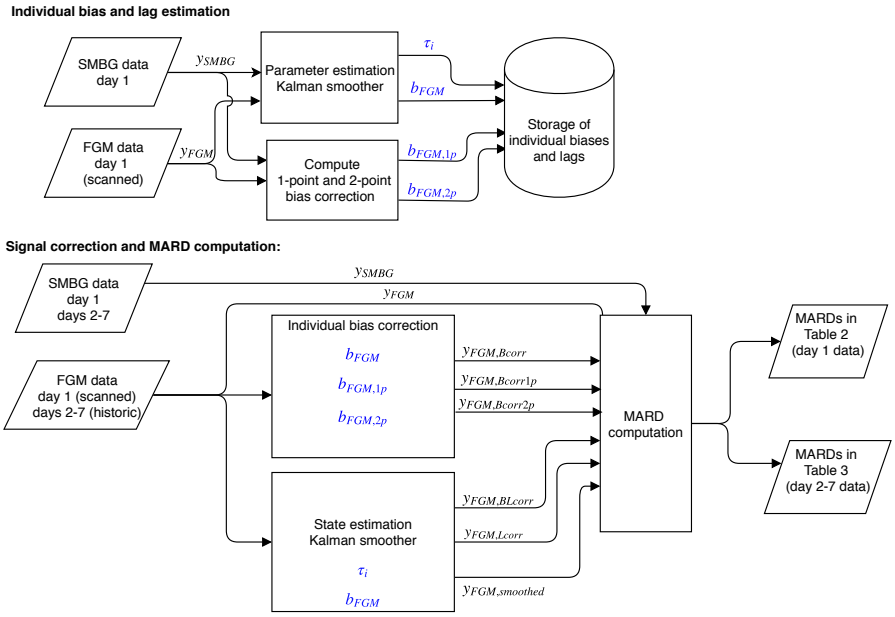



**Figure 2.** An overview of the signal processing. Items in blue are individual parameters estimated from day 1 data, and used to correct both day 1 and days 2–7 data for that individual. The parameters and signals shown are described in the main text, Section 2.2.

To investigate whether the biases seen in the first day of use persisted throughout the use of the sensor, we performed an analysis of the data from days 2–7. Participants were grouped based on the bias they had in the first day. Individuals experiencing more than 1 mmol/L bias in day 1 were assigned to the positive bias group. Those that had less than −1 mmol/L bias were assigned to the negative bias group. The remaining participants were assigned to the unbiased group. The data were plotted based on this grouping, and the mean and standard deviation was computed in each group, for both day 1 and day 2–7 data. A *t*-test was performed to determine if the grouped means in days 2–7 were significantly different from the unbiased group. We also used a t-test to determine if the group means changed from day 1 to days 2–7. A similar analysis for the lag was not possible, due to the sparsity of data in days 2–7.

Finally, we checked if participant characteristics were associated with the observed biases and lags. We investigated BMI, height, weight, age, sex, duration of diabetes, use of blood pressure medication, use of any other medication and whether or not the sensor fell off during days 2–7. For binary variables (sex, medication use, sensor fall-off) we used a *t*-test at $p = 0.05$ to test if there was a difference in the mean between groups. For continuous variables we performed a linear fit and estimated the 95% confidence interval (CI) of the correlation coefficient. If this CI spanned 0, the correlation was considered insignificant.

As reported by Pleus et al. [19], the FL generates two time series of glucose estimates that are available in its export file. Scanned glucose is the instantaneous glucose the sensor estimates at the time when the sensor is scanned. Historic glucose is generated by the sensor every 15 min independently of the scanning. Up to 8 h of historic data are stored in the sensor and transferred to the display unit as part of a scan. We used the scanned glucose values in the analysis of the high-frequency

sampling session (day 1), since SMBG measurements and FGM scans were performed simultaneously in this session. In the analysis of days 2–7, we used historic glucose values since there were too few occurrences of concurrent FL scans and SMBG measurements during the normal use in an everyday environment. The historic glucose values were interpolated as described by Staal et al. [12] to enable matching with the SMBG measurements.

## 3. Results

### *3.1. Overall MARD Analysis in Day 1 Data*

There were 1053 paired measurements from the frequent-sampling sessions in day 1, having an overall MARD of 12.3%. A Parkes/Consensus error grid (PEG) analysis found all measurements to be within zones A+B, with 81.7% in zone A. The PEG plot is shown in Figure 3.

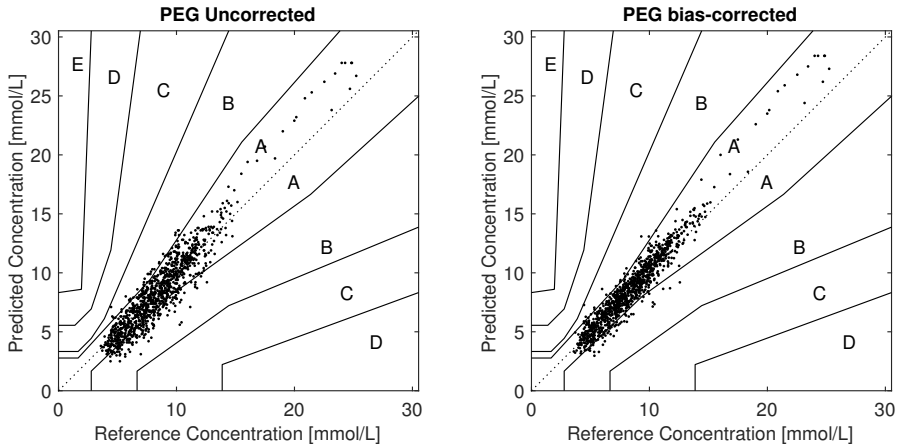



**Figure 3.** Parkes/Consensus Error Grid plots. **Left:** Uncorrected **Right:** Bias corrected Bias correction per individual makes 91.2% of the paired points lie in zone A.

