Artificial Intelligence and Physiologic Models for the Management of Type 1 Diabetes

By

Nichole Sahar Tyler, B.Sc.

A DISSERTATION

Presented to the Department of Biomedical Engineering
of the Oregon Health & Science University School of Medicine
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Biomedical Engineering

March 2021

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the PhD dissertation of

Nichole Sahar Tyler

has been approved

_____
Peter Jacobs, PhD, Mentor / Advisor

_____
Daniel Zuckerman, PhD, Committee Chair

_____
Jessica Castle, MD, Member

_____
Reid Thompson, MD-PhD, Member

_____
Dacian Daescu, PhD, Member

_____
Stuart Ibsen, PhD, Member

Acknowledgments

Graduate training requires constant effort, and leads to slow and steady changes – similar to cultivating a seedling.

Thank you Dr. Jacobs for being an attentive gardener. You are understanding, kind, and always make time to help your students to advance in their training. We met every week for 4 years and had many meaningful conversations. This special time together led to new ideas, learning opportunities, and contributions in fields of biosensors, machine learning algorithms, and physiology modeling.

The soil of a garden is also important. Thank you AIMS lab for providing a collaborative, safe and friendly work environment. Similar to a gardens biodiversity, our diverse backgrounds and skills enhance the growth and productivity of all of its members. Thank you Dr. Clara Mosquera Lopez, Dr. Leah Wilson, Dr. Robert Dodier, Brin, Brian, Debbie, Florian, Gavin, Joe L., Joe P., Wade, and former lab members Navid and Ravi for your friendship and support.

Thank you Dr. Jessica Castle and Dr. Joseph El Youssef for your guidance, for providing me the opportunity to work with you in clinic, and providing learning experiences in translational clinical research. Thank you Dr. Reid Thompson, Dr. Daniel Zuckerman and Dr. Dacian Daescu for overseeing the progress of my graduate studies.

I am grateful for the support and love from my friends (Leah and Addison and Leona), and my beautiful family (Nasrin and Scott), especially from my dear sister Dr. Yvonne Tyler who gives great advice.

Lastly, I am grateful to Dr. David Jacoby for the opportunity to join the M.D.-Ph.D. community at OHSU. In Iran, members of the Bahá'í religion are systematically banned from higher education, job permits, or even burial rights. I hope that through our actions, we can all plant seeds based on love, respect and equal opportunities.

> "Today the seed is sown, the grain falls upon the earth, but behold the day will come when it shall rise a glorious tree and the branches thereof shall be laden with fruit."
>
> - Abdul-Baha, Paris Talks

List of Abbreviations

AI	Artificial intelligence
AIDA	An interactive diabetes advisor system
ALPHA	Adaptive learning postprandial hypoglycemia algorithm
ARX	Autoregressive model with exogenous inputs
CBG	Capillary blood glucose
CBR	Case-based reasoning
CDE	Certified diabetes educator
EGP	Endogenous glucose production
FDA	United states food and drug administration
HbA1C	Hemoglobin A1C
HR	Heartrate
IIR	Insulin infusion rate
IOB	Insulin on board
KADIS	Karlsburg diabetes management system
LBGI	Low blood glucose index
MAE	Mean absolute error
MAPE	Mean absolute percent error
MARD	Mean absolute relative difference
MARE	Mean absolute relative error
MET	Metabolic expenditure
MMP	Multiple model predictor
NARMAX	Non-linear autoregressive moving average model with exogenous inputs
ODE	Ordinary differential equation
PVO_{2max}	Percent of maximal oxygen consumption
Ra	Rate of glucose appearance
Rd	Rate of glucose disposal
RMSE	Root mean squared error
RNN	Recurrent neural network
T1-Dexi	Type 1 diabetes exercise study

TDIR	Total daily insulin requirement
SMBG	Self-monitored blood glucose
SVM	Support vector machine
VARX	Vector autoregressive model with exogenous inputs
YSI	Plasma glucose measured by Yellow Springs Instrument 3200

Table of Contents

List of Tables

List of Figures

List of Equations

Abstract

Type 1 diabetes is a medical condition characterized by pancreatic β-cell dysfunction and insulin depletion. Management of type 1 diabetes requires precise injections of exogenous insulin in order to bring glucose within a target range (70-180 mg/dL), however this is difficult to achieve. Factors that affect optimal management of diabetes include incorrect insulin therapy parameters, difficulty to calculate insulin boluses, physiologic variations in insulin sensitivity, adherence with care, and infrequent visits with medical providers. Incorporation of regular exercise has been shown to improve short-term glycemic outcomes and reduce the risk of cardiovascular disease in type 1 diabetes. However, exercise also increases the risk of a medical complication called hypoglycemia (glucose < 70 mg/dL) which is difficult to predict and prevent. In this dissertation, I present artificial intelligence algorithms and physiologic model systems that are designed to provide decision support for the management of type 1 diabetes. The first contribution is the design of a K-nearest-neighbors decision support system (KNN-DSS) that identifies problematic glycemic patterns, and provides weekly recommendations regarding insulin therapy adjustments and dosing behaviors. The KNN-DSS achieves an overall agreement of 67.9% with endocrinologist recommendations, while overall agreement among endocrinologists was measured to be 55.9%. The KNN-DSS also improves glycemic outcomes during 12-week *in silico* clinical trials and reduces hypoglycemia in a human clinical study. The next contribution is the quantification of glucose changes during aerobic exercise using artificial intelligence models. We show definitively that people with higher aerobic fitness will experience significantly steeper glucose trends, and significantly lower glucose during and following exercise, than people with lower aerobic fitness. Adaptive machine learning algorithms

are designed to provide personalized estimations of exercise-related changes in glucose. The algorithms achieve a mean absolute error of 18.1 mg/dL and accuracy of 78% to predict hypoglycemia, and represent an upper limit for the predictive accuracy of exercise algorithm performance in real-world use cases. Lastly, this dissertation contributes new physiologic compartment models that estimate the impact of aerobic exercise on insulin and non-insulin mediated glucose disposal mechanisms. These models are incorporated into an existing virtual patient simulator for type 1 diabetes, and can be used to simulate glucose trends during aerobic exercise sessions across various exercise intensities and insulin loads. Taken together, this dissertation contributes tools for diabetes decision support in order to make adjustments to insulin therapy, and to better predict changes in glucose during aerobic exercise. These approaches are being incorporated into mobile decision support and automated insulin delivery algorithms.

1 Introduction

On December 26 2004, a magnitude 9 earthquake lasting 10 minutes was measured off the coast of Indonesia. Within hours, a series of 100-foot-tall tsunami waves swept across Southeast Asia and killed over 200,000 people. My friend Ryley and her family were vacationing in Phuket, Thailand when the earthquake occurred. A Tsunami alarm interrupted their breakfast, giving them minutes to ascend to the roof of the hotel. Ryley's father ran back to the hotel restaurant to retrieve their passports and wallets before the first wave swept through the ground level. He was successful and reunited with their family before a 2nd, more powerful wave was detected. The coastal areas underwent immediate evacuation. As Ryley's family proceeded with thousands of others towards high ground, her father became progressively disoriented and agitated. Against their pleading, he started to head back towards the destroyed hotel. *Screw the second wave. If I don't get food I'm going to die anyway*. He was experiencing a medical complication called hypoglycemia, which occurs when glucose levels in the blood drop below 70 mg/dL. For people with diabetes, hypoglycemia can occur when too much of a medication called *insulin* is administered, and also when there are changes in an individual's sensitivity to insulin, often due to exercise. In his case, physical activity to recover passports as well as dosing insulin before eating (his breakfast was interrupted by the Tsunami), contributed to the onset of this medical emergency. In their search for food, Ryley's family became separated from the crowd of evacuees. The sky grew dark, and soon Ryley's father would need additional medical treatment. English is not widely spoken outside of the tourist areas in Thailand, making it

extremely difficult to communicate for directions. Meanwhile, 20-foot waves destroyed their intended vacation destination, the Phi Phi Islands, leaving no means of escape for the residents and tourists. How could they have been fortunate to survive one of the deadliest natural disasters in history, only to be faced with the constant danger of diabetes management?

While the circumstances described in the story above are uncommon, medical complications surrounding diabetes management are a frequent occurrence. Diabetes is estimated to affect 463 million people worldwide [1], and is divided into three major subtypes: type 1 (autoimmune), type 2 (insulin resistant with relative insulin deficiency) and gestational (pregnancy-induced) diabetes. In 2019, the global impact of diabetes included 4.2 million deaths and 760.3 billion dollars in healthcare costs, with ~50% of these costs estimated to stem directly from the occurrence of diabetes-related complications [1]. Physicians and their patients must navigate a growing landscape of pharmaceuticals, complex treatment guidelines, and technologies to manage this health condition [2]. Proper management is critical to prevent short-term complications such as hypoglycemia or hyperglycemia, and long-term consequences including cardiovascular disease or nervous system disorders. However many people living with diabetes do not access treatment; an estimated 66.8% of people living with diabetes in low-income areas are undiagnosed, and upwards of 30% cannot access or afford medication [1].

While type 2 diabetes accounts for close to 90% of all cases, there is a growing incidence of type 1 diabetes (T1D) worldwide. Optimal management of type 1 diabetes requires aggressive and early insulin therapy to reduce the severity of long-term complications [3], however this can be difficult to adhere to because there is increased risk of hypoglycemia [4], and many people

experience difficulty to calculate insulin doses [5]. This difficulty magnified in type 1 diabetes due to constant variations in insulin sensitivity; regular changes in hormones, stress, illness, sleep cycle, and physical activity all affect an individual's response to insulin. The end result is a moving target for insulin therapy that requires regular titration and glucose monitoring.

Decision support systems (DSS) can be used by people living with diabetes, and their healthcare providers, to make treatment decisions in order to better manage diabetes. Commercial DSS are designed to provide communication between individuals and coaches or physicians, on-demand clinical guidelines, diaries to track daily activities, insulin bolus calculators, and diabetes education resources [6]. In clinic and hospital settings, DSS have also been integrated with electronic health record software in order to suggest modifications to pharmaceutical therapy for individuals with type 1 and type 2 diabetes [7-10].

The growing availability of smartphone technology has enabled researchers and clinicians to develop specific phone apps intended to help manage diabetes. In rural-care settings, mobile DSS have enabled screening of complications related to diabetes [11], and additional resources to community health workers regarding therapeutic strategies [12]. Commercially available mobile DSS for type 1 diabetes provide insulin bolus calculators, real-time alerts to predicted hyperglycemia or hypoglycemia, automated insulin dose advisors, and can display weekly glycemic trends and can send information to clinical providers. And while these approaches are promising, current commercial systems are not yet marketed for the management of glucose and prediction of hypoglycemia in the context of recent physical activity.

This chapter provides a brief introduction to diabetes, common issues in the management of insulin-dependent diabetes, and current state of the art approaches to assist with diabetes management. The methods that are used to develop computational model frameworks for diabetes management include artificial intelligence, mathematical models of glucose physiology, and Bayesian inference.

This dissertation contributes new computational algorithms and models for the management of type 1 diabetes in order to reduce medical complications surrounding insulin therapy and physical activity.

1.1 Diabetes

Diabetes is a chronic medical condition characterized by elevated plasma levels of a sugar molecule called *glucose*, due to the ineffective activity of a hormone called *insulin*. Normally, insulin is produced by the pancreas, and facilitates the cellular uptake of glucose into specific cell populations, such as skeletal myocytes, adipocytes, and cardiomyocytes. Glucose serves as a vital energy source for cellular metabolism and cellular function; but in diabetes, insulin-facilitated cellular uptake of glucose from blood is greatly diminished, and glucose accumulates in the blood causing *hyperglycemia* (glucose > 180 mg/dL). The consequence of chronic exposure to elevated glucose levels is organ damage; cardinal complications of loss of eyesight, loss of sensation and touch, kidney failure, cardiovascular disease and stroke [13], as well as increased risk of serious infections [14] and even cancer [15]. An estimated 463 million people worldwide live with diabetes [1]; the majority of these people fall into three diagnostic subtypes which include type 1, type 2, and gestational diabetes which

are described below. This dissertation describes new approaches in engineering designed to provide diabetes therapy recommendations and avoid common complications in insulin-dependent type 1 diabetes. And, chapter 2 and chapter 6 future discuss expanding these approaches to people living with all forms of diabetes.

1.1.1Type 2 diabetes

Type 2 diabetes accounts for 90% of all diabetes diagnoses [1]. Type 2 diabetes results from the decreased effectiveness of insulin to facilitate cellular glucose uptake, also known as *insulin resistance*. The pancreas compensates for body-wide insulin resistance by producing more insulin, which temporarily sustains normal levels of glucose uptake [16]. However, over time the insulin-dose response loses efficacy and hyperglycemia gradually worsens. It is estimated that close to 50% of people with type 2 diabetes are asymptomatic and undiagnosed until years after disease onset [1]. As such, type 2 diabetes is typically diagnosed in adults, and can even present with long-term diabetes-related complications such as retinopathy. However there is also an increased incidence of type 2 diabetes in children due to world-wide increase in childhood obesity and physical inactivity [1]. Management of type 2 diabetes is focused on bringing glucose levels to target % A1C, using lifestyle changes such as diet and exercise in order to reduce hyperglycemia and promote natural uptake of glucose into skeletal muscle; and additional pharmacologic therapy includes agents that decrease hepatic gluconeogenesis (*biguanides*), increase cellular glucose uptake in response to insulin (*thiazolidinediones*), increase insulin production (*sulfonylureas*), suppress glucagon release *(glucagon-like-peptide 1 agonists)*,

and also promote renal excretion of glucose (*sodium glucose co-transport-2 inhibitors*) [17].

1.1.2 Gestational diabetes

Gestational diabetes is characterized by hyperglycemia during pregnancy, due to the diverse hormonal milieu of the developing placenta and subsequent insulin resistance. Gestational diabetes can occur in individuals with pre-existing health conditions or a family history of gestational diabetes [1], and often resolves at the end of pregnancy. Management of this condition is critical to prevent complications during pregnancy and delivery, and often consists of alterations in diet and physical activity, as well as insulin therapy [18].

1.1.3 Type 1 diabetes

Type 1 diabetes (T1D) is characterized by the loss of pancreatic β-cell function and subsequent insulin depletion. Absence of circulating insulin causes drastic reductions in body-wide glucose uptake. The β-cell dysregulation also disrupts a delicate paracrine network between adjacent pancreatic α- and δ-cell populations [19], which results in the gradual loss of normal glucagon signaling and loss of normal hepatic glucose regulation that progresses in parallel to the onset of type 1 diabetes [19]. The causes of pancreatic β-cell dysregulation are diverse. In many cases, a combination of genetic and environmental factors induce autoimmune responses to β-cell populations [20]. Certain viral infections may serve as environmental triggers for autoimmune response [21]. Longstanding insulin resistance in type 2 diabetes can also cause gradual β-cell burnout and destruction [22]. In response to reduced cellular glucose uptake, the liver will metabolize fats for energy and release ketone bodies into the blood in

order to provide alternative energy to tissues [23]. Gradual β-cell loss and hyperglycemia will induce excessive thirst, hunger, and urination (*polydipsia*, *polyphagia*, and *polyuria*). However in early type 1 diabetes these symptoms can be easily overlooked, as overt symptoms may not present until an 80% loss in β-cell mass [20]. Eventually, the altered metabolic state and increased plasma ketones will induce a medical emergency called ketoacidosis, which presents with symptoms of nausea, vomiting, dehydration, and disorientation. Type 1 diabetes is frequently diagnosed in children, globally affecting 1.1 million children and youth under the age of 20, but can present at any age and is increasing in incidence worldwide [1]. Type 1 diabetes is managed by replacing the depleted insulin with doses of exogenous insulin, as well as frequent monitoring of blood glucose. This dissertation presents new strategies to manage type 1 diabetes.

