GLUCOSE PREDICTION DURING PHYSICAL ACTIVITY FOR TYPE 2 DIABETICS ON NON-INSULIN INTENSIVE THERAPIES

Lindsey Uribe

A REPORT

in

Data Science

Presented to the Data Science Faculties of the University of Oregon

in

Partial Fulfillment of the Requirements for the

Baccalaureate Degree of Science

2022

Faculty Advisor	Domain Advisor	Industrial Advisor
Joe Sventek	Courosh Mehanian	Mark Derdzinski
Data Science Program Head	Associate Research Professor	Senior Data Science Manager
University of Oregon	University of Oregon	Dexcom

ACKNOWLEDGMENT

"Prediction is very difficult, especially if it's about the future."
– credited to Niels Bohr

Over the course of this project, I have written over 29,367 lines of code, created 48 new scripts, processed over 49 queries, and created countless new plots (not all of them fruitful, but all educational). The simplest of these tasks I never would have imagined myself doing just a few years ago. None of this would have been possible were it not for the incredible infrastructure put into place at the University of Oregon to facilitate the creation of a Data Science program. As I told members of the University of Oregon Board of Trustees once before, I truly don't know what I would have done were it not for this program. I am eternally grateful.

The data supplied for this project was the result of a study funded by Dexcom Inc., a maker of continuous glucose devices. For the entirety of my time working with individuals at Dexcom, I have continued to be impressed by their leadership, their commitment to their core values as well as the incredible amount of organization exhibited in every facet of their operations. I am delighted to have had the opportunity to work with such amazing people, and such meaningful data.

To my friends and family, thank you for being supportive and understanding as I indulge in these academic pursuits.

To my advisors Courosh Mehanian, Mark Derdzinski, and Joe Sventek thank you for your kindness, patience, willingness to listen, and your sincere interest in the outcome of this project.

ABSTRACT

GLUCOSE PREDICTION DURING PHYSICAL ACTIVITY FOR TYPE 2 DIABETICS ON NON-INSULIN INTENSIVE THERAPIES

Lindsey Uribe

Physical activity has long been recommended by health experts for acutely managing blood sugars for individuals with diabetes. Although substantial research has shown that physical activity improves glycemic outcomes in the Type 2 diabetic, non-insulin intensive therapy (T2D NIIT) population, the relationship between physical activity and acute changes in glucose values population remains poorly understood. Many attempts at predicting glucose values for Type 1 Diabetes have been made with various levels of success, but little research exists to date regarding glucose prediction for the T2D NIIT population, and even less exists regarding the effects physical activity may have on glucose values. To predict changes in glucose associated with physical activity, a Long Short-Term Memory Recurrent Neural Network (LSTM RNN) is trained on glucose values detected via continuous glucose monitoring 2 hours prior to a bout of physical activity. The addition of a covariate feature containing activity-tracker-derived step cadences improved these predictions. Future glucose values are predicted for short-term time horizons, with RMSEs of 7.1, 11.1 and 13.9 mg/dL for 15-minute, 30-minute, and 60-minute prediction horizons respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

PREFACE

I have known three people who have lost a foot to diabetic peripheral neuropathy. Quite literally, I have witnessed loss of limb by amputation because of Type 2 diabetes. I take this disease very seriously not only because of the havoc I've seen it wreak on someone's life, or because of the metronome that could be set to the statistics that can be rattled off about this disease, but also because I know that I too am at risk. I see myself, my friends, and my family in each one of these bouts of activity, and in each row of every table. I hope that the following work helps in some small way to call attention to the need for further research and support for this community.

Chapter 1

Introduction

1.1 Type 2 Diabetes and Physical Activity: Motivating Questions

In 2019, the World Health Organization (WHO) estimated that roughly 2 million deaths were attributable to diabetes, and that over 95% of those with diabetes are individuals with Type 2 diabetes (T2D). In this same year, the International Diabetes Foundation (IDF) estimated that one death occurred every eight seconds due to diabetes-related complications. The IDF also notes that over 9% of the world's adult population is diabetic, and projects an increase in these numbers, year over year (IDF Diabetes Atlas, 2019). The pervasive rise of T2D is thus a health disorder of international concern.

The primary causes of T2D are two-fold: Chronic over-exposure to carbohydrate-rich foods paired with sedentary behavior results in a slow and subtle progression toward insulin resistance. This impaired cellular response to insulin marks a decline in the function of the metabolic pathways that facilitate the transport of glucose across cellular membranes. The resulting inability to effectively transport glucose from the bloodstream leads to various forms of tissue damage, often manifesting in the form of neuropathies, cardiovascular impairment, chronic kidney disease, and if left untreated, death.

According to the 2022 American Diabetes Association (ADA) Standards of Care, the prevention or delay of the onset of T2D and associated comorbidities relies heavily on increased physical activity. However, while exercise is most often associated with acute decreases in glucose, there are times when exercise can be associated with increases in glucose values. Adrenaline, stress, time of day, and the intensity of the activity can all cause glucose levels to increase, but exactly when these increases may occur in tandem with physical activity is difficult to predict (Exercise...|ADA, n.d.; Hinshaw et al., 2013; Tyler et al., 2022). These misalignments with expert advice and acute glycemic outcomes represent a cognitive burden and can be stressful and even frustrating for individuals with T2D (Beverly et al., 2011). Thus, a significant and frustrating knowledge gap stands in the way of many individuals living with T2D from understanding the impact that physical activity has on their glucose values.

One relatively recent advancement in diabetes management may serve to help bridge this gap. Continuous glucose monitoring devices (CGMs) have opened new pathways for exploration into the acute effects of physical activity on glucose values. Over the last 23 years, the increasing use of CGMs has been revolutionary for the management and monitoring of glucose values for both Type 1 and Type 2 diabetics. Concurrent with the development of CGM technology, the development and innovations of wearable fitness trackers have added new dimensionality to the plethora of questions that could be asked of glucose dynamics. These advances have resulted in troves of new data which can be used to fuel data-driven lines of questioning about physical activity and diabetes. Namely:

- Is it possible to predict an individual's glucose values from 15 minutes up to 1 hour after the start of a bout of activity?
- What are the relevant features of a bout of activity that might help to improve these predictions?

Although Type 2 diabetes and prediabetes are pervasive conditions, with dangerous implications for loss of life and a reduced quality of life for millions of individuals, these conditions are also preventable. Prevention is possible via behavior modification, with significant focus placed on increased physical activity. Understanding the impact that physical activity has on glucose values can serve to contribute to T2D prevention and treatment by:

- providing valuable insights into the subtle and nuanced interplay between physical activity and glucose values in the T2D population,
- providing individuals with T2D a pre-emptive estimation of how that physical activity will impact their future glucose values, up to a short time in the future,
- alleviating some of the frustration and confusion experienced by individuals with T2D to facilitate improved glycemic control, and
- serving as a basis upon which future predictive health optimization models might be built.

It is the author's hope that this work may contribute to each of these aims, and to educate and enlighten the reader as to the possibility of glucose prediction for the T2D population.

1.2 Research Goals

This project seeks to explore and quantify the relationship between physical activity and its acute effects on glucose values by predicting the changing glucose values for short-term time horizons. Thus, this project has three primary objectives:

1. Determine if a model can be developed which can utilize past glucose values collected prior to a bout of activity to predict future glucose values in the 15 minutes immediately following, and up to 1 hour after the start of a bout of activity.
2. Given a functional model capable of short-term glucose prediction, determine the relevant features germane to the individual or their bout of exercise that might further improve these predictions.
3. Assess the model's ability to generalize to never-before-seen data. This is the final stage of model development and is accomplished by evaluating the model on a holdout dataset composed of individuals' bouts of activity that are withheld from all prior analyses.

1.3 Dissertation Structure

Chapter 1 of this dissertation began with an introduction to the global problem of increasingly common diagnoses of Type 2 diabetes, the complex relationship between glucose values and physical activity, and an overview of the motivating questions driving the work herein. This chapter sets forth the objectives of this project and delineates the proposed approach. Chapter 2 soon follows with a review of past and pertinent works relevant to understanding the history of glucose measurement, glucose prediction, and the importance of physical activity for the T2D and prediabetic population. This literature review serves as the guiding foundation from which the methods for this study have been developed. Chapter 3 describes the methodologies used for the development of a long short-term memory recurrent neural network (LSTM RNN), which is used

to approach the problem of short-term glucose prediction. Chapter 4 of this dissertation delves into the evaluation of those methodologies and surveys the results of the model outputs. In Chapter 5, these results are discussed and interpreted. The limitations of this study are also described, and suggestions for further research are proffered. Chapter 6 concludes these works.

1.4 Project Scope

This project focuses on the acute, short-term effect of physical activity on glucose values in a subset of the T2D population, namely individuals with Type 2 diabetes or prediabetes that are not on any form of insulin-intensive therapies. The sub-setting of the T2D population, to the exclusion of individuals on any form of fast-acting or intermediate-acting insulin, is to ensure that short-term, acute decreases in glucose values are not the result of the immediate effects of fast-acting or intermediate-acting insulin doses. Thus, only pre-diabetic and T2D individuals on non-insulin-intensive therapies are included in this study, henceforth referred to collectively as the T2D NIIT population.

The observational data used for this project are the result of the T2 Help study, a 12-week study in which participants wore Fitbit activity trackers and Dexcom G6 CGMs and provided pertinent information related to their diabetic or prediabetic condition. Each individual included in this study was previously diagnosed as either prediabetic or Type 2 diabetic. With this backdrop of clinically significant demographic data, 199 individuals' physical activity and glucose values were recorded in tandem, for 12 weeks and used for predictive modeling.

This study focuses on short-term glucose predictions with time horizons no longer than 60 minutes. This is in accordance with the practices of current studies attempting to predict glucose values. Long-term predictions and patterns of daily exercise are not analyzed.

No bouts of physical activity lasting less than 10 minutes are included, and bouts of physical activity never achieving a minimum threshold of 60 steps per minute are not considered here. Thus, bouts of physical activity lasting 10 minutes or more, composed of step cadences of 60 steps per minute or higher are selected for and filtered to ensure non-overlap of bout data. Characteristics of each bout and participant are also recorded, including total bout duration, fluctuating step counts detected during the bout, as well as body mass index (BMI) and years since diagnosis of diabetes or prediabetes. CGM-derived glucose values are then selected up to 2 hours prior to each collected bout of physical activity, and up to 60 minutes after the start of physical activity. Participants with detected bouts of physical activity are then randomly selected for training, validation, and test sets.

An LSTM RNN is developed to predict glucose values. Model design allows for univariate and covariate inputs, with glucose predictions as outputs. Model refinement and hyperparameter tuning are performed on training and validation sets only. Upon final determination of the most performant model hyperparameters and covariate features using root mean squared error (RMSE) metrics, the final version of the developed model is run on the testing set.

Chapter 2

Literature Review

2.1 Type 2 Diabetes: An Overview

In nondiabetic individuals, cellular glucose uptake is mediated via insulin—a hormone produced by the pancreas. For individuals with diabetes, this mechanism of glucose uptake is impaired, resulting in overly persistent hyperglycemia (high blood glucose levels). There are two predominant forms of diabetes: Type 1 (T1D) and Type 2 diabetes (T2D). T1D is a relatively rare disorder, caused by an autoimmune response that results in an inability to effectively produce insulin. In comparison, T2D is remarkably common.

Of those living with diabetes, more than 95% have T2D (World Health Organization, 2022). In T2D, chronic over-exposure to high levels of glucose causes insulin resistance, and an impaired ability for cells to respond to insulin, which in turn results in an increased frequency of hyperglycemic states. Over time, hyperglycemia can cause tissue damage, typically manifesting in the form of neuropathies affecting radially dispersed nervous and soft tissues first, then slowly advancing to proximal tissues (Pop-Busui, 2010). Eventually, serious comorbidities can develop, involving damage to vital organs like the heart, kidneys, liver, eyes, and brain.

As shown in Figure 2.1 (Battelino et al., 2019), the ideal range of glycemic values lies within 70 – 180 mg/dl for all age groups. The percentage of time spent within this range, known simply as time in range (TIR), should be maximized. Although elderly individuals and individuals at greater risk of comorbidities may spend more time at values above 180 mg/dl, the amount of time spent above this range should be minimized. But diabetes management requires constant vigilance and many individuals with T2D find the stress and cognitive burden of controlling their glycemic values to be overwhelming (Kane et al., 2018). This diabetes-related distress is in turn related to poorer glycemic control, non-adherence to treatment plans, and a feedback loop of reduced health outcomes.

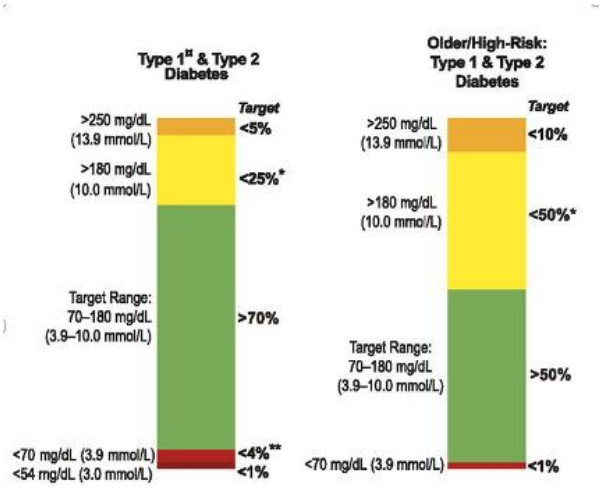


Figure 2.1 Target glucose values. Target ranges for all ages lie between 70-180 mg/dl. However, elderly individuals and individuals at high risk of comorbidities may often spend more time in higher ranges. Note. From "Clinical Targets for Continuous Glucose Monitoring Data Interpretation; Recommendations From the International Consensus on Time in Range" by Battelino et al., 2019, Diabetes Care, 42(8), p. 1593-1603 (https://diabetesjournals.org/care/article/42/8/1593/36184/Clinical-Targets-for-Continuous-Glucose-Monitoring).

T2D is a slowly advancing condition, characterized by a progression toward greater insulin resistance over long periods of time. Subsequently, prediabetes is a level of insulin resistance that also manifests in chronic high blood glucose values and typically arises before the onset of full-blown T2D. Both prediabetes and T2D are increasingly common. The Centers for Disease Control estimates that over one-third of the population of the United States is either prediabetic or diabetic, with a significant majority of undiagnosed cases (National Diabetics Statistics Report, 2022). The slow and subtle progression of diabetes makes this disease difficult to detect before significant tissue damage has occurred but also suggests many years in which individuals can make dietary and lifestyle changes to prevent (or possibly even reverse) T2D.

2.1.1 Diabetes Management

Methods of treatment and prevention of T2D typically take many multi-layered approaches but predominantly focus on either medication, lifestyle changes, or both. Roughly 70% of individuals living with T2D are on non-insulin intensive therapies (NIIT), meaning the condition is still treatable via lifestyle changes and medications other than insulin (Selvin et al., 2016). However, the typical progression of T2D treatments leads to a pattern referred to as "medication stacking", or polypharmacy. This means that upon diagnosis—although clinicians will typically prescribe the minimal level of medication possible—when improvements in glycemic metrics cease to improve, or initial improvements wane, new medications are prescribed, often without the removal of previous medications (Grant et al., 2004; Sharma et al., 2013). This slow increase in the number of add-on medications used to treat diabetes tends to increase as the disease progresses, each individual ages, and new comorbidities arise.