### *3.2. Day 1 Individual Participant MARD, Bias and Lag Analysis*

There were on average 27 paired measurements per individual. Individual participant MARDs varied between 4.0% and 25.5%. A total of 18 participants had MARD <10%, however five had MARD $\geq$ 20%. Biases varied between $-1.8$ mmol/L, and +1.4 mmol/L, with a mean of $-0.4$ mmol/L. Time constants ranged from 2 min to 24 min, with a mean of 9 min. Histograms of individual MARDs, biases and lags on day 1 are shown in Figure 4. The most extreme cases of bias and lag are shown as time series in Figure 1. Other effects besides bias and lag are seen in some sets, e.g., overestimation in periods of high glucose (data not shown).

If the data are corrected for the biases found, the overall MARD on day 1 falls significantly ($p < 10^{-6}$) from 12.3% to 9.2%. The participant with the largest MARD (25.5%) got a MARD of 6.4% with bias correction. The maximal individual MARD after bias correction is 17%. Of the 39 participants, 27 got a MARD below 10% after bias correction. For nine participants, the bias correction led to an increase in MARD, with one participant getting an increase of more than 3 percentage points. The MARD after bias correction is well correlated with the estimated lag; see Figure 5.

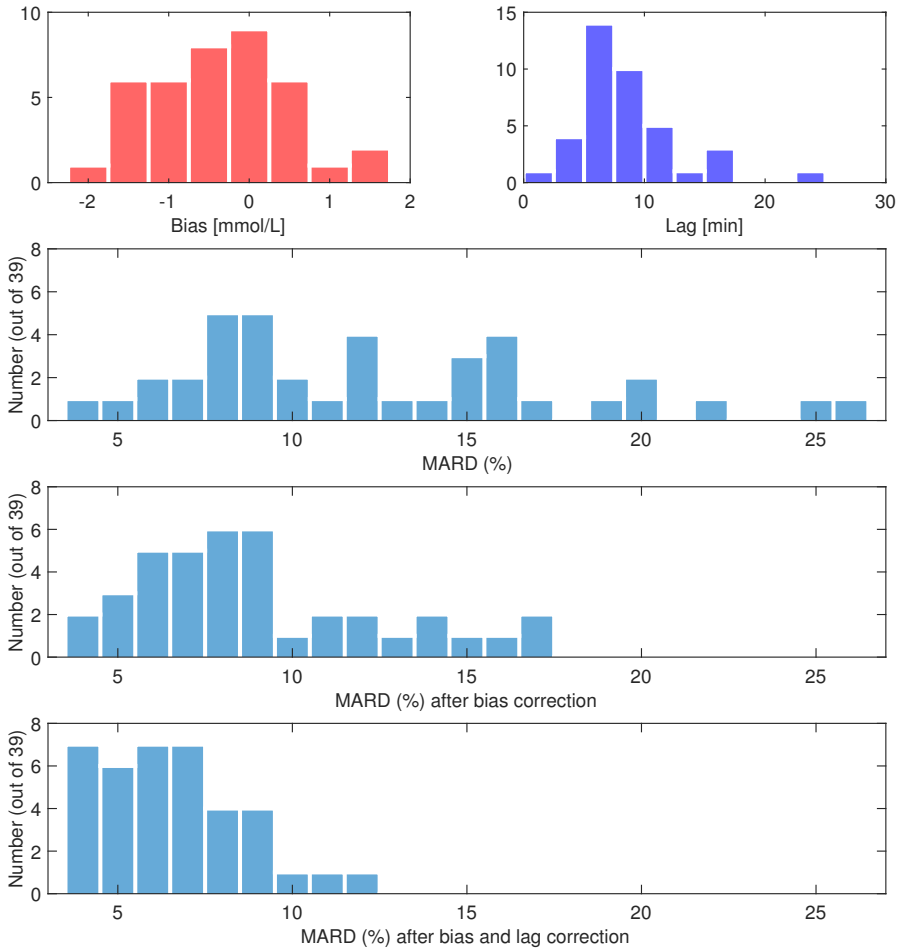



**Figure 4.** Histograms from day 1. **Upper left:** Estimated individual biases, $b_{FGM}$. **Upper right:** Estimated individual time constants, $\tau_i$. **Second row:** individual MARDs, uncorrected. **Third row:** individual MARDs, bias corrected. **Bottom row:** individual MARDs, bias and lag corrected.

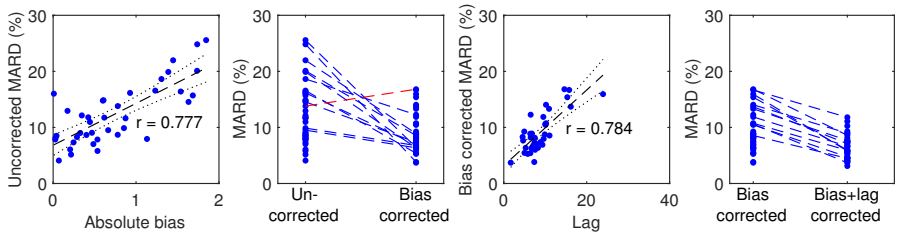



**Figure 5.** Effect of bias and lag correction. **Left:** Uncorrected MARD plotted against absolute bias. **Middle left:** MARD changes as a result of bias correction with individual participant tracing. Points belonging to the 12 participants whose MARD changed by more than 3 percentage points are connected by a dotted line (blue when MARD is reduced; red when MARD is increased). **Middle right:** Bias corrected MARD plotted against lag. **Right:** MARD changes as a result of lag correction. Regression lines are plotted in dashed black and 95% CIs in dotted black.

Compensating only for lag gave an overall MARD of 11.7%, while compensating for both bias and lag resulted in an overall MARD of 6.6%. The overall MARD, MAD and PEG zone A and A+B of uncorrected and corrected FGM data are listed in Table 2. Similar analyses for days 2–7 are given in Table A1 in Appendix B.