1.2 Management of type 1 diabetes

1.2.1 Insulin

An estimated 1.4 million adults in North America live with T1D [24], and must administer exogenous insulin in order to bring blood glucose within a target range (70-180 mg/dL) [2]. Insulin therapy can take diverse forms, including multiple daily injection therapy of fast-acting bolus insulin and long-acting basal insulin (MDI), or continuous infusion of insulin by an insulin pump. Newer formulations of insulin provide broad options of time-to-effect, with ultra-fast-acting insulin that impacts glucose and is disposed in a matter of minutes, as well as ultra-long-acting insulin that impacts glucose levels for over 4 days [25]. The combination of these varying insulin formulations through multiple daily injections

are intended to mimic basal levels of pancreatic insulin production and meal-time insulin response. Guidelines exist to help providers manage insulin therapy [2]. Major therapy parameters for type 1 diabetes include (1) basal insulin dosage, (2) a carbohydrate to insulin dosage ratio (carb:ins) that is used to calculate insulin doses for meals, and (3) an insulin sensitivity correction factor (CF) that is used to calculate insulin doses for general glucose excursions. Errors in the timing or dose of insulin injections are common, and can result in medical complications such as *hypoglycemia* (glucose <70 mg/dL), which, untreated, can lead to coma or death. In contrast, underdosing insulin can lead to *hyperglycemia* (glucose >180 mg/dL), to which long-term exposure can result in neuropathy, nephropathy, retinopathy, and microvascular disease [13]. Management of glucose with injections of insulin is further complicated by issues in numeracy to calculate precise doses for meal insulin and correction insulin [26], and natural variations in insulin sensitivity [27]. Normal physiological variations that affect insulin sensitivity include stress, illness, menstruation, pregnancy, sleep patterns, and exercise, all of which require modifications to insulin management. Parameters for insulin dosing can be adjusted by an endocrinologist and involves changing basal insulin doses, mealtime insulin dose settings (carb:ins), and correction insulin dose settings (CF); however provider adjustments may only occur every 12-24 weeks. The end result is a time-consuming, moving target for insulin therapy parameters (basal dosage, carb:ins ratio, correction factor). This dissertation contributes a new decision support system that provides weekly recommended adjustments in insulin therapy and dosing behaviors.

1.2.2 Hypoglycemia

Hypoglycemia is characterized by low plasma glucose. The occurrence of hypoglycemia in the context of existing diabetes is defined as blood glucose less than 70 mg/dL, and can present with neurogenic symptoms of sweating, shaking, palpitations, irritability and anxiety. Further decreases in plasma glucose lower than 70 mg/dL can deprive the brain of glucose, a condition called neuroglycopenia, and can present with confusion, neurological deficits, and seizures. While hypoglycemia can be attributed to a broad etiology, the main risks of hypoglycemia in people with diabetes are excessive and ill-timed insulin doses, decreased carbohydrate intake, and increased insulin sensitivity and glucose utilization.

A major target of insulin therapy algorithms is to predict, prevent and eliminate the occurrence of hypoglycemia. Chronic and repeated bouts of hypoglycemia are associated with autonomic failure and a blunted nervous system response to low glucose levels. Individuals with hypoglycemia unawareness may not experience traditional warning symptoms of neurogenic hypoglycemia, and therefore may not consume rescue carbohydrates to treat the hypoglycemia. The continued decline of glucose can potentially lead to loss of consciousness, coma, and death.

Recommended management of hypoglycemia includes consumption of 15g of rapid-acting carbohydrates every 15 minutes until glucose resolves to normal levels. In cases of severe hypoglycemia, injections of glucagon hormone are recommended. Glucagon is a hormone produced by pancreatic α-cells, and promotes hepatic glucose release and body-wide fatty acid metabolism. Glucagon is available in injectable syringe form, and also for use within dual-

hormone insulin pumps [28, 29]. This dissertation presents new algorithms to predict the occurrence of hypoglycemia during exercise, and an artificial intelligence decision support tool designed to reduce hypoglycemia through weekly adjustments to insulin therapy.

1.2.3 Exercise and type 1 diabetes

An important aspect of diabetes management is the incorporation of regular exercise into daily activities. However exercise is a major risk factor for hypoglycemia [30] and is oftentimes avoided, even though it has been shown to reduce the risk of cardiovascular disease in people with type 1 diabetes [31], and also improve glycemic control 24 hours following exercise in people with type 1 diabetes [32]. In the short- and long-term time-course following any bout of exercise, a diverse set of cellular mechanisms increase insulin sensitivity, and also facilitate secondary uptake of glucose into muscle tissue that is independent of insulin receptor activation [33]. These cellular processes induce drastic changes in glucose shortly following exercise and again in the 24 hours post-exercise [34]. In addition, there is considerable variability in changes in glucose that occur during exercise, both across people and in an individual's day-to-day activities. The fear of hypoglycemia is therefore a major barrier to incorporation of regular physical activity for people with type 1 diabetes [35]. In our recent survey of people living with T1D, we found 53% of responders were not confident in management of glucose during exercise, and there was considerable interest in a mobile phone application that can provide decision support in order to avoid hypoglycemia during exercise [36]. This dissertation explores new engineered methods to model the impact of aerobic exercise on glucose, in order to design new ways to prevent hypoglycemia during exercise.

1.2.4 Advancements in technology for type 1 diabetes

New technologies have been designed to improve the management of diabetes. Continuous glucose monitors (CGM) provide real-time measurements of glucose every 5 minutes and alert users to the occurrence of hypoglycemia or hyperglycemia thereby enabling people to treat with carbohydrates or insulin. This CGM data can be reviewed in order to inform users or providers of problematic glycemic patterns and make changes to insulin therapy. While use of CGM has been shown to improve %A1C by from 8.6% at baseline to 7.7% after 24 weeks in people with type 1 diabetes, it is not an adequate standalone therapy for everyone [37]. Continuous subcutaneous insulin infusion (CSII) by means of an insulin pump can deliver small amounts of fast-acting insulin every 5 minutes , and, when paired with CGM, has been shown to reduce %A1C from 8.3% at baseline to 7.5% after 1 year, as compared to a reduction to 8.1% in standard MDI therapy [38]. Continuous insulin infusion pumps and CGM systems can be further modified into automated insulin delivery (AID) systems. AID systems use control engineering algorithms [39, 40] to calculate and deliver precise doses of insulin in order to bring measured CGM glucose down to a target value in real-time. AID systems are designed to mimic *in vivo* pancreas insulin production and glucose sensing, and are aptly called artificial pancreas systems. AID systems have been shown to improve %A1C to 7.1% [41], and can also incorporate data about someone's physical activity in order to adjust insulin and glucagon doses [28, 42, 43]. For people who do not utilize insulin pumps,

new Bluetooth enabled insulin pens can record precise injection doses for MDI users in order to display insulin doses and glucose data together in a log book.

The ability of these devices to capture and log detailed data provides new opportunities for larger scale data collection. These large datasets enable data scientists to design new diabetes technologies and treatment algorithms. In this dissertation, I propose a new artificial intelligence system that utilizes CGM and insulin delivery devices to provide decision support for the management of type 1 diabetes.

1.2.5 Decision support for type 1 diabetes

Decision support systems for the management of type 1 diabetes are designed to provide recommendations for the adjustment therapies or behaviors, thereby bringing glucose within a recommended target range value [2]. As described in Tyler and Jacobs [44], these systems broadly fall into two categories: modifications to insulin dosing, and prediction of hypoglycemia. Systems have evolved considerably in the last four decades owing to the availability of commercial diabetes management technologies. Recently, the availability of larger data sets of time-series CGM data, physical activity and stress data, insulin dosage data, and contextual data, have improved our ability to design advanced control systems and artificial intelligence algorithms.

State of the art systems in decision support have addressed many aspects of diabetes management; algorithms have been developed to adjust basal insulin and bolus insulin therapy parameters [45-47]; predictive algorithms have forecasted the occurrence of hypoglycemia following meals [48] , overnight [49], and during exercise [50, 51]; artificial intelligence algorithms to detect insulin pump failure [52] or missing insulin doses associated with meals [53]; computer

vision to determine food content from meal photos [54]; and physiology models of glucose dynamics to simulate the outcomes of specific meals and insulin therapies [55-57]. Many of these systems are designed using insulin pump and CGM data and, as a result, will exhibit the optimal performance when evaluated in insulin pump users. However over 40% of people with type 1 diabetes in the United States use multiple daily injections (MDI) of insulin to manage their glucose [58], and cannot utilize the same control systems designed for insulin-pumps.

Automated DSS use artificial intelligence algorithms to analyze glycemic trends and generate therapy recommendations to people with T1D. Automated systems that are inclusive of MDI therapy represent a promising step in clinical care [45, 47, 59, 60], and have recently reported to improve HbA1C after 6 months of use [61]. In chapter 3 of this dissertation, a new artificial-intelligence decision support system is presented; this system provides insulin therapy and behavioral recommendations to people with T1D who use multiple daily injection therapy [62]. This decision support system demonstrates high agreement with endocrinologists' recommendations, and is shown to improve patient outcomes during *in silico* clinical trial evaluation and also in a preliminary 4-week human trial [62].

Prevention of hypoglycemia following exercise is a major aim of state of the art decision support designs. However, no commercial systems are marketed to assist people with type 1 diabetes to manage glucose during exercise. In this growing field of research, heuristic guidelines have been developed to recommend adjustments of insulin and carbohydrate intake [63], and are now being incorporated into a mobile phone application [64]. However these heuristic approaches require trial and error, and do not provide personalized predictions.

Artificial intelligence approaches have been designed to predict hypoglycemia in order to enable individuals to make treatment decisions; these include a decision tree algorithm [50], a logistic regression [45], and autoregressive models [65]. Still, more work is needed to personalize these population algorithms in order to better predict an individual's glycemic outcomes following exercise. In chapter 4 of this dissertation, a new adaptive artificial intelligence algorithm is introduced, and is personalized to better predict an individual's glucose changes during exercise. Chapter 4 also benchmarks published glucose forecasting algorithms using a common dataset. Because these algorithms were trained on data acquired under highly repeatable, well-controlled conditions, we can consider their performance to be an upper limit in terms of accuracy when evaluating on other real-world data sets. Additional algorithm training using real-world exercise data is forthcoming, and will enable predictive algorithms to better provide exercise decision support in diverse contexts.

1.3 Methods

This dissertation leverages artificial intelligence algorithms, mathematical models of glucose physiology, and Bayesian inference in order to (1) design and evaluate new decision support algorithms for management of insulin therapy in T1D, and (2) create new mathematical models describing changes in glucose during exercise.

1.3.1 Artificial intelligence models

Artificial intelligence is a broad field where machines are engineered to learn and perform complex tasks. Many applications in artificial intelligence are

designed to mimic human capabilities such as image recognition [66], however an even more powerful aspect of artificial intelligence is the ability to map uncharted medical quandaries. Machine learning is a branch of artificial intelligence in which algorithms are engineered to identify and learn complex patterns from data. For people living with type 1 diabetes, it can be difficult to predict if their glucose will become dangerously low after dosing insulin, during exercise, or overnight while they are sleeping. Machine learning models can address this uncertainty; these algorithms can identify the complex interplay of different environmental and physiological factors that impact glucose, and provide accurate estimates of glucose.

Machine learning models are data-driven, meaning that the algorithms require a *training dataset* that contains information related to the prediction task. For example, to design an algorithm that estimates the occurrence of hypoglycemia, the training dataset is comprised of input data to the algorithm that may include information about the most recent insulin doses, meals, glucose measurements, and also exercise sessions, in order to identify those important features and relationships that contribute to the occurrence of hypoglycemia. A caveat of this requirement is that an artificial intelligence model can only be expected to identify relationships present in the training dataset; if critical information is missing from the training dataset then it will also be missing from the artificial intelligence model. And, if error is present in the dataset (eg, misestimation of carbohydrate content in meals, or erroneous CGM data due to device connectivity issues), this can also impact the predictive performance of the artificial intelligence model.

To evaluate the performance of artificial intelligence models, the models are evaluated with a *testing set* which contains new examples of data that were

not used to train the algorithm, such as insulin doses, meal intake, glucose measurements, and exercise data from a new person. This type of evaluation is important, and reflects the generalizability of the artificial intelligence model.

There are many applications for machine learning in the field of diabetes. Artificial neural network models, support vector regression models, and autoregressive models [67] have all been designed to forecast future glucose levels. Models designed to predict hypoglycemia during exercise have also been designed and include a decision tree algorithm [50], a logistic regression [45], and autoregressive models [51, 65]. Another application in machine learning algorithms are the recommender systems that provide advice on insulin management [45-47, 62].

Artificial intelligence algorithms generally require considerable amounts of data in order to achieve accurate performance, and reflect the datasets from which they are trained. As such, many published algorithms in the field of diabetes are trained and evaluated using data collected during controlled, in-clinic study visits, and will likely perform poorly in scenarios outside of the clinical setting, including at-home scenarios. Furthermore, published systems are trained and evaluated on different datasets, making it difficult to perform a fair comparison. New datasets reflecting real-world living scenarios will soon be available for the development of more realistic machine learning algorithms, and can also provide opportunities for universal benchmarking approaches [68, 69]. This dissertation presents a new design paradigm for artificial intelligence systems; in chapter 3 of this dissertation, a virtual patient population that is composed of mathematical physiology models is used to generate realistic training data for the design of a k-nearest-neighbor (KNN) artificial intelligence model [70]. In addition, chapter 4 of this dissertation introduces a dataset that is

used to benchmark predictive algorithms to estimate changes in glucose during exercise. Chapter 4 also introduces a multivariate adaptive regression spline (MARS) model that is used to estimate the change in glucose during exercise, as well as describe how exercise history can be used to improve an individuals' predicted changes in glucose. A brief description of the MARS and the KNN algorithms used in this dissertation are introduced below.

1.3.1.1 Multivariate adaptive linear regression models

Linear regression is a powerful tool that is oftentimes used to investigate correlations and statistical significance between variables. Linear regression can also be used to design linear predictive models. The general linear regression form represents a predictive target (e.g., prediction of glucose 1 hour into the future) as a variable *Y*, with input data feature vector *X* (e.g., sex, age, insulin last dosed, last measured glucose level, etc.), and coefficients A, bias B, and random error ε.

$$Y = f(X) = AX + B + \varepsilon$$

Eq. 1.1 Linear regression

The machine learning model learns how to estimate glucose levels by analyzing the input features in X, and identifying numerical relationships between the predicted value y and the input features in X through the matrix A and bias B. For example, the matrix A might indicate that the last measured glucose value is 100 times more important than sex or age in order to predict hypoglycemia. Random observational error is represented by ε, and is assumed to be normally distributed noise that reflects variations in real-world data.

The goal of machine learning is to optimize coefficients in the matrix A in order to predict the target Y with high accuracy. The design of an accurate model requires a training dataset containing many example observations of glucose values at different times of day, from many different people, and under a variety of situations; *Y* therefore represents a column vector of N observations, and *X* represents a matrix of N observations and M unique variables. The bias term, B, is often included in *X* as a column of 1's, and *X* takes on dimensions of N by M+1.

Multivariate adaptive regression spline models [71] expand on the standard regression framework by taking into account ranges of the variable *X*. This approach enables non-linear mapping of variables within a linear regression framework. In a physiological application, this can be very useful; an individual's physiological state at a measured glucose of 70 mg/dL is likely different than at a glucose level of 350 mg/dL. Predictive models may be better tuned to reflect these varying physiologic states by enabling different ranges of glucose to yield different impacts on model predictions. The inputted features in *X* are first processed into paired hinge functions, φ. One of the hinge functions represents values of *X* above the hinge point, and the other hinge function represents` values of *X* below the hinge point.

$$\varphi_{upper} = \max(x - hingepoint, 0)\,, \qquad \varphi_{lower} = \max(hingepoint - x, 0)$$

Eq. 1.2 Hinge function

The *X* input features are modified to include the M original variables, and P hinge functions. The model bias can also be included as a column of 1's in the input matrix *X*. To solve for the coefficients in A, we used the analytic solution to

the weighted sum-of-squared-error cost function, Equation 1.3, where *X* is a matrix of N observations and M + P + 1 columns representing input features, hinge functions, and bias. *W* is a matrix of weights of dimensions N x N, and *Y* is a column containing the N target variables.

$$A = (X^T \cdot W \cdot X)^{-1} \cdot (X^T \cdot W \cdot Y)$$

Eq. 1.3 Sum of squared error solution

In a typical regression approach, the matrix *W* is set to the identity. In our weighted approach, hypoglycemia events are weighted by applying a constant C = 5 in the corresponding observation entry of *W*. The design and evaluation of this model is described further in chapter 4.

1.3.1.2 K-nearest-neighbors model

The k-nearest-neighbors algorithm (KNN) is a method for pattern classification [70]. In chapter 3 of this dissertation, the k-nearest-neighbors algorithm is used to analyze input features that represent an individual's weekly glycemic history in order to classify the individual as requiring "increases in insulin", "decreases in insulin" or "no change in insulin". KNN falls under the category of case-based-reasoning approaches in machine learning.