The array of T2D medications and their potential combinations are nearly as diverse as the heterogeneous community they aim to serve. Prescribers must take many multi-layered factors into consideration when determining the best methods of treatment since patients with T2D typically present with complex medical histories and more than one diabetes-related comorbidity. One medication, Metformin, is most prescribed in tandem with lifestyle modification and can also be considered a first line of defense against the onset of diabetes. Other commonly prescribed medications include SGLT2 Inhibitors (gliflozins), GLP-1 RAs, DPP-4 inhibitors, thiazolidinediones, sulfonylureas, and insulin. Each of these medications operates via a specific mechanism of action but also brings with it a new clutch of potential side effects and interaction effects. Figure 2.2 (American Diabetes Association, 2022) shows just how complicated the management of diabetes can be.

The medications needed to manage diabetes and prediabetes are not cheap. Low-income, racial, and ethnic minorities have been found to be at higher risk of developing T2D, and more likely to experience diabetes-related comorbidities. Across all races, the prevalence of cases of diagnosed diabetes increases with age, and the average number of medications an individual is prescribed also typically increases with age (Cowie et al., 2021). Because of this, the burdens that follow a diagnosis disparately affect minority, lower-income, and elderly individuals (McCoy et al., 2021; Peek & Thomas, 2021).

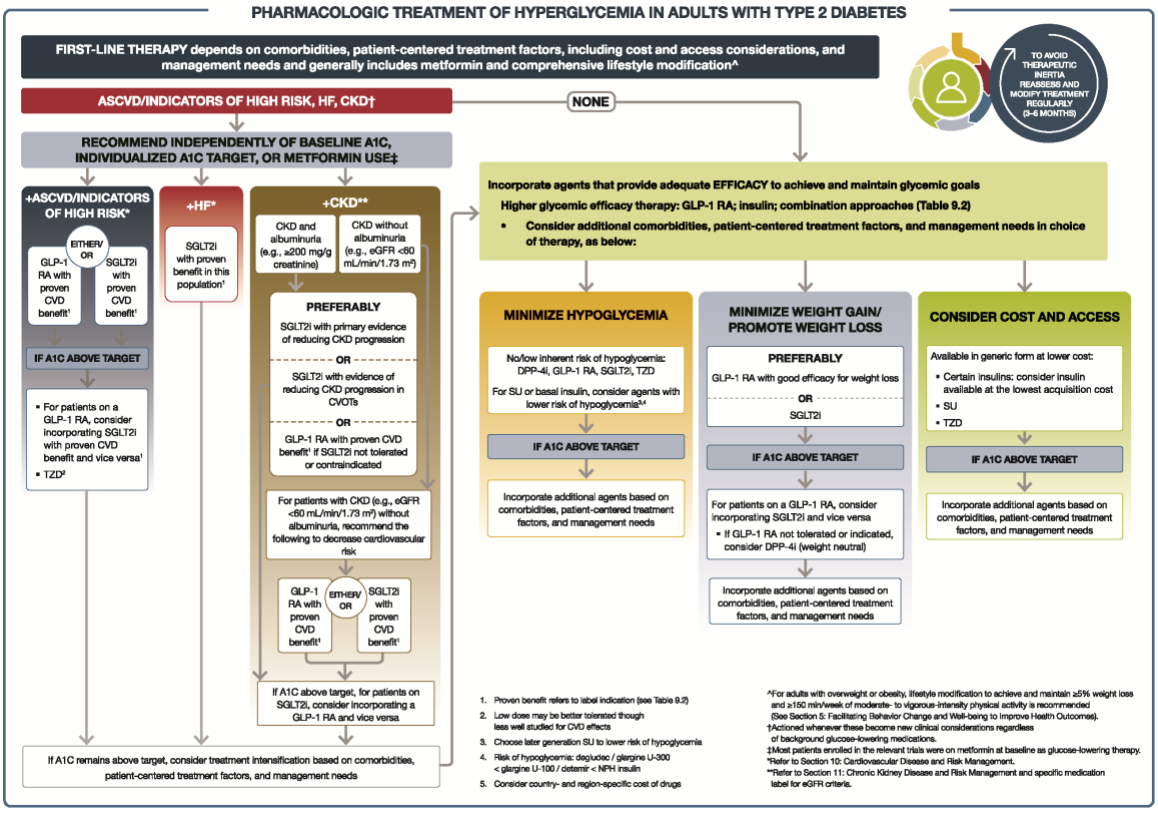


Figure 2.1.1 Pharmacologic Treatment. Diabetes management can be complicated, not only for patients but also for prescribers. There are many factors to consider when determining the right medication (or combinations of medications) that should be prescribed to an individual. Note. From "Standards of Medical Care in Diabetes—2022 Abridged for Primary Care Providers" by American Diabetes Association, 2022, 40(1), pp. 10-38 (https://doi.org/10.2337/cd22-as01).

The direct cost of diabetes, per person, can vary widely due to variability in insurance coverages, medication regimens, and the wide array of comorbidities that can often accompany a diagnosis. But for individuals with no comorbidities, controlling their diabetes with medication and diet alone, a 2003 estimate placed the annual average cost in the range of $1,900 a year, per person. When adjusting for inflation, this amounts to roughly $3,067.30 per year, per person as of the time of this writing. Any associated comorbidity increased these costs by a minimum of 10%. If insulin was prescribed, these costs increased by a minimum of 60%. If dialysis was required (necessary in cases with advanced chronic kidney disease), those costs increase 11-fold (Brandle et al., 2003).

As mentioned, according to the ADA Standards of Care (2022), the primary treatment measure for T2D is intensive lifestyle changes. The ADA definition of what might be considered an intensive lifestyle change is guided in large part by the Diabetes Prevention Program, a 3-year study carried out from 1996 – 2001. This study showed that, for individuals that exercised for a minimum of 150 minutes per week (ideally spread out across multiple days), for at least 10 minutes or more each session, the risk of developing diabetes was reduced by 58%. Intriguingly, the DPP found that although the primary goals of the study were focused on weight loss, simply increasing physical activity alone helped to reduce the incidence of T2D by 44% (Knowler et al., 2002). Notably, in comparison to the exorbitant cost of medication for the treatment and

prevention of diabetes, lifestyle changes remain markedly low-cost. Furthermore, some studies have suggested that, when adhered to, lifestyle changes and increased physical activity result in better long-term health outcomes (Hopper, et al., 2011; Knowler et al., 2002; Tuomilehto et al., 2001).

2.2 Diabetes and Physical Activity: A Set of Complex Interactions

Repeated engagement in physical activity has been shown to lead to the improvement of glycemic metrics, a decrease in insulin resistance, a reduction of inflammation, and a reduced presence of intra-abdominal fat—all risk factors for the development of T2D (Amanat et al., 2020). Acutely, physical activity leads to an immediate increase in blood flow to skeletal muscle, resulting in rapid uptake of glucose (Galicia-Garcia et al., 2020). Both the short- and long-term effects of physical activity are well-studied and heavily relied upon for glycemic control (Colberg et al., 2016; Krotkiewski et al., 1985; Prevention or Delay, 2022). Thus, in addition to being low-cost (or free), increased physical activity has long been suggested as a method to acutely reduce blood sugar levels, improve insulin sensitivity long-term, and has remained a standard of care in the treatment of pre-diabetes and T2D for decades.

While exercise is most often associated with acute decreases in glucose, there are times when exercise can be associated with increases in glucose values. Adrenaline, a hormone often associated with stressful stimuli, can cause an acute increase in blood glucose values (BGVs) via rapid glycogenolysis by the liver and skeletal muscle cells (Verberne, et al., 2016). High-intensity exercises, in which glucose production increases considerably faster than its uptake by muscles, are also known to increase BGVs (Adams, 2013; Marliss, 2002). The "dawn phenomenon," an abnormal rise in glucose due to circadian rhythms in patients with diabetes, can also contribute to increased BGVs during morning engagement in physical activity (Teo, et al., 2020). Furthermore, aside from the high degree of variability in glucose transport mechanisms across cellular membranes per individual, the T2D population represents a diverse range of prescribed medication regimens, all of which are relied upon to influence blood glucose values (Solomon, 2018). The commonly prescribed medication Metformin has also long been suspected to have a blunting effect on BGVs during physical activity (Malin et al., 2011; Terada & Boulé, 2019). Thus, although substantial research has shown that physical activity improves glycemic outcomes for all forms of diabetes, the relationship between physical activity and changes in glucose values remains poorly understood, especially for the T2D community.

2.3 The History of Glucose Prediction

Modeling glucose values is no new feat. Many attempts at model development for the dynamics of glucose have been in circulation in the scientific community since the late 1970s. Many of these early studies fall into a category that can be described as physiological or compartmental models (Aliberti et al., 2019; Contreras et al., 2017; Georga et al., 2011). The development of physiological models typically involves small, strictly controlled laboratory settings used to determine optimal model parameters as inputs to a set of differential equations. Models are designed to represent compartmental diffusion and uptake rates across tissue pools such as blood, vital organs, muscles, and interstitial fluid. Study subjects have ranged from human volunteers to "mongrel dogs", involving frequent injections and tediously repetitive blood draws. (Bergman et al., 1979; Ferrannini et al, 1985).

One of the first studies to contribute to a new class of data-driven glucose prediction models and achieve performant results was published in 1999, by Bremer & Gough. Their work marks one of the first studies to ask—and attempt to answer—whether present and future glucose values could

be predicted using only past glucose values. At the time of their writing, Bremer & Gough noted that it was currently unfeasible to perform finger-stick methods of blood glucose evaluation frequently enough to reliably generate enough data to predict out-of-range glucose excursions. But the authors also note that—given enough data—these kinds of predictions should theoretically be possible. Their first attempts relied on 22 datasets, composed of data from non-diabetic and T1D individuals, but as can be clearly inferred from the title of their work "*Is blood glucose predictable from previous values? A solicitation for data*.", one of the primary obstacles to model development was simply a lack of data. Coincidentally, their work was published in the same year as the FDA approval of the first CGM.

Bremer & Gough used data exclusively derived from individuals undergoing continuous enteral feeding (feeding tubes) to ensure their time series data exhibited stationary dynamics. Subjects undergoing continuous enteral feeding are not a particularly representative sample; however, their results are still impressive and worth examining. Using a simple linear autoregressive integrated model (ARIMA), they achieve an average root mean squared error (RMSE) of just under 9 mg/dl for a ten-minute prediction horizon. Since then, other models used for predicting glucose values have involved classical time series analysis methods like autoregression with exogenous inputs (ARX), and other variations on autoregressive moving average models adapted for exogenous inputs (deemed "MAX" models), such as ARIMAX and ARMAX (autoregressive integrated moving average with exogenous inputs and autoregressive moving average with exogenous inputs) (Georga et al., 2011; Xie & Wang, 2020). But the complexity of interactions between physical activity, carbohydrate input, insulin interactions, and sensitivity to inter-individual variability makes the problem of glucose prediction particularly well-suited to machine learning methods. Classical time series methods of glucose prediction can, at times outperform machine learning methods, but don't typically generalize as well to previously unseen data (Aliberti et al., 2019; Contreras et al., 2017; Xie & Wang, 2020). Still, machine learning methods require voluminous data input, which is not always easily obtained.

2.3.1 Glucose Prediction to Prevent Hypoglycemia

T1D and T2D are characterized by chronic hyperglycemia. So, efforts to treat both forms of diabetes aim to reduce blood sugar levels. But these treatments often come with their own hazards: treatment of hyperglycemia introduces the risk of hypoglycemia (glucose values less than 70 mg/dl), which can be just as dangerous. A significant portion of the studies undertaken to predict glucose values have done so with the aim of preventing hypoglycemia, and data for these efforts come primarily from the T1D population (Aliberti et al., 2019; Contreras et al., 2017; Jacobs & Mosquera-Lopez, 2021; Tyler et al., 2022; Xie & Wang, 2020). Individuals with T1D lack the ability to produce endogenous insulin and therefore require exogenous insulin to manage glucose values. Of course, insulin delivery can rapidly decrease blood glucose values, as it should (see Figure 2.3.1, (Dierks, 2021)). But when paired, insulin and physical activity can act synergistically to cause a rapid reduction of glucose values, which can lead to increased risk of hypoglycemia. Hypoglycemia can quickly become dangerous, thus efforts to predictively prevent hypoglycemia can help to save lives. This in part explains why most studies are focused on glucose prediction for T1D.

Notably, T2D NIIT individuals may also experience hypoglycemia. Elderly adults with multiple comorbidities, on any form of sulfonylurea (a commonly prescribed medication for T2D), or subject to food insecurity are especially at risk (Silbert et al., 2018; Tourkmani et al., 2018). Polypharmacy is also notably associated with a higher risk of hypoglycemia for T2D. Furthermore, engagement in physical activity—known to acutely decrease glucose values—is

also often noted as a risk factor for hypoglycemia for individuals with T2D (Silbert et al., 2018; Tourkmani et al., 2018). Long periods of time in which glucose values are not monitored will also put an NIIT individual at risk of hypoglycemia. Thus, for the T2D NIIT community on long-acting forms of antihyperglycemic medications, mortality caused by nocturnal hypoglycemia (dangerously low drops in blood sugar during sleep) is of great concern (Elliott et al., 2016). Unfortunately, even in daily life, symptoms of hypoglycemia for T2D individuals can often be confused with neurological conditions, especially in the elderly. Confusion, dizziness, and slurred speech are all behaviors typical of hypoglycemia that can be easily confused with dementia or stroke (Freeman, 2019).

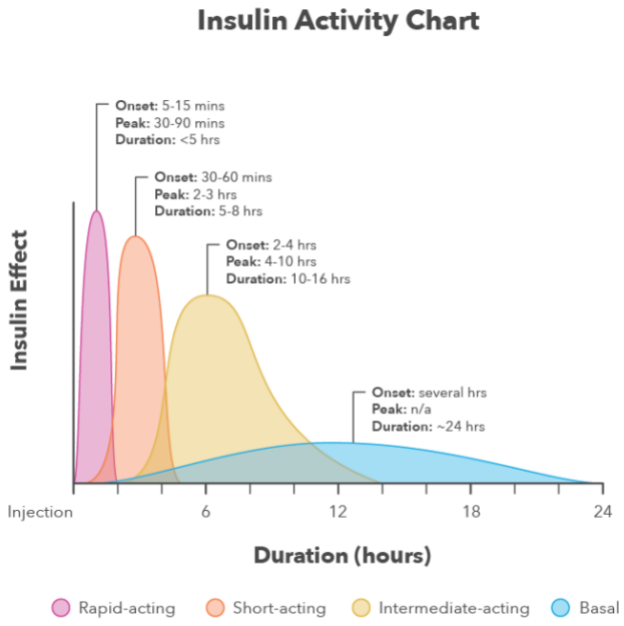


Figure 2.3.1. Insulin activity. Fast-acting insulins (shown above as Rapid-acting and Short-acting) and intermediate-acting insulins can cause drastic changes in glucose values that can put an individual at risk of hypoglycemia. Note. From "The Complete Guide to Insulin Types" by Dierks, M. (https://agamatrix.com/bloc/different-types-of-insulin/)

Clearly, the T2D NIIT community would be well-served by increased access to glucose monitoring, and further research regarding predictive glucose monitoring. To be most useful, predictive glucose monitoring for T2D should ideally, minimally be accurate for forecast horizons that allow an individual to take decisive action, prior to a glycemic event occurring. Previous predictive studies for T1D glucose management have settled on forecast horizon of at least 15 minutes or more as being a suitable length of time to allow for actions to be taken based on the predictive output of their models (Carillo-Moreno et al., 2020).