**Table 2.** Influence on overall MARD of different signal processing of day 1 data to correct for bias and lag.

| Signal Processing | Symbol | MARD (%) | MAD (mmol/L) | PEG zone A/A+B (%) |
|---|---|---|---|---|
| None (raw FGM scans vs SMBG) | $y_{FGM}$ | 12.3 | 1.0 | 81.7/100 |
| Bias corrected, multipoint | $y_{FGM,Bcorr}$ | 9.2 | 0.8 | 91.2/100 |
| Bias corrected, 1-point | $y_{FGM,Bcorr1p}$ | 11.4 | 0.9 | 83.0/100 |
| Bias corrected, 2-point | $y_{FGM,Bcorr2p}$ | 9.7 | 0.7 | 87.7/100 |
| Bias and lag corrected, multipoint | $y_{FGM,BLcorr}$ | 6.6 | 0.5 | 97/100 |
| Only lag corrected, multipoint | $y_{FGM,Lcorr}$ | 11.7 | 0.9 | 81.5/100 |
| Only smoothed, multipoint | $y_{FGM,smoothed}$ | 11.9 | 1.0 | 82.2/100 |

### 3.3. Persistence of Biases through Days 2–7

There were 356 data point pairs from days 2–7, these were plotted using the negative, positive and unbiased grouping based on day 1 biases, as described in Section 2.2. The plot is shown in Figure 6. The results of the grouped analysis of these data are given in Table 3. The difference of the mean between paired points from days 2–7 from each biased group compared to the paired points from the unbiased group is significant, judging by a *t*-test ($p < 10^{-6}$). Within each group, the difference of the mean between paired points in day 1 and day 2–7 changes insignificantly, except for the negative bias group, which has a significant change ($p < 10^{-6}$) moving towards zero.

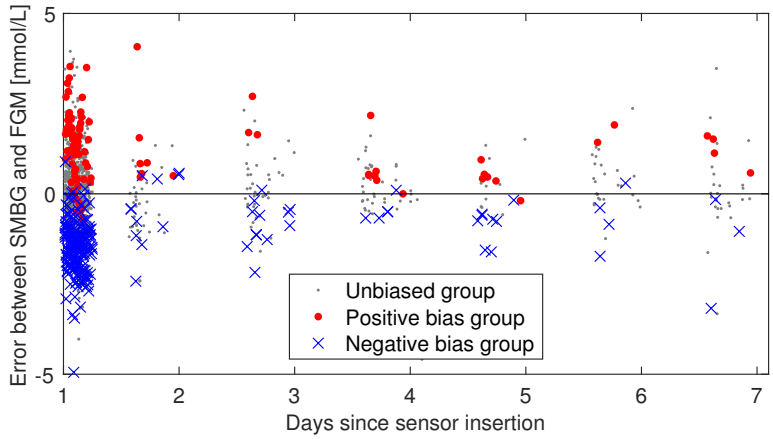



**Figure 6.** Persistence of bias in days 2–7. The individual sensor errors (SMBG–FGM) are plotted against time and colored according to what bias group the participant was in at day 1. If the bias observed in day 1 is not persistent, one would expect the blue and red data points to start mixing as time progresses. The day 1 session is included in this plot, but the analysis of bias persistence considered data from days 2–7 separately from day 1 data.

**Table 3.** Group summaries from bias persistence analysis. Mean: Mean of the errors [mmol/L]. SD: Standard deviation of the errors [mmol/L]. N: number of paired points. †: Group mean in days 2–7 is significantly different from the group mean in day 1. ∗: Group mean is significantly different from the unbiased group mean (days 2–7).

| Group | Number of Participants | Mean ± SD (N) Day 1 | Mean ± SD (N) Days 2–7 |
|---|---|---|---|
| Positive bias | 3 | 1.28 ± 1.03 (96) | 1.35 ± 1.44 (35) ∗ |
| Unbiased | 28 | 0.09 ± 1.06 (755) | 0.14 ± 1.00 (270) |
| Negative bias | 8 | −1.45 ± 0.87 (202) | −0.74 ± 0.73 (51) †∗ |

### *3.4. Participant Factors vs. Bias and Llag*

In the 18 comparisons we did to investigate correlation between bias, lag and the factors listed in Section 2.2, we found only one correlation with $p < 0.05$. This was between lag and age ($p = 0.046$). Applying a Bonferroni correction, the p-level needed to achieve significance is 0.0027, thus this lag-age correlation is also considered insignificant.

## 4. Discussion

Several studies have reported an overall performance of the FL comparable to that of other state-of-the-art CGM systems [7,20–29]. Our main results are in line with previous observations as we find an overall MARD of 12.3% and half of the participants having a MARD at or below 10%. However, five of our 39 participants experienced a MARD at or above 20%. Several of the previous FL performance studies also report individual MARDs, where some are as large as we observed in our study [7,22,27]. The studies by Ólafsdóttir et al. [21] and Alsaffar et al. [24] report a bias in the measurements, but on an overall level only. To the best of our knowledge, the present study is the first to report on individual bias and lag issues as reasons for the high variation in individual MARDs in FL.

We observe that a bias is present in several of the the worst-performing sensors, and that bias correction significantly improves MARDs, both overall and on the individual level for most participants, see Figure 5. We can speculate that either the factory calibration is not accurate for these sensors, or that these biases arise from person-sensor interaction effects that are not predictable from any of the participant factors we investigated.

For some participants, bias correction increased MARD slightly, e.g., the participant shown in red in Figure 5. There are several reasons why this may happen. Firstly, since we only correct the bias, and not the lag, the MARD may go up. Secondly, the bias is found in a way that does not try to optimize MARD. Thirdly, since only bias and lag are modeled in our method, the bias estimate may be inaccurate if other errors are present in the data. We do see effects in some of the data sets that are not explainable by only bias and lag. For instance, we saw some cases of overestimation of high glucose values that looks like a gain issue. The data from participant 3 shown in the upper right of Figure 1 shows this tendency. An increase in MARD after bias correction can happen in datasets where effects other than bias and lag are present, or in datasets where the lag is large. The former could lead to inaccurate bias/lag estimation by the Kalman smoother, because the model it uses (Equations (3) and (4)) accounts for no other effects than bias and lag. The latter, large lag, may be a cause of increased MARDs after bias correction, since the MARD computation penalizes deviations from low reference values more than the same deviation from high reference values. The bias correction may well lead to less alignment of the low glucose values to give better alignment of the high glucose values, which will increase MARD. In our investigation of participant factors we found only one barely significant correlation ($p = 0.046$) between age and lag, shown in Figure 5. We assume that this is a spurious correlation, firstly because we did enough comparisons to make it likely that one of them

shows significance at p = 0.05 even if the underlying data are truly uncorrelated. Secondly, the negative correlation goes against our intuition about how tissue develops with age; if anything, we would expect higher age to give more delay between blood and interstitial fluid, not less.

