



Modeling glucose dynamics during physical activity using a linear model for individuals with Type 1 Diabetes

Citation

Cho, Annabel Myung-Joo. 2023. Modeling glucose dynamics during physical activity using a linear model for individuals with Type 1 Diabetes. Bachelor's thesis, Harvard University Engineering and Applied Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37378265>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

**Modeling glucose dynamics during physical activity
using a linear model for individuals with Type 1
Diabetes**

A thesis presented by
Annabel Myung-Joo Cho

to

the Faculty of the

Harvard John A. Paulson School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the Bachelor of Arts degree with honors in

Engineering Sciences - Biomedical Sciences and Engineering Track

Faculty Advisor: Prof. Francis J. Doyle, III.

Harvard University

Cambridge, MA

March 24, 2023

Contents

List of Figures	4
List of Tables	6
1 Introduction	10
1.1 Background	10
1.2 Objective/Research Plan	14
2 Materials	16
2.1 Dataset collection	16
3 Methods	18
3.1 Dataset Preprocessing	18
3.2 Average Linear Time-Invariant Model Identification	22
3.3 Personalized Linear Time-Invariant Model Identification	27
3.3.1 Van Heusden Et Al. 2012 Model [1]	27
3.3.2 Adding the Identified HR-Glucose Transfer Function	27
4 Results	30
4.1 Identification and Testing Dataset	30
4.2 Model Performance	31
5 Discussion	37

List of Figures

3.1	HR data (top row) and Continuous Glucose Monitoring with Insulin data (bottom row) of resistance (left), aerobic (middle), and HIIT (right) physical activity sessions. The recorded sessions are shown by the vertical lines, with 10 minutes of data pre- and post-exercise.	19
3.2	Filtering data using a triangular and rectangular filter. On the left side, the applied convolution filters are plotted in blue. On the right side, the raw data are shown in black and the convoluted version in orange.	21
3.3	Raw data (blue) and its Kalman Estimate (orange) with a fixed R value of 1 and varying values of Q. Q = 1 is shown at the top (a), Q = 0.01 in the middle (b), and Q = 0.1 is at the bottom (c).	23
3.4	MATLAB Simulink representation of the system. HR data is Kalman filtered and inputted into the transfer function linear model with the insulin data. The model output of predicted Continuous Glucose Monitoring (CGM) data is plotted on the true CGM data to compare.	26
4.1	An R^2 (left) and RMSE (right) value comparison of the personalized model performance trained on 5, 10, 15, 20 subjects who performed aerobic exercise sessions. The testing set for this analysis consisted of all patients outside of the 20 subject training set for consistency	32

4.2	An example of the reference glucose data for an aerobic activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.	35
4.3	An example of the reference glucose data for an resistance activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.	35
4.4	An example of the reference glucose data for an HIIT activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.	36

List of Tables

3.1	Estimated parameters for equations 3.10 that were identified by minimizing	
	the prediction error between the model output $C\hat{G}M(n, \theta)$ and the observed	
	CGM data	25
3.2	Estimated parameters for equation 3.21 that were identified by minimizing the	
	prediction error between the model output $\Delta C\hat{G}M(n, \theta_P)$ and the observed	
	Δ CGM data	29
4.1	RMSE values for the personalized, average, and naive/simplistic model split	
	up by exercise type. The RMSE values for the first ten subjects of the testing	
	dataset are shown, as well as the mean and standard deviation for all subjects.	34
4.2	R^2 values for the personalized, average, and naive/simplistic model split up	
	by exercise type. The R^2 values for the first ten subjects of the testing dataset	
	are shown, as well as the mean and standard deviation for all subjects.	34
4.3	A summarized table of average \pm standard deviation RMSE and R^2 values	
	shown for the personalized, average, and naive/simplistic model. The mean	
	and standard deviation values for the RMSE and R^2 values for the first ten	
	subjects of the testing dataset are shown, as well as all subjects.	34

Abstract

For individuals with type 1 diabetes (T1D), regular exercise is beneficial as it can improve glycemic control, increase insulin sensitivity, improve body composition, and overall increase quality of life. However, physical activity can lead to complications for individuals with T1D, as it causes a fluctuation in glucose levels and increases the risk of hypoglycemia. This thesis focuses on the identification of both an average and patient-tailored linear time-invariant models that predict the glucose levels from the heart rate signal and the injected insulin records during a physical activity session of either aerobic, resistance, or interval exercise. Data from 177 adults with T1D who performed structured sessions of either aerobic, resistance, or high-intensity interval exercise at varying times of day were used. In this thesis, the simulation capabilities of the models were compared to a naive/simplistic model. The performance of the personalized model tested on all subjects in each exercise type for aerobic exercises: $RMSE = 4.132 \pm 6.480$ $R^2 = 0.776 \pm 0.197$, for resistance exercises: $RMSE = 6.026 \pm 24.815$ $R^2 = 0.752 \pm 0.239$, and HIIT exercises $RMSE = 6.723 \pm 35.676$ $R^2 = 0.737 \pm 0.239$. The personalized model had a lower RMSE and higher R^2 than both the average model and the naive/simplistic model.

Honor Code

In submitting this thesis to the Harvard John A. Paulson School of Engineering and Applied Sciences in partial fulfillment of the requirements for the degree with honors of Bachelor of Arts, I affirm my awareness of the standards of the Harvard College Honor Code.

Name: Annabel Cho

Signature: 

The Harvard College Honor Code Members of the Harvard College community commit themselves to producing academic work of integrity – that is, work that adheres to the scholarly and intellectual standards of accurate attribution of sources, appropriate collection and use of data, and transparent acknowledgement of the contribution of others to their ideas, discoveries, interpretations, and conclusions. Cheating on exams or problem sets, plagiarizing or misrepresenting the ideas or language of someone else as one’s own, falsifying data, or any other instance of academic dishonesty violates the standards of our community, as well as the standards of the wider world of learning and affairs.

Acknowledgements

This thesis would not have been possible without the support from the Doyle Research Group at the Harvard John A. Paulson School of Engineering and Applied Sciences.

I would especially like to thank Professor Francis J. Doyle III and Dr. Eleonora Aiello for their constant encouragement, thoughtful feedback, and unwavering support throughout this last year. Thank you for inspiring me and reigniting my love for learning and research. I always looked forward to our weekly research meetings, where I would feel excited to learn about the novel changes and discoveries being made in the field of Diabetes research.

Thank you to Dr. Linsey Moyer, my bioengineering undergraduate advisor, for supporting and guiding me through my 4 years at Harvard. You have simultaneously helped me make some of the hardest decisions while reassuring my interest in bioengineering through ES53 and BE128.

Thank you to my lovely roommates-turned-best-friends for supporting me and making Harvard truly feel like home. You have always been there for me every step of the way.

Thank you Mom and Dad for always believing in me. Words cannot express my gratitude for your continuous support, encouragement, and love. Thank you to Christopher for being the best older brother ever and for always cheering me on.

And of course, a big thank you to Bo-Mi, the absolutely best dog in the multiverse.

Chapter 1

Introduction

1.1 Background

Type 1 Diabetes (T1D) is a chronic autoimmune disease caused by the loss of insulin-producing pancreatic β cells [2]. For most T1D individuals, these β cells are thought to be destroyed by an autoimmune reaction mediated by T-cells, and can go unnoticed for weeks to years [3]. This autoimmune reaction is detected by the presence of autoantibodies that target the insulin-producing β cells. However, in a small subgroup of T1D individuals, the β cell loss cause is unknown, as no autoantibodies are found in the body [3]. For individuals with T1D, their pancreas does not produce adequate levels of insulin, a hormone that allows the body to convert glucose into energy [4]. This results in abnormal metabolism of carbohydrates, generating excess glucose in the bloodstream (hyperglycemia) which can cause symptoms such as polyuria, hunger, thirst, and weight loss [3].

Although usually diagnosed during childhood, T1D can be diagnosed at any age. The criteria for diagnosis can come from a fasting blood glucose concentration of above 126 mg/dL, a random blood glucose concentration of above 200 mg/dL with symptoms, or a 2-h plasma glucose (2-h PG) value above 200 mg/DL during a 75-g oral glucose tolerance test [5]. According to the International Diabetes Federation (IDF), over 537 million individuals

have diabetes, with 5 to 10% of these cases being T1D [6].

There is no cure for T1D and it requires lifelong treatment. Currently, many T1D individuals use insulin injections and insulin pumps to maintain the blood glucose levels in a safe range between 70 and 140 mg/dL [2]. Since the glucose levels need to be checked several times throughout the day, a continuous glucose monitoring (CGM) device is often used as it can provide information on speed and direction of the glucose levels. Additionally, a fingerstick is also used to calibrate the CGM or to inform treatment decisions such as the insulin bolus at mealtime. Automated insulin delivery (AID) systems close the loop between a glucose sensing device and an insulin delivery device to compute and deliver insulin (typically every 5 min) to achieve a desired glucose level while reducing the risk of extreme glucose variations below (hypoglycemia) or above desired range (hyperglycemia) in individuals with T1D. In addition to these treatments, individuals with T1D need to consider several aspects of physiology and metabolism. They need to understand the effects of insulin dosages, which vary depending on meal contents, physical activity, illness, and stress. Carbohydrate monitoring and counting is crucial for correct insulin dosage. High fiber (at least 14 g fiber/1,000 kcal) carbohydrate sources are recommended [7]. Typically, a multidisciplinary care team is provided to individuals with T1D to ensure optimal glycemic control [8].