2.3.2 Glucose Prediction for T2D

Interestingly, very few studies exist focusing on glucose prediction for individuals with T2D. Although somewhat puzzling given the pervasiveness of T2D, this lack of research on the glucose dynamics for Type 2 diabetics can be partially explained by the historically limited access to CGMs for treatment (Anderson et al., 2020; Galindo et al., 2021; Peek & Thomas, 2021). Most studies focused on glucose prediction are feasible only because of the tremendous amount of data available via the use of CGM devices. CMS, the largest government-sponsored insurer in the United States, has until very recently only offered CGM coverage for individuals on insulin-

intensive therapies, capable of providing proof of routine blood draws (four times a day or more) necessary to manage their glucose levels (Sullivan et al., 2009; Melillo, G., 2021). It is also reasonable to assume that, influenced by the policy terms of insurers, CGM device makers themselves may have historically focused their development and marketing efforts solely toward individuals with Type 1 diabetes and their care providers. Thus, CGM devices have historically been provided to individuals with T1D, and only offered to individuals with T2D given sufficient insurance coverage, or sufficiently severe need as determined by their use of insulin. As such, modeling efforts for data-driven methods are primarily found to be focused on T1D, not T2D.

2.3.3 Data-Driven Glucose Prediction

The new possibilities for data-driven analysis and prediction made possible by the advent of CGM technologies are augmented by their convenience. CGM devices are composed of three main components: A sensor, a transmitter, and a receiver. The sensor is comprised of a very small wire, inserted underneath the skin, which calculates the glucose values in the interstitial fluid of cells by calculating voltage gradients across the surface of the wire. Sensors are typically placed directly on the skin with a small adhesive patch. The transmitter receives the values detected by the sensor, stores those values, and transmits them to the receiver. A receiver can take the form of any Bluetooth-capable device. Currently, CGM sensors typically last anywhere from 7 – 10 days and are designed to be worn while exercising, sleeping, showering, swimming, and even SCUBA diving—essentially during most forms of everyday activity. Today, many CGM devices detect and calculate a glucose reading every 5 minutes, generating 288 readings per day. This means that the effect of everyday activities on an individual's glucose values can be monitored in near real time. This also means that large, heterogeneous data samples, entirely representative of free-living individuals can be gathered with relative ease for use by data-driven models.

Researchers have been making good use of the large, heterogeneous datasets generated by T1D CGM users. Just as the variety of classical time series methods and compartmental methods have been used in the past, a plethora of machine learning models have been used. Some notable approaches include basic feed-forward neural networks (FFNNs), artificial neural networks (ANNs), random forests, convolutional neural networks (CNNs), support vector regression, fuzzy logic, recurrent neural networks (RNNs) (Carrillo-Moreno et al., 2020; Jacobs & Mosquera-Lopez 2021; Li et al., 2019). These various methods can be grouped into two basic categories: sequential models and non-sequential models. Sequential models are those capable of handling ordered data in which adjacent values are inherently highly correlated. In other words, the relationship between data points lies not only in their values, but in the order in which those values appear (Diettrich, 2002). Non-sequential models like FFNNs, ANNs, and random forests are better suited to handle independent, identically distributed data in which each model input can be assumed to be entirely independent of the previous input. Because of their ability to retain and account for dependencies on the values of previous states, CNNs and RNNs are well-suited to handling sequential data (Borovykh, 2018).

Although each study approached the problem of glucose prediction with a variety of unique methods and objectives, a few other common themes can be drawn from the collective efforts of previous attempts at glucose prediction. First, and most importantly, it is possible to predict future glucose values using only past glucose (Aliberti et al., 2019; Bremer & Gough, 1999; Pérez-Gandía et al., 2010). One form of modeling, long short-term memory recurrent neural networks (LSTM RNNs), find common inclusion amidst the crowds of other models attempting to predict glucose for short-term time horizons. When an LSTM is used, the smallest input vector of previous glucose values used was 4 previous CGM readings (corresponding to 20 minutes of past

glucose values) (Sun et al., 2018). The longest model input used the previous 3 hours, or 36 previous CGM readings (Jacobs & Mosquera-Lopez, 2021). Others use past glucose values within this range, composed of roughly 2 hours' worth of past glucose values (Li et al., 2019; Aliberti et al., 2019). Since most of these models are reliant on data from T1D populations, exogenous inputs very commonly include insulin. Other covariate features found to increase the predictive performance of these models include heart rate, carbohydrate intake, physical activity, stress, and illness. Physical activity is a known source of increased error rates due to its highly variable impact on glucose metrics, but some studies have found that including heart rates as covariates may improve prediction accuracy by accounting for the impact of physical activity. Most of these efforts also focus on very short-term time horizons—up to a maximum of 60 minutes for most models, with few studies venturing into longer-term predictions of up to 90 minutes. Generally, the longer a prediction horizon is, the higher the error rates. Typical benchmark error rates for most of these studies range from 12 – 24 mg/dL for a 30-minute prediction horizon (see Table 2.3.3).

Table 2.3.3
Error Rates of Related Works

	Author(s)	Model Type	Model Inputs	RMSE, 30 min. Prediction Horizon
T1D	Pérez-Gandía et al., 2010	FFNN	CGM data	18
	Sun et al., 2018	LSTM	CGM data	21.5
	Li et al., 2019	CNN	CGM data, insulin and carbohydrates	21.07
	Aliberti et al., 2019*	LSTM	CGM data	5.93*
	Carrillo-Moreno et al., 2020	LSTM	CGM data, insulin delivery and carbohydrates	12.6
	Xie & Wang, 2020	ARX	CGM data, insulin, heart rate and carbohydrates	19.48
	Jacobs & Mosquera-Lopez, 2022	LSTM	CGM data, insulin delivery via closed-loop insulin delivery	19.8
		LSTM	CGM data, insulin delivery via sensor-augmented pumps	19.6
During Physical Activity (T1D)	Hobbs et al., 2019	PBSID*	CGM data, heart rates	26.3
	Romero-Ugalde et al., 2019	ARX	CGM data, (accelerometer- and heartrate-derived) energy expenditure*, insulin and carbohydrates	7.8 +/- 4.5
	Romero-Ugalde et al., 2019	ARX	CGM data, (accelerometer- and heartrate-derived) energy expenditure*, insulin and carbohydrates	16.7 +/- 15.6
T2D	Kim et al., 2019	FFNN	*unspecified*	37.18
		LSTM	*unspecified*	50.81

Table 2.3.3. Most glucose prediction studies which are not focused on predicting glucose during physical activity achieve RMSEs of 12 – 24 mg/dL for a 30-minute prediction horizon. Note: One notable outlier was found in the literature with an error rate of 5.93 mg/dL for a 30-minute horizon, published by Aliberti et al., 2019. Fewer studies focused on glucose prediction during physical activity. For those that did, results were mixed. A single study focused on glucose prediction for T2D, however the methods used in this study were unclear.

2.4 Defining and Detecting Physical Activity

Along with the development of CGM devices, fitness trackers have also begun to generate inordinate amounts of new data analysis opportunities. Device manufacturers like Fitbit, Garmin, Oura Ring, and others have become increasingly useful, not only for personal activity tracking but also for research and as FDA-approved medical devices (Lima et al., 2022; Lubitz et al., 2022). In 2020 Fitbit joined the ranks of Apple as a maker of FDA-approved smartwatches capable of detecting atrial fibrillation. When paired with CGM devices, the data generated by activity trackers can be used to fuel research focused on glucose metrics as they might occur for a diverse range of individuals experiencing the many facets of everyday life including stress, excitement, sleep, diet, and activity.

This study focuses on glucose prediction during and immediately following physical activity for the T2D NIIT population, using CGM- and Fitbit-derived data. As such, it is important to consider methods of detecting meaningful bouts of physical activity that cater to the well-established expert recommendations for physical activity for this population, and that work within the constraints of data availability. The minimum level of physical activity recommended by the American Diabetic Association is composed of regular sessions of "light-intensity ambulation", ideally occurring three times a week, lasting for at least 10 minutes per session (Colberg et al., 2016; Prevention or Delay, 2022). As mentioned, (section 2.3), the basis for these recommendations comes from the DPP, which implemented these constraints in a cohort of prediabetic individuals and showed that lifestyle changes were highly effective at promoting glycemic control and reducing the frequency of newly diagnosed cases of T2D. Thus, there are a variety of ways observational activity data can be filtered to identify meaningful bouts of activity, but there are a few key considerations that should be considered when doing so. Namely, any method of defining and detecting a bout of physical activity should account for both the duration of the activity and the intensity of the activity.

2.4.2 Step Cadence

Although each fitness tracker employs their own unique algorithms and methods for detecting physical activity and subsequently quantifying the intensity of the activities detected, many of the major activity-tracking device manufacturers make use of one common metric: step cadence. As a measure of physical activity, step cadence—the speed at which an individual walks—can be used to determine the intensity of an activity and can of course be measured by duration by accounting for the length of time an individual ambulates at a rate above some pre-defined threshold.

To understand step cadence and its utility, it is helpful to first understand METs or metabolic equivalents of tasks. METs are widely used and recognized in sports medicine. One MET is defined as the amount of oxygen consumed while at rest and can be easily calculated given an individual's height, weight, and basal metabolic rate. METs are an easy-to-calculate, widely used heuristic for assigning and evaluating the intensity at which an activity was performed. However, obtaining basal metabolic rates from observational, activity-tracker-generated data gathered from free-living subjects is unrealistic. In recognition of this limitation, much work has been done to determine common metrics for quantifying step-defined intensities in a wide range of demographics and to relate those intensities to METs. These studies have found that roughly 100 steps per minute correspond to moderate-intensity ambulation, or "brisk walking" (Marshall et al., 2009; Serrano et al., 2015). Of particular interest are the multiple studies carried out by Tudor-Locke et al. (2005, 2011, 2018, 2021) wherein: purposeful, activity-related movements have repeatedly been shown to map to step cadences of 60 steps per minute or higher; moderate-intensity exercise (≥3 METs) is shown to be associated with step cadences of 100 steps per minute or higher, and; the transition from walking to running or jogging (considered high-intensity activities) is shown to occur at rates of 120 and 140 steps per minute (see Table 2.4). Using these cadence bands, it is then possible to determine the lowest minimum step rate threshold that constitutes "light intensity ambulation", lasting ten minutes or more, and that can be used to begin tracking a bout of activity.

Table 2.4
Intensity-Defined Cadence Bands

Intensity	Step-Defined Intensity	MET Band	Step Cadence Band
Sedentary	No steps	0	0
Very Light	Incidental Movement	<2	1 - 19
	Sporadic Movement		20 - 39
Light	Purposeful steps	2 - 2.9	40 - 59
	Slow Walking		60 - 79
Moderate	Medium Walking	3 - 6	80 - 99
	Brisk Walking		100 - 119
Vigorous	Jogging	6 - 8.7	120 - 139
	Running		140 - 179
Near Maximal	Sprinting	>= 8.8	>= 180

One important factor to consider when detecting activity from pedometer-based data generated by the T2D NIIT population is that many of these individuals may have mobility-limiting comorbidities. Choosing a minimum level of step cadence that avoids the exclusion of any individuals under study will be an important consideration for any method of activity detection. Previous studies have shown that steps at rates of less than 40 steps per minute are typically associated with sporadic movement and that when intentional bouts of activity do occur, they occur in tandem with higher step cadences of 60 steps per minute or higher (Granat et al., 2015; Orendurff et al., 2008; Tudor-Locke et al., 2018). However, one study found that for elderly individuals with mobility-limiting comorbidities, minimum thresholds of 40 steps per minute or higher have been shown to be more inclusive (Webber et al., 2020). Thus, both thresholds should be evaluated.

2.4.1 A Note on Heart Rates

There are two ways heart rate values might be used for a project focused on glucose prediction during physical activity: as covariates, or as a method of detecting physical activity. As a covariate, there is little issue with using heart rates. A feature can simply be added to a model, evaluated, and discarded if no improvement is evidenced by error metrics. However, when used as a method of detecting physical activity, there are significant issues. But given the allure of a feature so clearly rooted in direct physiological measurement, an exploration of the potential pitfalls of heart rates is warranted.

Several previous studies attempting to predict glucose values have found heart rates to be a useful feature when used as covariate model inputs. When attempting to predict glucose values during physical activity in T1D adolescents, Hobbs et al. (2019) found that the inclusion of heart rate-derived metrics served to improve the performance of their model. Romero-Ugalde et al. (2019) also found that, when relative exertion was accounted for via the incorporation of heart rates, along with accelerometer data, they were able to achieve relatively robust predictions for models trained on individual data. Many activity trackers also include heart rate detection as a critical part of their activity-detection algorithms to estimate the intensity with which an individual is exercising. Moreover, recent advancements and FDA approvals for heart rate detection algorithms in activity trackers like Fitbit and Apple might persuade researchers to consider using

heart rate data as a potential method of detecting physical activity. But this is one area where care should be taken, as cardiac enervation in T2D populations may differ significantly from other, non-diabetic populations, or even T1D.

Cardiac autonomic neuropathy (CAN) is a subclinical complication of both T1D and T2D. The symptoms of CAN, when detected, can manifest in impaired heart rate variability (static heart rates), resting tachycardia (non-stimuli-associated rapid increases in heart rates), exercise intolerance, abnormal blood pressure regulation, orthostatic hypotension (low blood pressure upon standing), and syncope (fainting). It has been estimated that CAN is present in roughly 35% of individuals with Type 2 diabetes but these estimates are known to vary greatly depending on sampling methods (Al Olaiwi et al., 2019; Pop-Busui, 2010; Serhiyenko & Serhiyenko, 2018). Much like T2D, individuals with CAN are often unaware of their condition until a significant health event occurs, with negative impacts on their quality of life. As such, it is difficult to determine the frequency with which CAN is present in a study cohort without performing rigorous testing for its presence (Vinik & Ziegler, 2007). This is problematic for activity detection, and the T2Help population is no exception (see Figure 2.4.1). If used as a method of detecting a bout of physical activity, instances of resting tachycardia would result in numerous false positives. In cases where symptoms of CAN include static heart rates, this would result in numerous false negatives for activity detection. Thus, the implication of a pervasive frequency of CAN in the T2D population makes the use of heart rates a questionable choice as a method of detecting physical activity. Any metrics which might indirectly rely on heart rates would also be subject to these same pitfalls.

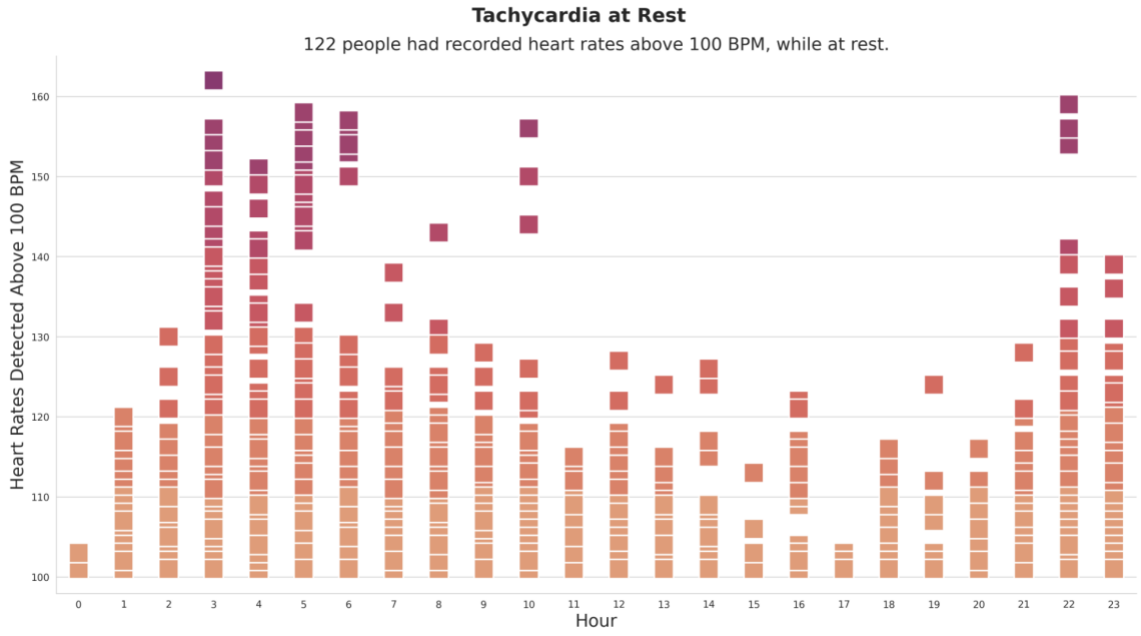


Figure 2.4.1 Recorded instances of resting tachycardia in the T2 Help NIIT population. Note, a period of rest was defined using Fitbit-detected sleep events (in which heart rates are continuously recorded, but no movement is detected). To remove potential false positives in which a participant may have abruptly ended a bout of physical activity prior to the start of a sleep event, heart rates detected within the first 30 minutes of a sleep event were not considered.