Our study was not designed to resolve the matter of whether the biases are linked to the sensor or the individual. This is an obvious follow up question to the present study. Answering this question would require a study with at least two FL sensors per person, for instance as in the study performed by Freckmann et al. [29], where each of the 20 participants wore two FL sensors and two DexCom G5 sensors (DG5), and SMBG was measured every hour during three clinic visits. A larger between-sensor discrepancy was seen in FL than in DG5, as measured by Precision Absolute Relative Difference (PARD), and it was seen that four of 20 participants had a PARD $\geq$ 15. This is supported by the biases we observed in our study, and suggests that the biases might follow the sensor rather than the individual. Further, we can speculate that since DG5 is a conventional CGM which requires calibration SMBG measurements twice daily, the increased PARD of DG5 over FL could be caused by factory calibration issues in the FL. In a study performed by other researchers in our group [30], it was seen that the in vitro responses of four different sensors had a tendency to be offset from each other. This result also points at the sensor as the source of the bias rather than some sensor/participant interaction effect.

Insights into the factory calibration process of FL is provided by Hoss and Budiman [31]. The FL factory calibration is achieved through low sensor-to-sensor variability within a sensor lot, and performing in vitro tests of a sample of sensors within a lot to produce a factory calibration valid for all sensors of that lot. The authors state that "The factory calibration process is based on the assumption that the in vitro sensor sensitivity predicts the in vivo sensor response". While this assumption may be true on an overall level where data from 10 or more sensors and individuals are pooled and averaged, it allows having significant errors on the individual level. This may be what we are observing.

A limitation of our study is that our data set did not include blood glucose measurements using lab glucose analyzers, e.g., YSI 2300 Stat Plus. Had such data been recorded we would have been able to eliminate SMBG measurement bias as a possible cause of the FGM vs SMBG biases. Since SMBG measurement errors have been reported to not be correlated in time [32], and SMBG measurements are not reported to exhibit biases of the magnitudes we observed [33,34], we are inclined to believe that the bias is a problem with the FL measurements, not the SMBG measurements.

Figure 6 and Table 3 indicate that the individual biases persist in days 2–7, however, the negative bias group seems to be moving towards zero. More frequent data sampling in days 2–7 from more participants would be needed to confirm that this is the case. To confidently answer the question of how the bias develops throughout a sensor session, bias must be accurately determined per day, requiring simultaneous estimation of the lag, like we did for day 1. To accomplish this we would have needed several frequent sampling sessions throughout the 14-day sensor lifetime. An alternative is to minimize the influence of lag by making sure that SMBG measurements used for calibration are taken in periods of low glycemic variation, as is the recommended practice in calibration of CGMs. Our data set contains too few such periods.

The performance of the bias correction on day 2–7 data indicate no significant improvement from any of the bias correction methods we applied, see Table A1 in Appendix B. The data from day 1 in our study is unsuitable for finding an accurate bias using only 1- or 2-point calibration, due to the lack of periods of low glycemic variation. This could be part of the reason why these corrections failed to improve MARD. However, the "best possible" multipoint bias estimated using a Kalman smoother also failed to improve MARD in days 2–7. This could be an indication that the bias changes over time, which seems to be the case for the negative bias group of participants (see Table 3). If so, a bias correction based on day 1 data would lead to over-correction in days 2–7, which seems to be the case at least for the negative bias group, see Figure A2 in Appendix B. The non-improving MARDs could also be caused by lag in the data. However, when we tried to also correct for the lag observed in day

1, this did not improve MARDs in days 2–7 (see Table A1), indicating that what is measured in day 1 is insufficient to correct the situation in days 2–7. Another explanation could lie in the difference between "scanned" and "historic" glucose values in the Freestyle Libre, as we based the bias estimate on scanned values from day 1 but correct historic values in days 2–7. Further research is needed to answer these questions, using more frequent SMBG sampling of data in days 2–7.

Assuming that the bias and lag observed in day 1 stays constant throughout the sensor lifetime seems like an invalid assumption to make. There are several physiological and technical reasons why both lag and bias may vary during the sensor lifetime, some of which are:

1. The insertion of the sensor into the interstitial fluid introduces local trauma to and/or minor bleeds in the tissue around the sensor, altering the glucose flow, thus making the sensor less accurate in the time immediately after insertion [1].
2. Biofouling of the sensor contributes to making sensor characteristic changes over the wear time likely.
3. On the technical side, the electrochemical sensor may suffer from non-physiological drift in the initial period of sensor wear [9].
4. Effects like Pressure Induced Sensitivity Attenuation (PISA) [35] may be present in the FGM data.

In addition, SMBG measurement outliers can occur, which will be difficult to detect and compensate for when there are only a few SMBG data points per person per day. Increasing the number of participants would also be advantageous in order to get sufficient statistical power to make a conclusion about the bias development with time.

Unless the factory calibration of FL can be improved to eliminate the bias errors we observe, voluntary user-supplied bias correction would be a desirable functionality in FL for patients and researchers alike. This correction could be done as in normal CGM systems, e.g., using one or two SMBG measurements per day (but in the case of FL, they could be voluntary). Such measurements should ideally be taken at periods of low glucose variability to minimize the influence of lag on the bias correction. The development of the bias over the 14-day life of the sensor is not known, so a bias correction of the whole period based on only one or two SMBG measurements from day 1 is not necessarily the right thing to do. Most studies of FL performance report that MARDs improve over time, suggesting that biases may be decreasing over time. This seems to be the case also in our data, at least for the group of participants that had a large negative bias in day 1. If so, a bias correction based only on data from the first day of use could be detrimental to the performance of the sensor in the subsequent days, as it could give an over-correction. The best way to include a voluntary bias correction in real-time use of the device is not clear. The methods that have been used by CGM systems for this purpose [9] are likely applicable, however these require daily calibration measurements. Researchers planning to use FL in studies should consider adding mechanisms for post-study bias correction of FL data in the study designs, for instance by including sufficient reference measurements to evaluate the bias at the beginning, middle and end of the study.