For individuals with T1D, regular exercise is beneficial as it can improve glycemic control, increase insulin sensitivity, improve body composition, and overall increase quality of life [9]. The American Diabetes Association (ADA) recommends that individuals with any type of diabetes should exercise daily, not allowing more than 2 days in between exercises [7]. In addition to daily exercise improving insulin sensitivity, it has also been shown to reduce postprandial blood glucose excursion. By keeping track of physical activity using wearable fitness trackers and measuring glucose area under curves (GAUCs), researchers were able to show that patients with a higher daily physical activity had lower average postprandial GAUC [10]. This relationship of blood glucose control is crucial to improving glucose control

with T1D treatment strategies. This also exemplifies the all day long effect of physical activity outside of just the immediate hypoglycemia possibilities. Monitoring unstructured physical activity and predicting its effects around the clock allows for several improvements in lifestyle for individuals with T1D.

However, physical activity can pose considerable challenges for individuals with T1D, as exercise can lead to dramatic fluctuations in glucose levels and heightened risk of hypoglycemia. Many individuals with T1D, due to the changes in glucose uptake and insulin sensitivity, are hesitant to exercise regularly because of their fear of hypoglycemia. Glucose response to exercise depends on initial glucose levels, the type/intensity of physical activity, how long they are active, and insulin dose changes [5]. In addition, each individual has a different glucose response to exercise which needs to be considered when defining an appropriate insulin therapy. This makes effective insulin dosing difficult, as there are many different types of exercise, which can be spontaneous or structured. Even activities like inclined walking, yard work, or even house work can cause fluctuations in the glucose response.

The activity type can be categorized into three main classes: high intensity interval training (HIIT), aerobic, and resistance, as these are the main activity characteristics that mediate the glucose outcomes during and following an exercise session [11]. Specifically, aerobic activity has an immediate impact on individuals' insulin sensitivity and glucose disposal rate [11], while resistance exercise and HIIT rely more on anaerobic metabolism and muscle glycogen utilization, which can lead to an immediate increase in blood glucose levels [12]. When individuals with T1D engage in any kind of activity, they have to use extra care to their insulin dosing changes, meals, and guidelines to the physical activity. In addition, the continued advancements in diabetes technology have facilitated the management of physical activity for individuals with T1D. To overcome this barrier to regular exercise, advanced and personalized strategies for managing physical activity, enhanced exercise-informed technology (e.g., automated insulin delivery algorithms and decision support systems), as well as

educational interventions are promising tools that can be beneficial to individuals with T1D [7], [11].

The primary goal of AID systems is to regulate glucose by rejecting previously noted disturbances on glucose to maintain an average glucose approximated by with a percent time in range (TIR) of 70-180 mg/dL >70%, percent time below (TB) 70 mg/dL <4%, and percent time above (TA) 180 mg/dL <25%. An AID system is subjected to various disturbances, of which the most significant are represented by the glucose variations induced by the meals intake and physical activity. It is important to note that these disturbances may be announced, approximately known, or even predictable.

Modern AID systems, even with limitations, such as requiring user-initiated meal and correction insulin boluses, can improve glycemic control overall over conventional open-loop therapy [13]. However, there is evidence that modern AID systems have been unable to prevent the sharp decreases in blood glucose levels that come with moderate exercise. If the user announces an exercise, the device is able to adapt and adjust insulin dose modulation. However, if the physical activity is unannounced or the activity type is not well-defined, it is difficult to accurately adapt the insulin therapy. However, when comparing these pre-existing AID systems to a new hybrid model predictive controller and modified system, which was designed to detect unannounced exercise and predict the effects, researchers have been able to decrease the number of hypoglycemic events [14]. This study highlighted the importance of understanding the glycemic impact of daily physical activity in improving all around significant health benefits. Therefore, advanced exercise-informed diabetes technologies are promising tools in facilitating treatment adjustment against exercise-induced glucose imbalances in individuals with T1D.

1.2 Objective/Research Plan

A model, which simulates the effect of the physical activity, can enhance the AID design to adapt insulin infusion in presence of a physical activity. For instance, basal rates may be reduced during and after prolonged aerobic exercise, due to increased insulin sensitivity and higher risk for hypoglycemia, while they may be increased to help treat or prevent hyperglycemia caused by high intensity short anaerobic exercise [9]. Such augmented feedback may also be used to reduce or suspend the prandial insulin doses administered at the meals after exercises, given the elevated insulin sensitivity after physical exercise. [9]

Several different research groups have worked on the identification of models that can predict changes in blood glucose during and after a physical activity [15], [16]. Varying inputs, and varying types of models have yielded different results. In [15], the authors utilized a pre-existing control-to-range (CTR) algorithm that was informed by heart rate (HR). When triggered by a change in HR that was 125% of an individual's resting HR, the risk of hypoglycemia would increase at higher predicted glucose level and at a faster speed. This would lead to an earlier basal attenuation and reduction in insulin dosage. The HR informed CTR in this study lead to a decrease in hypoglycemic events during exercise and a significant reduction in blood glucose decline during exercise for this ten patient cohort. In [16], the authors used a linearized and discretized derivation version of Bergman et al's minimal model [17]. This model had two inputs of absorbed carbohydrates and active insulin. In this study, they compared the model with an without a HR additional input. The coefficients for the model were population based, with the one personalized measure of adding body weight. By adding HR as an input to the model, there was a statistically significant improvement in the root-mean-square error compared to the model without HR.

In [18], the authors presented a multivariable simulator called mGIPsim, which uses meals, insulin, and physical activity on a simulation of virtual subjects with T1D for a glucose-insulin dynamic model. The exercise subsystem best categorizes low to medium intensity exercise and uses HR above resting HR. Generally, more intense exercises are cor-

related with having increased HR.

The main aim of this work is the identification of an average and a patient-tailored linear time-invariant models, which can simulate the effect of different types of physical activity on the glucose. The model predicts the glucose levels from the HR signal and the injected insulin records during a physical activity session. Three separate models have been identified to simulate the glucose excursion during an aerobic, resistance, or HIIT activity, respectively.

Chapter 2

Materials

2.1 Dataset collection

In this work, the model is identified from clinical data that were collected within The Type 1 Diabetes Exercise Initiative (T1DEXI), which is a real-world study designed to create a shareable dataset that will help researchers better understand the factors that may influence the glycemic responses to different types of exercise [19, 20]. This dataset includes a one-month observational study of exercise-related glycemia from 497 adults in the US living with type 1 diabetes (n=183 on standard pump therapy; n=226 on hybrid closed loop; n=88 on multiple daily injections). Participants self-reported physical activity events, including randomized assignment to study-designed aerobic, resistance or interval type exercise, and food intake using a custom smartphone application. Data collection also included insulin delivery and activity monitors (Polar HR chest strap, Verily Health Watch) to contextualize each activity event and relate to exercise-associated changes in glycemia as assessed by continuous glucose monitoring (Dexcom G6).

For this thesis, only structured exercise video sessions were included to ensure clean and clear data for model creation. Each exercise video was approximately 30 minutes in duration, led by a certified exercise professional and designed with activities so that participants could

achieve a target HR of 70%–80% of age-predicted maximal HR (aerobic video), or with intervals of HR up to 80%–90% based on their age predicted maximal HR (HIIT video), or with resistance activities that elicited major muscle group fatigue after three sets of eight repetitions (resistance video), similar to what is recommended in current guidelines. Each participant was assigned either resistance, aerobic, or HIIT when conducting their study video sessions. Additionally, only patients that were using insulin pumps as their insulin source were included, as we aim to integrate this model in a AID system. Analyzable study exercise sessions were determined as sessions that had more than 15 minutes of CGM readings and more than 10 minutes of HR readings. The data were split between identification, and testing sets on a participant basis. For identification, 10 patients were chosen as 5, 10, 15, and 20 patients were tested, however 10 patients allowed for a larger testing data set while still having significant improvement in error measurement compared to 5. As such, the training, and testing datasets were mutually exclusive to assess the prediction performance of the model against the interpatient variability.

For all preprocessing, computation, modeling, analysis, and figures, MATLAB and Simulink were utilized.

Chapter 3

Methods

3.1 Dataset Preprocessing

In order to work with the large dataset effectively, a preprocessing phase was required including:

1. the identification of the relevant data portion for modeling purposes
2. the management of different time zones in the data, including the conversion from UTC to local time
3. the design of a filter to remove the high-frequency noise component in the HR data

To analyze each patient's physical activity sessions, the following data were selected including HR, CGM, insulin records for each logged exercise session. As shown in Figure [3.1](#), 10 minutes pre- and post-exercise was also included to obtain initial glucose level data and continued data during exercise cool down. For model training and testing, one hour pre- and post-exercise was included. All exercise-related details were stored in UTC time. Specifically, the HR data was recorded in UTC time which was converted to Local time for consistency across all data. However, the insulin and CGM data was recorded in Local time,

which needed to be converted to UTC in alignment with the exercise-related information, and then converted back to local time for our purposes.