CHAPTER 3

Methodology

3.1 Data Pre-Processing

The data used for this project come from the T2 Help study and has been provided by Dexcom Inc., a maker of continuous glucose monitoring devices. The T2 Help dataset is composed of CGM-derived glucose values and fitness tracker-generated step counts for 199 individuals with Type 2 diabetes and prediabetes collected over 12 weeks (see Table 3.2.1). In this study, each participant wore a Dexcom G6 CGM for 12 weeks. The CGM provided 288 readings, corresponding to a glucose reading every 5 minutes, per day. Each participant also wore a Fitbit, also for 12 weeks, which provided minute-by-minute step counts. To ensure that any acute changes in glucose values immediately following the initiation of a detected bout of activity were the result of the activity, and not the result of any form of insulin, individuals on any form of fast or intermediate-acting insulin were excluded from this analysis.

Table 3.1
Characteristics of the T2 Help NIIT Cohort

Summary Data	Value ± SE
Subjects, N	199
Age (years)	56.37 ± .82
Biological sex (Female/Male)	117/82
Years since diagnosis	9.45 ± 0.65
HbA1C score (percent)	6.85 ± 0.08%
BMI	33.52 ± 0.57
Comorbidities	Percent of Cohort
Atherosclerotic cardiovascular disease	10.05%
High blood pressure	65.33%
Hypertension	63.82%
Retinopathy	11.56%
Mobility limitation	2.51%

3.1.1 Interpolation

To account for small intervals of time (intervals lasting less than 2 hours—up to 24 CGM readings) in which CGM data was missing (this can happen due to sensor sessions ending or signal loss), cubic spline interpolation was used to fill missing values. Longer intervals of time in which CGM readings were unavailable were excluded from this analysis.

3.1.2 Bout Detection

There are two dimensions to the weekly recommendations of the ADA regarding physical activity: duration and intensity. The minimum duration should be ten minutes. The minimum intensity should correspond to “light intensity ambulation”. Weber et al., (2020) found that for

older individuals (N=25, median age: 77) with limited mobility, the most inclusive minimum thresholds for step cadences should be 40 steps per minute. Several other studies found that slightly higher cadences of 60 steps per minute or more were suitable for a wide range of demographics (Granat et al., 2015; Orendurff et al., 2008; Tudor-Locke et al., 2018). But upon filtering for any bouts of activity lasting 10 minutes or more, it was found that, in the T2 Help NIIT population, step cadences of 60 steps per minute or higher were amply inclusive (see Figure 3.1.2). To prevent potential mixed interaction effects of multiple bouts of activity, any bouts with a fresh bout detected within 120 minutes of another bout were excluded. This ensures that all glucose values used for training, validation, and testing are solely influenced by a single bout of activity. Finally, data were collected 2 hours prior to the start of a bout, and up to 60 minutes after the start of a bout of physical activity for each individual.

Thus, discrete periods of time composed solely of pre-bout glucose values and glucose values collected during and after bouts of activity were defined and collected.

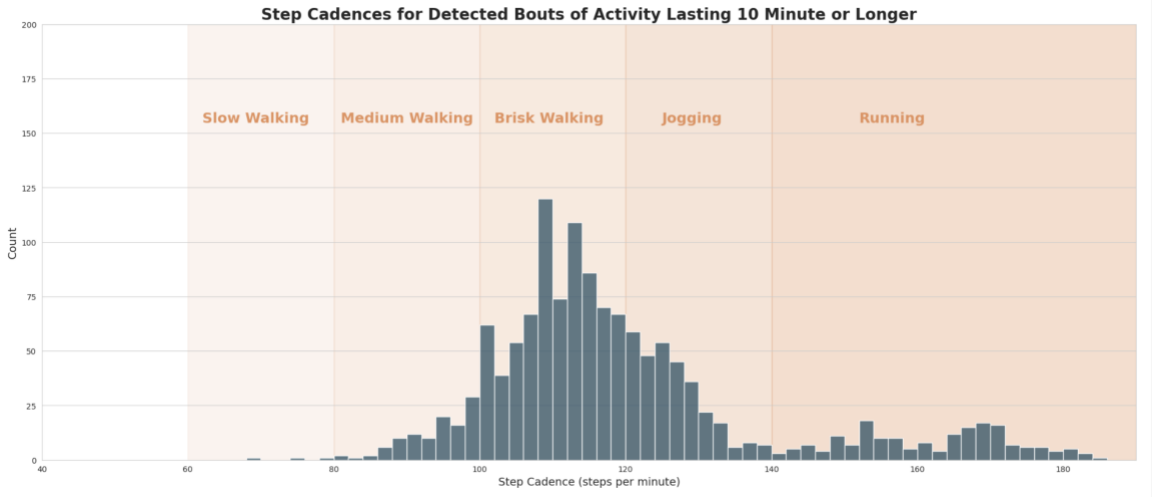


Figure 3.1.2 For walking bouts lasting 10 minutes or more, subjects in the T2Help NIIT cohort were predominantly walking at a pace faster than 60 steps per minute.

3.1.3 Normalization

To normalize values, the first three weeks of glucose values were used to establish a baseline median, per person. Any glucose values detected during or up to 1 hour after the end of a bout of physical activity in these first three weeks were excluded. All glucose values collected after the initial three weeks were then normalized to this baseline.

3.1.4 Stratified Sampling

Due to the high degree of variability in the number of bouts detected for each individual, stratified sampling was used to ensure even random sampling across groups with low counts of detected bouts as well as those with high counts of detected bouts (see Figure 3.4.1). Training, validation, and testing: 70% of the individuals with detected bouts of activity were randomly assigned to a training dataset, 10% were randomly assigned to a validation dataset, and the remaining 20% were assigned to testing (see Table 3.1.4). The testing set was withheld from all model development efforts.

Table 3.1.4

Characteristics of the Training, Validation and Testing Datasets

Characteristic	Value ± SE
Training	
Subjects, N	86
Bouts of physical activity, N	692
Age (years)	56.49 ± 1.30
Biological sex (female/male), N	51/35
Years since diagnosis	9.05 ± 0.96
HbA1C score (percent)	6.75 ± 0.11%
Body mass index	32.91 ± 0.95
Validation	
Subjects, N	10
Bouts of physical activity, N	83
Age (years)	53.60 ± 4.26
Biological sex (female/male), N	4/6
Years since diagnosis	9.4 ± 2.53
HbA1C score (percent)	6.78 ± 0.33%
Body mass index	32.63 ± 1.13
Test	
Subjects, N	23
Bouts of physical activity, N	210
Age (years)	56.09 ± 2.18
Biological sex (female/male), N	10/13
Years since diagnosis	7.0 ± 1.86
HbA1C score (percent)	6.36 ± 0.15%
Body mass index	32.39 ± 1.53

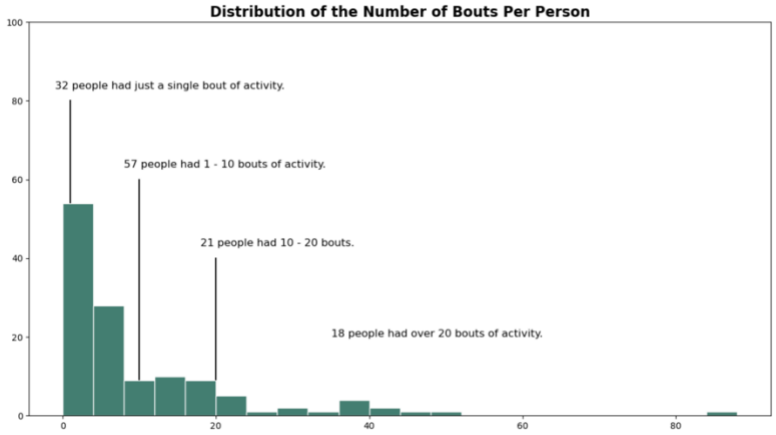


Figure 3.1.4. Due to the highly skewed distribution of bouts stratified sampling was used to ensure a roughly even mixture of highly active and less active individuals.

3.3 Model Development

The task of glucose prediction for the T2 Help NIIT population is treated as a supervised learning problem. Darts package version 0.22.0, Block Recurrent Neural Network architecture was used to develop a basic LSTM RNN composed of one hidden layer. PyTorch Lightning callbacks Early Stopping was used during model development to prevent overfitting. All model inputs were scaled using Scikit Learn MinMaxScaler. The LSTM RNN was initially developed and evaluated as a one-to-one model. Forward feature selection is used to develop and evaluate further, many-to-one models. Each model is trained and fit using timestamped, labeled glucose values from the training data set then evaluated on the validation set. The holdout set is withheld from all model development efforts.

3.3.1 LSTM Model Architecture

Hochreiter & Schmidhuber (1997) are the original developers of the LSTM RNN. Their novel network architecture was specifically designed to overcome issues of exploding and vanishing gradients. This meant that their LSTM RNN could overcome the two main problems associated with what are now deemed "vanilla RNNs"--that is, recurrent neural networks that rely on conventional backpropagation for computing error metrics during the learning phase. Their solution allows a model to react to both long and short-term dependencies in sequences and allows for processing of very long sequences—something vanilla RNNs perform poorly with. The key components of an LSTM network architecture that allow for such reactivity are the three so-called "gates" that make up an LSTM memory cell (also known as a block). Each gate is composed of a small neural network that makes use of either a hyperbolic tangent function (Equation 1) or a sigmoid function (Equation 2) or a combination of the two.

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \qquad \textit{Equation 1}$$

$$g(x) = \frac{e^{x}}{e^{x} + 1} \qquad \textit{Equation 2}$$

Each memory cell receives three inputs: X_t(the input from the current timestep *t)*, h_{t-1}(the output from the previous cell), and C_{t-1} (the previous cell state, or "memory"). The three gates of a memory cell are the forget gate, the input gate, and the output gate. The forget gate utilizes a sigmoid function to determine f_t, what percentage of the long-term memory is retained (Equation 3). The input gate (Equations 4, 5 & 6), which is composed of two sub-gates, generates a new cell state, C_t via an input gate layer (Equation 4) and a hyperbolic tangent-defined layer (Equation 5). The input gate layer controls which values are updated, whereas the hyperbolic tangent layer creates a new vector of potential new values to be added to the cell state. Lastly, the new cell state C_t is created via element-wise multiplication of f_t (the vector of values to be forgotten or retained) and the previous cell state $C_{t-1}$and the addition of $i_t \cdot \hat{C}$ (the set of new values and scaled weights assigned in the input gate layer) (Equation 6). The output gate is similarly composed of two sub-gates: a sigmoid layer (Equation 7) and a hyperbolic tangent layer (Equation 8). The sigmoid layer determines what portion of the cell state to output, while the tanh layer uses multiplication of the output of the sigmoid layer with scaled version of the current cell state. Thus, the output gate serves to control the output prediction. The flow of data through an unrolled LSTM network can be visualized in Figure 3.3.1.

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t \right] + b_f \right)$$ *Equation 3*

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t \right] + b_i \right)$$ *Equation 4*

$$\widehat{C_t} = \tanh\left(W_C \cdot \left[h_{t-1}, x_t \right] + b_C \right)$$ *Equation 5*

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widehat{C_t}$$ *Equation 6*

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t \right] + b_o \right)$$ *Equation 7*

$$h_t = o_t \cdot \tanh\left(C_t \right)$$ *Equation 8*

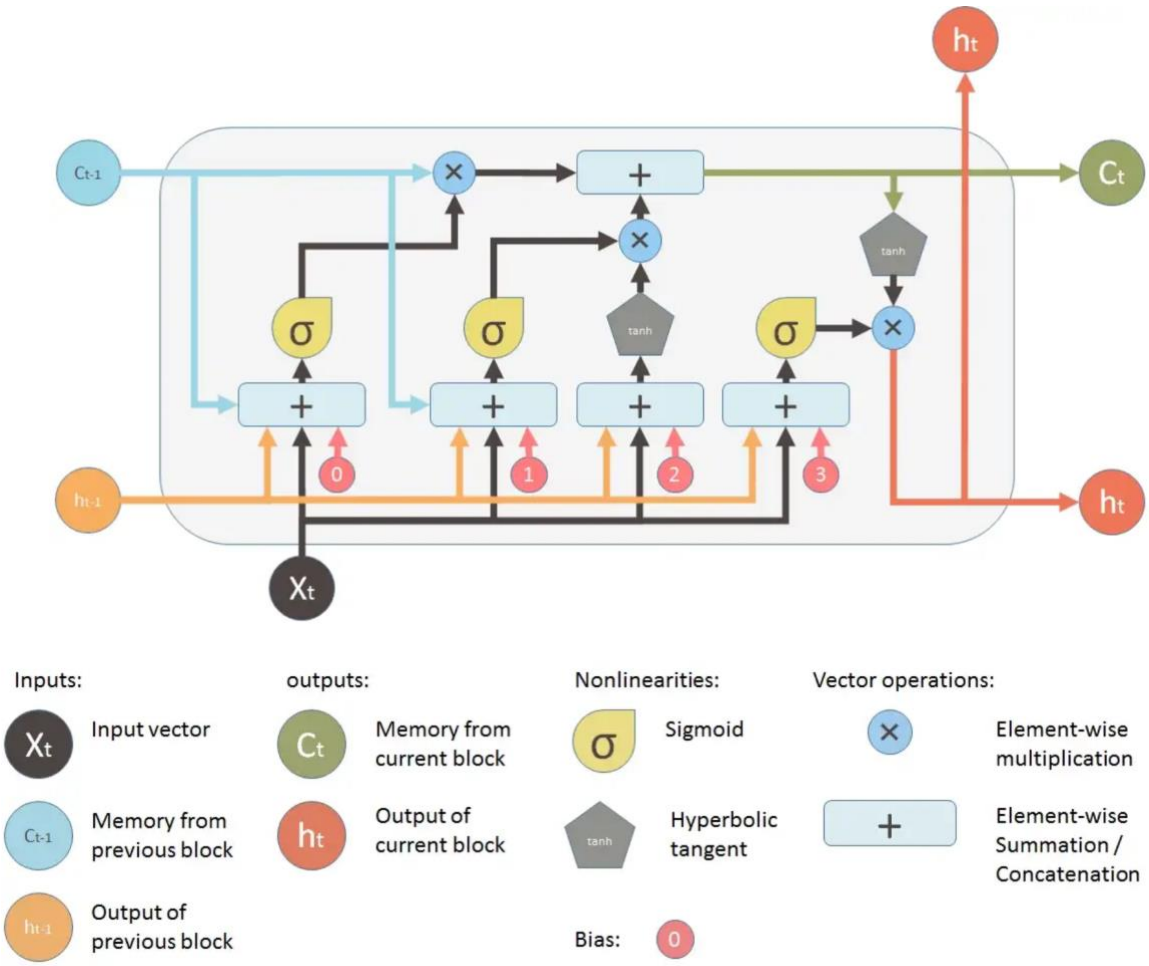


Figure 3.3.1. The flow of data through an unrolled LSTM network. Note. From "Long Short-Term Memory [Review of Long Short-Term Memory]" by Sood, A. (n.d.) In University of Wisconsin-Madison School of Computer, Data & Information Sciences. Pp. 7-15. (https://pages.cs.wisc.edu/~shavlik/cs638/lectureNotes/Long%20Short-Term%20Memory%20Networks.pdf).