Bias is emphasized in this work, since it is potentially easily corrected for in real time. However, lag is also a significant reason for high MARDs, as seen in Figure 5. The individual MARDs remaining after bias correction are well correlated with the estimated lags. Correction of the lag gives a further improvement in overall MARD, however the bias contribution to MARD dominates, and lag correction only has an effect on MARD when the bias is also removed. Correction of bias has been commonplace in CGMs, through their daily calibration against SMBG measurements. Real-time correction of lag is possible in theory, but only if the lag is a pure time constant that does not change significantly over time. If instead a pure time delay is present, causality prohibits real-time correction of the lag. Practically feasible correction of a time constant requires quite noise free FGM data to avoid introducing new error by the lag correction. Lag correction also needs knowledge of the lag per participant and sensor, which requires combining FGM/CGM data with SMBG data with a similar sampling frequency as we used in day 1 of our study. This is practically unfeasible and not acceptable to most users, especially if it must be repeated for every new sensor.

Finally, it should be acknowledged that most would consider CGMs with MARD values greater than 20% unsuited for use by patients to guide them on their actual glucose levels, and even less so in assisting on deciding insulin doses. Consequently, both patients and health care personnel should be informed that some of the FL sensors have this limitation. Large positive biases as those experienced by three of our participants are particularly problematic, since they represent a risk of failing to detect hypoglycemia or impending hypoglycemia. Since our analysis is based mainly on data from day 1 of the sensor wear, which is known to be less accurate than data from subsequent days of wear, in a sense it represents a worst case analysis of the Freestyle Libre performance. It is still important for users and caretakers to know in what way the sensor could be inaccurate also in the first day of wear, since the FL presents glucose estimates to the user 60 min after sensor insertion without any warning to the user that the results are more inaccurate in day 1.

## 5. Conclusions

The observed overall MARD between SMBG and FL in the first day of use was 12.3%. However, MARDs in individual participants ranged from 4% to 25%. Many of the high individual MARD cases are caused primarily by bias. Lag and other effects are also present. The biases seem to persist beyond day 1 of wearing the FL. The FL is factory calibrated, and manual bias correction by the user is not possible. Our data indicate that the manufacturer and some patients could benefit from introducing a *voluntary* calibration mechanism in the FL, which could result in an improved MARD for some users. This calibration should likely be based on the same kind of principles as those in conventional CGM systems, i.e., using SMBG measurements from periods of low glycemic variation. Researchers using the FL may gain from designing their studies to allow for an external bias correction. Patients and health care personnel should be informed about the risks of measurement error in FGM devices, and how these errors may manifest themselves on an individual level. Further research is needed to determine if the bias follows the sensor or the patient, to investigate in more detail how the biases and lags evolve over the lifetime of the sensor, and possibly to exploit methods for detection and mitigation of biases and lags.

**Author Contributions:** Prediktor Medical, S.C.C. and H.M.U.H. collected data; O.M.S. analyzed the data and wrote the paper; Ø.S., A.L.F and S.M.C. supervised the research activities; All authors revised the paper.

**Funding:** This research was funded by Norges Forskningsråd (grant numbers 242167 and 248872) and Prediktor Medical. Norges Teknisk-Naturvitenskapelige Universitet (NTNU) covered the costs of open access publishing.

**Acknowledgments:** This research was part of the Double Intraperitoneal Artificial Pancreas project, which is a Centre for Digital Life Norway (digitallifenorway.org) project. The project is led by the Artificial Pancreas Trondheim research group (apt-norway.com).

**Conflicts of Interest:** The authors declare no conflict of interest. Prediktor Medical funded the REK Midt 2016/1172 study, and had a role in the study design and data collection. The funding sponsors had no influence on the study design, result interpretation or decision to publish the results of the present sub-study.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| MARD | Mean Absolute Relative Difference |
| PARD | Precision Absolute Relative Difference |
| FGM | Flash Glucose Monitor |
| CGM | Continuous Glucose Monitor |
| SMBG | Self Monitoring of Blood Glucose |
| NFC | Near Field Communication |
| FL | Freestyle Libre |
| DG5 | DexCom G5 |
| CI | Confidence Interval |

## Appendix A. Kalman Smoothing to Correct for Bias and Lag

Given an FGM signal with a known bias $b_{FGM}$ and a lag $\tau_i$, we would like to compute a signal that is corrected for both effects. We use the model of glucose dynamics given as Model 2 in the paper by Staal et al. [12], and augment it with the plasma-interstitial model given by Equations (3) and (4)), where $\tau_i$ and $b_{FGM}$ are now fixed quantities. The overall model becomes:

$$\dot{G_p} = C_r \tag{A1}$$

$$\dot{C_c} = -\frac{1}{T_d}C_c + v_c(t) \tag{A2}$$

$$\dot{C_r} = \frac{1}{T_d}(C_c - C_r) \tag{A3}$$

$$\dot{G_i} = \frac{1}{\tau_i}(G_p - G_i) \tag{A4}$$

where $T_d$ is a parameter describing the flow between the internal compartments $C_c$ and $C_r$, set to 10 min. $v_c(t)$ is process noise affecting state $C_c$. For more information about the model, and choice of parameters, process and measurement noises, see [12].

This system is linear, and can be put on the form $\dot{\mathbf{x}} = A(\tau_i)\mathbf{x}$, and a discretized version of the system is $\mathbf{x}_{k+1} = \Phi\mathbf{x}_k$, where $\Phi = e^{A(\tau_i)\Delta t}$. We used $\Delta t = 0.5$ min.