Unlike a linear regression which identifies a matrix A that represents the linear relationship between the input features *X* and the predicted output *Y*, the KNN algorithm is a nonlinear method for mapping the input features to a predicted output based on how similar the input features are to examples present inn a training dataset. To do this, the KNN algorithm uses the training dataset to define a look-up table database. Any new input observation, i.e. a new week of glycemic data, is compared to other examples of weekly glucose data in the look-up table. The glucose observation(s) from the look-up table that most closely

match the new glucose observation is presumed to be the best observation(s) on which to make inferences.

The k-nearest-neighbors decision support system (KNN-DSS) described in chapter 3 compares a week of glycemic outcomes against tens of thousands of existing glycemic profiles stored in a database look-up table. Each example observation in the KNN-DSS look-up table is composed of weekly glycemic features that are paired with a recommended change in basal or bolus insulin that will improve % time-in-range and % time-in-hypoglycemia. The KNN-DSS evaluates the new week of glycemic features, and locates similar examples of weekly glycemic features in the look-up table. The KNN-DSS then determines the most common recommended changes to insulin therapy among those similar examples.

$$\{d\} = \{\|x_i - X\|_2\}_{i=1}^{N}$$

Eq. 1.4 Euclidian distance

$$\{n\} = \{sort(d)\}_{i=1}^{K}$$

Eq. 1.5 Sorted Euclidian distance

In Equation 1.4, the mathematical Euclidian distance is calculated between the glycemic features of the new user, *X*, and all N glycemic examples in the look-up table, $x_1..x_N$. Then, in Equation 1.5, the examples are sorted by Euclidian distance to obtain the closest K neighbors in the look-up table. Distance and class-based weighting improve classification by increasing the weight of neighbors with more similar features

$$w_d = \frac{1}{|D_i|}, \qquad D_i = \frac{d_i}{d_{K+1}} \qquad for\ i = 1, \dots, n$$

Eq. 1.6 Distance-based weighting

$$w_c = \frac{1}{\left(Num(C_i^c)/\min_c\{Num(C_l^c)\middle| l = 1, \dots, K\}\right)^{1/2}}$$

Eq. 1.7 Class-based weighting

In Equation 1.6, those neighbors that are more similar to the new example than all other neighbors are normalized to the most distant neighbor and upweighted using distance-based weighting [72, 73]. In Equation 1.7, those neighbors that are less similar to the new example but are more numerous in the look-up table database are counted and scored lower through class-based weighting[74]. After scoring each neighbor by their class and distance, the KNN-DSS returns the recommendation most commonly found among those highly scored neighbors.

$$R_c = arg\max_c \left[\sum_{i=1}^{K} w_d w_c I_c\,(n_i = c)\right]$$

Eq. 1.8 K-nearest-neighbors classification

In Equation 1.8, R_c is the recommended adjustment to insulin dosage based on the neighboring examples in the dataset. The evaluation of this system is described further in chapter 3, and the recommendations generated by this system achieve high agreement with insulin adjustments suggested by endocrinologists.

1.3.2 Mathematical models of glucose physiology

Mathematical representations of glucose physiology have been under development for nearly three decades. Glucose levels in plasma change dynamically in response to circulating insulin, the absorption of nutrients during digestion, hepatic glucose storage or breakdown, changes in insulin sensitivity, and non-insulin-mediated glucose transfer or transport. The contributions of each

of these mechanisms have been quantified and represented mathematically in order to simulate glucose trends. Mathematical compartment models can be used to represent these physiological systems (Figure 1.1). In this approach, a compartment Q_1 can describe the amount of glucose in the plasma, and a second compartment Q_2 is defined to represent glucose in interstitial tissue over time. Arrows are drawn between compartments to represent the rate at which glucose moves from one compartment to another, and arrows can be drawn into or out of compartments to describe appearance and disposal of glucose (Figure 1.1). These compartment models are formally represented by systems of ordinary differential equations (see chapter 5). We have shown that mathematical compartment models of glucose physiology can be engineered to represent diverse patient populations and various insulin regimens in type 1 diabetes, and are referred to as *virtual patient populations* (VPP) [57].

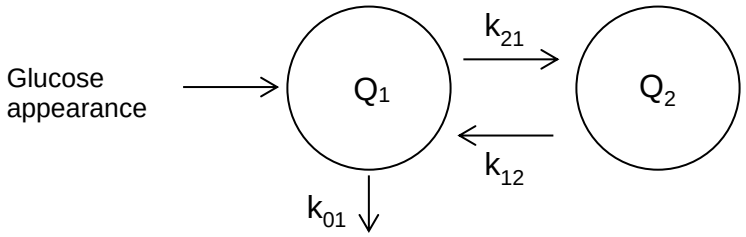


Figure 1.1: Compartment model describing glucose transfer and clearance.
Glucose resulting from carbohydrate digestion is inputted into compartment Q_1. Glucose moves between compartments at defined rates of k_{21} and k_{12}, and is disposed from the system at a rate of k_{01}.

Virtual patient populations serve many functions. Virtual patient models can be used to perform *in silico* clinical trials in order to evaluate artificial pancreas algorithms and decision support algorithms for insulin therapy adjustment. These virtual patient populations can mimic the results of clinical trials in humans [57], and have been accepted by the FDA to evaluate algorithms

prior to human clinical trials [75]. Another use of virtual patient populations is the design of model-predictive control (MPC) AID systems. MPC algorithms use the glucose physiology models to forecast glucose and calculate insulin doses necessary to bring the forecasted glucose to target [39, 40]. Yet another use of virtual patient models is retrospective simulation of glucose outcomes to determine the impact of modifying insulin dosage parameters [56], and as an educational tool that can be used to demonstrate how meal and insulin patterns affect glycemic control [55].

The Oregon Health and Science University Type 1 Diabetes Simulator (OHSU T1D Simulator) [57] models body-wide glucose as existing in either an observable state within plasma (Q_1), or in an unobserved state within cell bodies or tissue interstitium (Q_2) (Figure 1.2).

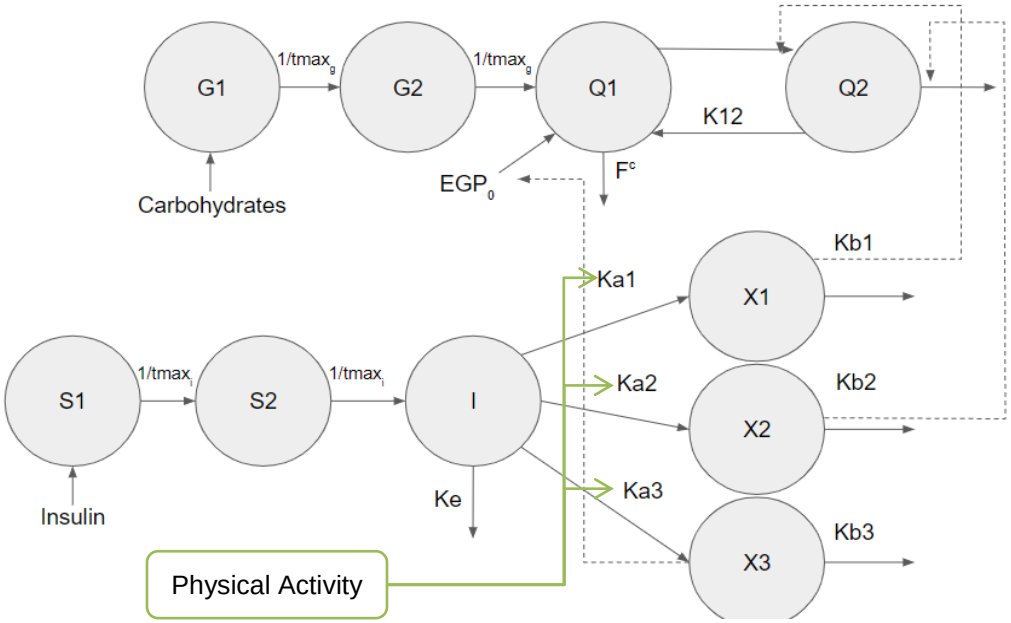


Figure 1.2: Mathematical compartment model of glucose physiology.

Glucose absorption and transfer can be modeled as a 10-compartment system representing digestion of carbohydrates (G_1, G_2), endogenous glucose production (EGP_0), transfer of glucose (Q_1 and Q_2), and insulin absorption (S_1, S_2, I) and targeted insulin action (X_1, X_2, X_3). Physical activity in the OHSU T1D simulator is currently modeled as impacting insulin appearance.

Glucose is modeled to enter the plasma through the digestion of carbohydrates in a two-step diffusion process (G_1 and G_2), and also through hepatic glucose production (EGP_0).

The impact of insulin is modeled as a three-function effect on glucose transport (X_1), glucose metabolism (X_2), and hepatic glucose production (X_3). The injection of exogenous insulin is modeled as a two-step diffusion process (S_1 and S_2) and eventual appearance in the plasma (I).

The OHSU T1D simulator models variations in insulin sensitivity parameters (ka_{1-3} and kb_{1-3}) to represent diverse weights and total daily insulin requirements, as well as circadian variations in insulin sensitivity, and the impact

of aerobic exercise on insulin sensitivity and glucose disposal. In the OHSU T1D simulator, estimated percent active muscle mass and are modeled to impact insulin appearance, endogenous glucose suppression, insulin-mediated transfer and clearance (Figure 1.2).

Glucose trends are approximated by physiology models using simulated or empirical data that are input into the model structures. These inputs are typically comprised of insulin syringe doses or insulin pump infusion rates, meal content and timing, and physical activity data obtained from accelerometer monitors. However, models can be designed to take into account new inputs. Figure 1.3 shows a comparison of the OHSU T1D model predictions as compared to real-world CGM data using empirical study inputs to estimate glucose.

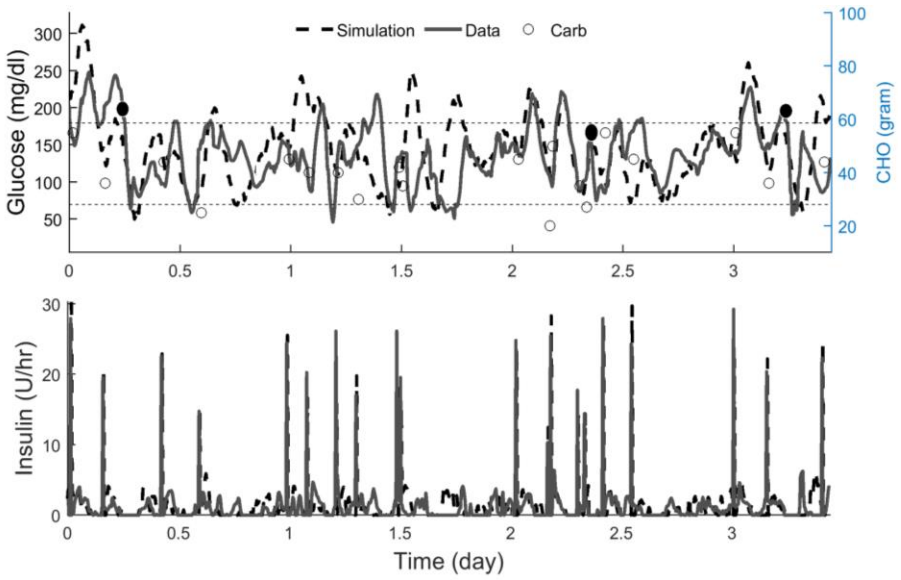


Figure 1.3: Physiologic model estimates of glucose data

Simulated vs. actual glucose and insulin profiles of one representative subject in a single-hormone artificial pancreas trial. Both experiments were initialized at 8:00 am. Carbs are shown with circles. Filled circles show the start of exercise. Figure and caption originally published in *Resalat, Navid, et al. "A statistical virtual patient population for the glucoregulatory system in type 1 diabetes with integrated exercise model." PloS one 14.7 (2019): e0217301.*

The difference between forecasted glucose values and the study data shown in Figure 1.3, referred to as the *residual error*, varies across different times of day. This error can be caused by many reasons; first, the model structure of glucose and insulin physiology may not account for all processes that affect glucose control; second, the parameters of the existing model structure may not be optimized to reflect an individual's data; third, the input data may be incomplete or corrupted.

The work presented in chapter 5 of this dissertation contributes new elements to the OHSU T1D virtual patient model that estimate the impact of aerobic exercise on insulin and non-insulin mediated glucose disposal, and

endogenous glucose production. The proposed models use data obtained directly from wearable physical activity monitors, and the model estimates track well with empirical glucose. Another contribution of this dissertation is presented in chapter 3, where we demonstrate that virtual patient simulators can be used to generate training data for the design of artificial intelligence model.

1.3.3 Bayesian inference for model parameter estimation

Artificial intelligence models and physiology compartment models must undergo parameter optimization in order to accurately represent the glucose trends. While the data-driven machine learning models introduced in this chapter are optimized using minimization of mean squared error, other methods must be used to identify unknown parameters of non-linear ordinary differential equations (ODEs) of glucose physiology, such as the OHSU T1D simulator. In this dissertation, Bayesian inference is used to optimize the parameters of the ODE models proposed in chapter 5 in order to accurately reflect the impact of aerobic exercise on endogenous glucose production and disposal. Bayesian inference can be used to generate point estimates of optimizable parameters, and the parameter distributions can also be used to make inferences on the model topology and the feasibility of the proposed model architecture and relationships.

In a Bayesian inference approach, the model parameters of the proposed ODEs are optimized using a probability framework. A joint probability distribution is defined, written as $p(y, x, \theta)$, that represents distribution of probable parameter values, θ, given the model structure and the empirical data with input x and predictor y.

While this probability distribution has a complex analytical form that, in many cases, may not be explicitly defined, the general form can be deconstructed using Bayes theorem of conditional probability [76].

$$p(y, x, \theta) = p(y, x|\theta) * p(\theta) = p(\theta|y, x) * p(y, x)$$

Eq. 1.9 Bayes conditional probability theorem

This theorem defines the general joint probability function as being equivalent to the likelihood of the data given the parameters,$p(y, x|\theta)$, and prior knowledge about the parameters, $p(\theta)$, the prior probability. This expression is also equivalent to the conditional probability of the parameters given the study data, $p(\theta|y, x) * p(y, x)$.

The posterior probability, $p(\theta|y, x)$, represents the probable parameter values given the study data, and is proportional to the likelihood of the data given the current model and our prior distribution of the parameters [76].

$$p(\theta|y, x) = \frac{p(y, x|\theta) * p(\theta)}{p(y, x)}$$

Eq. 1.10 Bayes conditional probability estimation

For complex ODE systems, these probability functions do not take on explicit formulations, and therefore $p(\theta|y, x)$ is not solved for analytically. Instead, Hamiltonian Monte Carlo sampling is a method that can be used to approximate the posterior probability of the parameters of interest. In this approach, parameter values are iteratively searched and evaluated [77]. Over time, this iterative search and sampling helps to map the posterior probability of the parameters. The posterior probability distributions of the parameters are then used to make inferences about the proposed model structures. The median value of the posterior distribution represents point estimates of model parameters.

This dissertation proposes new ODE models that describe the impact of aerobic exercise on glucose appearance and disposal, for use in the OHSU T1D virtual patient population. Bayesian inference is used to optimize model parameters to fit a new clinical dataset describing glucose trends during moderate and intense aerobic exercise, and various insulin clamp values. The proposed models are evaluated through analysis of the model-forecasted glucose, as well as the Bayesian estimates of parameter distributions. The models presented in chapter 5 track well with real-world glucose data and improve upon the existing OHSU T1D exercise simulator.

1.4 Contributions

This dissertation contributes new approaches to decision support for the management of type 1 diabetes. In chapter 2, we highlight that advances in artificial intelligence approaches to decision support largely consist of insulin therapy advisors and systems to predict hypoglycemia. State of the art systems for clinical decision support include artificial intelligence approaches to modify bolus insulin [46, 78] and basal insulin [79-81] therapy for both insulin pump and MDI users [47, 59]. Additional state of the art systems utilize physiological models of glucose metabolism to capture changes in insulin sensitivity and modify basal insulin therapy [45], and forecast glucose during daily activities and exercise [51]. The contributions of this dissertation to advance the field of decision support are described below.

1. The third chapter of this dissertation describes an artificial intelligence decision support system that is designed to assist people using multiple daily injection therapy to adjust their insulin therapy settings. State of the

art approaches to decision support require considerable amounts of real-world data in order train artificial intelligence algorithms, and many are not evaluated on real-world human data. In addition, many automated decision support systems do not provide recommendations regarding insulin dosing behaviors. The work presented in chapter 3 contributes a new design paradigm that uses mathematical models of physiology to simulate human data in order to train an artificial intelligence decision support system. This system is shown to deliver insulin and dosing behavior recommendations similar to those of physicians, improve glycemic outcomes *in silico*, and helps to reduce hypoglycemia in a 4-week proof-of-concept human study.