3.3.2 Performance Metrics

Prior studies have unanimously made use of root mean squared error (RMSE) as a method of assessing the overall accuracy of model predictions. Thus, the primary performance indicator used in this study is the standard error metric, RMSE (Equation 9, note $\hat{g}_i$ refers to predicted glucose values, while g_i $refers\ to\ actual\ glucose\ values$). To recover interpretable glucose values, inverse scaling is applied prior to evaluation. Overall model performance across all time horizons (used first for initial model development, then for forward selection of covariate features) is accomplished by averaging RMSE across all prediction horizons (from 15 minutes to 60 minutes).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{g}_i - g_i\right)} \qquad \textit{Equation 9}$$

3.3.3 Base-Class Model Development

Initial model development focused on using only past glucose values to predict future values. This first phase of testing involved two possible base-class models. Model A utilized the first ten minutes of glucose values (two CGM readings) during a detected bout of activity; Model B did not utilize the first ten minutes of a bout. Both models used glucose values detected during the 2 hours before the start of a bout of activity. Thus, Model A was trained and fit on timestamped input chunk lengths of size 26, with a vector of length 10 composed of labeled and timestamped glucose values for learning. Model B was trained and fit on timestamped input chunk lengths of size 24, with a vector of length 12 composed of labeled and timestamped glucose values. Both models generated autoregressive glucose predictions for short-term time horizons of 15 minutes up to 60 minutes and every 5-minute interval between. These initial models were built using a single hidden layer with 25 neurons, with Pytorch Adam optimization. PyTorch Early Stopping was used to prevent overfitting and set to stop when loss on the labeled data ceased to exceed 0.0001 scaled units over the course of 25 training epochs. Both models were trained and fit using data from the training dataset. Evaluations were performed on the validation set.

3.3.4 Hyperparameter Tuning

Hyperparameter tuning was accomplished via grid search. Parameters tuned include the number of epochs ([100, 150, 200, 300]), minimum change in validation loss ([0.001, 0.0001, 0.00001]), patience (the number of epochs requiring a change greater than or equal to the minimum change in validation loss) ([10, 15, 20, 25]), as well as the number of neurons in the hidden dimension ([10, 15, 20, 25]).

3.3.5 Covariate Testing

Several studies have found significant performance benefits from the use of covariate features (Carillo-Moreno et al., 2020; Li et al., 2019; Romero-Ugalde et al., 2019; Xie & Wang, 2020). The T2 Help dataset involved the collection of demographic information and diabetes-specific information that might also serve to improve prediction accuracy. Forward feature selection was used to determine if prediction accuracy over the initial base-class model could be improved. Each new bivariate model was trained and fit using the training data and evaluated on the validation data. A total of five covariate features were tested including step cadence, years since diagnosis, HbA1C, bout duration, and BMI. Each feature was chosen for its suspected impact on glucose values during physical activity.

3.3.5.1 Step Cadence

The speed at which an individual is walking, in steps per minute, serves to provide a measure of the intensity of engagement in physical activity. Of particular interest are the changes in glucose values during moderate to high-intensity exercise. It is well known that high-intensity exercise has been shown to cause acute increases in blood sugar values, rather than decreases. This is because, during high-intensity exercise, glucose production via rapid glycogenolysis by the liver and skeletal muscle cells can increase blood glucose values considerably faster than its uptake by muscles (Adams, 2013; Marliss, 2002). But low-intensity and moderate-intensity activities are also well-studied, and heavily relied upon as methods for acutely decreasing glucose values (Colberg et al., 2016; Prevention or Delay, 2022). Thus, this potential for bi-directional changes in glucose values as a result of the intensity of physical activity makes the addition of step cadence a natural choice as a covariate feature.

3.3.5.2 Years Since Diagnosis

Medication regimens are another set of features of the T2 Help dataset suspected to influence glucose value during physical activity. It isn't clear whether different forms of medication might have a buffering or modulating effect (as is suspected of the biguanide Metformin), or whether another medication might act synergistically to decrease glucose values during physical activity (as is suspected of sulfonylureas) (Malin et al., 2011; Terada & Boule, 2019; Tourkmani et al., 2018). Medication regimens in the T2D NIIT community are highly diverse, making an analysis of their effects while predicting glucose values unwieldy at best. But one trend commonly described in the literature is that of polypharmacy (Dobrică et al., 2019; Freeman, 2019; Lipska et al., 2016). The likelihood of polypharmacy is described as increasing as individuals age and hence as the number of years since diagnosis increases. In the T2 Help study, rates of polypharmacy follow these same trends (see Figures 3.3.5.2.1 and 3.3.5.2.2). So, while the number of medications and medication combinations for each individual in the T2 Help study might represent a large, multi-dimensional matrix of one-hot encoded values, the use of a single variable, years since diagnosis, was suspected to serve as a sufficient proxy variable for antihyperglycemic medication regimens.

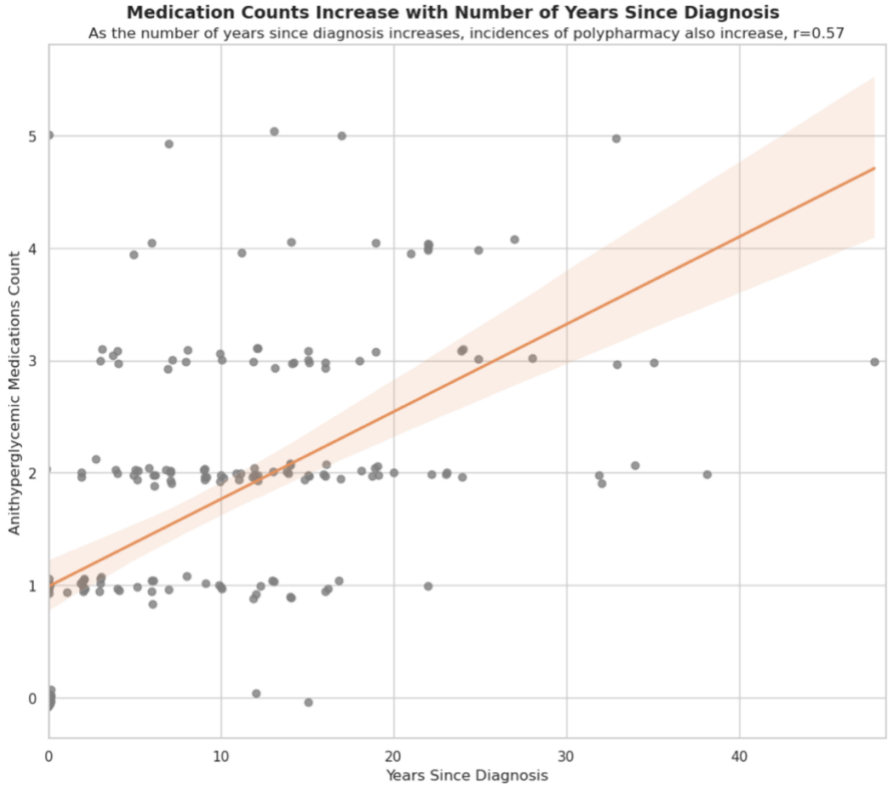


Figure 3.3.5.2.1. An indication of the presence of polypharmacy in the T2 Help NIIT cohort. As the number of years since diagnosis increases and an individual ages, the number of medications typically increases.

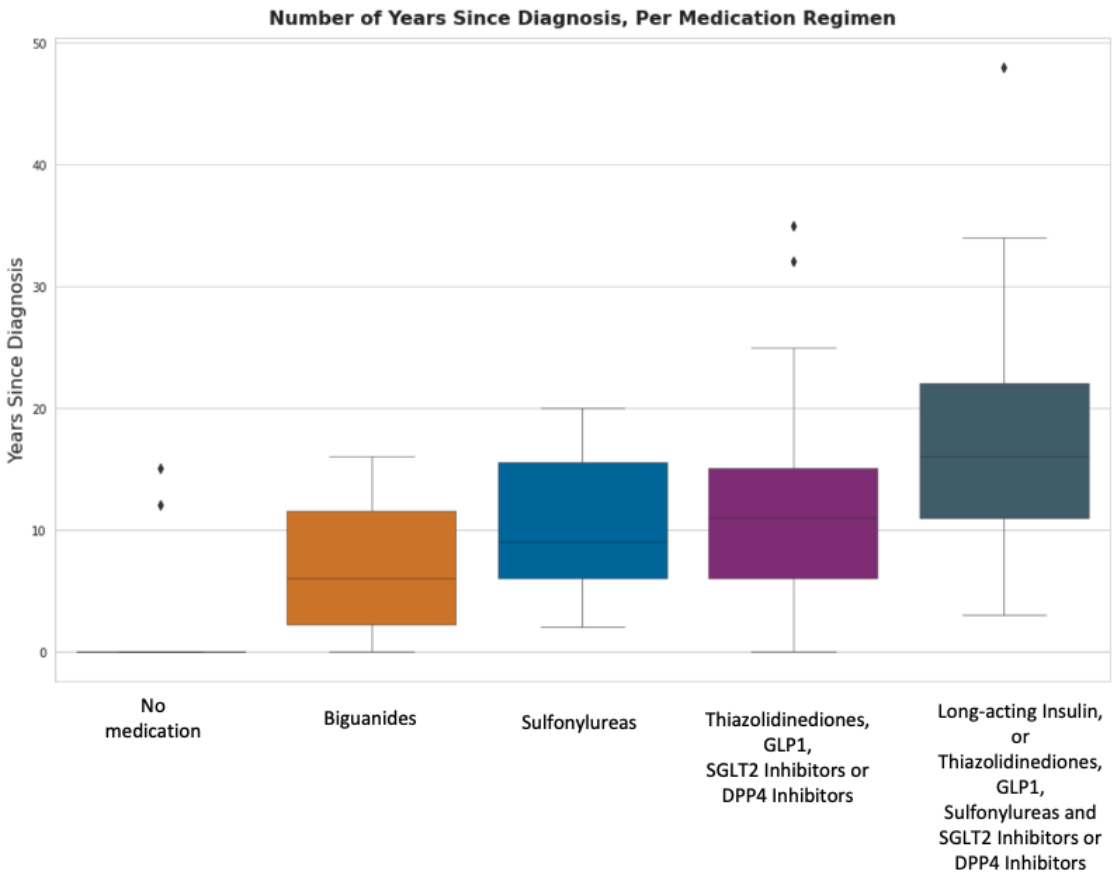


Figure 3.3.5.2.2 Further evidence of polypharmacy in the T2 Help NIIT cohort. Notice individuals making lifestyle changes only (no medication) have been diagnosed for the shortest periods of time. Individuals with the greatest number of medications, and the most complicated combinations of medications, have been diagnosed the longest.

3.3.4.3 HbA1C

All participants from the T2 Help study provided HbA1C test results at the start of the study. An HbA1C test measures the percentage of hemoglobin that is covered by glucose in a blood sample (CDC, 2018). Since glucose present in the bloodstream tends to stick to hemoglobin for the lifespan of a red blood cell and the length of time a red blood cell remains in circulation is roughly 8 weeks, HbA1C test results serves as a good indication of how well-controlled an individual's glucose values have been for roughly the last three months (CDC, 2018). HbA1C scores are also used to diagnose prediabetes and diabetes, with scores of 5.7% or higher indicating prediabetes or T2D. Individuals with high HbA1C scores are more likely to have poor glycemic control, and thus may differ in their response to exercise, hence the selection of this feature for covariate testing.

3.3.4.4 Bout Duration

Bout duration, the total length of time spent actively ambulating at rates above 60 steps per minute, was suspected to influence changes in glucose values. If physical activity were to influence glucose values, then that effect might reasonably be suspected to last longer or be more

pronounced for longer durations of engagement in physical activity. Thus, bout duration was added as a static covariate to the base model.

3.3.4.5 BMI

The CDC defines categories of body mass indices relative to an individual's weight (kilograms) and the square of their height (in meters). Although this metric is not a perfect indicator of body fat percentage, it is well correlated (CDC, 2021). The average BMI in the T2 Help T2D NIIT cohort was 33.53 ± 0.57 kg/m^2, with roughly 63% falling into a category classified as obese. Due to its potential to impact gait, and thus relative exertion during ambulation, BMI was also included as a feature for covariate testing.

CHAPTER 4

Results

4.1 Base-Class Model Evaluation

Base-Class model evaluation proceeded in two stages. First, the most performant model (evaluated using average RMSE across all prediction horizons) was selected from the two possible glucose-only models: Model A and Model B. This model is then used for hyperparameter tuning. Next, covariate features were forward selected and tested using the parameters determined by hyperparameter tuning. All feature selection evaluations were performed on the validation set.

4.1.1 Glucose-Only Model Selection & Hyperparameter Tuning

Model A achieved an average RMSE of 13.41 mg/dl across all prediction horizons, outperforming Model B (RMSE: 13.46 mg/dl). Thus, Model A was selected for further development. Both models achieved satisfactory performance on the 30-minute time horizon, with RMSEs of 11.6 mg/dl and 11.9 mg/dl for Model A and B, respectively. Model performances are shown in Figure 4.1.1.

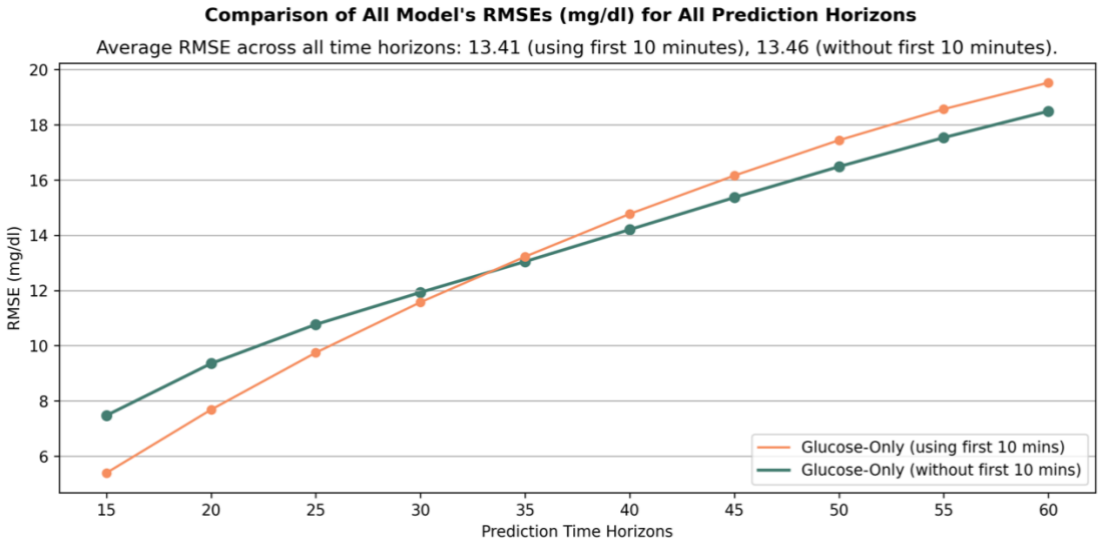


Figure 4.1.1. Model A (utilizing the first ten minutes of glucose detected during a bout, shown in orange) outperforms Model B (reliant on solely 2 hours of pre-bout data, shown in green). Average RMSE for Model A: 13.41 mg/dl. Average RMSE for Model B: 13.46 mg/dl.