The Kalman filter equations are [15]:

$$\bar{\mathbf{x}}_k = \Phi_{k-1}\hat{\mathbf{x}}_{k-1} \tag{A5}$$

$$\bar{P}_k = \Phi\hat{P}_{k-1}\Phi^T + Q \tag{A6}$$

$$K_k = \bar{P}_k H^T(H\bar{P}_k H^T + R_k)^{-1} \tag{A7}$$

$$\hat{\mathbf{x}}_k = K_k(\mathbf{y}_k - b_{FGM} - H_k\bar{\mathbf{x}}_k) \tag{A8}$$

$$\hat{P}_k = (I - K_k H)\bar{P}_k \tag{A9}$$

here $H$ is a measurement matrix, $Q$ is the process noise covariance matrix, and $R$ is the measurement matrix. $P$ matrices are state covariance matrices propagated by the filter. We used $H = [\,0\;0\;0\;1\,]$, and we subtracted the bias $b_{FGM}$ from each measurement as part of the Kalman filter measurement update Equation (A8). The $Q$ matrix we used was all zero, except for the diagonal element corresponding to $C_c$, this was set to 0.025. $R$ (scalar in our case) is set for each measurement based on ISO 15197 limits, as described in [12].

The Rauch-Tung-Striebel (RTS) algorithm [14] makes use of stored data from the Kalman filter forward pass, namely the sequences of a priori and a posteriori state estimates $\bar{\mathbf{x}}_k$, $\hat{\mathbf{x}}_k$ and state covariance matrices $\bar{P}_k$ and $\hat{P}_k$. These are input to a backward pass that computes the smoothed estimates $\hat{\mathbf{x}}^s_k$ and $\hat{P}^s_k$ as follows [15]:

$$C_k = \hat{P}_k\Phi\bar{P}^{-1}_{k+1} \tag{A10}$$

$$\hat{\mathbf{x}}^s_k = \hat{\mathbf{x}}_k + C_k(\hat{\mathbf{x}}^s_{k+1} - \bar{\mathbf{x}}_{k+1}) \tag{A11}$$

$$\hat{P}^s_k = \hat{P}_k + C_k(\hat{P}^s_{k+1} - \bar{P}_{k+1})C^T_k \tag{A12}$$

After smoothing is completed, the bias and lag corrected signal is available as the first element in the $\hat{\mathbf{x}}^s$ signal, corresponding to the $G_p$ state. Examples of the result of the state estimation Kalman smoother is given in Figure A1.

The smoother does not introduce any additional delay in the smoothed signal, by virtue of its forward-backward nature.

The Matlab code of the Kalman smoother for use in glucose applications is publicly available, see [12].

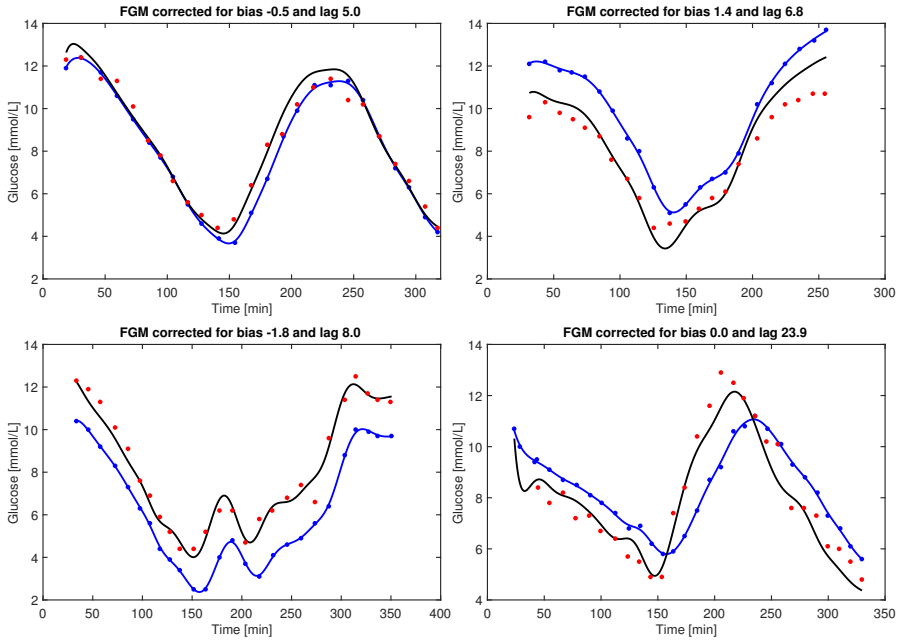



**Figure A1.** Bias and lag correction of the same data as shown in Figure 1. Here, only the FGM (**blue**) points, a bias and a lag are used as inputs to the smoother, and the black line shows the corrected signal. Red points are SMBG measurements. We see that the bias and lag correction brings the corrected FGM signal close to the SMBG measurements in all cases seen here.

## Appendix B. MARD Analysis of Day 2–7 Data Using Different Bias Correction Approaches

**Table A1.** Influence on overall MARD of different signal processing for days 2–7, using biases and lags estimated from day 1 data.

| Signal Processing | Symbol | MARD (%) | MAD (mmol/L) | PEG zone A/A+B (%) |
|---|---|---|---|---|
| None (FGM historic data vs SMBG) | $y_{FGM}$ | 11.3 | 0.8 | 87.3/99.2 |
| Bias corrected, multipoint | $y_{FGM,Bcorr}$ | 12.5 | 0.8 | 89.0/97.7 |
| Bias corrected, 1-point | $y_{FGM,Bcorr1p}$ | 13.8 | 0.9 | 82.4/99.7 |
| Bias corrected, 2-point | $y_{FGM,Bcorr2p}$ | 12.2 | 0.8 | 89.2/99.7 |
| Bias and lag corrected, multipoint | $y_{FGM,BLcorr}$ | 12.7 | 0.8 | 86.7/97.5 |

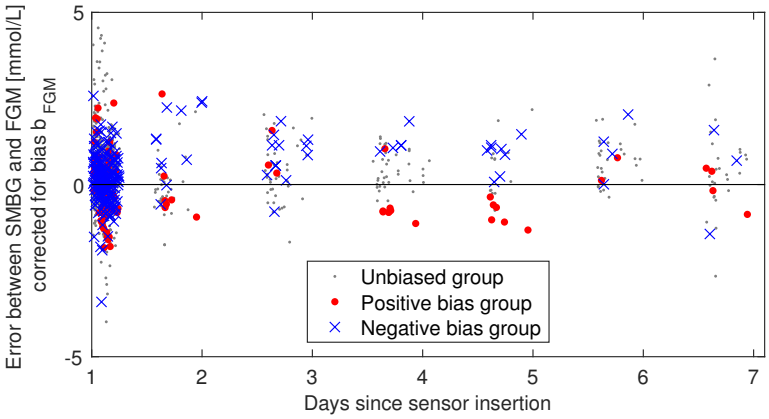



**Figure A2.** Same plot as in Figure 6, but with individual bias correction ($b_{FGM}$) applied. Using biases found from day 1 to correct day 2–7 data seems to give over-correction, at least in the negative bias group.