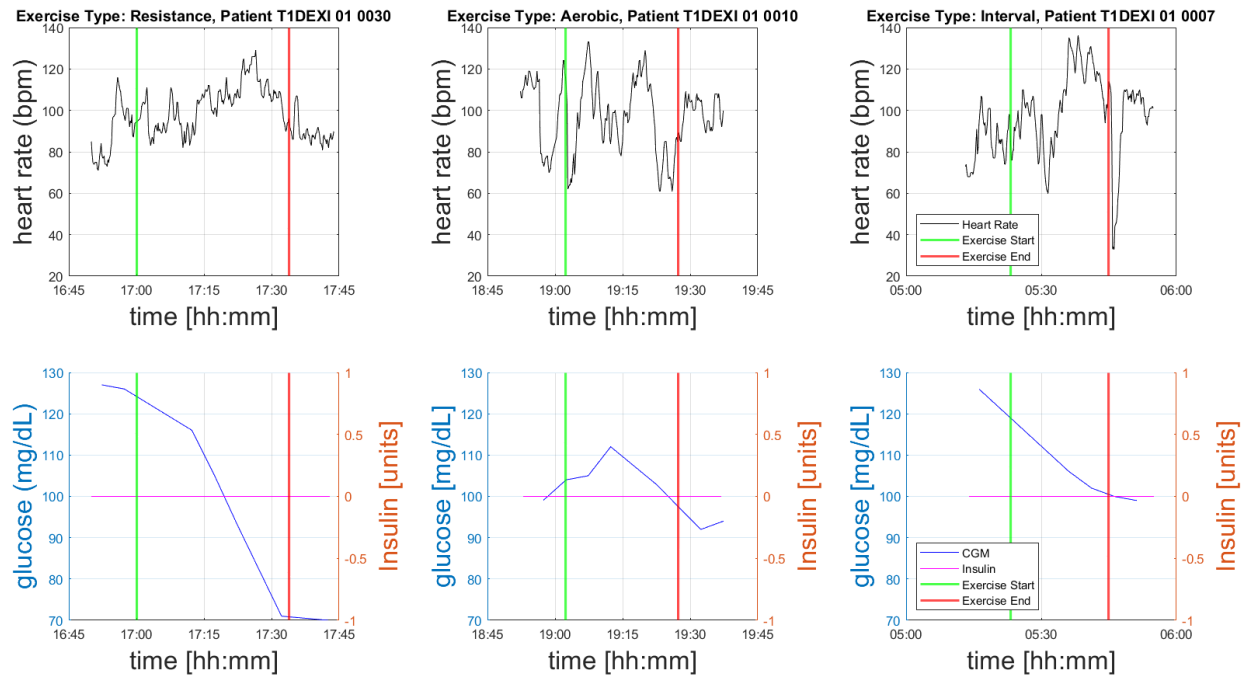


Figure 3.1: HR data (top row) and Continuous Glucose Monitoring with Insulin data (bottom row) of resistance (left), aerobic (middle), and HIIT (right) physical activity sessions. The recorded sessions are shown by the vertical lines, with 10 minutes of data pre- and post-exercise.

While HR data points were collected every 10 seconds, CGM and Insulin records were only collected every 5 minutes. After the datasets were created, the HR data was filtered to remove the high frequency noise component. Two different filtering methods were applied to the HR data based on the use of a Kalman filter and convolution filters.

A triangular and a rectangular filter were applied, i.e., the traces of the signal were convoluted with a triangle- and rectangular-shaped filter, respectively. A width of 30 and 14 HR measurement intervals were selected for the triangular and a rectangular filter, respectively. A width of 30 seconds was chosen for the triangular filter to weigh the center data point in a 30 second time frame. A triangular shaped filter is essentially a rectangular shaped filter performed twice, which is why the width is approximately double the width of the rectangular shaped filter. The doubled width will ensure "smoother" filtering since the

weight is on the center point. The filtered signal is obtained as follows:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \quad (3.1)$$

where n is the discrete time instant with a sampling time of 10 seconds, x is the HR signal, and h is the filter. In case of a triangular filter, h is designed as follows:

$$h_{triangle}(n) = \begin{cases} \frac{1}{15^2}x(n) & 0 \leq x(n) \leq 15 \\ -\frac{1}{15^2}x & 15 < x(n) \leq 30 \end{cases} \quad (3.2)$$

while

$$h_{rectangle}(n) = \begin{cases} 0 & x(n) = 0 \text{ or } x(n) = 14 \\ \frac{1}{12} & 15 < x(n) \leq 30 \end{cases} \quad (3.3)$$

The use of the triangular filter with a wide width results in having many of the peaks filtered out, while the rectangular filter with half the width allows to capture the peak trends, but not all of the extremes in amplitude, as shown in Figure [3.2](#).

The use of a Kalman filter was also investigated because it supports a real-time model for filtering the signal. The Kalman filter is based on a first order model whose state space representation is reported below:

$$\begin{aligned} x(n+1) &= Ax(n) \\ y(n) &= Cx(n) \end{aligned} \quad (3.4)$$

where x is the state vector of the HR signal, A is the discrete-time model matrix, and C is the discrete-time measurement matrix of the measured state y , with A and C both equal to 1 and $x(0)$ is set to the HR record at time $n=0$. The discrete-time state-space equations for

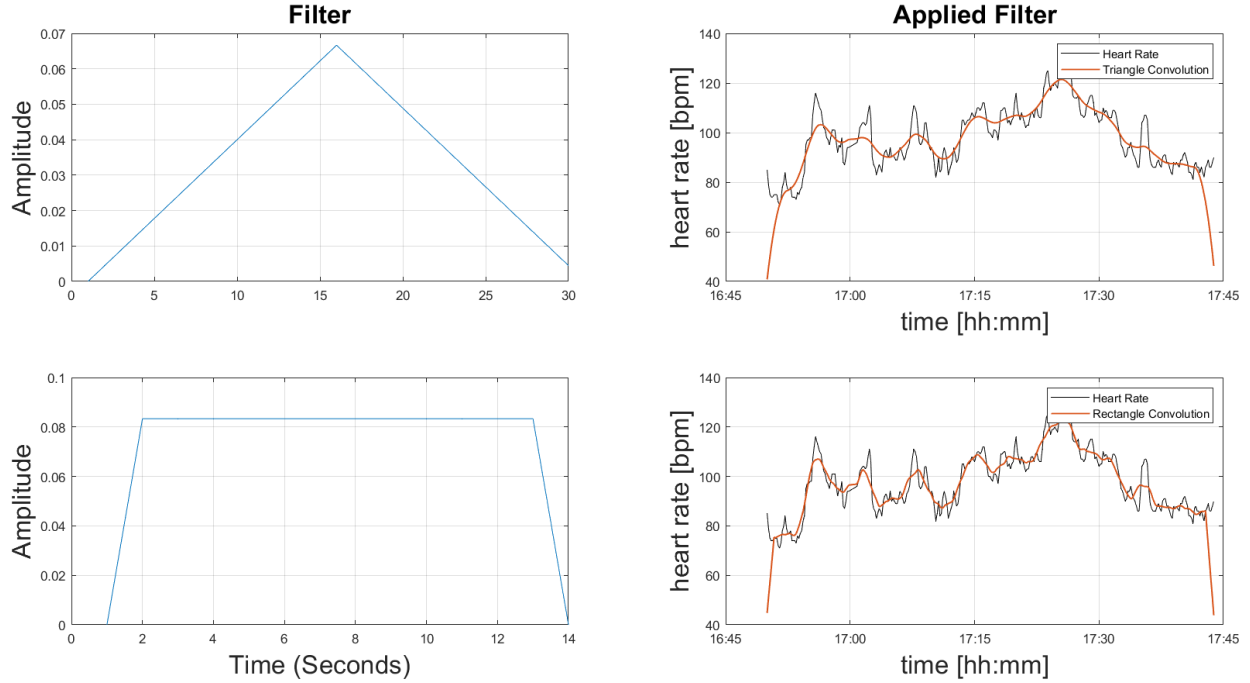


Figure 3.2: Filtering data using a triangular and rectangular filter. On the left side, the applied convolution filters are plotted in blue. On the right side, the raw data are shown in black and the convoluted version in orange.

the Kalman filter formulation are:

$$\begin{aligned} x(n+1) &= Ax(n) + w(n) \\ y(n) &= Cx(n) + v(n) \end{aligned} \tag{3.5}$$

where w and v are discrete zero-mean white noise sources, which are assumed to be Gaussian as reported below:

$$\begin{aligned} w &\sim N(0, Q) \\ v &\sim N(0, R) \end{aligned} \tag{3.6}$$

where Q is the process noise covariance and R is the measurement noise covariance [21]. The Kalman filter prediction step can be carried out using the standard equations [21]:

$$\begin{aligned} \hat{x}(n+1 | n) &= A\hat{x}(n | n) + 0 \\ P(n+1 | n) &= AP(n | n)A' + Q \end{aligned} \tag{3.7}$$

where $\hat{x}(n+1|n)$ represents the propagated state at the current time step, $\hat{x}(n|n)$ is the propagated and updated state from the previous time step. The state covariance matrix P is propagated using the PK model matrix and Q . The measurement, or correction, step for the state estimate and covariance is then performed as follows [21]:

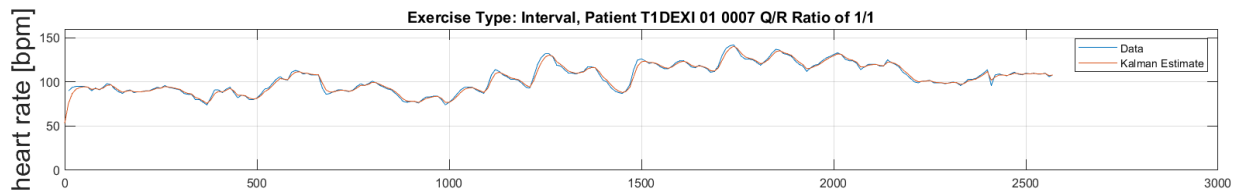
$$\begin{aligned}
 K(n+1) &= P(n+1|n) C' [C P(n+1|n) C' + R]^{-1} \\
 \hat{x}(n+1|n+1) &= \hat{x}(n+1|n) + K(n+1) [y(n+1) - C \hat{x}(n+1|n)] \\
 P(n+1|n+1) &= [I - K(n+1) C] P(n+1|n)
 \end{aligned} \tag{3.8}$$

where K is the optimal Kalman filter gain, which minimizes the residual error.