Hyperparameter tuning via grid search revealed most performant parameters of 300 epochs, a 0.0001 minimum change in validation loss, and patience of 25 epochs for early stopping (Figure 4.1.2). The ideal number of neurons in the hidden dimension was determined to be 10. All subsequent models tested utilized these parameters during covariate evaluation.

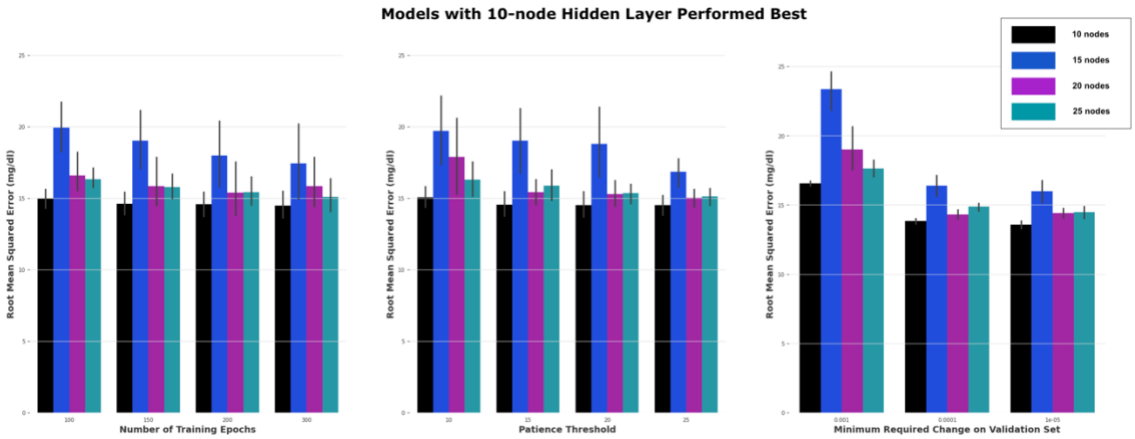


Figure 4.1.2. Base-Class Model A's error rates were lowest with hyperparameters of 300 training epochs, 10 neurons in the hidden layer, and when early stopping was triggered when validation loss ceased to decrease by more than 0.0001 over a period of 25 epochs.

4.2 Covariate Testing

Each covariate tested yielded improvements to the selected base-class glucose-only model. After initial bivariate models were tested, the most performant models were selected for further testing. Models with more than one covariate showed no further improvement in error metrics when run on the validation set, so model development ended with the final testing of two covariate features.

4.2.1 Forward Feature Selection of Single Covariates

Each covariate added to the glucose-only model served to improve error rates. When BMI was accounted for, model predictions improved modestly to 12.72 mg/dl across all time horizons. Bout duration and HbA1C also improved metrics, with an average RMSE of 12.72 mg/dl and 11.92 mg/dl across all prediction horizons, respectively. The feature representing the number of years since diagnosis served to improve error metrics, with an average RMSE of 11.69 mg/dl across all prediction horizons. Lastly, the addition of step cadence improved model prediction error rates the most, yielding an average RMSE of 11.35 mg/dl across all time horizons (See Table 4.2.1 and Figure 4.2.1). Performance at the 30-minute prediction horizon for all covariates also fell within the range of benchmark error rates.

Table 4.2.1

Average Error Rates for Covariate Feature Testing

Target Feature	Covariate Feature	Average RMSE (mg/dl) across all prediction horizons
Glucose	Step cadence (steps p/minute)	11.35
	Years since diagnosis	11.69
	HbA1C test results (gathered at the start of the T2Help study)	11.92
	Bout duration	12.72
	BMI	12.79
Glucose-only base model (for reference)		*13.41*

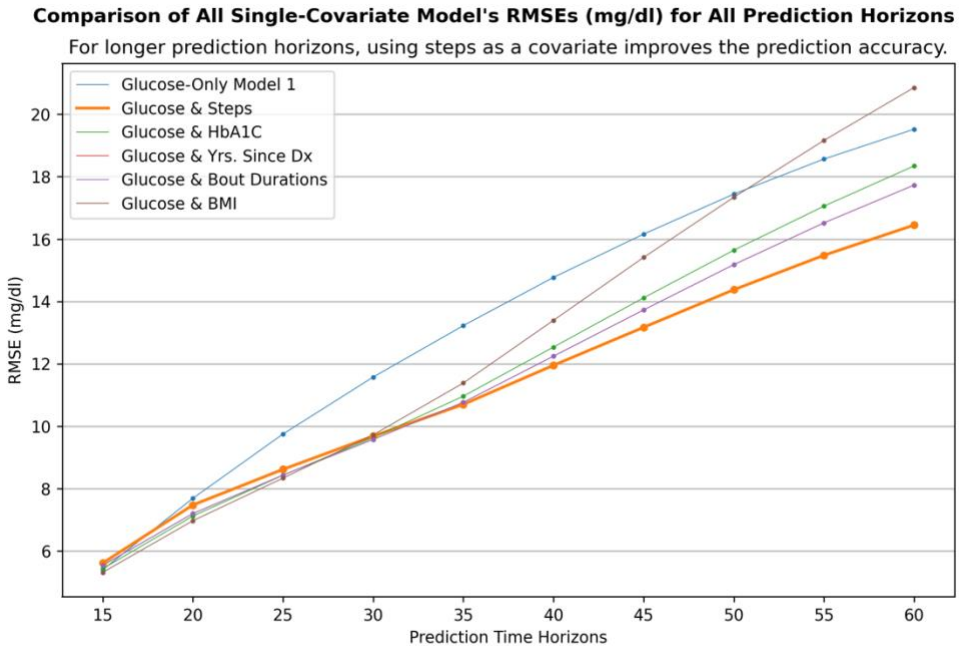


Figure 4.2.1 Error rates from the validation set across all prediction horizons for the bivariate models. Note the model utilizing glucose and step cadences outperforms other models for longer prediction horizons.

4.2.2 Forward Feature Selection of Two Covariates

Because the bivariate model utilizing past glucose and step cadence yielded the lowest average error rates across all prediction horizons, this feature was selected for the next phase of model development. Thus, additional features were added to the set already containing glucose and step values. Each new model composed of three features was trained and fit on the training dataset and evaluated on the validation data set. This is where improvements ceased (see Table 4.2.2 and Figure 4.2.2).

Table 4.2.2

Average Error Rates for Covariate Feature Testing (Two Covariates)

Target Feature	Covariate Features	Average RMSE (mg/dl) across all prediction horizons
Glucose	Step cadence & HbA1C	11.41
	Step cadence & BMI	12.43
	Step cadence & years since diagnosis	12.53
	Step cadence & bout duration	14.17

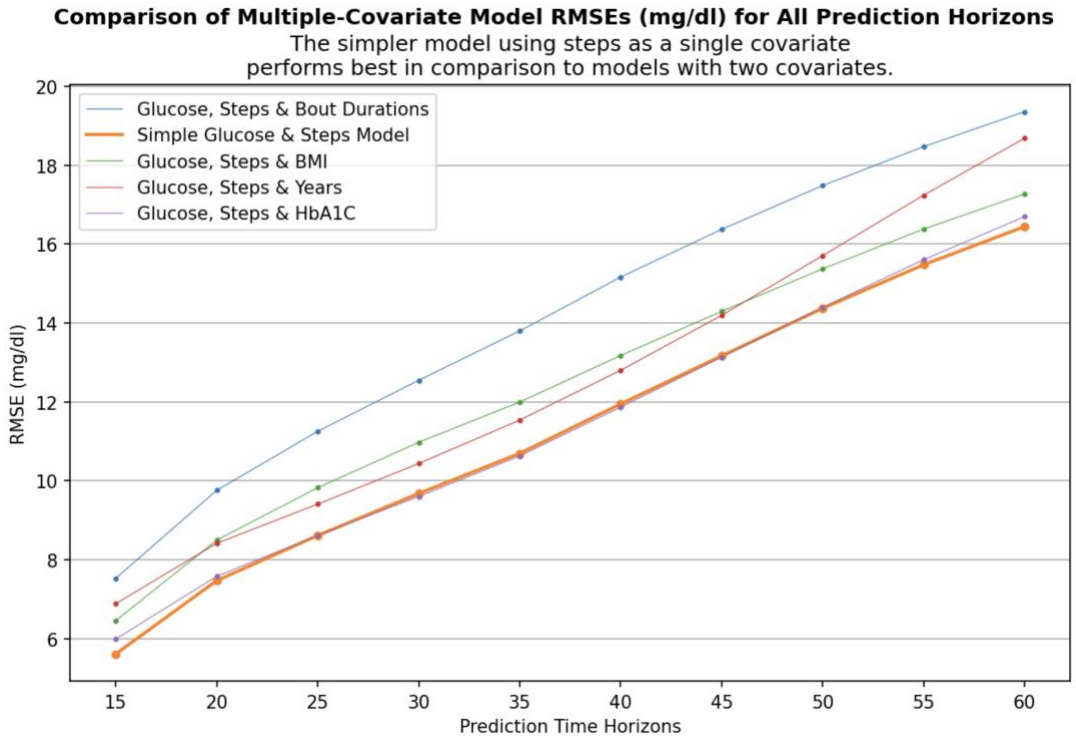


Figure 4.2.2. The previously tested model using glucose and step cadence (shown in orange) outperforms the more complex models utilizing glucose, step cadence and an additional covariate.

4.3 Final Model Evaluation

Model performance on the holdout set yielded an average RMSE of 11.32 ± 0.67 mg/dl across all forecast horizons (see Figure 4.3.1). RMSE for the predictions at the 30-minute horizon was 11.1 ± 0.88 mg/dl. Some predictions were markedly close to the true glucose values, but most decreased in accuracy for longer prediction horizons. Figures 4.3.2 & 4.3.3 show predictions for bouts from the testing set (see Appendix for additional glucose predictions from the testing set). However, glucose predictions made for the longer prediction horizons were markedly lower for the testing set than for the validation set

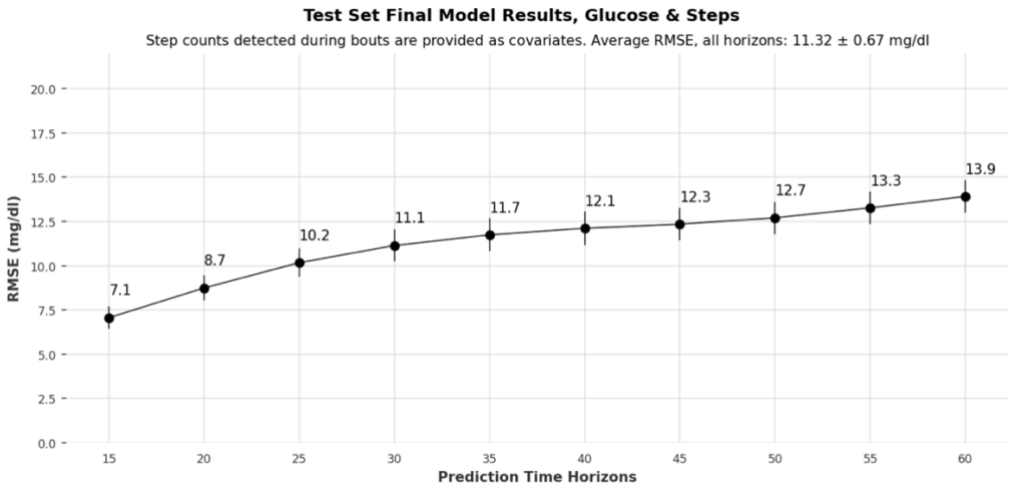


Figure 4.3.1. Final model evaluation on the testing set. Given prior performance on the validation set, the model performs as expected on the testing set.

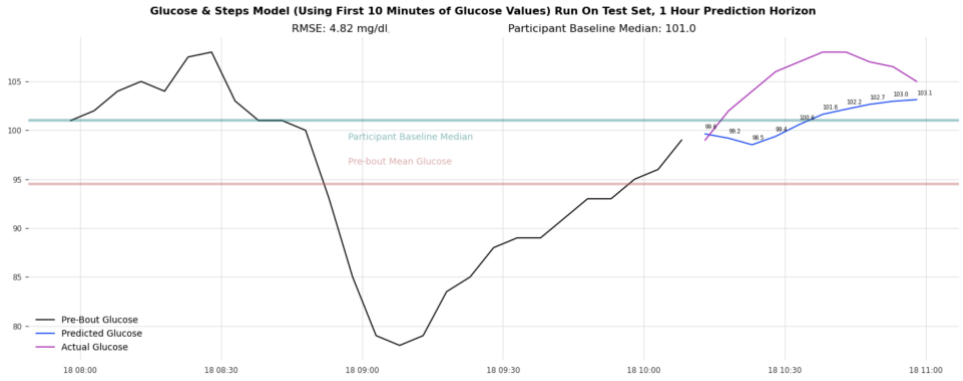


Figure 4.3.2. An example prediction for a bout of activity in which significant changes occurred prior to the bout.

Figure 4.3.3 An example prediction for an individual's bout of activity. Pre-bout glucose values are relatively stable.

CHAPTER 5

Discussion

5.1 Discussion

The most performant model was found to be that which utilized the first ten minutes of glucose values collected during a bout and the step rates coinciding with each timestep (Figures 4.2.2 and 4.3.1). The error rates achieved by this model are within the range of previously published studies attempting to predict glucose for short-term time horizons, and thus fit nicely within previous benchmarks. This shows that when a measure of the intensity of physical activity such as step counts can be taken into consideration, the changes to glucose values can be predicted with better accuracy than simply by using glucose values alone (Figure 4.2.1). Furthermore, predictions for forecast horizons of 30 minutes or more are particularly useful. Forecasts such as these can provide valuable insights to individuals with T2D to help them understand how their activities are likely to impact glucose values, while still giving them enough time to change their behavior. This can also help to provide clarity and reduce the cognitive burden associated with managing glucose values. As a basis for future works, this work also shows the utility of data gathered by passive health monitoring. Activity tracker-derived step counts and CGM data served well as sole model inputs for glucose prediction. Future studies may thus expect high utility in similarly derived datasets from free-living individuals.

In terms of potential productization, this work proffers a key finding: glucose values and step values alone can provide valuable insights to T2D NIIT CGM users. This implies a potentially useful partnership between activity tracker manufacturers and CGM makers in which activity data and CGM data are paired in real-time. Data streams thus composed of glucose data and step data can then be used to preempt future glucose values for short-term timeframes and allow users to employ strategies to attain better glycemic outcomes. Application features such as these might serve to alleviate some elements of diabetes-related distress by providing clarity regarding the impacts of physical activity.