2. Chapter 4 of this dissertation explores the impact of aerobic exercise on blood glucose concentration. Changes in glucose during exercise are difficult to manage and difficult to predict. State of the art AID systems and decision support systems employ manual physical activity announcements to modify basal insulin and target glucose [42, 43], heuristic guidelines for modification of insulin doses or carbohydrate consumption [63, 64, 82], physiological models to recommend modifications to bolus insulin [83], and predictive hypoglycemia models [45, 50, 51, 65]. However the majority of these algorithms do not provide personalized recommendations or predictions, and it is not clear how effective these will be in real-world scenarios. Chapter 4 explores the repeatability of changes in glucose across identical exercise sessions, across participants and also across physical fitness levels. We contribute definitive findings that people with higher aerobic fitness will experience lower minimum glucose during and following exercise, than people with

lower aerobic fitness. Next, a new adaptive machine learning MARS model is designed to predict changes in glucose in aerobic exercise, and is personalized to participants. Lastly, chapter 4 contributes new perspective for AID development by benchmarking the predictive accuracy of algorithms designed to estimate glucose following exercise.

3. Chapter 5 of this dissertation contributes physiologic model topologies that can be used estimate glucose trends during exercise. Current state of the art virtual patient populations use varying architectures to describe the impact of aerobic exercise on insulin secretion, hepatic glucose production, and insulin-independent and insulin-mediated glucose disposal [84-86]. This dissertation proposes new models to estimate insulin- and non-insulin-mediated glucose disposal and endogenous glucose production. We show that the proposed models improve upon the existing OHSU T1D virtual patient simulator by simulating glucose trends that track well with changes in glucose during exercise in real-world human data. In addition, the proposed models use simple activity metrics obtained from wearable devices as inputs.

These approaches contribute to a growing field of computational algorithms that are designed to assist people to manage type 1 diabetes. And, these contributions may soon be expanded to Type 2 diabetes. The underlying design strategies described in this dissertation are likewise relevant to a broad field of health conditions.

2 Artificial Intelligence in Decision Support Systems for Type 1 Diabetes

Summary:

- This chapter is a comprehensive review of decision support approaches designed to improve the management of diabetes, with a focus on systems that use artificial intelligence frameworks.
- Current Artificial intelligence decision support systems are designed to help people with diabetes make adjustments to their insulin therapy, and also to avoid hypoglycemia.
- There is a need for systems that can assist people manage glucose during exercise, in the context of varying insulin loads and exercise types. In addition, systems currently do not exist to provide decision support to women with diabetes who are pregnant.
- Systems perform well *in silico*, but few have been evaluated in human trials. There is additionally a need for standardized big datasets that can be used to perform decision support system benchmarking.

This manuscript was published in Sensors in June 2020: Tyler NS, Jacobs PG. Artificial Intelligence in Decision Support Systems for Type 1 Diabetes. Sensors (Basel). 2020 Jun 5;20(11):3214. doi: 10.3390/s20113214. PMID: 32517068; PMCID: PMC7308977.

Abstract: Type 1 diabetes (T1D) is a chronic health condition resulting from pancreatic beta cell dysfunction and insulin depletion. While automated insulin delivery systems are

now available, many people choose to manage insulin delivery manually through insulin pumps or through multiple daily injections. Frequent insulin titrations are needed to adequately manage glucose, however, provider adjustments are typically made every several months. Recent automated decision support systems incorporate artificial intelligence algorithms to deliver personalized recommendations regarding insulin doses and daily behaviors. This paper presents a comprehensive review of computational and artificial intelligence-based decision support systems to manage T1D. Articles were obtained from PubMed, IEEE Xplore, and ScienceDirect databases. No time period restrictions were imposed on the search. After removing off-topic articles and duplicates, 562 articles were left to review. Of those articles, we identified 61 articles for comprehensive review based on algorithm evaluation using real-world human data, *in silico* trials, or clinical studies. We grouped decision support systems into general categories of 1) those which recommend adjustments to insulin and 2) those which predict and help avoid hypoglycemia. We review the artificial intelligence methods used for each type of decision support system, and discuss the performance and potential applications of these systems.

Keywords: Type 1 Diabetes; Decision Support; Artificial Intelligence; Insulin Advisor

2.1 Management of Type 1 Diabetes

Type 1 diabetes (T1D) is a medical condition caused by deficient insulin production and results in dysregulation of blood glucose. Maintaining blood glucose in a target range (70–180 mg/dL) can help to prevent complications related to hyperglycemia (>180 mg/dL) and hypoglycemia (<70 mg/dL), however, this is difficult to achieve for most people with T1D. Existing treatment strategies have evolved over the last 100

years from one-time daily insulin injections, into multiple injections of modern long-acting and rapid-acting insulin formulations, and the automated insulin delivery (AID) systems that are available today [41, 87]. While these AID systems are now commercially available, many people choose not to use them for various reasons including cost, inconvenience, issues with form factor, etc. People with T1D may instead prefer to manage their glucose using an insulin pump that delivers fast-acting insulin through continuous subcutaneous insulin infusion (CSII) prior to meals and continuously throughout the day. Or they may instead prefer to use multiple daily injection (MDI) therapy whereby they self-administer fast-acting insulin before meals and long-acting insulin once or twice per day using a needle syringe. MDI therapy continues to be the primary therapy for many people with T1D in the US and worldwide [58].

People who manage their glucose levels using either CSII or MDI therapy must navigate a complicated landscape of heuristic guidelines for the maintenance of basal insulin doses and meal and correction bolus doses. This can be challenging because there are a number of therapy parameters that impact insulin dosing: pre-prandial glucose level, the grams of carbohydrate that they will consume, their insulin sensitivity, their specific insulin-to-carbohydrate ratio, and the current insulin-on-board (IOB). People may also need to consider insulin variations that can occur throughout the day, their current glucose trend, and the activity context under which an insulin dose is being taken (e.g., prior to exercise, during an illness, etc.). This is particularly difficult for people using MDI therapy, as compared to a person using a pump with a bolus calculator, although more recent smart insulin pens have recently made bolus calculation possible for MDI users [88]. Current approaches at heuristic-based guidelines for patients to calculate their insulin dosing can be overwhelming for many people with T1D. In addition, some people may lack the numeracy skills necessary to accurately calculate their insulin prior to meals and throughout the day [26, 89]. Ahola et al. found

that 64% of patients miscalculate their prandial insulin need, often resulting in repeated hypoglycemia and hyperglycemia [90]. People may also struggle with the challenges of accounting for daily and seasonal variations in their insulin needs, and may wait 12–24 weeks to receive insulin dosage adjustments from their care providers.

The increasing ubiquity of health-based mobile computing and the growing usage of wireless continuous glucose monitors (CGM) [91] has created an opportunity for development of automated decision support systems (DSSs) for people with T1D. A growing number of mobile applications have been developed to provide people with diabetes access to on-demand decision support for their glucose management. Decision support can be provided to a person with T1D either through a health professional who provides the recommendations through a mobile interface, or alternatively directly to the patient using a smart-phone app that automatically generates the recommendations. While some smart-phone-based DSSs have been evaluated in clinical studies [6, 92], the outcomes have been mixed with some showing a benefit in terms of improving glycemic outcomes [93, 94], while others do not [95, 96]. The approaches that have demonstrated the best performance typically involved some form of direct contact between the person with T1D and their care provider [93, 94].

Artificial intelligence (AI) is the study of computational approaches that can be used to enable machines to perform intelligent problem solving and accomplish sophisticated tasks. Machine learning is a subset of AI and is specifically the study of how algorithms running on machines can learn and improve performance on a task through past experience and without being specifically instructed to do that task. AI applications that provide decision support to people with T1D have been developed [45, 47, 59, 62, 97]. These applications can provide far more frequent insulin dosage adjustment recommendations in between less frequent physician visits. These systems can provide personalized, on-demand insulin bolus calculators. When combined with

CGM sensors that provide real-time glucose estimations, AI-based DSS algorithms can also provide personalized hypoglycemia prediction and prevention. Recent studies have demonstrated that AI-recommended insulin dosage adjustments agree with physician opinion with a level of accuracy approaching that of inter-physician agreement [47, 62]. Other AI-based systems have demonstrated the ability to reduce hypoglycemia after short-term use [45, 47, 59, 62, 97]. In this manuscript, we review AI DSS frameworks that have been evaluated in clinical studies, *in silico* trials, and retrospective analysis of real-world human data. We discuss emerging trends and future opportunities in AI-based DSSs, which may hold promise for improving glycemic outcomes in people with T1D.

2.2 Common Objectives for Decision Support Systems for People with Type 1 Diabetes

DSSs are designed to help people better manage their diabetes by advising on medication management, alerting to medical complications, providing data visualization, simplifying carbohydrate counting, providing diabetes education, and helping to adjust daily behaviors and lifestyles. In this section, we introduce two types of DSSs: insulin therapy adjustments and hypoglycemia prevention. Adjustments to insulin therapy is one of the most common forms of AI-based decision support. For people using MDI therapy, a DSS can provide guidance on injections of long-acting insulin, and also provide guidance on injections of fast-acting insulin related to meals to help avoid hypo- and hyperglycemic excursions. For people using CSII pump therapy, a DSS can provide guidance on the basal rate of fast-acting insulin for different time windows during the day, as well as boluses related to meals or hypo and hyperglycemic excursions.

The amount of insulin dosed for a meal is proportional to the amount of meal carbohydrates, and is calculated by dividing the estimated carbohydrates by a

carbohydrate ratio (g/unit). However, it is quite common for people with T1D and even nutritionists to misestimate the amounts of carbohydrates in a meal, leading to suboptimal estimates of meal insulin. As such, some DSSs covered in this review provide guidance on carbohydrate estimation and also carbohydrate ratio adjustment.

When a person's glucose is too high, a correction dose of insulin is required. The amount of correction insulin is determined by (1) the person's current glucose level g(t), (2) their target glucose level g_T, (3) the current insulin in their body or IOB, and (4) the person's correction factor (mg/dL/unit). To calculate how much insulin should be dosed when glucose is higher than its target, Equation 2.1 can be used. Some of the DSSs described in this review provide guidance on correction bolus doses.

$$\text{Advised Correction Dose} = \frac{\mathrm{g(t)} - \mathrm{g_T}}{\text{correction factor}} - \mathrm{IOB}$$

Eq. 2.1 Correction insulin dose calculation

Hypoglycemia may be particularly challenging for people with T1D. DSSs can provide guidance to help people avoid hypoglycemia in general and also avoid exercise-induced hypoglycemia. Recently developed DSSs are oftentimes closely integrated with CGM sensors, which provide near real-time (typically every 5 min) estimates of interstitial glucose. As a result, these DSSs can utilize CGM to forecast glucose to help people with T1D avoid hypoglycemia. We will discuss both short-term glucose prediction algorithms (i.e., 30–60 min in the future), algorithms that predict glucose during and following exercise, and algorithms that predict glucose overnight, prior to bedtime when hypoglycemia can be particularly dangerous.

2.3 Common Outcome Measures for Assessing Performance of Decision Support Systems

The primary clinical objective of AI-based DSS is to improve glycemic outcomes and prevent medical complications related to diabetes such as hypoglycemia [6]. In this review we focus on AI-based DSS and the measures used to assess DSS performance including accuracy of the underlying algorithms and the clinical impact of using the algorithms.

2.3.1 Clinical Measures of Decision Support Systems

An effective DSS is one which can increase the percent of time that the person with T1D spends in a target glucose range or reduce the percent of time spent in hypoglycemia. The person's mean glucose is also a measure of glycemic performance whereby a lower mean glucose may mean that they have better glucose control. Hemoglobin A1C (HbA1c) is another commonly used measure that reflects a 3-month estimate of glucose control. Measured HbA1c less than 7% is considered to be within the target range for people with T1D, while values greater than 8.0% are considered very poor glucose control. In addition to these metrics, the high blood glucose index and low blood glucose index (LBGI) by Kovatchev and colleagues has been used to assess the performance of DSSs and AID systems [98].

The benefit of a DSS is typically measured either with respect to the person's own standard of care or in comparison with a control group population that has not received the DSS intervention. If a DSS is capable of increasing the amount of time (as a percentage of their day) that a person's glucose is within a target glucose range (70–180 mg/dL) from 50 to 60%, or reducing HbA1c from 8 to 7%, without impacting time in hypoglycemia, that would be considered an impressive performance. A typical

commercial AID is capable of increasing absolute time in target range by 5.0% for the Medtronic 670g [99] and 10% for the Tandem Control-IQ system [41], while reducing the percent time in hypoglycemia from 6.4% to 3.4% and 3.6% to 1.6%, respectively. It remains to be seen whether a DSS can match the level of performance achieved by AID systems.

2.3.2 Measures Used to Evaluate Accuracy of Decision Support Systems

While clinical outcomes can indicate the effectiveness of DSS use, other metrics may be used to assess whether a DSS recommendation is in agreement with recommendations provided by a physician. For a DSS that provides insulin dosage adjustments, the recommendations provided by the DSS can be compared with recommendations provided by a physician. The agreement with a physician can be reported as percent agreement of the DSS recommendations with physician recommendations. The Pearson's correlation coefficient of the provider recommendation with the DSS recommendation can also be used to evaluate the accuracy of a recommendation.

For DSSs that forecast glucose levels and provide advanced warning of impending hypoglycemia, other traditional metrics are used. A common set of accuracy metrics for algorithms that predict glucose in the future include the root mean squared error (RMSE) and the mean absolute percent error (MAPE). These are typically presented over different prediction horizons (e.g., 30 min to several hours). RMSE and MAPE are defined in equation 2 and equation 3, respectively. The true CGM or glucose value is Y, and the model forecasted value is $\hat{Y}$, and N is the number of predicted observations.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2}$$

Eq. 2.2 Root-mean-squared-error

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100$$

Eq. 2.3 Mean absolute percent error

For quantifying the accuracy of hypoglycemia prediction, traditional measures like sensitivity, specificity, and area under the curve are typically used. In addition, the Clarke error grid [100] and, more recently, the consensus error grid [101] are used to assess the clinical impact of predictions on decisions regarding insulin dosing or carbohydrate treatment. The Clarke error grid (Figure 2.1) is a plot of true or reference glucose values (x-axis) compared to predicted values (y-axis). The Clarke error grid is divided into clinically relevant regions of A, B, C, D, and E. Regions A and B are considered safe predictions, while C, D and E indicate potentially dangerous predictions where glucose was predicted to be higher than the true value, thereby missing hypoglycemic events or alternatively, glucose predictions were lower than the true value, causing unnecessary carbohydrate consumption. Commercial CGMs typically report 98% percent of values in the A and B regions with no values in the D or E regions as reported from various companies on the US Food and Drug Administration (FDA) web site [102] ("United States Food and Drug Administration. Available online: https://www.fda.gov/ ")

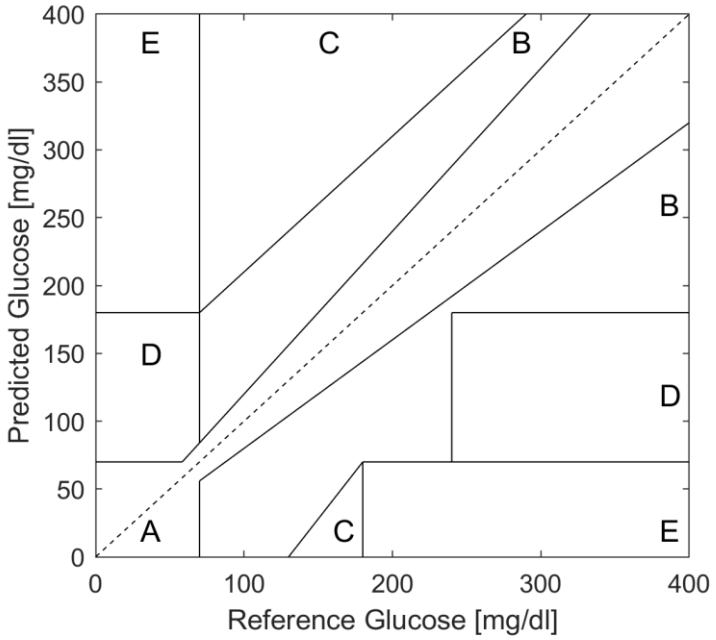


Figure 2.1: Clarke Error Grid.