## References


1. Schrangl, P.; Reiterer, F.; Heinemann, L.; Freckmann, G.; del Re, L. Limits to the Evaluation of the Accuracy of Continuous Glucose Monitoring Systems by Clinical Trials. *Biosensors* **2018**, *8*, 50. [CrossRef] [PubMed]
2. Cengiz, E.; Tamborlane, W.V. A tale of two compartments: interstitial versus blood glucose monitoring. *Diabetes Technol. Ther.* **2009**, *11*, S11–S16. [CrossRef] [PubMed]
3. Rebrin, K.; Sheppard, N.F.; Steil, G.M. Use of subcutaneous interstitial fluid glucose to estimate blood glucose: revisiting delay and sensor offset. *J Diabetes Sci. Technol.* **2010**, *4*, 1087–1098. [CrossRef] [PubMed]
4. Facchinetti, A.; Del Favero, S.; Sparacino, G.; Castle, J.R.; Ward, W.K.; Cobelli, C. Modeling the glucose sensor error. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 620–629. [CrossRef] [PubMed]
5. Schmelzeisen-Redeker, G.; Schoemaker, M.; Kirchsteiger, H.; Freckmann, G.; Heinemann, L.; Del Re, L. Time delay of CGM sensors: Relevance, causes, and countermeasures. *J. Diabetes Sci. Technol.* **2015**, *9*, 1006–1015. [CrossRef] [PubMed]
6. Shah, V.N.; Laffel, L.M.; Wadwa, R.P.; Garg, S.K. Performance of a Factory-Calibrated Real-Time Continuous Glucose Monitoring System Utilizing an Automated Sensor Applicator. *Diabetes Technol. Therapeutics* **2018**, *20*, 428–433. [CrossRef] [PubMed]
7. Bailey, T.; Bode, B.W.; Christiansen, M.P.; Klaff, L.J.; Alva, S. The performance and usability of a factory-calibrated Flash Glucose Monitoring system. *Diabetes Technol. Ther.* **2015**, *17*, 787–794. [CrossRef] [PubMed]
8. FreeStyle Libre | FreeStyle Blood Glucose Meters. Available online: https://abbottdiabetescare.co.uk/our-products/freestyle-libre (accessed on 20 December 2015).
9. Acciaroli, G.; Vettoretti, M.; Facchinetti, A.; Sparacino, G. Calibration of minimally invasive continuous glucose monitoring sensors: state-of-the-art and current perspectives. *Biosensors* **2018**, *8*, 24. [CrossRef] [PubMed]
10. Facchinetti, A.; Sparacino, G.; Cobelli, C. Modeling the error of continuous glucose monitoring sensor data: critical aspects discussed through simulation studies. *J. Diabetes Sci. Technol.* **2010**, *4*, 4–14. [CrossRef] [PubMed]
11. Pleus, S.; Heinemann, L.; Freckmann, G. Blood Glucose Monitoring Data Should Be Reported in Detail When Studies About Efficacy of Continuous Glucose Monitoring Systems Are Published. *J. Diabetes Sci. Technol.* **2018**, *12*, 1061–1063. [CrossRef] [PubMed]
12. Staal, O.M.; Sælid, S.; Fougner, A.L.; Stavdahl, Ø. Kalman smoothing for objective and automatic preprocessing of glucose data. *IEEE J. Biomed. Health Inform.* **2018**. [CrossRef] [PubMed]
13. Parkes, J.L.; Slatin, S.L.; Pardo, S.; Ginsberg, B.H. A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose. *Diabetes Care* **2000**, *23*, 1143–1148. [CrossRef] [PubMed]