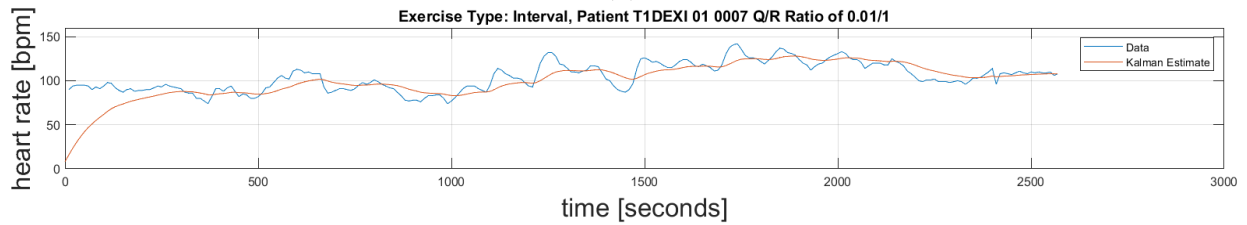
A key challenge when designing a Kalman filter is correctly characterizing the uncertainties in both the measurements and state dynamics. The performance of the filter, i.e., the speed of convergence of the state estimate, which is our filtered signal, to the true value, depends on the optimal Kalman filter gain. If measurement covariance can be assumed to be known, a large state covariance may lead to a very noisy estimate, while a small state covariance will tend to smooth the estimate. The measurement covariance was set equal to 1, then several values for Q were experimented with the Kalman Filter tests, and a value of 0.1 was decided upon as it preserves the data the best while removing the unnecessary noise. A Q equal to 0.01 caused a too much smoothed estimate, almost flat lining the data across, removing all peaks, as shown in Figure 3.3. A Q equal to 1 tended to maintain the noise on the output, with a state estimate almost identical to the original data (Fig. 3.3).

3.2 Average Linear Time-Invariant Model Identification

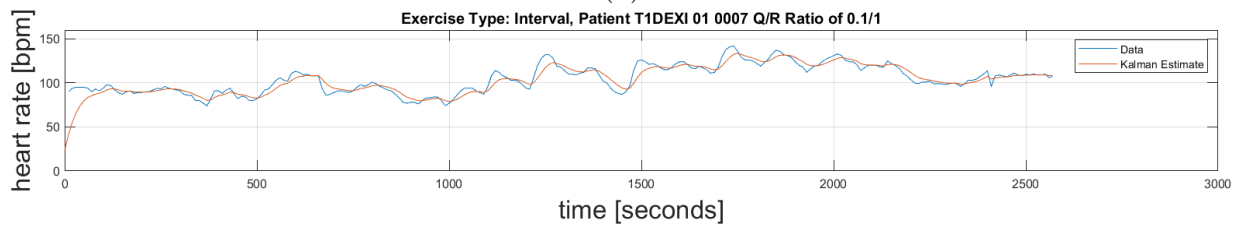
To predict the behavior and changes in the glucose levels, insulin and HR are utilized in a linear time-invariant model. The measurable inputs of the average model are the injected insulin, $i(t)$, and the recorded HR, $hr(t)$. The model output is the glucose concentration



(a)



(b)



(c)

Figure 3.3: Raw data (blue) and its Kalman Estimate (orange) with a fixed R value of 1 and varying values of Q. Q = 1 is shown at the top (a), Q = 0.01 in the middle (b), and Q = 0.1 is at the bottom (c).

measured by the CGM sensor, $CGM(t)$. Denoting with $I(z)$, $HR(z)$ and $CGM(z)$ the Zeta transformations of inputs and output transformations of inputs and output, the model has the following structure:

$$CGM(z) = G_i(z)I(z) + G_{hr}(z)HR(z) \quad (3.9)$$

where $G_i(z)$ and $G_{hr}(z)$ are transfer functions to be estimated from the data. The two transfer functions were defined as follows:

$$\begin{aligned} G_i(z) &= \frac{\mu_i z^{-1}}{(1-z^{-1}p_{i1})(1-z^{-1}p_{i2})(1-z^{-1}p_{i3})} \\ G_{hr}(z) &= \frac{\mu_{hr} z^{-1}}{(1-z^{-1}p_{hr1})(1-z^{-1}p_{hr2})(1-z^{-1}p_{hr3})} \end{aligned} \quad (3.10)$$

where p_{i1}, p_{i2}, p_{i3} and $p_{hr1}, p_{hr2}, p_{hr3}$ are the poles of the two transfer functions, and μ_i and μ_{hr} are the corresponding gains. The HR data goes through the Kalman filter before being inputted into the model, while the insulin data is processed by a zero-order-hold to be synchronized with the HR sampling size. The identification phase was performed by estimating the model parameters by minimizing prediction error $\epsilon(n, \theta)$ between the model output $CGM(n, \theta)$ and the observed CGM data below: [22]:

$$\hat{\theta} = \arg \min \epsilon(n, \theta) \quad (3.11)$$

where

$$\epsilon(n, \theta) = CGM(n) - C\hat{G}M(n, \theta) \quad (3.12)$$

where $\hat{\theta}$ is the vector of the optimal parameter values and θ is defined as follows:

$$\theta = [\theta_i \quad \theta_{hr}]' \quad (3.13)$$

with

$$\theta_i = [\mu_i \quad p_{i1} \quad p_{i2} \quad p_{i3}]' \quad (3.14)$$

$\hat{\theta}$	Activity		
	Aerobic	Resistance	HIIT
μ_i	-0.0014	0.0079	-4.6855e-05
p_{i1}	-0.9972	1.000	1.000
p_{i2}	1.000	0.9831	0.7651
p_{i3}	0.6913	-0.5209	0.2384
μ_{hr}	0.0005	-0.0002	-0.0001
p_{hr1}	-0.9851	-0.9977	-1.000
p_{hr2}	0.9909	0.9817	1.969
p_{hr3}	0.7835	0.8328	-0.9692

Table 3.1: Estimated parameters for equations [3.10](#) that were identified by minimizing the prediction error between the model output $C\hat{G}M(n, \theta)$ and the observed CGM data

and

$$\theta_{hr} = [\mu_{hr} \quad p_{hr1} \quad p_{hr2} \quad p_{hr3}]'. \quad (3.15)$$

There are separate transfer functions identified for each exercise type of aerobic (G_{i_a}, G_{hr_a}), resistance (G_{i_r}, G_{hr_r}), and HIIT (G_{i_i}, G_{hr_i}).

Using these transfer functions identified by exercise type, whose parameters are reported in Table [4.3](#), the predicted glucose values are obtained as shown in Figure [3.4](#) where the Simulink representation is reported. First, the HR data goes through a Kalman Filter with $Q=0.1$ and $R=1$. The insulin data is put through a zero order hold to match the HR data's sampling size of 10 seconds (originally 5 minutes). The two data inputs are put through a multiplexer to format into the transfer function equations (Eq. 3.9). The initial condition of the model states is calculated using Knudsen's backcast method [\[23\]](#). The output of the predicted CGM data is then compared to the original CGM data.

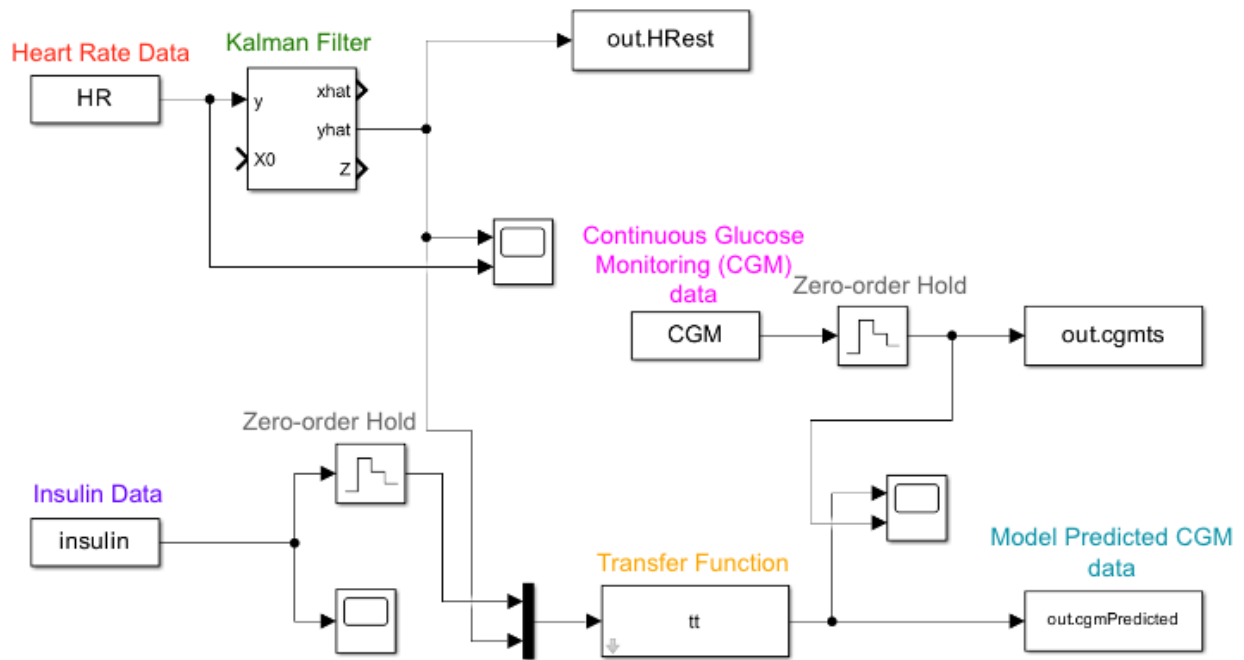


Figure 3.4: MATLAB Simulink representation of the system. HR data is Kalman filtered and inputted into the transfer function linear model with the insulin data. The model output of predicted Continuous Glucose Monitoring (CGM) data is plotted on the true CGM data to compare.