Clarke error grid showing predicted (y-axis) vs. reference (x-axis) glucose. The dotted diagonal line shows perfect prediction of glucose. The A region is considered clinically accurate. The B region is considered a clinically safe region of prediction, though not accurate. The C, D and E regions are considered progressively more clinically dangerous regions of prediction.

2.4 Models and Simulations Used in T1D Decision Support Systems

2.4.1 Physical Models of Glucose-Insulin Dynamics Using Differential Equations

A certain class of the DSSs described in this review relies on physical models of glucose and insulin metabolism dynamics. Such physical models, described below, are comprised of linear and also non-linear differential equations that use compartment models to describe the digestion of carbohydrates, the subcutaneous absorption of

injected insulin, and the insulin action effects on glucose metabolism [103]. While these dynamic models can be used to forecast glycemic outcomes [45, 104, 105], many of these physical models have been developed to simulate the glucose dynamics of T1D populations [57, 75] and to design and evaluate DSS algorithms [62].

An interactive diabetes advisor (AIDA) simulator was developed in the early 1990s for use by diabetes educators, clinicians, and their patients [55]. The AIDA simulator consists of a four compartment model describing insulin absorption and elimination, insulin action, carbohydrate absorption, and plasma glucose response. The AIDA software is open-access with an online interface that allows users to input virtual patient features, meal pattern, and insulin doses. The simulator then returns the estimated glucose trends. This system is available as an educational tool [55].

Another physical model of glucose dynamics was published by Hovorka and Wilinska out of Cambridge [103, 106]. Hovorka and Wilinska described a non-linear eight-compartment model describing subcutaneous insulin absorption into plasma, the action of insulin on glucose uptake, disposal, and hepatic glucose production, prandial carbohydrate absorption, and plasma glucose levels. Later in 2010, Wilinska et al. [107] outlined an additional two compartment meal model describing the absorption of carbohydrates into plasma glucose, which became a part of the Cambridge Simulator.

The UVA-Padova simulator [75] is one of the most widely used T1D simulators and is described as being accepted as an evaluation tool for artificial pancreas algorithms by the FDA. This system includes multiple compartments to model carbohydrate absorption, glucose availability, insulin transport dynamics and body-wide insulin action, as well as body-wide glucagon production and action. This simulator also includes models for CGM measurement noise, insulin sensitivity variations, subject-specific rescue carbohydrate, and hypoglycemia unawareness.

Other authors have developed models specifically to describe the effects of physical activity on glucose. In 2009, Dalla Man et al. described an exercise model that utilized heart-rate as an input that impacts insulin-independent and insulin-dependent glucose uptake [85]. Roy and Parker modified the Bergman minimal model, a three-compartment model of insulin absorption and activity in plasma and glucose response, to include the effect of maximal oxygen consumption (percent VO2) on insulin excretion and hepatic glucose uptake [86].

Resalat et al. [57] modified the eight-compartment model designed by Hovorka and Wilinska et al. [103, 106], and created a stochastic virtual patient population by sampling from a distribution of possible insulin sensitivities. Resalat et al. [57] further incorporated an aerobic model of exercise developed by Hernandez–Ordonez et al. [84], describing the impact of physical activity on active muscle mass and insulin sensitivity into their open source simulator. The code for the Resalat et al. simulator is available on Github [108]. In addition to simulating the impact of aerobic exercise on glucose, the simulator includes circadian variations in insulin sensitivity, models for CGM noise, algorithms for simulating response to rescue carbohydrates, administering correction boluses, and dosing behaviors.

2.4.2 Data-Driven Models of Glucose-Insulin Dynamics

While physical models provide a physiologically realistic interpretation of the glucose dynamics, data-driven models have also been used to model glucose–insulin dynamics, especially for shorter horizon (e.g., 30 min) estimations of glucose.

Xie and Wang [109] developed an empirical non-linear autoregressive moving average model with exogenous inputs (NARMAX) that models the bimodal effects of exercise on glucose changes including short-term acute glucose changes and long-term insulin sensitivity changes. This model was trained and evaluated on the Dalla Man et al.

[85] model, but not on real-world glucose data. The Dalla Man et al. model is an updated version of the UVA-Padova model [75] that includes a model of physical activity. However, the NARMAX model was shown to effectively model the short and long-term impacts of simulated exercise, achieving a MAPE of 12%, and reported 87% of predictions in Zone A of the Clarke error grid. Further evaluation on real-world human data is needed.

Xie and Wang [67] further evaluated 11 different model approaches to predict glucose using prior glucose, insulin, carbohydrates, and exercise as inputs to the models. They compared less complex autoregressive regression models with exogenous inputs (ARX) with other machine-learning models such as support vector machines (SVM); ElasticNet; gradient-boosted trees; and deep learning models including recurrent neural networks (RNN), long-short-term-memory networks, and temporal convolution networks. While all models had fairly comparable performance in terms of RMSE for predicting short-term glucose dynamics within a 30-min prediction horizon, interestingly, the simpler ARX model had the lowest RMSE at 19.48 mg/dL. However, they also found that the ARX model was more sensitive to spurious noise and tended to under-predict peaks and over-predict minima.

Likewise, adaptive approaches and reinforcement learning are now being utilized to improve glucose predictions. Most recently, He et al. [110] reported that a RNN algorithm with an adaptive learning strategy achieved an 8.46 mg/dL RMSE when predicting glucose 30-min in the future. The evaluation set included people without diabetes, with T1D, and with type 2 diabetes, and further reporting of T1D-specific results are needed.

2.5 Early Approaches at Decision Support System Design

Researchers have been pioneering algorithms used to titrate insulin delivery since the early 1960s [111], however, these approaches were limited by lack of accurate glucose sensor data and short battery life of the devices running the algorithms [112]. Over the years, the algorithms developed from heuristic approaches, to model-based control system approaches, and AI approaches. The goal of the early phases of algorithm development was to build AID systems. More recently, AID and DSS algorithms have become sophisticated and reliable enough to become commercially viable. With the growth of new technologies including smart insulin pens, CGM devices, smart phones, health-related apps, and activity trackers, there is a growing opportunity for the development of DSS algorithms that has not previously been possible.

One of the first DSS algorithms developed was a piece-wise linear algorithm developed by Peterson, Jovanovic, and Chanoch [113, 114]. Their algorithm provided basal insulin recommendations using glucose as an input. The algorithm also provided adaptive bolus recommendations and enabled time-dependent dosage profiles. The algorithm was able to reduce the HbA1c in seven people with T1D by 1.6%, from 7.8% to 6.2% after 6 weeks of use, as compared to a 1.5% reduction, from 8.4% to 6.9%, in the standard of care group. One particular set of heuristics published in 1981 by Skylar et al. [115] became a basis for subsequent heuristic approaches to DSSs. These heuristic recommendations were integrated onto portable computers and the authors reported an average Pearson's correlation coefficient of 0.61 with physician recommendations on real-world human data [116]. The model was further modified and evaluated by Chiarelli and Albisser in a 24-week study in children with T1D. While hypoglycemia increased in both the experimental and standard of care group over the

course of the study, the experimental group ended with significantly lower hypoglycemia compared with standard of care at both the 16-week (−1.1%) and 24-week (−1.4%) time points. [117].

Around the same time these heuristic approaches were being developed and tested, model-based DSSs were being harnessed to predict user-specific glycemic responses to insulin regimens. These models were used to approximate a person's glucose based on a variety of inputs including meals, insulin, and even exercise. In the early 90s, a descriptive study of the Glucoject model was published, which modeled the insulin plasma of different injected formulations for the purposes of treatment replay [118]. Another model approach described by Hauser et al. [119] was developed as an educational tool to help users see the effect of insulin or exercise on their glucose regimens. The authors compared the model predictions of glucose trends to physician-predicted glucose trends, and reported a Pearson's correlation coefficient of 0.97.

The Karlsburg Diabetes Management System (KADIS) was a model-based DSS that utilized user-entered carbohydrate, insulin, and exercise inputs. Exercise physiology models were not available at the time, and the KADIS algorithm modeled exercise in terms of equivalent insulin units. After fitting a user-specific model from fingerstick glucose, insulin, carbohydrate and exercise data, the model was then used to simulate anticipated meals, planned insulin doses and exercise [120, 121]. In 1994, it was marketed as an educational tool for people with T1D and for providers who were less experienced with diabetes management [122]. By 2007, the algorithm had evolved to utilize CGM technologies, and was evaluated in 2007 and 2010 in adults with T1D and T2D. While KADIS-augmented physician support was shown to significantly decrease HbA1c from 7.10 to 6.73% ($p < 0.01$, where p is the statistical significance of a hypothesis test) regardless of diabetes type, the system is designed to be utilized by providers once per year to help them adjust their patients' insulin during visits [123].

DIABETEX was a model-based approach that utilized Bayes decision theory to determine short-acting insulin recommendations. The DIABETEX system was tested on data from 12 adults with T1D and demonstrated a 66% full agreement with endocrinologist-recommended adjustments for meal-related doses, and a 42% full agreement for basal-related doses [124]. The DIABETEX program was later evaluated in children and was reported to improve HbA1c from 5.85% to 5.0%, and reduced the number of measured hypoglycemic events in children from 143 to 58 during a 12-month evaluation period [125].

In the late 1990s, nascent AI and data-driven algorithms were beginning to emerge in the field of diabetes DSSs. Early artificial neural network algorithms used patient demographics and glycemic targets to predict an optimal insulin therapy regimen. Authors reported that the classifier was able to classify 92% of insulin regimens in agreement with health care providers [126, 127].

DIAS was a model-based DSS platform pioneered by Carson et al. that utilized probabilistic causal-networks to predict the 24-h glucose profiles for adults with T1D, and then performed model-based replay to determine insulin dosage and carbohydrate intake strategies to minimize hypoglycemia and maximize time-in-range [128]. Evaluation in 20 adults with T1D whereby users were given recommendations by the DIAS system demonstrated equivalent clinical efficacy as compared to recommendations delivered by a diabetes specialist nurse [129].

All of these early algorithms were developed prior to the age of mobile computing. As a result, use of these algorithms required extensive user foresight and planning because they relied on the person with T1D to access a personal computer. While many of these methods proved impractical for real-time DSSs, they set the stage for the explosion of algorithms that occurred when accurate CGM technologies became FDA approved and when mobile computing became ubiquitous

2.6 More Recent Decision Support Systems

Mobile DSSs are now becoming available for use by people with T1D who use either CSII or MDI and can potentially provide improvements in glycemic outcomes. Existing automated DSSs are designed to provide recommendations to people with T1D regarding insulin doses, anticipated hypoglycemia, and modifications to daily behaviors that may improve their glycemic outcomes. Recent publications on DSS algorithms described advanced control system and traditional AI approaches to deliver personalized recommendations to people with T1D, and have shown promise in providing recommendations that agree with physician opinion. We categorized DSSs into those that provide advice on insulin adjustment and those that provide guidance on hypoglycemia prediction and prevention.

2.6.1 Decision Support Systems for Adjustment of Insulin Therapy

A number of DSSs have recently been developed that may be used to recommend changes to insulin dosing. They generally fall into the category of (1) physiologic model-based algorithms; (2) clustering, or case-based algorithms; (3) heuristic rule-based algorithm; (4) data driven model-based algorithms; and (5) carbohydrate estimations and meal detection algorithms.

2.6.1.1 Physiologic Model-Based Algorithms

In 2008, Palerm et al. [79] described automated titration of basal insulin pump infusion rates using run-2-run. A run-to-run methodology implies that a model will adapt specific parameters after a run, whereby a run could be considered a day, a meal, an exercise period, an overnight period, etc. If hypoglycemia is observed after one or more

runs, the model will adapt and adjust the model parameters to improve future recommendations. Herrero et al.[80] and Toffanin et al.[81] also evaluated run-2-run for adaptation of basal insulin rates, showing an improvement in time in range of 20%[81] and 28% [80], during *in silico* evaluation in adults in the UVA-Padova simulator. Their run-2-run algorithm was additionally utilized to modify insulin dosage settings. In a study by Zesser et al. in adult participants with T1D, run-2-run bolus adaptation was shown to improve glycemic excursions in the 2 h following meals from 149.1 mg/dL down to 109.4 mg/dL after algorithm convergence [78]. Herrero et al. also utilized run-2-run insulin bolus adaptation, and demonstrated a reduction in hypoglycemia in an *in silico* study using 10 adults from the UVA-Padova simulator [46].

Physiological models provide a natural basis for insulin dosage adjustment by allowing glucose replay to optimize bolus strategies. In 2008, Wong et al. described a model-based approach for modification of a two-bolus regimen through iterative insulin sensitivity estimations. The authors evaluated this method in an *in silico* trial using the AIDA simulator and reported a 58%–91% reduction in hypoglycemia with use of the adaptive meal bolus calculator, as compared to standard bolus calculator use [130]. Rosales and Garelli described a model-based approach to a meal bolus DSS whereby constrained optimization is used to calculate bolus amount and basal rate during postprandial periods. When evaluated *in silico* using the UVA-Padova simulator, this approach improved time-in-range from 81.9 to 89.5%, and reduced hypoglycemia from 5.92 to 0.97%, as compared to standard bolus calculator use [131]. Revert et al. also described a method to optimize bolus and basal rate adjustment for different meals using interval analysis [132]. Rosetti et al. evaluated elements of this approach in a clinical study of 12 adults with T1D, and demonstrated that use of the algorithm nominally reduced postprandial area under the glucose curve in the 5 h following a 40 g meal by 103.6 mg*hr/dL, as compared to a standard bolus calculator. However, these

results did not translate to a 100 g meal and, therefore, did not achieve statistical significance [133].

Breton et al. also developed a DSS for people utilizing MDI and CSII therapy. This comprehensive DSS included three main algorithms: 1) a model-based insulin-replay for basal recommendations, 2) a Kalman-filter approach for estimation of insulin sensitivity and real-time bolus recommendations, and 3) a logistic regression hypoglycemia prediction algorithm for exercise decision support. During a cross-over study in 24 adults with T1D, participants underwent one 48-h in-patient session per study arm, with standardized meals and aerobic exercise. The authors reported a statistically significant reduction of percent time-in-hypoglycemia following use of the DSS (3.2% [1.3, 4.8] control vs. 0.9% [0.4 2.3] DSS) but did not report significant changes in percent time-in-range [45].

Most recently, Goodwin et al. [104] reported the results of a model based approach, whereby an individual's glucose dynamics are modeled by fitting a simple glucose model to at-home data to create a digital twin. This personalized model is then modified using stochastic error and disturbances to form an envelope of models. Next, this envelope of models is used to estimate the best bolus shape to be delivered, with specific focus on whether dual-wave bolus, a split bolus, or single bolus should be utilized. This proof of concept study was trained on real-world data from 12 adults, and evaluated 2 years later on two subjects. The results indicated that 54%–74% of the real-world data was captured by enveloped predictions. This approach required considerable amounts of subject data with prescribed meal scenarios to train the algorithm, and further evaluation of glycemic outcomes is needed to determine the generalizability of this approach.

2.6.1.2 Clustering Algorithms

Clustering algorithms are a powerful way to identify glycemic patterns, or groups of similar people with T1D in order to provide user-specific recommendations and glucose predictions. Case-based-reasoning (CBR) DSS algorithms use a database of cases (or examples) of glycemic responses corresponding to specific insulin, meal, exercise, or other inputs. If a new set of observations matches a given case from the database, then the recommendation corresponding to the best-matched case in the database is returned by the algorithm. New cases are added to the database as they are obtained. The k-nearest neighbors approach [70] is a common way of implementing a CBR DSS.

An early CBR-DSS was described in 2002 by Bellazzi et al. [134] as part of the T-IDDM telehealth system. This system was designed to aggregate data from people with T1D, and provide recommendations to physicians who were helping with glucose management. The CBR approach was modified in 2008 by Schwartz et al. [135], who designed a 6-week clinical study in 20 adults with T1D for the purposes of data collection and development of a case-base to be used by physicians. Later in 2010, this system was introduced as the 4 Diabetes Support System [136], but while the authors reported metrics regarding problem identification, we did not find published participant glycemic outcomes.

Herrero et al. [46] described a CBR DSS (named ABC4D) that used an adaptive run-2-run algorithm for real-time insulin bolus support. The ABC4D system was evaluated *in silico* using the UVA-Padova simulator, and demonstrated improved percent time-in-range in 10 virtual adults from 75.2 ± 11.7% to 81.9 ±13.4% ($p < 0.05$) and a reduction in percent time-in-hypoglycemia from 0.3 ± 0.5% to 0% ($p = 0.17$) after 4 weeks of use. Reddy et al. [59] further evaluated the ABC4D algorithm in a real-world, 6-

week clinical study. The authors reported non-significant improvements in percent time-in-range from 55.0% to 60.9%, and also a non-significant reduction in percent time-in-hypoglycemia from 5.0 to 3.6%.