14. Rauch, H.E.; Striebel, C.; Tung, F. Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **1965**, *3*, 1445–1450. [CrossRef]
15. Gustafsson, F. *Statistical Sensor Fusion*; Studentlitteratur AB: Lund, Sweden, 2010; pp. 178–179.
16. Stavdahl, Ø.; Fougner, A.L.; Kölle, K.; Christiansen, S.C.; Ellingsen, R.; Carlsen, S.M. The artificial pancreas: A dynamic challenge. *IFAC-PapersOnLine* **2016**, *49*, 765–772. [CrossRef]
17. Del Favero, S.; Facchinetti, A.; Sparacino, G.; Cobelli, C. Improving accuracy and precision of glucose sensor profiles: Retrospective fitting by constrained deconvolution. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 1044–1053. [CrossRef] [PubMed]
18. Guerra, S.; Facchinetti, A.; Sparacino, G.; Nicolao, G.D.; Cobelli, C. Enhancing the accuracy of subcutaneous glucose sensors: A real-time deconvolution-based approach. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 1658–1669. [CrossRef] [PubMed]
19. Pleus, S.; Kamecke, U.; Link, M.; Haug, C.; Freckmann, G. Flash Glucose Monitoring: Differences Between Intermittently Scanned and Continuously Stored Data. *J. Diabetes Sci. Technol.* **2018**, *12*, 397–400. [CrossRef] [PubMed]
20. Fokkert, M.J.; van Dijk, P.R.; Edens, M.A.; Abbes, S.; de Jong, D.; Slingerland, R.J.; Bilo, H.J. Performance of the FreeStyle Libre Flash glucose monitoring system in patients with type 1 and 2 diabetes mellitus. *BMJ Open Diabetes Res. Care* **2017**, *5*, 1. [CrossRef] [PubMed]
21. Ólafsdóttir, A.F.; Attvall, S.; Sandgren, U.; Dahlqvist, S.; Pivodic, A.; Skrtic, S.; Theodorsson, E.; Lind, M. A clinical trial of the accuracy and treatment experience of the flash glucose monitor FreeStyle Libre in adults with type 1 diabetes. *Diabetes Technol. Ther.* **2017**, *19*, 164–172. [CrossRef] [PubMed]
22. Sekido, K.; Sekido, T.; Kaneko, A.; Hosokawa, M.; Sato, A.; Sato, Y.; Yamazaki, M.; Komatsu, M. Careful readings for a flash glucose monitoring system in nondiabetic Japanese subjects: individual differences and discrepancy in glucose concentrarion after glucose loading [Rapid Communication]. *Endocr. J.* **2017**, *64*, 827–832. [CrossRef] [PubMed]
23. Ancona, P.; Eastwood, G.M.; Lucchetta, L.; Ekinci, E.I.; Bellomo, R.; Martensson, J. The performance of flash glucose monitoring in critically ill patients with diabetes. *Crit Care Resusc* **2017**, *19*, 167–174. [PubMed]
24. Alsaffar, H.; Turner, L.; Yung, Z.; Didi, M.; Senniappan, S. Continuous Flash Glucose Monitoring in children with Congenital Hyperinsulinism; first report on accuracy and patient experience. *Int. J. Pediatr. Endocrin.* **2018**, *2018*, 3. [CrossRef] [PubMed]
25. Boscari, F.; Galasso, S.; Facchinetti, A.; Marescotti, M.C.; Vallone, V.; Amato, A.M.L.; Avogaro, A.; Bruttomesso, D. FreeStyle Libre and Dexcom G4 Platinum sensors: Accuracy comparisons during two weeks of home use and use during experimentally induced glucose excursions. *Nutr. Metab. Cardiovasc Dis.* **2018**, *28*, 180–186. [CrossRef] [PubMed]
26. Scott, E.M.; Bilous, R.W.; Kautzky-Willer, A. Accuracy, user acceptability, and safety evaluation for the FreeStyle Libre Flash glucose monitoring system when used by pregnant women with diabetes. *Diabetes Technol. Ther.* **2018**, *20*, 180–188. [CrossRef] [PubMed]
27. Bonora, B.; Maran, A.; Ciciliot, S.; Avogaro, A.; Fadini, G. Head-to-head comparison between flash and continuous glucose monitoring systems in outpatients with type 1 diabetes. *J. Endocrinol. Invest.* **2016**, *39*, 1391–1399. [CrossRef] [PubMed]
28. Aberer, F.; Hajnsek, M.; Rumpler, M.; Zenz, S.; Baumann, P.M.; Elsayed, H.; Puffing, A.; Treiber, G.; Pieber, T.R.; Sourij, H.; et al. Evaluation of subcutaneous glucose monitoring systems under routine environmental conditions in patients with type 1 diabetes. *Diabetes Obes Metab.* **2017**, *19*, 1051–1055. [CrossRef] [PubMed]
29. Freckmann, G.; Link, M.; Pleus, S.; Westhoff, A.; Kamecke, U.; Haug, C. Measurement Performance of Two Continuous Tissue Glucose Monitoring Systems Intended for Replacement of Blood Glucose Monitoring. *Diabetes Technol. Ther.* **2018**, *20*, 541–549. [CrossRef] [PubMed]
30. Bösch, P.C.; Åm, M.K.; Stavdahl, Ø.; Fougner, A.L.; Ellingsen, S.M.C.R.; Hjelme, D.R. Setup and procedure for intraperitoneal glucose monitoring in anaesthetised animals. Poster Presentation at Advanced Technologies & Treatments for Diabetes (ATTD), Paris, France, 2017.
31. Hoss, U.; Budiman, E.S. Factory-calibrated continuous glucose sensors: the science behind the technology. *Diabetes Technol. Ther.* **2017**, *19*, S44–S50. [CrossRef] [PubMed]
32. Vettoretti, M.; Facchinetti, A.; Sparacino, G.; Cobelli, C. A model of self-monitoring blood glucose measurement error. *J. Diabetes Sci. Technol.* **2017**, *11*, 724–735. [CrossRef] [PubMed]

33. Link, M.; Schmid, C.; Pleus, S.; Baumstark, A.; Rittmeyer, D.; Haug, C.; Freckmann, G. System accuracy evaluation of four systems for self-monitoring of blood glucose following ISO 15197 using a glucose oxidase and a hexokinase-based comparison method. *J. Diabetes Sci. Technol.* **2015**, *9*, 1041–1050. [CrossRef] [PubMed]
34. Freckmann, G.; Baumstark, A.; Schmid, C.; Pleus, S.; Link, M.; Haug, C. Evaluation of 12 blood glucose monitoring systems for self-testing: system accuracy and measurement reproducibility. *Diabetes Technol. Ther.* **2014**, *16*, 113–122. [CrossRef] [PubMed]
35. Baysal, N.; Cameron, F.; Buckingham, B.A.; Wilson, D.M.; Chase, H.P.; Maahs, D.M.; Bequette, B.W.; Buckingham, B.A.; Wilson, D.M.; Aye, T.; et al. A novel method to detect pressure-induced sensor attenuations (PISA) in an artificial pancreas. *J. Diabetes Sci. Technol.* **2014**, *8*, 1091–1096. [CrossRef] [PubMed]

# Appendix B

# Unpublished work

## B.1 Glucose-insulin metabolism model reduction and parameter selection using sensitivity analysis

Conference article submitted to *American Control Conference (ACC) 2019*

O. M. Staal, A. L. Fougner, S. Sælid, and Ø. Stavdahl, "Glucose-insulin metabolism model reduction and parameter selection using sensitivity analysis," *ACC*, Submitted to American Control Conference (ACC) 2019.

Full text is included on the next page.

## B.2 Meal input estimation from Continuous Glucose Monitor data using Kalman filtering and hypothesis testing

This article was not complete at the time of submitting the thesis, and is included here as a draft manuscript.

The remaining tasks planned for the article are:

- Tests on more real data sets, ideally ones where the meal time is known to better accuracy than that given by the user.
- Comparison against other methods on real data sets

Full text is included on the next page.