3.3 Personalized Linear Time-Invariant Model Identification

3.3.1 Van Heusden Et Al. 2012 Model [1]

To personalize the linear time-invariant model, the Van Heusden Et Al. 2012 model was built upon by adding heart rate and patient body weight and basal rates. Each model has two measurable inputs, the injected insulin, and the HR, and one output, the glucose concentration measured by the CGM sensor, collected every 5 minutes, described in Eq. 3.9. To achieve an individualization of the model, the glucose-insulin transfer function is based on the discrete-time, linear time invariant model of insulin–glucose dynamics proposed in [1], with sample-period equal to 5 minutes. This model is defined as:

$$G_{P_i}(z) = \frac{k_i F c z^{-3}}{(1 - p_1 z^{-1})(1 - p_2 z^{-1})^2} \quad (3.16)$$

$$k_i = \frac{1800}{0.6 * BW} \quad (3.17)$$

$$c = -60(1 - p_1)(1 - p_2)^2 \quad (3.18)$$

where $BW = 55\text{kg}$, $F = 1.5$ is a unitless safety factor, c is for correcting gain and unit conversion, and $p_1 = 0.98$ and $p_2 = 0.965$ are the poles.

3.3.2 Adding the Identified HR-Glucose Transfer Function

Starting from the glucose-insulin model obtained in [1], the goal of the this step is to identify the HR-glucose transfer function ($G_{P_{hr}}$). The glucose-insulin model is linearized around the steady-state of the subject-specific, time-dependent basal input rate i_B (eq. 3.19), achieving a blood-glucose output of 120 mg/dL [1], while the glucose-HR is linearized around the

individual's deviance from the basal HR (hr_B) of 72 bpm [24].

$$i_B = i(t) * \frac{5}{60} \quad (3.19)$$

Denoting with $\Delta I(z)$, $\Delta HR(z)$ and $\Delta CGM(z)$ the Zeta transforms of inputs and output, the model has the following structure:

$$\Delta CGM(z) = G_{P_i}(z)\Delta I(z) + G_{P_{hr}}(z)\Delta HR(z) \quad (3.20)$$

where $\Delta HR(t) = HR(t) - HR_B$, $\Delta I(t) = I(t) - I_B(t)$, and $\Delta CGM(t) = CGM(t) - CGM_B$, G_{P_i} is the glucose-insulin model proposed in [1] and $G_{P_{hr}}$ is a transfer function to be estimated from the data. The transfer function $G_{P_{hr}}$ is defined as:

$$G_{P_{hr}} = \frac{\mu_{P_{hr}} z^{-1}}{(1 - z^{-1}p_{P_{hr1}})(1 - z^{-1}p_{P_{hr2}})(1 - z^{-1}p_{P_{hr3}})} \quad (3.21)$$

where $p_{P_{hr1}}, p_{P_{hr2}}, p_{P_{hr3}}$ are the poles of the two transfer functions, and μ_i and $\mu_{P_{hr}}$ are the corresponding gains.

Similar to the average model, the HR data are filtered to remove the high-frequency noise filter before being inputted into the model, while the insulin data is processed by a zero-order-hold to be synchronized with the HR sampling size. The identification phase was performed by estimating the model parameters by minimizing prediction error $\epsilon(t, \theta)$ between the model output $\Delta \hat{CGM}(n, \theta)$ and the observed ΔCGM data below: [22]:

$$\hat{\theta}_P = \arg \min \epsilon(n, \theta_P) \quad (3.22)$$

where

$$\epsilon(n, \theta_P) = \Delta CGM(n) - \Delta \hat{CGM}(n, \theta_P) \quad (3.23)$$

$\hat{\theta}_P$	Activity		
	Aerobic	Resistance	HIIT
$\mu_{P_{hr}}$	1.884e-05	0.0002	0.0018
$p_{P_{hr1}}$	-1.0000	-1.0000	1.0000
$p_{P_{hr2}}$	1.0000	0.9991	0.3350+0.9422i
$p_{P_{hr3}}$	0.9476	0.9407	0.3350-0.9422i

Table 3.2: Estimated parameters for equation [3.21](#) that were identified by minimizing the prediction error between the model output $\Delta\widehat{CGM}(n, \theta_P)$ and the observed ΔCGM data

where $\hat{\theta}_P$ is the vector of the optimal value of θ_P , which is defined as follows:

$$\theta_P = [\theta_{P_{hr}}]' \quad (3.24)$$

with

$$\theta_{hr} = [\mu_{P_{hr}} \quad p_{P_{hr1}} \quad p_{P_{hr2}} \quad p_{P_{hr3}}]'. \quad (3.25)$$

There are separate HR transfer functions identified for each exercise type of aerobic ($G_{P_{hra}}$), resistance ($G_{P_{hr_r}}$), and HIIT ($G_{P_{hr_i}}$).

Chapter 4

Results

4.1 Identification and Testing Dataset

For the model identification, different training datasets per exercise type were created. First, the sample size (N) per activity per was investigated. A small subset of subjects data could compromise the generalization capability of the model, but a large subset could produce a flat prediction due to the large inter-subject variability. A reference sample size is derived from the UVA/Padova simulator, which is equipped with a cohort of 10 subjects [25]. This metabolic simulator is accepted by Food and Drug Administration (FDA) as a substitute to animals trials for subcutaneous glucose sensing and insulin delivery [25]. In this thesis, we investigated $N = 5, 10, 15, 20$ and selected the best sample size based on the R^2 and Root Mean Square Error ($RMSE$) on the training data. The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{m} \sum_{n=1}^m (CG\hat{M}(n) - CGM(n))^2} \quad (4.1)$$

and the R^2 is defined as:

$$R^2 = \frac{\sum_{n=1}^m (CG\hat{M}(n) - \overline{CGM})^2}{\sum_{n=1}^m (CGM(n) - \overline{CGM})^2} \quad (4.2)$$

where \hat{CGM} is the predicted values, CGM is the observed values, and \overline{CGM} is the mean of the values.

Figure 4.1 shows the R^2 values (left) and the RMSE values (right) for the average aerobic model trained on 5, 10, 15, and 20 subject data. The performance was tested on the remaining subjects outside of the 20 patient training dataset for consistency. This figure shows the best performance in both R^2 and RMSE for the model trained on 10 subjects, as there is the highest R^2 value and minimalized RMSE value. Therefore, the rest of the analysis was computed on 10 subjects. This 10 subject training data set was decided based on diminishing returns as models trained on with 15 and 20 subjects dropped in performance. We selected a sample size $N = 10$.

The training datasets included HR, CGM, and insulin data for all the study video sessions of a cohort of 10 subjects in a merged data object structure (MATLAB iddata), for a total of 46 aerobic, 48 resistance, and 51 HIIT sessions. The identification data was used to identify both $\hat{\theta}$ and $\hat{\theta}_P$.

To validate the models, a testing dataset consisting of 10 different subjects per activity was used, consistent with the identification sample size. The individual and average RMSE and R^2 values were calculated for each of these 10 subjects. In addition, to validate the generalization capability of the identified models, the average RMSE and R^2 values were calculated for all of the subjects per exercise type, again excluding the identification dataset subjects. This consisted of 49 aerobic subjects, 48 resistance subjects, and 50 HIIT subjects, for a total of 238 aerobic, 223 resistance, and 254 HIIT sessions.

4.2 Model Performance

To assess the performance of the personalized and average models, a third “naive” model was introduced. This “naive” model is a first order hold defined as follows:

$$\hat{CGM}_{FOH}(n * Ts + t) = a_1 t + CGM(n * Ts) \quad 0 < t < T \quad (4.3)$$

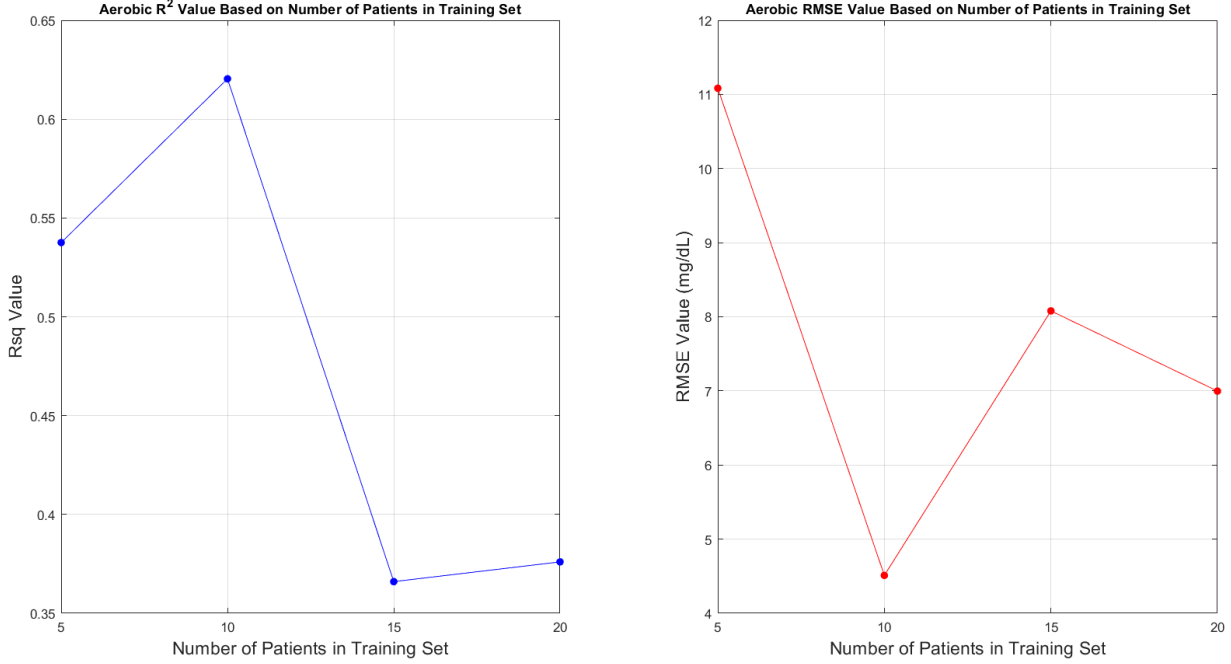


Figure 4.1: An R^2 (left) and RMSE (right) value comparison of the personalized model performance trained on 5, 10, 15, 20 subjects who performed aerobic exercise sessions. The testing set for this analysis consisted of all patients outside of the 20 subject training set for consistency

with

$$a_1 = \frac{CGM(n * Ts) - CGM((n - 1)Ts)}{Ts} \quad (4.4)$$

where \hat{CGM}_{FOH} is the first order hold output, CGM is the recorded signal, and $T_s = 10$ seconds is the sampling time. The initial state of the output was set to 120.