Soon after Reddy et al.'s ABC4D clinical study was published, other approaches to CBR with slight modifications to optimize the adaptive nature of the algorithm were proposed. Similar to the method proposed by Herrero, Torrent-Fontbona and Lopez [137] incorporated "concept drift" into their CBR, effectively replacing old examples in the case-base with newer examples that better reflect the user's physiologic state. The authors evaluated the algorithm *in silico* using 11 adults from the UVA-Padova simulator, and reported a time-in-range of 84.0% after 90-days of use with optimal basal settings.

More recently, Tyler and Jacobs [62] published a paper describing a k-nearest neighbors DSS (KNN-DSS) recommender engine that was trained on the OHSU *in silico* simulator to provide weekly recommendations for updating carbohydrate ratios, correction factors, and basal rates for people with T1D using MDI therapy. The recommendations provided by the KNN-DSS agreed with board-certified endocrinologists 67.9% of the time, which was found to be comparable with inter-physician agreement. When evaluated *in silico*, percent time in target range increased from 59.5% to 79.8% while maintaining hypoglycemia less than 2%. When evaluated in a small feasibility study on 16 people with T1D whereby physicians provided the recommendations from the engine on a weekly basis to the study participants across 4 weeks, a statistically significant reduction of hypoglycemia events by 43% overnight ($p = 0.04$) and a 25% decrease in hypoglycemia events overall ($p = 0.051$) was observed from the first week of the study to the final week of the study. This KNN-DSS is now being evaluated in a larger clinical trial using a mobile app called DailyDose.

A new system called PEPPER has been proposed by Liu and Herrero [60]. The PEPPER system was designed to analyze user data to deliver real-time hypoglycemia

alerts, CBR bolus recommendations, predicted low glucose insulin suspension, and carbohydrate consumption recommendations. In this system, the run-2-run algorithm, which is typically used to recommend new insulin settings, is instead utilized to modify a carbohydrate sensitivity factor for rescue carb recommendations. Six adult participants took part in an 8-week study to evaluate the predictive-low alert system and carbohydrate recommendation components of the PEPPER system, with regular physician decision-support for insulin dosage adjustments. The study found that use of PEPPER resulted in significant decreases in time below 60 mg/dL from 1.8% to 0.7% ($p = 0.05$), and improvements in % time-in-range from 52.8% to 61.3% ($p = 0.02$).

Biagi et al. also described a compositional data analysis k-means clustering algorithm in order to group 24-h glucose profiles. The algorithm was trained using data from six adults with T1D undergoing CSII treatment over 8 weeks. The authors showed that the algorithm returned profile clusters exhibiting high variability, high hypoglycemia, high hyperglycemia, and adequate control. While this algorithm is in the preliminary stages, the approach may be used to categorize glucose profiles for specific activities or behaviors. [138].

2.6.1.3 Rule-Based Algorithms

Nimri et al. [47] evaluated an AI fuzzy-logic system to adjust insulin pump settings, including basal rate, carbohydrate ratio, and correction factor. The insulin dosage adjustments achieved an agreement with physician recommendations that was comparable to inter-physician agreement. The system was evaluated in a 12-week clinical study whereby adults with T1D underwent engine-augmented physician decision support. The authors reported a reduction in hypoglycemia and improvement in time-in-range after 12 weeks of use, but no significance measures were reported [47].

A more discrete rule-based algorithm called VoiceDiab was proposed by Pankawska et al. [139]. This system utilizes voice recognition for meal entries, and given these verbal recordings of meals, provides estimated nutritional meal content and ultimately recommendations for insulin bolus size and amounts given across time. When evaluated in 12 subjects with T1D during a cross-over study comparing standard bolusing with VoiceDiab bolusing, postprandial glucose excursions were reduced but an increase in hypoglycemia in many subjects was also observed. While this was a heuristic algorithm, this group brings to light two critical areas of consideration for DSSs: exploitation of existing nutritional databases for meal content estimation, and bolus shape optimization.

2.6.1.4 Other AI Algorithms

Sun et al. [140] described an advisory system for MDI users called ABBA. This system utilizes actor-critic reinforcement learning to retrospectively analyze user data, and supply daily basal insulin dosage suggestions. This system was designed to utilize CGM or self-monitoring blood glucose data (from finger-stick glucose meters) as inputs. An *in silico* study using 100 adult subjects from the UVA-simulator demonstrated reduction in hypoglycemia from 2.5% to 1.0% while maintaining % time-in-range above 85% after 13 weeks of use.

Srinivasan et al. [141] developed a particle swarm optimization method to determine the optimal meal bolus timing and bolus shapes for meals of different sizes and carbohydrate and fat content using *in silico* data, and devised a set of heuristics to use for insulin bolusing; however, this method has not been tested yet in humans.

2.6.2 Decision Support Systems for Carbohydrate Estimations and Meal Detection

2.6.2.1 Computer Vision Algorithms

One of the challenges for people with T1D when dosing meal insulin is estimating the correct amount of carbohydrates within any given meal. While these methods are in their early stages of development, several groups have attempted to leverage smart phone cameras to attempt to automate carbohydrate estimations for people with T1D. The GoCARB system by Anthimopoulos and Mougiakakou [142] used image processing and they trained machine learning algorithms on photographs of food to analyze meal content and provide carbohydrate estimations. The system demonstrated a MAPE of 10% in carbohydrate estimation when evaluated on test images, not used in the training. However, this study was done using a closed set of only 24 meals, with prescribed lighting conditions, making it less applicable to real-world usage. When evaluated by Vasiloglou et al. [54] on a larger set of 54 prescribed meals that again followed specific formulations (e.g., three meal types of food per plate), the system was able to achieve an accuracy of 14.8 g while the estimation by nutritionists was comparable at 14.9 g

2.6.2.2 Physiologic Model-Based Algorithms

Many people with T1D fail to announce their meals to DSSs or AID systems and may also simply forget to bolus insulin prior to a meal. Several groups have attempted to utilize CGM patterns to detect meals for the purpose of alerting the person with T1D and reminding them to deliver insulin in case they forget to do so.

Mahmoudi et al. [143] developed a meal-detection algorithm utilizing a Kalman filter, and evaluated it in the UVA-Padova *in silico* simulator. When their algorithm detected a meal, they used a bolus calculator to administer the insulin determined by the

bolus calculator to the virtual subject. In the UVA-Padova simulation, it was demonstrated that the algorithm required 40 min for meal detection, and use of the algorithm improved time-in-range from 53% to 83% for virtual subjects as compared to no meal announcement. This algorithm needs to be evaluated on real-world data. In addition, while this approached utilized an MPC algorithm for use in an AID, evaluation of this meal-detection algorithm and subsequent bolusing may eventually be evaluated for CSII and MDI subjects for use in decision support.

2.6.2.3 Rule-Based Algorithms

Samadi and Cinar also reported on the design of a fuzzy logic estimated controller for unannounced meal-detection. This algorithm utilized glucose trends and insulin dosed to perform shape identification of glycemic profiles and meal content estimation, resulting in an 87% sensitivity when evaluated *in silico*, and a 93% sensitivity when evaluated on human clinical data, with a mean time to meal detection of 34.8 min [53]. Though these systems will likely exhibit the best performance within AID or CSII systems where insulin data can be easily tracked, they may also be integrated into MDI-based DSSs to help with meal reminders and carbohydrate estimation.

2.6.3 Decision Support Systems for Hypoglycemia Prediction

While insulin dose adjustment is a critical component of decision support, people with T1D also need to be notified about impending acute glucose changes that could lead to dangerous hypo- or hyperglycemia. Hypoglycemia, if left untreated, can cause coma or death, and even a small number of exposures to extreme hypoglycemic episodes can lead to long term damage to the brain and the heart [144]. In this section, we describe some approaches to preventing hypoglycemia using AI approaches.

Hypoglycemia prediction is largely accomplished through modeling of glucose and CGM trends. Prediction algorithms may be augmented with additional inputs such as insulin data, meal information if available, and physical activity that is either announced or available from fitness tracking devices. An effective DSS would be able to accurately anticipate low glucose levels and notify or alert the person with T1D in advance such that hypoglycemia can be avoided. If the person is using an insulin pump, these predictive algorithms can trigger the insulin pump to stop delivering insulin.

2.6.3.1 Physiologic Model-Based Algorithms

One Kalman-filter-based approach described by Cameron and Bequette uses 30–70 min prediction horizons. The algorithm has been used to shut off basal insulin if the blood glucose is forecasted to go below 80 mg/dL [105]. This algorithm was evaluated extensively in subsequent clinical trials, and was shown by Calhoun et al. in youth with T1D, and Buckingham et al. in adults with T1D to reduce the number of nights with hypoglycemic events by upwards of 25% in adults [145] and youth [146]. However the overall reported time in hypoglycemia was not reported or found to be significant. By 2017, the algorithm was commercially available in Medtronic pumps under the name SmartGuard and was shown to significantly reduce frequency and duration of hypoglycemia in pediatric participants [147].

2.6.3.2 Data-Driven Algorithms

In early 2007, Sparacino and Cobelli [148] presented a hypoglycemia prediction model using polynomial models and single-order autoregressive models to predict hypoglycemia with a 30 min prediction horizon window. The models achieved an RMSE of 17–18 mg/dL. In 2010, Perez Gandia et al. [149] further reported on the predictive accuracy of the neural network approach and reported a 17–20 mg/dL RMSE on a 30-

min prediction horizon when evaluated on real-world data. Zecchin et al. likewise utilized a neural network prediction strategy and showed through *in silico* evaluation that hypoglycemia could be significantly reduced through alert-based carbohydrate treatments triggered by the hypoglycemia prediction algorithm [150].

In 2013, Daskalaki and Mougiakakou [151] compared an ARX algorithm, a RNN algorithm, and a fusion approach whereby outputs of both the ARX and RNN were used to improve prediction accuracy. When evaluated on clinical trial data gathered from adults with T1D, the authors reported an RMSE of 18.9 mg/dL using the RNN. The authors also developed an early warning system for hypoglycemia and reported a 100% sensitivity to hypoglycemic events, with a 16.7 min predictive horizon, but with a false alarm rate of 0.8 per day.

2.6.3.3 Clustering Algorithms

Clustering algorithms, discussed above for CBR approaches to insulin titration, have also been used to improve glucose predictions by Contreras et al. [152]. Specific glycemic profiles were grouped using normalized compression distance clustering, and then cluster-specific grammatical evolution reinforcement learning models were trained. The authors found that the model achieved an RMSE of 4.27 mg/dL for 60-min predictions using *in silico* data; however, further evaluation on real-world human data is needed as simulator predictions are notoriously far more accurate than those on human data under free-living conditions.

2.6.4 Postprandial Hypoglycemia Avoidance

Recently, groups have developed AI approaches for the purpose of postprandial hypoglycemia avoidance. Montaser et al. [48] developed a seasonal autoregressive integrative moving average with exogenous inputs model of glucose dynamics during the

postprandial period. To train the model, data from real-world human subjects were clustered based on exercise type and postprandial glycemic response. CGM data, insulin data, and energy expenditure data were used as inputs to the model. The error reported for the model across all datasets was 6.29 mg/dL for 30 min prediction horizons after the start of meals. It is important to consider that this model was designed and evaluated only on post-meal windows of data. Additionally, it was evaluated on a data set whereby study participants all performed the same exercise and maintained a consistent eating schedule while using closed-loop insulin therapy to manage their glucose levels.

Toffanin et al. [153] studied how data-driven multiple-model predictor (MMP) frameworks could predict post-prandial glycemic patterns across multiple time horizons. While the MMP model predictions were reported to correlate well with human data, the greatest accuracy was observed with morning postprandial glucose responses to meals. This system has not yet been utilized to provide decision support with regards to bolusing.

Oveido et al. [154] developed a predictive model of postprandial hypoglycemia using real-world data from adults with T1D. The authors trained an SVM and demonstrated a 71% specificity for prediction of hypoglycemia in the 4 h period following meals. *In silico*, Oveido and Vehi [155] demonstrated that these predictive measures could be implemented in a bolus-reduction calculator, and significantly reduced hypoglycemia from 7.6% to 4.68% in the UVA-Padova simulator. Further evaluation using human data has not yet been performed.

Cappon et al. described an extreme gradient-based tree algorithm to predict postprandial glycemic outcomes and modify insulin boluses. The algorithm utilized CGM data, carbohydrate data and insulin data and trained an extreme gradient-based tree algorithm to predict three conditions 6 h after meal consumption: hyperglycemia, in

target range, and hypoglycemia. The algorithm was trained and evaluated using UVA-Padova simulated data, and demonstrated a 97% area under the ROC curve accuracy for predicting hypoglycemia following meals. It was demonstrated also within the simulator that use of an algorithm to modify boluses based on predicted probability of hypoglycemia improved time-in-range from 62% to 67%, but no significant reduction in hypoglycemia with use of the bolus modification. This method has not been validated on human data [156].

2.6.5 Nocturnal Hypoglycemia Prediction

Nocturnal hypoglycemia is a dangerous complication of diabetes. Severe episodes may cause coma or death. Prediction and prevention of nocturnal hypoglycemia following normal activity or high physical activity days is a critical area of DSSs.

2.6.5.1 Linear-Regression Algorithms

One of the earliest algorithms was described by Schiffrin et al. [157] before CGM was available to improve the accuracy of the prediction. Schiffrin et al. utilized linear regressions and logistic regressions to define the probability of hypoglycemia and hyperglycemia given the users glucose at bedtime, and defined a heuristic rule for consuming a carbohydrate before bed if glucose is less than 120 mg/dL. Use of this heuristic was found to reduce the incidence of nighttime hypoglycemia from 13% in control group, down to 0% in the heuristic intervention group.

2.6.5.2 Support Vector and Other Data-Driven Algorithms

Mosquera–Lopez and Jacobs [49] developed a nocturnal hypoglycemia algorithm trained using CSII data and trained a support vector regression algorithm using data

from 124 people (22,804 nights) with T1D from the Tidepool Big Data Donation Dataset [68]. When validated on data from 10 people with T1D on CSII pump therapy across 4 weeks of free-living, and utilizing an announced bedtime of 11 pm and the preceding 15 h of data, the algorithm reported a 94% sensitivity and 72% specificity for predicting nocturnal hypoglycemia. The group also reported a decision-theory approach for selecting the threshold at which to predict low glucose. Similar to Schiffrin et al., they found that adherence to a simple heuristic metric of consuming a carbohydrate if bedtime glucose was less than 149 mg/dL could achieve a 94.1% sensitivity but a lower specificity of 61% specificity for predicted nocturnal hypoglycemia.

Guemes and Herrero [158] utilized an available dataset, OhioT1DM, to compare different machine-learning algorithms to predict the glycemic status of an individual prior to going to bed. After processing input data of CGM, insulin dosed, and self-reported meals, the group developed three classifiers to predict hypoglycemic events (<70 mg/dL), hyperglycemia (>180 mg/dL), and within target range (70–180 mg/dL). The group compared Random forest, artificial neural network, vector machines, linear logistic regressions, and extended tree classifiers. Using an announced bedtime of 11 pm and the preceding 18 h of glycemic data, the authors reported that the SVM achieved the best results, with a sensitivity of 68%, and specificity 71% during cross validation. Hyperglycemia prediction resulted in a 59% sensitivity and 65% sensitivity.

Vehi et al. [159] developed a bimodal hypoglycemia prediction algorithm. An SVM was used to predict postprandial hypoglycemia, while an artificial neural network was developed to predict nocturnal hypoglycemia. The group utilized three databases to train and validate their data, two databases containing data from adults with T1D on insulin pumps, and an additional *in silico* dataset from the UVA-Padova dataset. The model achieved a 44% sensitivity for predicted nocturnal hypoglycemic events on real-world human data, using a 6 hr prediction window.

Bertachi et al. [160] developed a nocturnal hypoglycemia algorithm for MDI utilizing an SVM and multilayer perceptron neural network, trained on 10 adults during a 12-week at-home study. Using data from the 6 hrs preceding sleep, the algorithm achieved a sensitivity of 78% and specificity of 82% for nocturnal hypoglycemia.

