

# **Master Thesis - Project Specification**

**HbA1c Impact Estimation of Detected Events from  
Continuous Glucose Monitoring Data**

THONY PRICE

Master in Computer Science

Date: October 26, 2018

Supervisor: Pawel Herman

Examiner: -

Swedish title: Estimerad HbA1c Inverkan Av Detekterade Event ur  
Kontinuerlig Glukos Matardata

School of Electrical Engineering and Computer Science



## **Abstract**

About this document...

Diabetes is on the rise... Proper care comes from good data... Diet is of major concern... CGM allows easy data collection... Meal detection is... Carbohydrate estimation...

# Contents

|          |                                      |           |
|----------|--------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                  | <b>1</b>  |
| 1.1      | Project Introduction . . . . .       | 1         |
| 1.2      | Project Aim . . . . .                | 2         |
| 1.2.1    | Objective . . . . .                  | 2         |
| 1.2.2    | Research Question . . . . .          | 3         |
| <b>2</b> | <b>Background</b>                    | <b>4</b>  |
| <b>3</b> | <b>Method</b>                        | <b>5</b>  |
| 3.1      | Data . . . . .                       | 5         |
| 3.2      | Implementation . . . . .             | 6         |
| 3.2.1    | Wavelet Filter . . . . .             | 6         |
| 3.2.2    | Qualitative Representation . . . . . | 7         |
| 3.2.3    | Event Detection . . . . .            | 7         |
| 3.2.4    | Intervention Analysis . . . . .      | 8         |
| 3.3      | Evaluation . . . . .                 | 9         |
|          | <b>Bibliography</b>                  | <b>10</b> |

# Chapter 1

## Introduction

This chapter intends to provide the reader an overview of the current state of diabetes healthcare as well as explain the outline and aim of this thesis project.

### 1.1 Project Introduction

In every country the disease burden related to diabetes is already high, and it is steadily increasing [8]. In 2017 the estimated prevalence of diabetes was 451 million people globally and approximately 5 million deaths were attributed to diabetes [3]. Aside from reduced life expectancy, diabetes increase the risk of multiple other conditions such as heart disease, stroke and peripheral vascular diseases [8]. With a projected prevalence of 693 million diabetes patients in 2045 and given the seriousness of diabetes, proper medical care for patients are of utmost importance [3].

Diabetes is a group of metabolic diseases that is characterized by hyperglycemia. It resulting from defects in insulin secretion, insulin action, or both [1]. Insulin is necessary to maintain normal blood glucose concentration (BGC) internally by facilitating cellular glucose uptake and regulating carbohydrate metabolism [12]. The vast majority of cases of diabetes fall into two broad categories, type 1 diabetes (T1D) and type 2. T1D is caused by an absolute deficiency of insulin secretion, thus patients need to induce exogenous insulin on a regular basis

to maintain a balanced BGC [1].

Maintaining balanced BGC is an every day challenge of T1D patients. Treatment guidelines put heavy emphasis on self management activities that benefit a balanced regulation. This include recommendations of eating patterns, exercise, carbohydrate consumption and much more. Patients favorably consult a clinician regularly to get evaluate these activities and plan for the continued self management process [4].

Continuous glucose monitoring (CGM) sensors are wearable devices that measures the blood glucose frequently (usually every 1-5 minutes). CGM sensors have revolutionized the ability for clinicians to review a patients data and deliver care driven by extensive data [6]. Care based CGM data have proven effective in lowering patients long term average BGC, known as HbA1c [5]. The data enable precise monitoring of immediate processes, such as accuracy of a specific insulin dosing, BGC peak after a certain meal etc. However, there is a lack of research in deriving insights for long term medical advice autonomously from CGM data.

## 1.2 Project Aim

Given CGM data of a patient, autonomously generate insight of which events and patterns in the data indicates a certain progression of long term HbA1c. For example, if provided 8 week's worth of data some segments of the data may infer a healthy development, such as healthy meals and exercise. The aim is to detect these events in the CGM data in order to classify which activities have a positive impact on a patient's HbA1C.

### 1.2.1 Objective

The objective of this thesis is to investigate how an autonomous system, which fulfills the certain criterias below, can be implemented and in which configuration it performs optimally. The system should:

- Detect events in from a batch CGM time series data (such as meals, exercises or sleep).

- Estimate intervention caused by one or more events (what impact did the event(s) have on the continued time series).

### **1.2.2 Research Question**

1. Is it possible to detect events from a sequence of qualitative representations generated from a batch CGM data using a hidden Markov model?
2. Given one or more events in (1.), can intervention analysis estimate the long term impact of a single, or a set of events?

## Chapter 2

### Background

This chapter sets out define diabetes and the relevant aspects of T1D care enough to understand the content in latter chapters. It further describes theory of the techniques in the suggested method and reviews related work that should assist putting this thesis into context of the current field.