As seen in Tables [4.1](#) and [4.2](#), the RMSE and R^2 values are shown for all three models to compare, split up by exercise type, the first ten subjects of the testing dataset, and the entire dataset average. Table [4.3](#) shows the summarized results, which averages the different types of exercise to strictly compare the models for the first ten subjects of the testing dataset and all subjects. It was expected for the personalized model to perform the best, followed by the average model and then the naive model. This deemed true, as across the board, the personalized model performed consistently better than the average model and naive model per exercise type in both R^2 and RMSE values for both the 10 subjects but also all subjects.

Additionally, it was expected for Aerobic models to perform the best, as aerobic exercises consist of greater consistent heart rate changes, especially compared to resistance exercises. HIIT was suspected to be the worst performing model, as HIIT exercises alternate between high heart rate to resting periods. These expectations proved true when assessing the personalized model on R^2 values for all subjects, but not the 10 subject cohort. In the 10 subject cohort, the resistance model performed the best, followed by the aerobic, then HIIT for both RMSE and R^2 .

A visual reference of the average and personalized model performance are shown for three different patients per exercise type in figures [4.2](#), [4.3](#), and [4.4](#). In each of the figures, the raw CGM data is shown in blue, while the model CGM estimate is displayed in orange. With [4.2](#) we see that for the average model (left), there is difficulty in approximating the initial CGM value, and fails to closely follow and predict the changes in the true CGM data. However, consistently, the personalized model (right) is able to predict the trends and changes of the true CGM data. This pattern is consistent across [4.3](#) and [4.4](#) as well. Notably, in [4.3](#) we see that the average model commonly plateaus in its predictions, failing to predict the ending values of CGM. However, the personalized model is able to continuously increase or decrease to predict the quick increase and decrease changes in CGM, likely due to the nature of HIIT exercises. We see with the middle patient, that the average model fails to predict a decrease in the CGM values followed by an increase. However, the personalized model is able to capture and predict this dip in CGM values more rapidly and accurately.

RMSE	Personalized Model			Average Model			Naive/Simplistic Model		
	Aerobic	Resistance	HIIT	Aerobic	Resistance	HIIT	Aerobic	Resistance	HIIT
Subject #1	2.806 ± 2.636	11.908 ± 16.155	13.916 ± 13.588	6.430 ± 7.637	21.805 ± 29.799	20.170 ± 16.440	13.521 ± 7.589	59.322 ± 76.508	38.552 ± 27.387
Subject #2	3.939 ± 5.694	0.676 ± 0.632	4.802 ± 6.024	5.876 ± 7.040	0.590 ± 0.243	8.111 ± 13.721	14.129 ± 13.461	7.202 ± 6.868	32.160 ± 41.326
Subject #3	9.576 ± 6.272	1.423 ± 0.795	4.669 ± 6.571	19.216 ± 18.346	1.262 ± 0.856	14.050 ± 23.338	86.476 ± 78.298	10.762 ± 7.201	29.746 ± 34.089
Subject #4	7.962 ± 5.532	5.770 ± 4.666	1.424 ± 1.627	20.632 ± 17.368	6.994 ± 6.631	3.367 ± 5.865	41.600 ± 28.938	19.124 ± 11.885	16.605 ± 29.182
Subject #5	6.419 ± 3.278	2.539 ± 1.534	4.140 ± 2.588	18.684 ± 13.931	6.642 ± 5.578	9.390 ± 10.581	47.984 ± 35.748	19.837 ± 9.281	21.537 ± 13.438
Subject #6	1.478 ± 1.570	4.778 ± 5.271	2.987 ± 2.902	2.160 ± 2.031	9.894 ± 14.177	3.365 ± 4.527	15.089 ± 18.700	28.382 ± 25.712	31.562 ± 36.233
Subject #7	3.606 ± 4.998	2.893 ± 3.075	2.165 ± 2.620	6.928 ± 9.161	5.127 ± 8.748	2.410 ± 4.375	25.041 ± 31.127	29.833 ± 27.541	13.774 ± 25.054
Subject #8	4.692 ± 4.168	5.843 ± 4.724	4.101 ± 3.419	9.338 ± 10.449	6.283 ± 5.624	5.324 ± 4.720	25.955 ± 21.863	18.541 ± 16.998	12.610 ± 5.802
Subject #9	2.508 ± 1.535	3.782 ± 2.965	4.750 ± 3.581	5.144 ± 3.278	8.423 ± 7.524	16.324 ± 17.923	11.501 ± 5.742	22.099 ± 22.597	30.478 ± 23.879
Subject #10	3.448 ± 2.955	2.253 ± 1.908	3.444 ± 3.024	7.023 ± 8.328	3.622 ± 4.087	2.098 ± 2.428	19.286 ± 4.650	18.643 ± 15.029	23.473 ± 12.626
Mean ± std for subjects 1-10	4.588 ± 4.572	3.514 ± 4.367	4.786 ± 6.329	9.971 ± 11.688	5.795 ± 8.589	8.526 ± 12.340	28.642 ± 34.318	20.461 ± 21.772	24.997 ± 24.762
Mean ± std for all subjects (aerobic: 49 resistance: 48 HIIT: 50)	4.132 ± 6.480	6.026 ± 24.815	6.723 ± 35.676	7.857 ± 10.370	10.169 ± 29.109	6.934 ± 10.233	22.080 ± 25.879	19.539 ± 22.299	22.163 ± 23.086

Table 4.1: RMSE values for the personalized, average, and naive/simplistic model split up by exercise type. The RMSE values for the first ten subjects of the testing dataset are shown, as well as the mean and standard deviation for all subjects.

R ²	Personalized Model			Average Model			Naive/Simplistic Model		
	Aerobic	Resistance	HIIT	Aerobic	Resistance	HIIT	Aerobic	Resistance	HIIT
Subject #1	0.816 ± 0.097	0.829 ± 0.129	0.831 ± 0.096	0.395 ± 0.264	0.547 ± 0.139	0.494 ± 0.350	0.012 ± 0.014	0.017 ± 0.021	0.007 ± 0.009
Subject #2	0.740 ± 0.286	0.841 ± 0.067	0.855 ± 0.087	0.406 ± 0.419	0.717 ± 0.343	0.824 ± 0.131	0.011 ± 0.010	0.017 ± 0.007	0.008 ± 0.008
Subject #3	0.776 ± 0.341	0.716 ± 0.274	0.885 ± 0.044	0.635 ± 0.373	0.779 ± 0.221	0.706 ± 0.400	0.007 ± 0.008	0.011 ± 0.007	0.015 ± 0.012
Subject #4	0.806 ± 0.192	0.573 ± 0.409	0.660 ± 0.357	0.375 ± 0.239	0.678 ± 0.389	0.606 ± 0.357	0.006 ± 0.010	0.017 ± 0.012	0.013 ± 0.004
Subject #5	0.762 ± 0.250	0.901 ± 0.078	0.740 ± 0.324	0.353 ± 0.287	0.652 ± 0.291	0.574 ± 0.316	0.007 ± 0.008	0.010 ± 0.010	0.008 ± 0.008
Subject #6	0.805 ± 0.060	0.837 ± 0.171	0.876 ± 0.050	0.509 ± 0.336	0.677 ± 0.292	0.901 ± 0.055	0.018 ± 0.011	0.004 ± 0.004	0.013 ± 0.009
Subject #7	0.738 ± 0.283	0.842 ± 0.080	0.529 ± 0.399	0.423 ± 0.308	0.834 ± 0.114	0.690 ± 0.160	0.014 ± 0.011	0.015 ± 0.012	0.010 ± 0.005
Subject #8	0.787 ± 0.171	0.576 ± 0.344	0.601 ± 0.183	0.552 ± 0.164	0.511 ± 0.434	0.432 ± 0.311	0.008 ± 0.008	0.014 ± 0.010	0.012 ± 0.012
Subject #9	0.829 ± 0.097	0.884 ± 0.060	0.866 ± 0.108	0.415 ± 0.272	0.536 ± 0.289	0.571 ± 0.392	0.003 ± 0.004	0.006 ± 0.010	0.008 ± 0.010
Subject #10	0.867 ± 0.025	0.882 ± 0.114	0.755 ± 0.112	0.606 ± 0.340	0.848 ± 0.106	0.901 ± 0.106	0.011 ± 0.011	0.009 ± 0.004	0.018 ± 0.007
Mean ± std for subjects 1-10	0.788 ± 0.197	0.789 ± 0.221	0.756 ± 0.232	0.455 ± 0.290	0.687 ± 0.284	0.663 ± 0.304	0.010 ± 0.010	0.011 ± 0.009	0.011 ± 0.008
Mean ± std for all subjects (aerobic: 49 resistance: 48 HIIT: 50)	0.776 ± 0.197	0.752 ± 0.239	0.737 ± 0.239	0.445 ± 0.278	0.606 ± 0.302	0.656 ± 0.288	0.012 ± 0.010	0.011 ± 0.010	0.011 ± 0.009

Table 4.2: R^2 values for the personalized, average, and naive/simplistic model split up by exercise type. The R^2 values for the first ten subjects of the testing dataset are shown, as well as the mean and standard deviation for all subjects.