2.6.6 Exercise-Induced Hypoglycemia Prediction

Exercise is known to substantially impact glucose levels in people with T1D as covered extensively by Riddell et al. [63]. Steady, moderate intensity, aerobic exercise [161] in particular is known to cause steep drops in glucose in people with T1D. Complicating efforts at decision support on glycemic management during exercise, however, is the fact that other factors can impact glucose dynamics in many different ways. Such factors include the type of exercise, the time of day of exercise, the IOB, the competitive aspect of the exercise, and the person's level of physical fitness for example. Resistance training and high intensity interval exercise are known to cause less of a drop in glucose during exercise and may even result in increases in glucose, especially when done in the morning in the fasted state, when insulin levels are the lowest [162]. Published guidelines and consensus statement have been developed to assist the management of glucose during exercise [63]. More recently, these guidelines are now being evaluated through clinical trials on high intensity interval exercise to determine how modifications of basal insulin may be helpful in preventing nighttime and mealtime hypoglycemia [163, 164]. However, other studies have shown that basal reductions may not significantly prevent hypoglycemia [165]. It has also been shown that exercise-related hypoglycemia can be nominally reduced through use of online educational materials preceding exercise [166], however, these results were not significant. Further development and studies are required to learn how we can best assist people with T1D during and after exercise through DSSs.

While there are several recent publications providing guidance on how and when people with T1D should consider exercising safely, there are not many decision support tools currently available to guide people with T1D during exercise. This represents an opportunity for future AI-based DSS algorithms and apps. If a DSS could predict hypoglycemia during or prior to exercise, it would be able to notify or alert the user in advance so that action could be taken to avoid hypoglycemia. For example, if the person was notified in advance that hypoglycemia was likely, they could consume a carbohydrate prior to exercising to prevent their glucose from dropping too low during exercise. There have been several algorithms published recently that may be used to predict glucose changes or hypoglycemia during exercise.

2.6.6.1 Linear Regression Algorithms

Ben Brahim et al. [167] use a linear regression to perform a secondary analysis of exercise data collected from 51 people with T1D. They looked at correlations of glucose trends with insulin data, age, total daily insulin requirement, and body weight. The authors determined that the most predictive factors of exercise-related glucose changes were the glucose measured at the start of exercise and the ratio of IOB at the start of exercise to total daily insulin requirement. These features were used in a subsequent hypoglycemia prediction algorithm within an automated DSS. The algorithm provided guidance on carbohydrate consumption prior to exercise session if hypoglycemia was predicted.

2.6.6.2 Decision Tree Algorithms

Reddy and Jacobs [50] developed a random forest decision tree algorithm that was designed to be used to predict hypoglycemia during exercise. The algorithm was trained and tested using three datasets comprised of CGM data, insulin data, and

physical activity data collected from adults with T1D. A random forest model was trained to predict the occurrence of hypoglycemia during either aerobic or resistance exercise, and demonstrated an 86% accuracy. The algorithm was used in a clinical study to automate the shut-off of insulin and bolus glucagon within a bi-hormonal closed-loop system [168].

2.6.6.3 Data-Driven Models

Hajizadeh and Cinar [169] developed a vector autoregressive model with exogenous inputs (VARX) including CGM, insulin, energy expenditure, and other data collected from a fitness watch. The model's coefficients were identified in real time and the model was then used for glucose predictions. Real-world data was collected from people with T1D age 19–39 who were participating in a 60-h study of closed-loop AID glucose control. The algorithm yielded a prediction accuracy RMSE of 25.5 mg/dL. This approach was further explored by Hobbs and Cinar [51] to improve glucose predictions during physical activity in adolescents. Data were obtained from adolescents utilizing experimental AID or standard of care CSII devices at a study at ski and snowboarding camps. The group developed an ARX model utilizing carbohydrate, insulin, and heartrate inputs. The group found that utilizing heartrate inputs resulted in better accuracy in terms of a lower RMSE 26.25 mg/dL, as compared to a physical model with a Kalman filter, which had an RMSE of 29.18 mg/dL. However, during the physical activity, larger errors were observed (46.16 mg/dL RMSE). While the authors highlighted the improved prediction accuracy using the VARX model, the results indicate that exercise-specific models may be needed to improve accuracy of predicted glucose-related changes during physical activity.

Romero–Ugalde et al. [65] developed an ARX model to predict glucose dynamics specifically during and after exercise. This group utilized in-clinic data from adults with

T1D collected during an exercise study. The inputs to the algorithm included smoothed CGM, insulin, carbohydrate, and activity data. The ARX model that was trained to predict glucose 30, 60, and 120 min after exercise was initiated. They developed both a population model and individualized models specific for each study participant. They found that user-specific models could yield a predictive accuracy of 7.75 mg/dL for a 30-min prediction horizon, which exceeded the accuracy of the population model. While these studies are promising, the authors indicated that the training and evaluation set was comprised of best-case data scenarios, and further evaluation is needed before utilization in a real-time DSS for exercise. However, the results indicate the benefit of using personalized models that can adapt over time to improve accuracy of glucose predictions during exercise.

2.6.7 Exercise-Induced Hypoglycemia Prevention

Physiologic Model-Based Algorithms

Fabris et al. [170] discussed the design of a bolus calculator that incorporates an activity-on-board metric, designed to adjust calculated meal boluses given the user's physical activity history directly preceding the meal. *In silico*, this was shown to reduce postprandial hypoglycemia in meals following physical activity from 13.4% with standard bolus calculator use to 3.9% with activity-on-board calculator use; however, there are no reports of further evaluation in human studies.

The following year, Fabris and Breton [45] demonstrated that real-time insulin sensitivity estimation for smart bolusing using Kalman filtering could effectively reduce postprandial hypoglycemia in the 4 h period following aerobic exercise. Though this paper did not utilize the activity-on-board algorithm described earlier by Fabris et al., it did demonstrate the utility of insulin sensitivity estimation on glycemic outcomes as described in Breton's earlier published automated DSS. In a clinical study of this

exercise bolus advisor, 15 adults who utilized insulin pumps were enrolled. A 4-week run-in was utilized to gather data and perform subject-specific insulin sensitivity estimation. After the data collection, the subjects underwent two in-clinic exercise sessions and underwent aerobic exercise followed by a standardized dinner. The results showed that the bolus algorithm adjusted for exercise significantly reduced hypoglycemia by nearly 50% in comparison to standard bolus calculations (8.33 vs 14.58, $p < 0.05$), and a significant reduction in postprandial LBGI as compared to standard bolus calculators (1.16 vs 2.86, $p < 0.05$) following exercise [83]. This algorithm shows promise for use in a real-time DSSs following aerobic exercise.

A comprehensive exercise DSS published by Ramkissoon et al. [171] describes use of automated aerobic exercise detection using only CGM data to help reduce hypoglycemia. This algorithm first performs a Kalman filter analysis of historical data collected on non-exercise days to define an activity threshold. Real-time analysis can determine if this threshold has been crossed, at which point the algorithm suggests reduction of basal, reduction of subsequent meal boluses, and a calculated carbohydrate suggestion. Evaluation *in silico* using the UVA-Padova simulator showed reduction in hypoglycemia from 2.4% to 0.0%, and avoidance of serious hypoglycemia, as compared to AID systems without exercise announcement. In addition, the system effectively reduced the frequency and severity of hypoglycemia in the 2–4 h following aerobic exercise from 2.0% to 0.0%, compared to AID systems with exercise announcement. While these are impressive *in silico* results, this algorithm has not yet been evaluated in a clinical study.

Further work in the area of preventive carbohydrate consumption was reported by Beneyto and Vehi [172], who developed a feedback proportional-derivative controller to suggest 15 g doses of carbs in real-time if hypoglycemia was predicted by an ARX model. *In silico*, utilization of this algorithm with simulated aerobic exercise showed a

statistically significant decrease in daytime hypoglycemia from 2.2% to 0.9%, and a statistically significant increase in nighttime nadir CGM from 42.8 to 59.2 mg/dL that was significantly different from standard AID use.

Garcia–Tirado and Breton utilized the net-effect replay method to estimate the physiologic disturbance of exercise during *in silico* studies. This estimate effect was then used as a disturbance input to the MPC controller during announced exercise, showing reductions in hypoglycemia from 3.8% to 0.8% as compared to standard MPC control *in silico* using the UVA-Padova simulator [173]. Although designed for use within an AID system, similar approaches may be used to estimate short-term and long-term glycemic effects of exercise on real-world human data for use in DSSs. This approach also suggests that exercise may be critical for informing bolus calculators and other components of DSSs in the future.

2.7 Combining Certified Diabetes Education with Decision Support Systems

While many groups strive for the design of fully automated DSSs, combining mobile technology and automated DSSs with a human intervention from a certified diabetes educator (CDE) may lead to optimal improvements in outcomes. Kirwan et al. [94] showed in a randomized trial of 72 people with T1D that weekly feedback from a CDE through text messaging was able to reduce their HbA1C from 9.08 (SD 0.75) to 7.9 (SD 0.75), compared with the control group which showed no significant change. Ideally, someone with T1D may benefit from obtaining regular feedback from a CDE on their insulin dosing strategies without waiting 3–6 months between specialist visits.

CDE systems rely on accurate and innovative ways of displaying their patients' data that help them provide helpful feedback and guidance. Zhang et al. [174] developed

a framework for optimizing the review of glucose data by diabetes educators. Their method automatically aggregates and reconstructs data, allowing for easy display of glucose patterns and insulin dosing behaviors. This system was reported to reflect specialist workflow when reviewed by diabetes care specialists.

While there are many studies that have discussed the impact of diabetes educators for people with T1D, there are far fewer studies evaluating how automated mobile DSSs can best augment CDE-based interventions. A recent meta-analysis [175] has indicated that there needs to be longer, structured studies to evaluate the outcomes of CDE-based support via mobile app or other electronic means.

One such study that integrated a DSS with a CDE intervention involved the Diabeo system. The Diabeo system described by Franc and Charpentier utilizes a heuristic algorithm to advise subjects undergoing CSII and MDI [176]. This mobile Diabeo system included a logbook, bolus calculator, and provides adaptive bolus adjustment for meals, as well as basal adjustment algorithm. Healthcare professionals could also access their patients' data from a web interface and review glycemic history. Early studies indicated reductions in HbA1C of 0.91% when subjects utilized Diabeo and also received regular decision support from a CDE. The benefits of combining automated mobile decision support and CDE feedback and consultations were described in the TeleDiab I and II studies [93], whereby individuals with type 1 or type 2 diabetes used the Diabeo mobile app with CDE check-ins. The addition of a DSS was reported to have a greater effect on reduction of HbA1c (0.91%) as compared to mobile DSS used alone (0.67%), or CDE intervention alone.

2.8 Potential Future Directions in Decision Support Systems in Type 1 Diabetes

2.8.1 Exercise Decision Support Systems and Exercise as an Adjunct Therapy

This review has identified the need for more extensive tools to help people with T1D better manage their glucose during exercise. There have been relatively few DSS developed that can be used to provide advice to people with T1D across a variety of exercise types, durations, intensities, times of day, insulin loading conditions, and also for people with different fitness levels. The Leona M. and Harry B. Helmsley Charitable trust has recently invested in the funding of a large clinical study called 'Type 1 Diabetes in Exercise' (T1-Dexi study) that will yield a large publicly available dataset with time-matched CGM, insulin, food, and exercise data [69]. With new and growing publicly available data sets like this, the area of DSS design in exercise for T1D is likely to be an important future area that will yield many algorithmic advances and new mobile tools in the years ahead.

Exercise as a glycemic intervention is something that almost no one has discussed in the literature. While most of the studies described in this review have focused on how to avoid glycemic excursions caused by exercise, it may be important to consider how exercise can help to achieve glycemic targets. After all, exercise increases glucose uptake and may help reduce postprandial glucose excursions if the correct exercise is done at the right time, at the right intensity, and for the right duration. Xie and Wang [177] explored this concept by designing a non-linear ARMAX model that would recommend the optimal carbohydrate intake, insulin dosage, and target exercise heartrate required to optimize a person's daily time in glucose target range. *In silico*

evaluation, using the 30 subjects from the UVA-Padova Simulator, indicated use of the recommender system significantly reduced LBGI from 2.57 to 0.42, as compared to standard of care. Clearly, this approach may require considerable planning and foresight, which is not practical for most people with T1D. Nonetheless, this system is unique in that it recommends prophylactic measures that may be taken by a person to exercise in an optimal way to avoid hypoglycemia and maximize glycemic outcomes.

2.8.2 Optimizing Meal Bolus Timing and Other Time-Varying Dosing Parameters

The meal bolus DSSs described in this review have all focused on pre-meal insulin dosing. However, prior work [178] has shown that the timing of insulin delivery both before and following a meal (for CSII users) can help reduce postprandial hyper- and hypoglycemia. For example, for CSII users, a DSS may recommend consuming 50% of the meal insulin bolus prior to eating, and taking the remaining 50% over the next several hours. These bolus insulin suggestions may be adaptive and personalized using run-to-run based outcome measurement assessments to achieve the best glycemic performance possible for a given person.

Adapting the time windows for insulin management parameters like the glucose target, correction factor, and carbohydrate ratio may also show a benefit in improving glycemic outcomes. A person with T1D, whether on CSII or MDI, maintains different glucose targets, carbohydrate ratios and correction factors for specific times of day. Recent positions by the American Diabetes Association have indicated that people with T1D will benefit from different glycemic targets for different times of day and different contexts such as exercise, while sleeping, etc.

Eissa et al. [179] reported a k-means clustering method for determining the optimal time-blocks for bolus calculations. Data from 70 participants with T1D were processed and analyzed using k-means clustering to determine the optimal time windows for insulin dosing parameters. The time-blocks determined by the k-means clustering algorithm were then compared to the time-blocks suggested by clinicians, specialist nurses, and dieticians, and exhibited a 39.1% agreement. This was consistent with the agreement measured between specialists and participants of 36.1%. The algorithm could also identify optimal time blocks for the weekend and also for specific days of the week. This algorithm was not validated in a human study, but simply identifies an interesting method of automating the selection of time blocks for insulin settings based on similar patterns.

However, there has not been a lot of research done in this area to date.

2.8.3 Pregnancy

Use of available technologies such as CSII, CGM, and more recently AID systems [180] have been shown to improve outcomes in pregnant users. Currently, there are no DSSs designed for pregnancy. While the algorithms described here show promise in reducing hypoglycemia and may improve time-in-range in an out-patient setting, these systems require further evaluation in medically complex populations and have not been evaluated in women who are pregnant [181].

2.8.4 Integrating Decision Support Systems with AID

Recent publications by Kovatchev et al. [182] and others have shown that the primary benefit of AID systems is during the overnight period. The primary reason for this is that meals and exercise are very challenging for even an AID system to handle. For example, hypoglycemia that results from exercise may occur even in automated

closed-loop systems even when glucagon is delivered in anticipation of exercise [28]. Integrating DSSs with AI may provide a way to improve and optimize AID systems by leveraging patterns observed in glucose and insulin data to help people with T1D make better choices about the settings on the AID systems.

2.9 Conclusions

AI techniques provide a powerful means to address many challenges in diabetes care, and these techniques may be used effectively in the design of DSSs. We have provided a comprehensive review of (1) DSS algorithms that provide insulin dosing recommendations to people using MDI or CSII therapy and (2) DSS forecasting algorithms that provide real-time alerts and notifications with regards to predicted glucose excursions and especially hypoglycemia. While AI is rapidly developing within the field of medical informatics, many of the systems presented in this article utilize traditional machine learning algorithms to adjust insulin therapy or predict hypoglycemia. Cutting-edge machine learning algorithms, such as deep-learning algorithms, have been used in glucose forecasting but have not yet been applied to estimations of insulin dosage adjustments or hypoglycemia prevention.

A common theme on the insulin dosing DSSs which have been evaluated in human studies is that, thus far, they have been effective at reducing glycemic variability and reducing hypoglycemia with short-term use, but they have not yet shown similar improvements in HbA1c, time in range, or mean glucose that AID systems have demonstrated. Like many medical interventions, DSSs require consistent usage to impact clinical outcomes, and so longer studies are required to assess the impact of these DSSs on outcomes such as HbA1c, mean glucose and time-in-range.