**[This chapter is under construction. It should be continued when the method is decided upon.]**



# Chapter 3

## Method

### 3.1 Data

The data is collected during a 4,5 month pilot study includes 20 patients. All patients are at least 18 years old and have been injecting insulin at least 1 year prior to the study. Each patient are equipped with a CGM sensor for the entirety of the study. For each patient, the CGM sensor measures the BG concentration (BGC) at intervals of 1-5 minutes. The time-BGC constitutes the structure of each patient's data set. Additionally, other events such as meals and exercise are logged manually by and are also associated with a certain time. A summary of collected data points are presented in table 3.1.

| Notation    | Field                               | Format  |
|-------------|-------------------------------------|---------|
| $t$         | Time                                | Date    |
| $t_{BG}$    | Glucose value at $t$                | Float   |
| $t_{ACT}$   | Physical activity at $t$            | Integer |
| $t_{INS}$   | Injected insulin at $t$             | Float   |
| $t_{IMG}$   | Food image at $t$                   | Float   |
| $t_{EVENT}$ | Manually reported by patient at $t$ | Text    |

Table 3.1: The data for each patient include continuous measurements at time steps  $t$  of intervals between 1-5 minutes. Each measurement at  $t$  *always* include BG value and *may* contain other field presented in the table.

## 3.2 Implementation

The objective of the proposed system is to analyse the data for a closed time interval, identify meaningful events and classify them accordingly with respect to their influence over future measurements. The analysis is not performed in real time, thus should be considered a batch problem (that is, all data is available immediately). The proposed steps of implementation can be overviewed in figure 3.1. Each step is described in more detail in its corresponding section below.

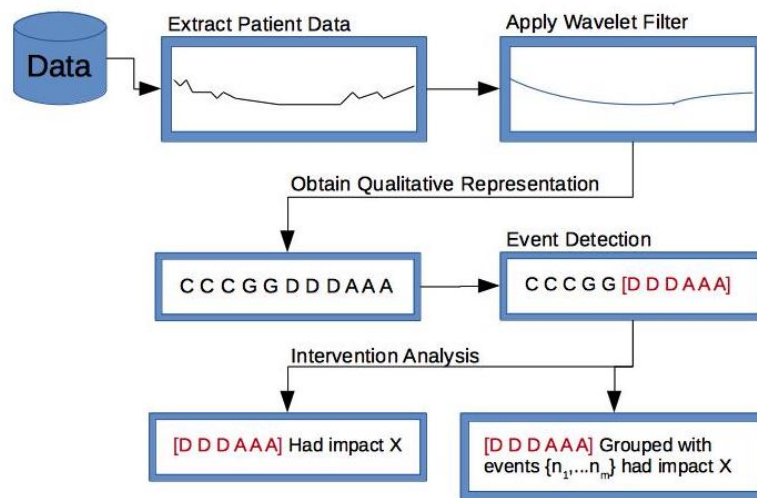


Figure 3.1: Schematics of implementation.

### 3.2.1 Wavelet Filter

Studies have shown data from CGM is subject to distortion. This is caused by diffusion processes and by time-varying systematic under/overestimations due to calibrations and sensor drifts [7]. Noise can trigger false positives in event detection because abrupt fluctuations overrides the true underlying derivatives of the curve [6]. Wavelet filters have been used repeatedly with CGM data and proved successful in reducing noise while retaining events such as spikes [9], [6], [10]. Figure 3.2 displays an example of wavelet filtering applied to CGM data.

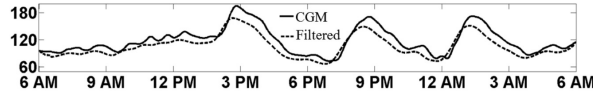


Figure 3.2: Wavelet filter applied on CGM data. Vertical axis represents glucose concentration [mg/dl]. Image courtesy of Samadi et al. [10].

### 3.2.2 Qualitative Representation

To identify events in the denoised CGM data feature, extraction is used. Feature extraction can be achieved by either a qualitative or quantitative method. The qualitative method offer benefits such as more transparent reasoning and ability to provide explanations for for solutions provided [11].

In qulitative representation by triangular shapes, a CGM data segment can take seven shape variables. Figure 3.3 shows the different shapes. Each is a unique combination of the first and second order derivaive on the curve of the current segment. The derivates can be read from segment adjacent points allowing a CGM data series to be presented as a sequence of shapes describing fluctuations in BG concentration.

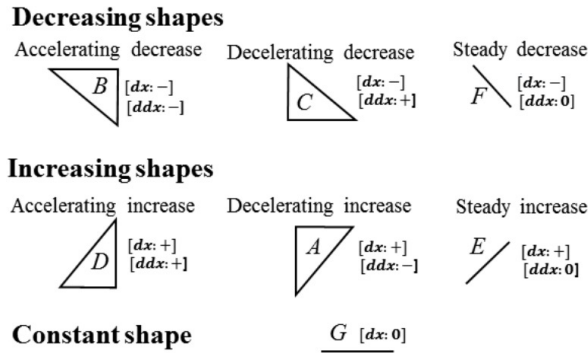


Figure 3.3: Scheme of the qualitative variables A-G. Image courtesy of Samadi et al. [10].