RMSE	Personalized Model	Average Model	Naive/Simplistic Model
Mean for Patients 1-10 across all exercise types	4.296 ± 5.165	6.431 ± 10.995	24.700 ± 27.477
Mean for all subjects across all exercise types	5.627 ± 25.368	8.320 ± 18.794	21.261 ± 23.805
R ²	Personalized Model	Average Model	Naive/Simplistic Model
Mean for Patients 1-10 across all exercise types	0.778 ± 0.217	0.602 ± 0.293	0.011 ± 0.009
Mean for all subjects across all exercise types	0.755 ± 0.226	0.569 ± 0.290	0.011 ± 0.010

Table 4.3: A summarized table of average ± standard deviation RMSE and R^2 values shown for the personalized, average, and naive/simplistic model. The mean and standard deviation values for the RMSE and R^2 values for the first ten subjects of the testing dataset are shown, as well as all subjects.

Aerobic Patients Model Performance
Average (left) and Personalized (right)

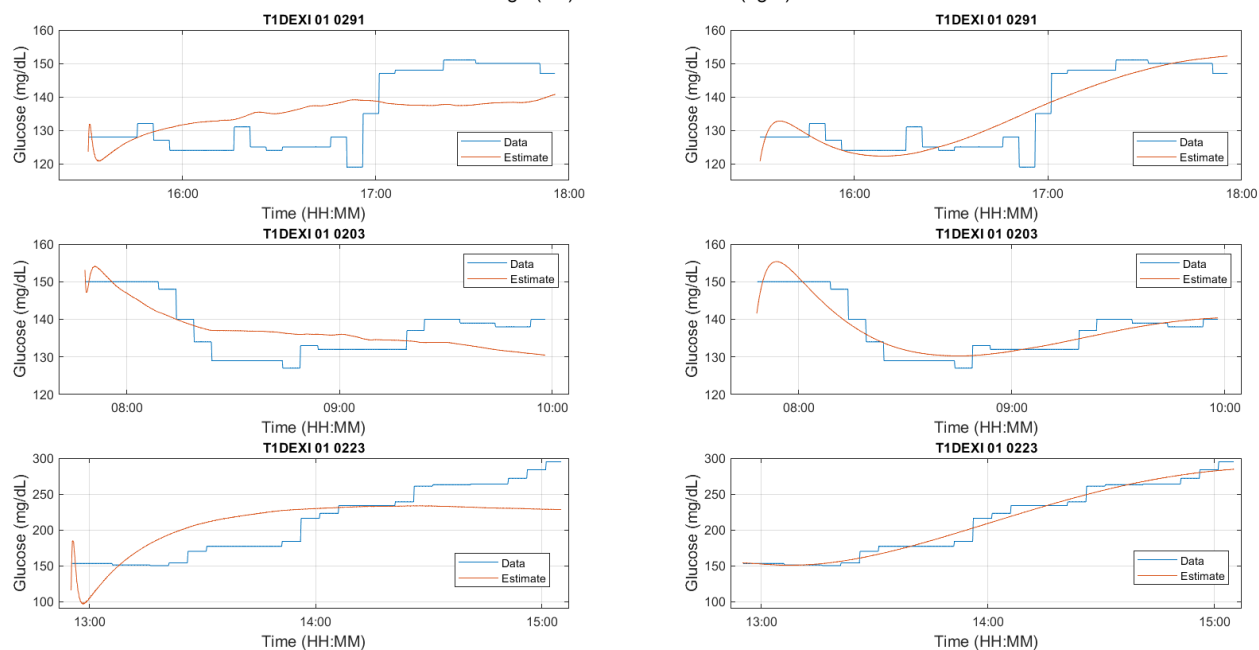


Figure 4.2: An example of the reference glucose data for an aerobic activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.

Resistance Patients Model Performance
Average (left) and Personalized (right)

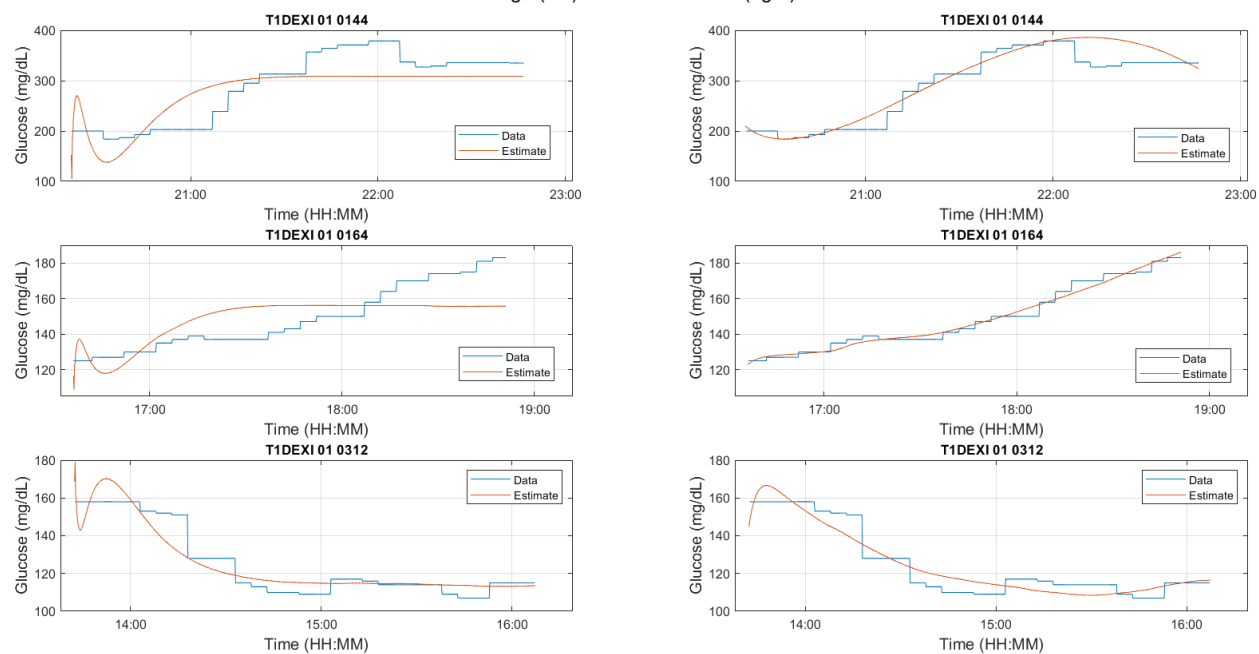


Figure 4.3: An example of the reference glucose data for a resistance activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.

Interval Patients Model Performance
Average (left) and Personalized (right)

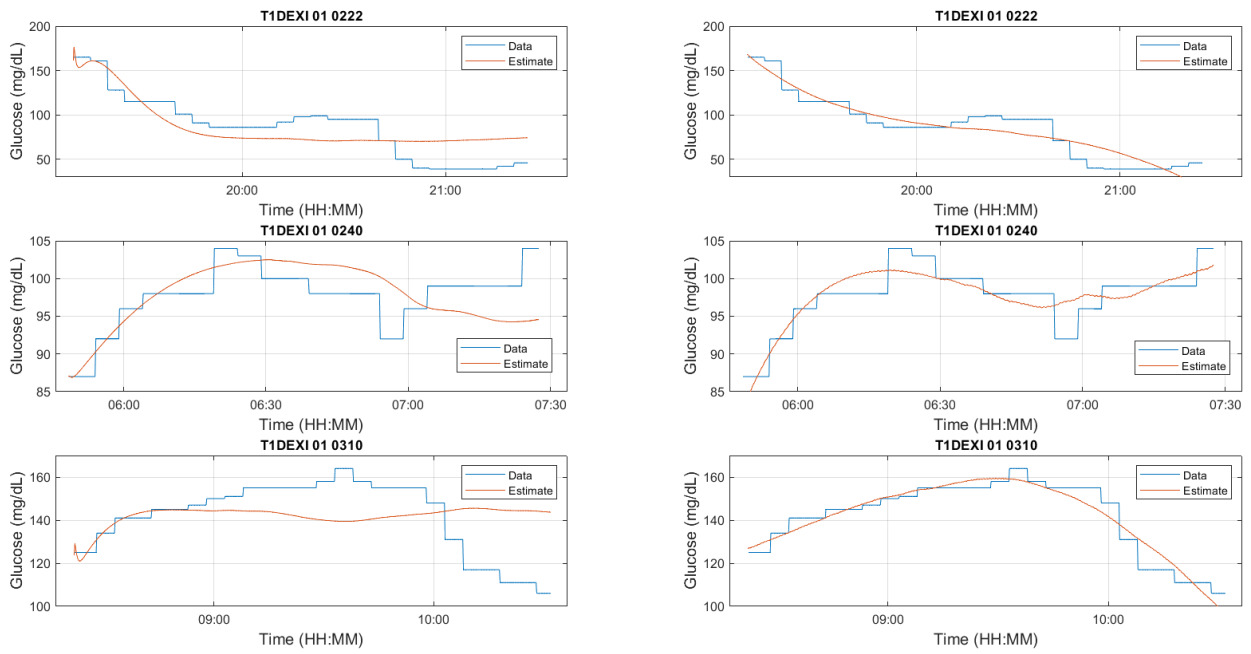


Figure 4.4: An example of the reference glucose data for an HIIT activity (blue) compared with the average (left) and personalized (right) model predictions are reported for three subjects of the testing dataset.