It is also worth pointing out that while the model that utilized the glucose readings from the first 10 minutes of a bout of activity (Model A) outperformed the model trained and fit to predict glucose values solely from the glucose detected 2 hours prior to a bout (Model B), the performances of these two models were very close. Model A had an average RMSE of 13.41 mg/dl, while Model B had an RMSE of 13.46 (Figure 4.1.1). Model A performed best for horizons longer than 35 minutes, while model B performed best for short-term forecast horizons. So, while this study progressed with the development of further models using the same structure of inputs as Model A, it is not unlikely that Model B could provide similarly useful outcomes. The potential utility of Model B is worth pointing out because it could serve to provide useful predictions, given no information except prior glucose values.

The performances of all models were on par with the results of studies attempting to predict glucose for Type 1 diabetes. However, a few key differences could contribute to this. The T2D NIIT population may have slightly different glucose-response patterns associated with physical activity. When studying the changes in glucose values during physical activity in a T1D cohort,

Tyler et al. (2022) found that even under strictly controlled conditions, inter-individual and intra-individual variability in responses to physical activity were vast, making prediction very difficult. Hobbs et al. (2019) also found that, even with reliably gathered heart rates, attempting to predict glucose during physical activity resulted in higher error rates. Romero-Ugalde (2019) found that they were able to achieve lower error rates (7.8 ± 4.5 mg/dl) for the 30-minute time horizon when a model was trained on data from several days' worth of previous glucose values, then used to predict glucose values for that single individual. But once a model was used for population-wide predictions, error rates were relatively high (16.7 ± 15.6 mg/dl for the 30-minute horizon) (Table 2.3.3). But these studies were all focused on T1D patients. The T1D population requires insulin. Insulin alone introduces a new dimension of variability to glucose metrics and thus, this might explain why these studies yielded higher error rates. Despite the differences in populations under study, this work shows that the same mathematical methods used to predict glucose for T1D can be applied to glucose prediction for T2D NIIT individuals.

5.1.2 Limitations of this Study

The T2 Help dataset is composed of 199 T2D NIIT individuals. After filtering out overlapping bouts, imposing the constraints for bouts lasting 10 minutes or more, and excluding any bouts of activity with large gaps of missing glucose values, there were 119 individuals left. Many of these individuals had just a single bout of activity. Sedentary lifestyles are associated with T2D, but this also means that by modern machine learning standards, the sizes of the training, validation, and testing datasets were very small. Furthermore, most studies involving short-term glucose prediction solely involved predictions for T1D individuals. A single study was found involving T2D patients, with a reported RMSE of 37.18 mg/dl for a 30-minute forecast horizon, but this represents a very limited set of prior works to use for benchmarking (Table 2.3.3). Thus, small sample sizes and a lacking body of knowledge pertaining to this specific population mean that the results of this study are largely incommensurate with previous works.

Other limitations of this study include the lack of context for other relevant facets of daily life that might affect glucose values. Inspection of the predicted glucose in comparison to the actual glucose values shows a pattern: the predicted values are composed of very smooth, sinusoid traces while the actual values are much more variable (Figure 5.1.2.1). Predicted changes in glucose values are very similar—smooth, short-lived decreases with gradual increases, while the actual changes in glucose values are, again, much more variable (Figure 5.1.2.2). This indicates an inability of the model to capture the true relationship between glucose values from past timesteps to future timesteps. Possible reasons for this abound, but carbohydrate intake is one dimension of data that assuredly affects glucose values and is not accounted for here. Other features potentially worth investigating include stress responses, body temperature, and medication intake.

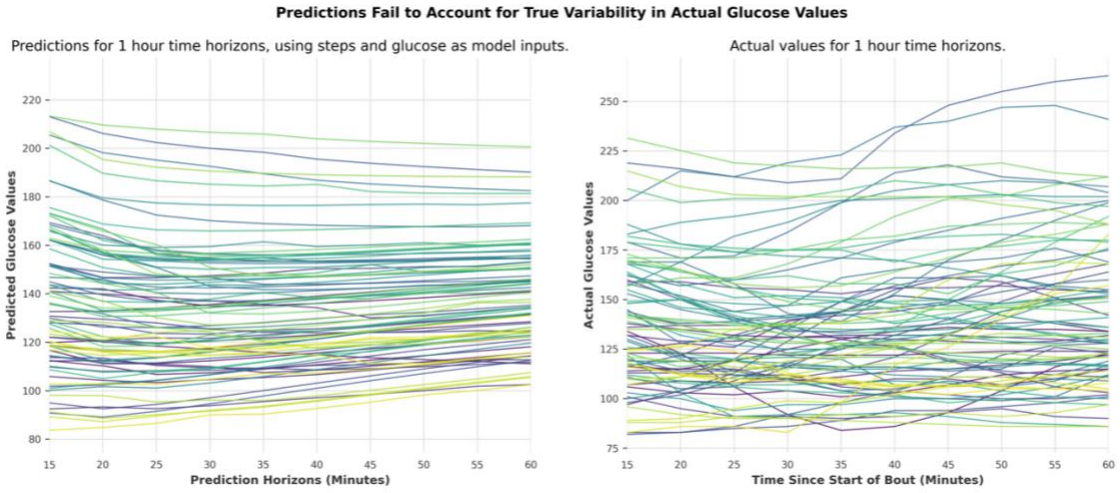


Figure 5.1.2.1 Predicted glucose values (left) in comparison to actual glucose values (right). Actual glucose values exhibit a high degree of variability, while model-generated values exhibit smooth traces.

Additionally, this study focused on step cadences of 60 steps per minute or higher, which worked for the small sample size, but it is possible that this step rate threshold may not be adequately inclusive for larger sample sizes with more diverse age ranges and comorbidities. Further studies might find improved performance by applying individually defined cadence thresholds based on age and mobility. Other improvements that might be made to these works include the additional testing of hyperparameters. A limited set of parameters were tested, two of which yielded improved error metrics, but at either end of the extreme set of parameters (Figure 4.1.2). Further hyperparameter tuning might reveal improved results with parameters selected from further extremes.

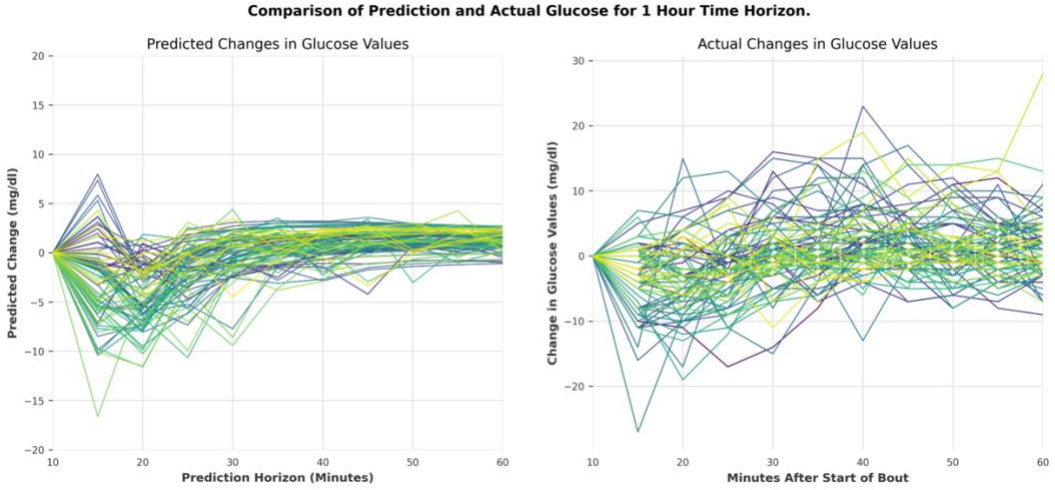


Figure 5.1.2.2 Predicted changes in glucose values (left) compared to actual changes in glucose values. The difference between patterns in the two plots suggest further features(s) relevant to a subject's physiological response to physical activity that is not accounted for by this model.

This study employed machine learning methods for what is inherently a system of physiological processes. This means that, although the predictions may be sufficiently "good" and the outputs might be useful, the model hasn't provided insights into the relationships between the physiological processes that it attempts to describe. This model has no constraints on the outputs it generates, and could easily predict improbable, or even impossible physiological glucose responses to exercise. Designing an algorithm that could operate within the constraints of a deterministic system, while still facilitating the essential stochastic process that allows the model to learn the ideal weights to generate useful predictions would be an interesting, useful endeavor that may serve to deepen the understanding of how glucose dynamics in the T2D population change during physical activity.

5.1.3 Potential Future Works

A potential continuation of this work might follow in the footsteps of the many other researchers who have tackled glucose prediction for T1D. Many of these works evaluated numerous models including classical timeseries models, physiological models, other machine learning methods and hybrid models, and chose the model that worked best. In this work, a single LSTM model architecture was developed and tested. But an LSTM is a relatively complex machine learning method, and it is entirely possible that a simple model could perform just as well. Other models that have shown promising results include FFNs, ARX models and many others not explicitly stated here. Novel predictive glucose models for the T2D NIIT population are sorely needed. Future works might contribute to a standard set of benchmarks upon which future projects might build and improve.

This study focused solely on short-term predictions for glucose values, but another important tenet of prevention and treatment for T2D is the long-term benefits of physical activity. Solomon (2018) found a high degree of variability in glucose transport mechanisms across cellular membranes, per T2D individual and showed that there were many T2D individuals for whom exercise had little or no beneficial effect on their glucose control. The author points out yet another critical knowledge gap in this regard for the T2D population, bemoaning a complete lack of studies focused on the "impact of exercise dosage, frequency, type, or anti-hyperglycemic drugs on changes in blood glucose control for individuals with T2D or prediabetes". The distribution of detected bouts in the cohort for this study suggests that there are groups of individuals in this population that exercise very infrequently. It is possible that a different approach could utilize clustering as an additional pre-processing step to identify characteristics of individuals or of bouts of physical activity that might have similar glucose-responses. Using this method, it may be possible to identify individuals that have uncharacteristic or unexpected glycemic responses and to train separate sequential models on each cluster of individuals. This approach could be applied to short-term predictions, as was done in this study, or to long-term predictions. A study focused on changes to long-term patterns of glycemic control with increasing frequencies of activity might focus on metrics such as increased time in range or decreased time above range to gain further insight and contribute much-needed research on this topic.

CHAPTER 6

Conclusion

My primary objectives in this study were to determine if a model could be developed to predict future glucose values during a bout of physical activity for an individual in the T2D NIIT population. Furthermore, I sought to determine the relevant features germane to the individual or their bout of exercise that might further improve these predictions. Then of course, I needed to ensure that this model should be useful, and thus generalize sufficiently well to never-before-seen data. I found that not only was it possible to rely solely on past glucose values to predict future glucose values, but that given the added feature of step cadences, these predictions could be improved. I used the results of previous studies—which primarily focused on glucose prediction for T1D—as benchmarks to determine the relative performance of this model. As a proof of concept, this work shows that step cadence, which serves to account for the intensity of an individual's engagement in physical activity is a significant determining factor affecting the outcome of changes in glucose values. This work also shows that it may be possible to generate insights that illustrate the impact of what current actions might be on future glucose values.

While the results of this study illustrate some utility for future works and for T2D NIIT glucose prediction, one major limitation of this study was simply a lack of data and subsequently small sample sizes. The T2 Help study and the data it generated show that it is possible to glean an incredible amount of insight from individuals wearing a simple activity tracker with their CGM. But much more is needed to improve these predictions, especially if more advanced methods like clustering are to be performed to train different models on different individuals. Thus, it can be said that over the course of this study, a constant theme has emerged: a lack of research surrounding glucose metrics for the T2D community, particularly regarding the impacts of physical activity. It seems fitting that, in closing, this narrative should echo the plea from Bremer & Gough (1999) for more data, but to tailor that plea more specifically for the T2D community.

BIBLIOGRAPHY

Aliberti, A., Pupillo, I., Terna, S., Macii, E., Di Cataldo, S., Patti, E., & Acquaviva, A. (2019). A Multi-Patient Data-Driven Approach to Blood Glucose Prediction. IEEE Access, 7, 69311–69325. https://doi.org/10.1109/access.2019.2919184

Al Olawai, L., Alharbu, T., & Tourkmani, A. (2018). Prevalence of Cardiovascular Autonomic Neuropathy and Gastroparesis Symptoms among Adults with Type 2 Diabetes Who Attend a Primary Health Care Center. Diabetes, 67(Supplement 1), 559-P. https://doi.org/10.2337/db18-559-p

American diabetes association. (2021). Standards of Medical Care in Diabetes—2022 Abridged for Primary Care Providers. Clinical Diabetes, 40(1). https://doi.org/10.2337/cd22-as01

Battelino, T., Danne, T., Bergenstal, R. M., Amiel, S. A., Beck, R., Biester, T., Bosi, E., Buckingham, B. A., Cefalu, W. T., Close, K. L., Cobelli, C., Dassau, E., DeVries, J. H., Donaghue, K. C., Dovc, K., Doyle, F. J., Garg, S., Grunberger, G., Heller, S., & Heinemann, L. (2019). Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care, 42(8), 1593–1603. https://doi.org/10.2337/dci19-0028

Borovykh, A., Bohte, S., & Oosterlee, C. W. (2018). Dilated convolutional neural networks for time series forecasting. Journal of Computational Finance. https://doi.org/10.21314/jcf.2019.358

Carrillo-Moreno, J., Pérez-Gandía, C., Sendra-Arranz, R., García-Sáez, G., Hernando, M. E., & Gutiérrez, Á. (2020). Long short-term memory neural network for glucose prediction. Neural Computing and Applications. https://doi.org/10.1007/s00521-020-05248-0

Cohen FJ, Neslusan CA, Conklin JE, Song X. Recent antihyperglycemic prescribing trends for US privately insured patients with type 2 diabetes. Diabetes Care. 2003 Jun;26(6):1847-51. doi: 10.2337/diacare.26.6.1847. PMID: 12766121.

Contreras I, Oviedo S, Vettoretti M, Visentin R, Vehí J. Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models. PLoS One. 2017 Nov 7;12(11):e0187754. doi: 10.1371/journal.pone.0187754. PMID: 29112978; PMCID: PMC5675457.

Dierks, M. H. (2021, July 15). The Complete Guide to Insulin Types [w/ Downloadable Charts]. AgaMatrix. https://agamatrix.com/blog/different-types-of-insulin/

Dietterich, T. G. (2002). Machine Learning for Sequential Data: A Review. Lecture Notes in Computer Science, 15–30. https://doi.org/10.1007/3-540-70659-3_2

Dobrică, E.-C., Găman, M.-A., Cozma, M.-A., Bratu, O. G., Pantea Stoian, A., & Diaconu, C. C. (2019). Polypharmacy in Type 2 Diabetes Mellitus: Insights from an Internal Medicine Department. Medicina, 55(8). https://doi.org/10.3390/medicina55080436

Elliott, L., Fidler, C., Ditchfield, A., & Stissing, T. (2016). Hypoglycemia Event Rates: A Comparison Between Real-World Data and Randomized Controlled Trial Populations in Insulin-Treated Diabetes. *Diabetes Therapy*, *7*(1), 45–60. https://doi.org/10.1007/s13300-016-0157-z

Freeman, J. (2019). Management of hypoglycemia in older adults with type 2 diabetes. Postgraduate Medicine, 131(4), 241–250. https://doi.org/10.1080/00325481.2019.1578590

Georga, E., Vasilios C., Fotiadis D., Fotiadis, P. (2011). Glucose Prediction in Type 1 and Type 2 Diabetic Patients Using Data Driven Techniques, Knowledge-Oriented Applications in Data Mining, Prof. Kimito Funatsu (Ed.), ISBN: 978-953-307-154-1, InTech, Available from: http://www.intechopen.com/books/knowledge-oriented-applications-in-data - mining/glucose-prediction-in-type- 1-and-type-2-diabetic-patients-using-data-driven-techniques

Granat, M., Clarke, C., Holdsworth, R., Stansfield, B., & Dall, P. (2015). Quantifying the cadence of free-living walking using event-based analysis. Gait & Posture, 42(1), 85–90. https://doi.org/10.1016/j.gaitpost.2015.04.012

Grant RW, Pirraglia PA, Meigs JB, Singer DE. Trends in complexity of diabetes care in the United States from 1991 to 2000. Arch Intern Med. 2004 May 24;164(10):1134-9. doi: 10.1001/archinte.164.10.1134. PMID: 15159272.