### 3.2.3 Event Detection

With the qualitative representation event detection can be performed by analysing the sequence of shapes. In figure 3.4 an event could be

triggered by identifying a continuously accelerating increase (for example the four sequential D's from timestep 12).

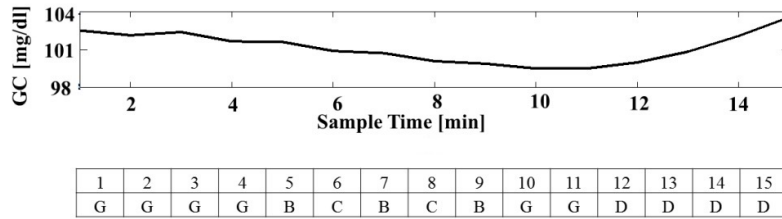


Figure 3.4: Shape sequence representation of CGM curve. Image courtesy of Samadi et al. [10].

Given a sequence of observations (i.e. shapes) with some underlying state (i.e. eating, sleeping, exercising) a hidden Markov model can provide an estimate of what state a current time step most probably corresponds to.

### 3.2.4 Intervention Analysis

Intervention analysis provides a tool to assess how much a given event has changed the series (if at all) [2]. The analysis is able to detect 4 patterns:

1. Permanent constant change to the mean level.
2. Brief constant change to the mean level.
3. Gradual increase or decrease to a new mean level.
4. Initial change followed by gradual return to previous mean level.

Intervention analysis compares the mean average before the event and its progression afterwards. The characteristics of the progression curve define which of the four patterns it corresponds to. Because changes in mean BG concentration are subtle and changes naturally take place over a longer timespan, an alternative to the approach is suggested.

An event by itself may not cause an observable change in the mean average but combined events might. The intervention analysis is also evaluated on sets of adjacent events. The output can hint on when the aggregated data implies a positive development of the mean. In other

words, the analysis can suggest when combined activities by a patient results in a change of his or her long term A1C value.

### **3.3 Evaluation**

Training/test...

Padova...

Manual labeling...

# Bibliography

- [1] American Diabetes Association. "Diagnosis and Classification of Diabetes Mellitus". In: *Diabetes Care* 33.Supplement 1 (2010), S62–S69. ISSN: 0149-5992. DOI: 10.2337/dc10-S062.
- [2] George EP Box et al. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- [3] N. H. Cho et al. "IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045". In: *Diabetes Research and Clinical Practice* 138 (Apr. 2018), pp. 271–281. ISSN: 0168-8227. DOI: 10.1016/j.diabres.2018.02.023.
- [4] Debbie Cooke et al. "Structured Type 1 Diabetes Education Delivered Within Routine Care". In: *Diabetes Care* 36.2 (2013), pp. 270–272. ISSN: 0149-5992. DOI: 10.2337/dc12-0080.
- [5] Kathleen Dungan and Neha Verma. "Monitoring technologies—continuous glucose monitoring, mobile technology, biomarkers of glycemic control". In: *Endotext [Internet]*. MDText. com, Inc., 2018.
- [6] Andrea Facchinetti. "Continuous Glucose Monitoring Sensors: Past, Present and Future Algorithmic Challenges". In: *Sensors (Basel)* 16.12 (Dec. 2016). PMC5191073[pmcid], p. 2093. ISSN: 1424-8220. DOI: 10.3390/s16122093.
- [7] Andrea Facchinetti et al. "Modeling the glucose sensor error". In: *IEEE Transactions on Biomedical Engineering* 61.3 (2014), pp. 620–629.
- [8] Nita Gandhi Forouhi and Nicholas J. Wareham. "Epidemiology of diabetes". In: *Medicine (Abingdon)* 42.12 (Dec. 2014). 25568613[pmid], pp. 698–702. ISSN: 1357-3039. DOI: 10.1016/j.mpmed.2014.09.007.

- [9] Nicolas Magdelaine et al. "Wavelets for CGM off-line denoising". In: (Feb. 2016). DOI: 10.13140/RG.2.1.2048.2326.
- [10] Sediqeh Samadi et al. "Meal detection and carbohydrate estimation using continuous glucose sensor data". In: *IEEE journal of biomedical and health informatics* 21.3 (2017), pp. 619–627.
- [11] Venkat Venkatasubramanian et al. "A review of process fault detection and diagnosis: Part III: Process history based methods". In: *Computers and Chemical Engineering* 27.3 (2003), pp. 327–346. ISSN: 0098-1354. DOI: [https://doi.org/10.1016/S0098-1354\(02\)00162-X](https://doi.org/10.1016/S0098-1354(02)00162-X).
- [12] Gisela Wilcox. "Insulin and insulin resistance". In: *Clin Biochem Rev* 26.2 (May 2005). PMC1204764[pmcid], pp. 19–39.