Chapter 5

Discussion

Through this thesis, a personalized model predicting changes in CGM levels during aerobic, resistance, and HIIT physical activity for T1D individuals was identified. Upon testing on all subjects across all exercise types, the personalized model had an R^2 value of 0.755 ± 0.226 and an RMSE value of 5.627 ± 25.368 , as reported in Table 4.3. This model performed under the circumstances of pre-identified exercise types, and would need to be further developed to identify the exercise type without manual user input.

Based on both RMSE and R^2 values, for the first 10 subjects, the personalized model performed best on resistance subjects, then aerobic subjects, and then HIIT subjects. However, for all patients in the dataset, the personalized model performed best on aerobic subjects, then resistance, and then HIIT subjects. For the first 10 subjects, the average model performed the best on resistance subjects, then, HIIT subjects, and then aerobic subjects. For all subjects, based on RMSE, the average model performed best on HIIT subjects, then aerobic subjects, and then resistance subjects. However, for all subjects, based on R^2 , the average model performed best on HIIT subjects, then resistance subjects, and then aerobic subjects. On average, the personalized model consistently outperformed both the average and naive/simplistic model, as summarized in table 4.3.

For aerobic subjects, the personalized model would consistently match up with the true

data especially during the exercise segment compared to the average model predictions (Fig. 4.2). For resistance subjects, the average model tended to plateau lower than the climbing glucose values, while the personalized model continued to climb (Fig. 4.3). For HIIT subjects, the average model tended to mispredict the ending glucose value of the exercise segment (Fig. 4.4).

Although the personalized model consistently performed better than the average and naive/simplistic model, there were very high values of RMSE standard deviation that need to be improved to be reliable for individuals with T1 Diabetes. Adding in an additional input of step count may also improve the model, to clearly differentiate an aerobic and resistance exercise. Resistance exercises are difficult to solely predict based on HR, as many weight-lifting activities don't cause an obvious disruption to HR compared to aerobic exercises, which step count could assist in identifying. An additional way to improve the personalized model is taking into account age and gender of the individual, as the average HR may vary based on demographics. The baseline value can be adjusted based on the individual to improve the personalized model to become more specific.

There needs to be more research into the identification of the start of an exercise, as in practice, the model needs to be able to differentiate between the different exercise types. In order to employ the correct model between the aerobic, resistance, and HIIT models, the observer needs to be able to identify exercise types. However, to implement these three different models in a preliminary test, the user would need to input which exercise that they are about to perform. By doing so, there is a potential impact of decreasing hypoglycemic episodes due to improved glucose prediction and control.

In addition, further specifics need to be explored such as the time of day affecting the model in addition to an individual's sleep chronotype. The time that a physical activity is performed, whether it is early morning, afternoon, or night may affect glucose control. Nutrition may also affect an individual's response to physical exercise, as taking into account a meal or snack before an exercise can drastically change the behavior of glucose levels.

This study was also limited to training on study video data, which is controlled and clearly separated between exercise types of aerobic, resistance, and HIIT exercise. Individuals in this study participated by following exact exercise videos, however, the personal motivation of each study participant may have contributed to variance in performing the same exercise. In current AID systems, users announce an 'exercise'. However, with the model developed in this thesis, the exercise type can be announced more specifically, with a selection of aerobic, resistance, or HIIT. However, current limitations still persist for categorically ambiguous activities such as skiing or various sports.

Bibliography

- [1] K. van Heusden, E. Dassau, H. C. Zisser, D. E. Seborg, F. J. Doyle III, Control-relevant models for glucose control using a priori patient characteristics, *IEEE Transactions on Biomedical Engineering* 59 (7) (2011) 1839–1849.
- [2] M. A. Atkinson, G. S. Eisenbarth, A. W. Michels, Type 1 diabetes, *The Lancet* 383 (9911) (2014) 69–82.
- [3] A. Katsarou, S. Gudbjörnsdóttir, A. Rawshani, D. Dabelea, E. Bonifacio, B. J. Anderson, L. M. Jacobsen, D. A. Schatz, Å. Lernmark, Type 1 diabetes mellitus, *Nature Reviews Disease Primers* 3 (1) (2017) 1–17.
- [4] M. S. Rahman, K. S. Hossain, S. Das, S. Kundu, E. O. Adegoke, M. A. Rahman, M. A. Hannan, M. J. Uddin, M.-G. Pang, Role of insulin in health and disease: an update, *International Journal of Molecular Sciences* 22 (12) (2021) 6403.
- [5] A. D. Association, 2. classification and diagnosis of diabetes: standards of medical care in diabetes—2018, *Diabetes Care* 41 (Supplement 1) (2018) S13–S27.
- [6] N. H. Cho, J. Shaw, S. Karuranga, Y. Huang, J. da Rocha Fernandes, A. Ohlrogge, B. Malanda, *Idf diabetes atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045*, *Diabetes Research and Clinical Practice* 138 (2018) 271–281.
- [7] A. D. Association, et al., 4. lifestyle management: standards of medical care in diabetes-2018, *Diabetes Care* 41 (Supplement 1) (2018) S38–S50.

- [8] A. Meglio, B. Molina, R. Oram, Type 1 diabetes, *The Lancet* 391 (10138) (2018) 2449–2462.
- [9] M. Cescon, D. Choudhary, J. E. Pinsker, V. Dadlani, M. M. Church, Y. C. Kudva, F. J. Doyle III, E. Dassau, Activity detection and classification from wristband accelerometer data collected on people with type 1 diabetes in free-living conditions, *Computers in Biology and Medicine* 135 (2021) 104633.
- [10] B. Ozaslan, S. D. Patek, M. D. Breton, Impact of daily physical activity as measured by commonly available wearables on mealtime glucose control in type 1 diabetes, *Diabetes Technology & Therapeutics* 22 (10) (2020) 742–748.
- [11] M. C. Riddell, I. W. Gallen, C. E. Smart, C. E. Taplin, P. Adolfsson, A. N. Lumb, A. Kowalski, R. Rabasa-Lhoret, R. J. McCrimmon, C. Hume, et al., Exercise management in type 1 diabetes: a consensus statement, *The Lancet Diabetes & Endocrinology* 5 (5) (2017) 377–390.
- [12] S. R. Colberg, R. Laan, E. Dassau, D. Kerr, Physical activity and type 1 diabetes: time for a rewire?, *Journal of Diabetes Science and Technology* 9 (3) (2015) 609–618.
- [13] E. M. Aiello, S. Deshpande, B. Özaslan, K. L. Wolkowicz, E. Dassau, J. E. Pinsker, F. J. Doyle III, Review of automated insulin delivery systems for individuals with type 1 diabetes: Tailored solutions for subpopulations, *Current Opinion in Biomedical Engineering* 19 (2021) 100312.
- [14] J. Garcia-Tirado, S. A. Brown, N. Laichuthai, P. Colmegna, C. L. Koravi, B. Ozaslan, J. P. Corbett, C. L. Barnett, M. Pajewski, M. C. Oliveri, et al., Anticipation of historical exercise patterns by a novel artificial pancreas system reduces hypoglycemia during and after moderate-intensity physical activity in people with type 1 diabetes, *Diabetes Technology & Therapeutics* 23 (4) (2021) 277–285.

- [15] M. D. Breton, S. A. Brown, C. H. Karvetski, L. Kollar, K. A. Topchyan, S. M. Anderson, B. P. Kovatchev, Adding heart rate signal to a control-to-range artificial pancreas system improves the protection against hypoglycemia during exercise in type 1 diabetes, *Diabetes Technology & Therapeutics* 16 (8) (2014) 506–511.
- [16] N. Hobbs, I. Hajizadeh, M. Rashid, K. Turksoy, M. Breton, A. Cinar, Improving glucose prediction accuracy in physically active adolescents with type 1 diabetes, *Journal of Diabetes Science and Technology* 13 (4) (2019) 718–727.
- [17] R. N. Bergman, Y. Z. Ider, C. R. Bowden, C. Cobelli, Quantitative estimation of insulin sensitivity., *American Journal of Physiology-Endocrinology And Metabolism* 236 (6) (1979) 667.
- [18] M. Rashid, S. Samadi, M. Sevil, I. Hajizadeh, P. Kolodziej, N. Hobbs, Z. Maloney, R. Brandt, J. Feng, M. Park, et al., Simulation software for assessment of nonlinear and adaptive multivariable control algorithms: glucose–insulin dynamics in type 1 diabetes, *Computers & Chemical Engineering* 130 (2019) 106565.
- [19] M. Riddell, Exercise and type 1 diabetes: Preliminary results from the type 1 diabetes exercise initiative (t1dexi), in: *Diabetes Technology & Therapeutics*, Vol. 24, 2022, pp. A6–A6.
- [20] M. C. Riddell, Z. Li, R. L. Gal, P. Calhoun, P. G. Jacobs, M. A. Clements, C. K. Martin, F. J. Doyle III, S. R. Patton, J. R. Castle, et al., Examining the acute glycemic effects of different types of structured exercise sessions in type 1 diabetes in a real-world setting: The Type 1 Diabetes and Exercise Initiative (T1DEXI), *Diabetes Care* (2023) 221721.
- [21] J. B. Rawlings, D. Q. Mayne, M. Diehl, *Model predictive control: theory, computation, and design* 2 (2017).
- [22] L. Ljung, *System identification toolbox* (1995).

- [23] T. Knudsen, A new method for estimating armax models, IFAC Proceedings Volumes 27 (8) (1994) 895–901.
- [24] Y. Ostchega, K. S. Porter, J. Hughes, C. F. Dillon, T. Nwankwo, Resting pulse rate reference data for children, adolescents, and adults; united states, 1999-2008, National Health Statistics Reports 41 (2011) 1–16.
- [25] C. D. Man, F. Micheletto, D. Lv, M. Breton, B. Kovatchev, C. Cobelli, The uva/padova type 1 diabetes simulator: new features, Journal of Diabetes Science and Technology 8 (1) (2014) 26–34.