Hobbs, N., Hajizadeh, I., Rashid, M., Turksoy, K., Breton, M., & Cinar, A. (2019). Improving Glucose Prediction Accuracy in Physically Active Adolescents With Type 1 Diabetes. Journal of Diabetes Science and Technology, 13(4), 718–727. https://doi.org/10.1177/1932296818820550

Jetté, M., Sidney, K., & Blümchen, G. (1990). Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clinical Cardiology, 13(8), 555–565. https://doi.org/10.1002/clc.4960130809

Kane, N. S., Hoogendoorn, C. J., Tanenbaum, M. L., & Gonzalez, J. S. (2018). Physical symptom complaints, cognitive emotion regulation strategies, self-compassion and diabetes distress among adults with Type 2 diabetes. Diabetic Medicine, 35(12), 1671–1677. https://doi.org/10.1111/dme.13830

Li, K., Daniels, J., Liu, C., Herrero-Vinas, P., & Georgiou, P. (2019). Convolutional Recurrent Neural Networks for Glucose Prediction. IEEE Journal of Biomedical and Health Informatics, 1–1. https://doi.org/10.1109/jbhi.2019.2908488

Lima, F. V., Kadiyala, V., Huang, A., Agusala, K., Cho, D., Freeman, A. M., & Druz, R. (2022). At the Crossroads! Time to Start Taking Smartwatches Seriously. The American Journal of Cardiology, 179, 96–101. https://doi.org/10.1016/j.amjcard.2022.06.020

Lipska, K. J., Krumholz, H., Soones, T., & Lee, S. J. (2016). Polypharmacy in the Aging Patient. JAMA, 315(10), 1034. https://doi.org/10.1001/jama.2016.0299

Lubitz, S. A., Faranesh, A. Z., Selvaggi, C., Atlas, S. J., McManus, D. D., Singer, D. E., Pagoto, S., McConnell, M. V., Pantelopoulos, A., & Foulkes, A. S. (2022). Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. Circulation, 146(19), 1415–1424. https://doi.org/10.1161/circulationaha.122.060291

Marshall, S. J., Levy, S. S., Tudor-Locke, C. E., Kolkhorst, F. W., Wooten, K. M., Ji, M., Macera, C. A., & Ainsworth, B. E. (2009). Translating Physical Activity Recommendations into a Pedometer-Based Step Goal. American Journal of Preventive Medicine, 36(5), 410–415. https://doi.org/10.1016/j.amepre.2009.01.021

Mosquera-Lopez, C., & Jacobs, P. G. (2021). Incorporating Glucose Variability into Glucose Forecasting Accuracy Assessment Using the New Glucose Variability Impact Index and the Prediction Consistency Index: An LSTM Case Example. Journal of Diabetes Science and Technology, 16(1), 7–18. https://doi.org/10.1177/19322968211042621

National Diabetes Statistics Report [Review of National Diabetes Statistics Report]. (2022). In cdc.gov. Centers for Disease Control. https://www.cdc.gov/diabetes/data/statistics-report/index.html?ACSTrackingID=DM72996&ACSTrackingLabel=New%20Report%20Shares%20Latest%20Diabetes%20Stats%20&deliveryName=DM72996

Orendurff, M. S. (2008). How humans walk: Bout duration, steps per bout, and rest duration. The Journal of Rehabilitation Research and Development, 45(7), 1077–1090. https://doi.org/10.1682/jrrd.2007.11.0197

Pérez-Gandía, C., Facchinetti, A., Sparacino, G., Cobelli, C., Gómez, E. J., Rigla, M., de Leiva, A., & Hernando, M. E. (2010). Artificial Neural Network Algorithm for Online Glucose

Pop-Busui, R. (2010). Cardiac Autonomic Neuropathy in Diabetes: A clinical perspective. Diabetes Care, 33(2), 434–441. https://doi.org/10.2337/dc09-1294

Prediction from Continuous Glucose Monitoring. Diabetes Technology & Therapeutics, 12(1), 81–88. https://doi.org/10.1089/dia.2009.0076

Romero-Ugalde, H. M., Garnotel, M., Doron, M., Jallon, P., Charpentier, G., Franc, S., Huneker, E., Simon, C., & Bonnet, S. (2019). ARX model for interstitial glucose prediction during and after physical activities. Control Engineering Practice, 90, 321–330. https://doi.org/10.1016/j.conengprac.2019.07.013

Serhiyenko, V. A., & Serhiyenko, A. A. (2018). Cardiac autonomic neuropathy: Risk factors, diagnosis and treatment. World Journal of Diabetes, 9(1), 1–24. https://doi.org/10.4239/wjd.v9.i1.1

Serrano, F., Slaght, J., Sénéchal, M., Duhamel, T., & Bouchard, D. R. (2017). Identification and Prediction of the Walking Cadence Required to Reach Moderate Intensity Using Individually-Determined Relative Moderate Intensity in Older Adults. Journal of Aging and Physical Activity, 25(2), 205–211. https://doi.org/10.1123/japa.2015-0262

Silbert, R., Salcido-Montenegro, A., Rodriguez-Gutierrez, R., Katabi, A., & McCoy, R. G. (2018). Hypoglycemia Among Patients with Type 2 Diabetes: Epidemiology, Risk

Factors, and Prevention Strategies. Current Diabetes Reports, 18(8). https://doi.org/10.1007/s11892-018-1018-0

Sharma M., Nazareth I., Petersen I. Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study BMJ Open 2016;6:e010210. doi: 10.1136/bmjopen-2015-010210

Sun, Q., M. V. Jankovic, L. Bally and S. G. Mougiakakou, "Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network," 2018 14th Symposium on Neural Networks and Applications (NEUREL), 2018, pp. 1-5, doi: 10.1109/NEUREL.2018.8586990

Tourkmani, A. M., Alharbi, T. J., Rsheed, A. M. B., AlRasheed, A. N., AlBattal, S. M., Abdelhay, O., Hassali, M. A., Alrasheedy, A. A., Al Harbi, N. G., & Alqahtani, A. (2018). Hypoglycemia in Type 2 Diabetes Mellitus patients: A review article. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 12(5), 791–794. https://doi.org/10.1016/j.dsx.2018.04.004

Tudor-Locke, C., Camhi, S. M., Leonardi, C., Johnson, W. D., Katzmarzyk, P. T., Earnest, C. P., & Church, T. S. (2011). Patterns of adult stepping cadence in the 2005–2006 NHANES. Preventive Medicine, 53(3), 178–181. https://doi.org/10.1016/j.ypmed.2011.06.004

Tudor-Locke, C., Han, H., Aguiar, E. J., Barreira, T. V., Schuna Jr, J. M., Kang, M., & Rowe, D. A. (2018). How fast is fast enough? Walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. British Journal of Sports Medicine, 52(12), 776–788. https://doi.org/10.1136/bjsports-2017-097628

Tudor-Locke, C., Mora-Gonzalez, J., Ducharme, S. W., Aguiar, E. J., Schuna, J. M., Barreira, T. V., Moore, C. C., Chase, C. J., Gould, Z. R., Amalbert-Birriel, M. A., Chipkin, S. R., & Staudenmayer, J. (2021). Walking cadence (steps/min) and intensity in 61–85-year-old adults: the CADENCE-Adults study. International Journal of Behavioral Nutrition and Physical Activity, 18(1). https://doi.org/10.1186/s12966-021-01199-4

Tudor-Locke, C., Sisson, S. B., Collova, T., Lee, S. M., & Swan, P. D. (2005). Pedometer-Determined Step Count Guidelines for Classifying Walking Intensity in a Young Ostensibly Healthy Population. Canadian Journal of Applied Physiology, 30(6), 666–676. https://doi.org/10.1139/h05-147

Tyler, N. S., Mosquera-Lopez, C., Young, G. M., El Youssef, J., Castle, J. R., & Jacobs, P. G. (2022). Quantifying the impact of physical activity on future glucose trends using machine learning. IScience, 25(3), 103888. https://doi.org/10.1016/j.isci.2022.103888

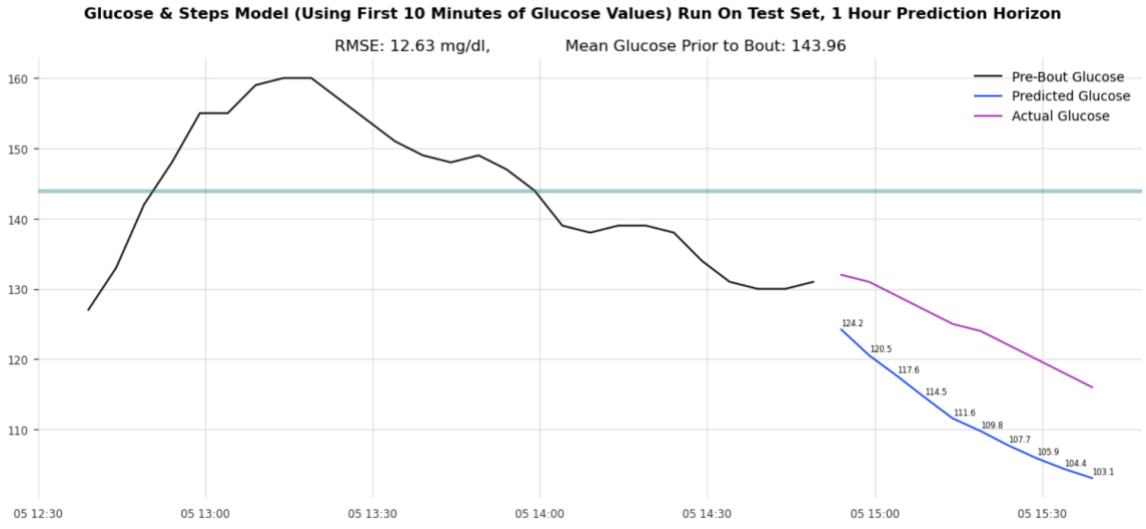
Vinik, A. I., & Ziegler, D. (2007). Diabetic Cardiovascular Autonomic Neuropathy. Circulation, 115(3), 387–397. https://doi.org/10.1161/circulationaha.106.634949

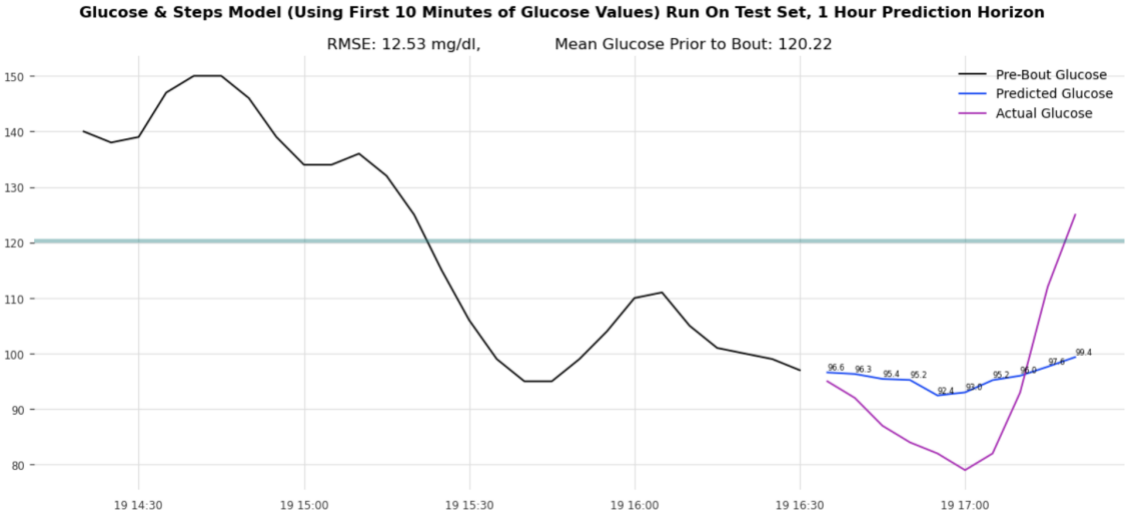
Webber, S. C., Hahn, F., Lix, L. M., Tittlemier, B. J., Salbach, N. M., & Barclay, R. (2020). Accuracy of Thresholds Based on Cadence and Lifestyle Counts per Minute to Detect Outdoor Walking in Older Adults With Mobility Limitations. Journal of Aging and Physical Activity, 28(5), 782–786. https://doi.org/10.1123/japa.2019-0235

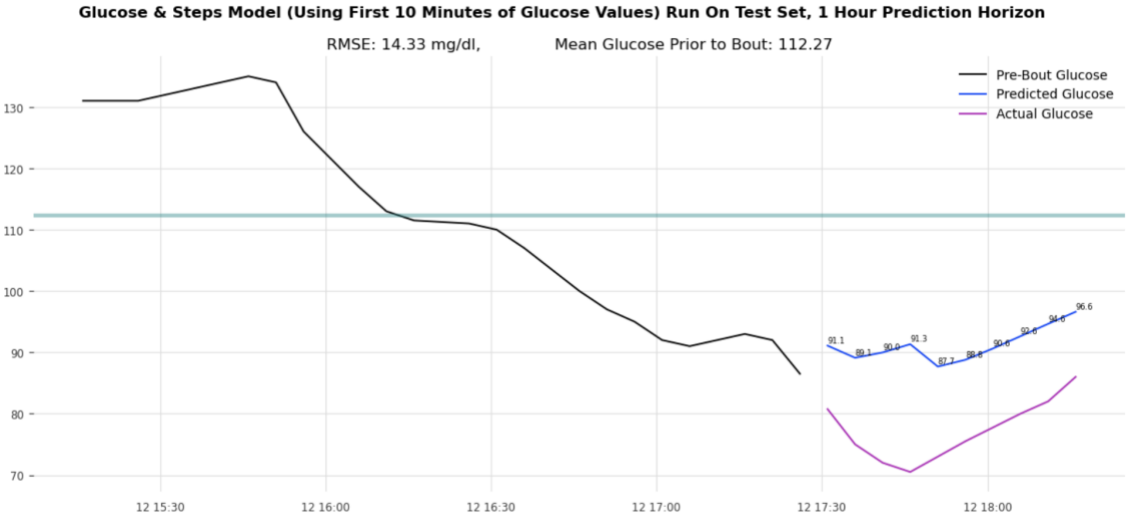
Xie J, Wang Q. Benchmarking Machine Learning Algorithms on Blood Glucose Prediction for Type I Diabetes in Comparison With Classical Time-Series Models. IEEE Trans Biomed Eng. 2020 Nov;67(11):3101-3124. doi: 10.1109/TBME.2020.2975959. Epub 2020 Feb 24. PMID: 32091990.

APPENDIX

Below are some examples of the predictions generated by the final version of the LSTM RNN. Each plot shows an individual's, pre-bout glucose values (in black), with the predicted glucose values (in blue, with labeled, predicted values), in comparison to their actual glucose values (purple). Each person's mean glucose values, detected in the two hours preceding the bout, are shown as a reference line in teal.







Glucose & Steps Model (Using First 10 Minutes of Glucose Values) Run On Test Set, 1 Hour Prediction Horizon