One theme from the review of the hypoglycemia prediction algorithms is that there is a wide range of prediction accuracy reported in the literature, and this may be because many of the data sets used for evaluation were either *in silico* data sets or data sets acquired under prescribed settings within a clinical study. Evaluation of these algorithms with real-world human data under free-living conditions is critical for obtaining a realistic estimate of the algorithm's accuracy. New big data sets acquired from organizations like Tidepool (San Francisco, CA, USA) and the Tidepool Big Data Donation Data Set [68] are now becoming available that include real-world, free-living CGM, insulin, and exercise data from people using CSII and more recently AID systems. The Leona M. and Harry B. Helmsley Charitable Trust is sponsoring a large study to generate time-matched CGM, insulin, nutrition, and exercise data in 600 people with T1D [69]. These data sets should be leveraged so that algorithms can be evaluated, compared, and benchmarked on a common, real-world data set.

Lastly, many of the insulin-dosing DSS algorithms and glucose prediction DSS algorithms reviewed here included only an *in silico* evaluation, making it challenging to assess how well the systems will ultimately perform in actual humans. While the *in silico* simulators have been absolutely critical and transformative in the design and preliminary evaluation of DSSs and AID systems, human testing should be a goal for any group that is serious about having their DSS algorithm translated to use by people with diabetes.

2.10 Materials and Methods

We performed a comprehensive search of PUBMED, IEEE Xplore, and ScienceDirect utilizing combinations for the following terms: "exercise OR Physical", "hypoglycemia", "type 1 diabetes", "prediction", "decision support", "insulin adjustment", "insulin management", "decision algorithm", "exercise adjustment", and "carbohydrate

intake”. No time span was imposed on these papers. After removing duplicate articles, we were left with 562 primary research articles, conference articles, and book chapters. Of these 562 items, we prioritized approaches that include validation via human pilot studies or clinical trials, and secondly considered novel approaches that evaluated their algorithms through *in silico* trials.

We did not include papers that described a system framework without algorithm descriptions, or without preliminary results on *in silico* or human data. While many glucose-prediction approaches were returned by our search, we only include the most recent state-of-the-art approaches. We likewise excluded AID and MPC control algorithm papers that did not relate to DSSs. All papers reviewed are included in Table 2.S1.

2.11 Supplementary data

Table 2.S1: Supplementary Table. Summary of decision support strategies evaluated in this review.

Citation	Purpose of Decision Support	Algorithm	Evaluation strategy	Study type / Dataset	Reported Outcome
Bellazzi et al	Insulin dose adjustment	Case-based reasoning	Real-world use	Clinical Study	% HbA1c
Schwartz et al	Insulin dose adjustment Carbohydrate intake	Case-based reasoning	Retrospective analysis of real-world data	Data collection Secondary data analysis	Accuracy of identified cases
Herrero et al	Insulin dose adjustments	Case-based reasoning Run-2-Run	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Reddy et al	Insulin dose adjustment	Case-based reasoning Run-2-Run	Real-world use	Clinical Study	% Time-in-range % time-in-hypoglycemia
Torrent Fontbona et al	Insulin dose adjustment	Case-based reasoning Concept drift	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Liu et al	Hypoglycemia alarm Carbohydrate intake	Run-2-Run	Real-world use	Clinical study	% Time-in-range % time-in-hypoglycemia
Tyler et al	Insulin dose adjustment	K-nearest-neighbors Model-based bolus adjustment Heuristic quality-control	Real-world use	Clinical study	% Time-in-range % time-in-hypoglycemia
Biagi et al.	Glycemic Pattern Identification	K-means clustering	Retrospective analysis of real-world data	Secondary data analysis	Glycemic profiles
Nimri et al	Insulin dose adjustment	Rule-based fuzzy logic	Real-world use	Clinical study	% Time-in-range % time-in-hypoglycemia
Pankawska et al	Insulin dose adjustment Carbohydrate estimation	Voice recognition Rule-based heuristics	Real-world use	Clinical study	% Time-in-range
Palerm et al	Basal insulin dose adjustment	Run-2-Run	Simulated use	Proof of concept	Time to convergence
Herrero et al	Basal insulin dose adjustment	Run-2-Run	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Toffanin et al	Basal insulin dose adjustment	Run-2-Run	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Zisser et al	Bolus insulin dose adjustment	Run-2-Run	Real-world Use	Clinical Study	Mean glucose preceding and following meals

Wong et al	Insulin dose adjustment	Model-based simulated replay	Simulated use	*In silico* clinical study: AIDA simulator	% HbA1C % time-in-hypoglycemia
Rosales et al	Insulin bolus dose and shape	Constrained optimization	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Revert et al Rosetti et al	Insulin bolus dose and shape	Interval analysis	Real-world Use	Clinical Study	Postprandial AUC
Breton et al	Insulin dose adjustment Exercise hypoglycemia prevention	Kalman-filter state estimation Model-based simulated replay Logistic regression	Real-world use	Clinical study	% Time-in-range % time-in-hypoglycemia
Goodwin et al	Insulin bolus dose and shape	Model-based forecasting	Retrospective analysis of real-world data	Data collection Secondary data analysis	% of subject data lying within the prediction envelope
Sun et al	Insulin dose adjustment	Actor-critic reinforcement learning	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Perez-Gandia et al	Insulin dose adjustment Carbohydrate intake	Artificial neural network	Real-world use	Clinical Study	Kovatchev's risk index
Srinivasan et al	Insulin bolus dose and shape	Particle swarm optimization	Simulated use	Proof of concept study, *In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Anthimopoulos et al Vasiloglou et al	Carbohydrate Estimation	Computer vision	Real-world use	Proof of concept study	Mean absolute error
Mahmoudi et al	Missed meal detection	Kalman-filter state estimation	Simulated use	*In silico* clinical study: UVA-Padova	% Time-in-range % time-in-hypoglycemia
Samadi et al	Missed meal detection	Rule-based fuzzy logic	Retrospective analysis of real-world data	Secondary data analysis	% Sensitivity % false positive rate for meal detection
Zhang et al	Glycemic pattern identification	Heirarchical task abstraction	Retrospective analysis of real-world data	Qualitative proof-of-concept	Physician feedback
Charpentier et al	Insulin dose adjustment	Rule-based heuristics Clinical diabetes educator	Real-world use	Clinical trial	% HbA1c
Cameron et al Calhoun et al Buckingham et al Biester et al	Basal insulin suspension	Kalman-filter state estimation	Real-world use	Clinical study	% HbA1c % time-in-hypoglycemia
Sparacino et al	Glucose forecasting and hypoglycemia prediction	Data-driven ARX	Retrospective analysis of simulated data	*In silico* clinical study: UVA-Padova	RMSE of forecasted glucose
Perez Gandia et al	Glucose forecasting and hypoglycemia prediction	Artificial neural network	Retrospective analysis of real-world data	Secondary data analysis	RMSE of forecasted glucose

Zecchin et al	Glucose forecasting and hypoglycemia prediction	Artificial neural network	Simulated use	*In silico* clinical study: UVA-Padova	% Time in hypoglycemi
Daskalaki et al	Glucose forecasting and hypoglycemia prediction	Data-driven cARX Recurrent neural network	Retrospective analysis of real-world data	Secondary data analysis	RMSE of forecasted glucose % Sensitivity for hypoglycemia prediction
Contrares et al	Glucose forecasting	Clustered grammatical evolution Reinforcement learning	Retrospective analysis of simulated data	*In silico* clinical study: UVA-Padova	RMSE of forecasted glucose
Montaser et al	Postprandial hypoglycemia prediction	Data-driven ARIMAX	Retrospective analysis of real-world data	Secondary data analysis	RMSE of forecasted glucose
Toffanin et al	Postprandial hypoglycemia prediction	Data-driven state-space	Retrospective analysis of real-world data	Secondary data analysis	FIT and Coefficient of determination of forecasted glucose
Oveido et al Oveido et al	Postprandial hypoglycemia prediction Insulin bolus adjustment	Support vector regression	Retrospective analysis of real-world data Simulated use	*In silico* clinical study: UVA-Padova	% Sensitivity for hypoglycemia prediction % time-in-hypoglycemia following meals
Cappon et al	Postprandial hypoglycemia prediction Insulin bolus adjustment	Xtreme gradient-boosted tree	Simulated use	*In silico* clinical study: UVA-Padova	AUROC % time in range % time in hypoglycemia
Schiffrin et al	Nocturnal hypoglycemia prevention, Carbohydrate intake	Linear regression Decision theory	Real-world use	Clinical study	Incidence of hypoglycemia % HbA1c
Mosquera-Lopez et al	Nocturnal hypoglycemia prediction and prevention	Support vector regression Decision theory	Retrospective analysis of real-world data Simulated use	Secondary data analysis *In silico* clinical study: OHSU T1D	% Sensitivity and % Specificity of predicted nocturnal hypoglycemia
Guemes et al	Nocturnal hypoglycemia prediction	Support vector regression	Retrospective analysis of real-world data	Secondary data analysis	% Sensitivity and % Specificity of predicted nocturnal hypoglycemia
Vehi et al	Nocturnal hypoglycemia prediction	Support vector regression Artificial neural network	Retrospective analysis of real-world and simulated data	Secondary data analysis	% Sensitivity and % Specificity of predicted nocturnal hypoglycemia
Bertachi et al	Nocturnal hypoglycemia prediction	Support vector regression, Multilayer perceptron neural network	Retrospective analysis of real-world data	Secondary data analysis	% Sensitivity and % Specificity of predicted nocturnal hypoglycemia
Fabris et al	Exercise hypoglycemia prevention, insulin bolus adjustment	Model-based activity on board adjustment	Simulated use	*In silico* clinical study: UVA-Padova	% time in range % time in hypoglycemia

Fabris et al	Exercise hypoglycemia prevention, insulin bolus adjustment	Kalman-filter state estimation	Real-world use	Clinical study	% time in range % time in hypoglycemia
Ramkissoon et al	Exercise hypoglycemia prevention, insulin dose adjustment, carbohydrate intake	Kalman-filter state estimation	Simulated use	*In silico* clinical study: UVA-Padova	% time in range % time in hypoglycemia
Beneyto et al	Exercise hypoglycemia prevention, carbohydrate intake	Proportional derivative controller	Simulated use	*In silico* clinical study: UVA-Padova	% time in range % time in hypoglycemia
Garcia-Tirado et al	Exercise hypoglycemia prevention	Model-based characterization of exercise	Simulated use	*In silico* clinical study: UVA-Padova	% time in range % time in hypoglycemia
Ben Brahim et al	Exercise-related glucose forecasting	Linear regression	Retrospective analysis of real-world data	Secondary data analysis	Pearson's correlation, Predictive features
Hayeri	Exercise-related glucose forecasting	Gradient boosted decision trees Support vector regression	Retrospective analysis of real-world data	Secondary data analysis	Clarke error grid
Reddy et al	Exercise hypoglycemia prediction	Random forest decision tree	Retrospective analysis of real-world data	Secondary data analysis	% Sensitivity and % Specificity of predicted exercise hypoglycemia
Hajizadeh et al Hobbs et al	Exercise-related glucose forecasting	Data-driven ARX	Retrospective analysis of real-world data	Secondary data analysis	RMSE of forecasted glucose
Romero-Ugalde et al	Exercise-related glucose forecasting	Data-driven ARX	Retrospective analysis of real-world data	Secondary data analysis	RMSE of forecasted glucose
Eissa et al	Adjustment of time-blocks used for insulin dosing	K-means clustering	Retrospective analysis of real-world data	Data-collection Secondary data analysis	Agreement with specialists
Avila et al	Adjustment of glycemic target	Decision tree, Recommender system	Retrospective analysis of simulated data	Secondary data analysis	RMSE of predicted glycemic variability

3 An artificial intelligence decision support system for the management of type 1 diabetes

Summary:

- This chapter presents a KNN artificial intelligence system designed to assist people who manage insulin through multiple daily injection therapy
- The system shows high agreement with endocrinologist recommendations.
- The system can improve patient outcomes *in silico* after 12 weeks of use, and reduces hypoglycemia in a 4-week human study.

This manuscript was published in June 2020: Tyler NS, Mosquera-Lopez CM, Wilson LM, Dodier RH, Branigan DL, Gabo VB, Guillot FH, Hilts WW, El Youssef J, Castle JR, Jacobs PG. An artificial intelligence decision support system for the management of type 1 diabetes. Nat Metab. 2020 Jul;2(7):612-619. doi: 10.1038/s42255-020-0212-y. Epub 2020 Jun 1. PMID: 32694787; PMCID: PMC7384292.

Abstract: Type 1 diabetes (T1D) is characterized by pancreatic beta cell dysfunction and insulin depletion. Over 40% of people with T1D manage their glucose through multiple injections of long-acting basal and short-acting bolus insulin, so called multiple daily injections (MDI).[183] Errors in dosing can lead to life-threatening hypoglycaemic events (< 70 mg dL^-1) and hyperglycaemia (> 180 mg dL^-1), increasing the risk of retinopathy, neuropathy, and nephropathy. Machine learning (artificial intelligence) approaches are being harnessed to incorporate decision support into many medical

specialties. Here we report an algorithm that provides weekly insulin dosage recommendations to adults with T1D using MDI therapy. We employ a unique virtual platform[57] to generate over 50,000 glucose observations to train a K-nearest-neighbours [70] decision support system (KNN-DSS) to identify causes of hyperglycaemia or hypoglycaemia and determine necessary insulin adjustments from a set of 12 potential recommendations. The KNN-DSS algorithm achieves an overall agreement with board-certified endocrinologists of 67.9% when validated on real-world human data, and delivers safe recommendations per endocrinologist review. A comparison of physician-recommended adjustments to insulin pump therapy indicates full agreement of 41.2% among endocrinologists, which is consistent with previous measures of inter-physician agreement (41-45%) [47]. *In silico* [57, 75] benchmarking using a platform accepted by the U.S. Food and Drug Administration for evaluation of artificial pancreas technologies, indicates substantial improvement in glycaemic outcomes after 12-weeks of KNN-DSS use. Our data indicate that the KNN-DSS allows for early identification of dangerous insulin regimens and may be used to improve glycaemic outcomes and prevent life-threatening complications in people with T1D.

3.1 Main

Optimal management of type 1 diabetes requires precise insulin administration to maintain glucose within safe ranges. Dosage regimens are complicated by day-to-day changes in insulin sensitivity, which can cause large excursions in glucose. Failure to dose insulin properly can result in diabetic ketoacidosis and hypoglycaemia, which may lead to coma or death. Intensive insulin regimens can enhance glycaemic outcomes in people with T1D who use MDI therapy [184], but a number of factors confound adherence to insulin dosing. Fear of hypoglycaemia, challenges with numeracy to

calculate meal or correction boluses, changes in insulin sensitivity during exercise, illness, stress and menstruation, and the psychological toll of this chronic disease make it difficult for people with T1D to adhere to these regimens.[185-188]

Whereas many smartphone apps are available to help people better manage their diabetes, most of these are not validated and have not shown clinical efficacy. A recent review indicated that out of hundreds of such applications, only 12 were validated in clinical trials and few of these significantly improved glycated haemoglobin (HbA1c) in people with T1D. [6, 92-94] Apps shown to improve glycaemic outcomes provided users with weekly or biweekly feedback from health professionals on insulin dosage adjustments. [93, 94] Continuous glucose monitoring (CGM) has been shown to significantly improve HbA1C, but as a sole intervention does not bring everyone to goal. [37] CGM-informed advisory systems [189] range from machine learning to physiologic models and heuristic approaches for the adjustment of basal [79-81] and bolus therapies.[46, 78] Nimri et al. developed a system to guide adjustment of insulin pump settings using capillary blood glucose or CGM data.[47] Perez-Gandia et al. developed a predictive neural-network to assist with real-time insulin administration or carbohydrate consumption.[97] MDI-inclusive approaches include a model-based decision support system for titration of insulin prior to exercise by Breton et al.[45], and an adaptive KNN case-based reasoning approach for titration of short-acting insulin by Reddy et al.[59]

The KNN-DSS that we describe provides up to four optimally selected dosing and behavioral recommendations once per week to adults with T1D who use MDI therapy. Recommendations are selected to manage insulin dosed for meals andsnacks to bring glucose to within a target range. A virtual patient simulator platform was implemented to design the KNN-DSS algorithm. The virtual patient simulator [57] is a mathematical representation of the glucoregulatory response to food, insulin, and exercise in people with T1D. The virtual patient simulator was used to train a machine-

learning KNN [70] model to predict optimal insulin recommendations that improve glycaemic outcomes. Input data for the algorithm are acquired from CGM data, insulin data obtained from Bluetooth-enabled capture devices, and physical activity metrics obtained through wearable sensors (Figure 3.1a). The KNN-DSS then classifies glycaemic features and delivers recommendations to improve percent time in ange (70-180 mg dL^-1) and reduce percent time in hypoglycaemia (Figure 3.1b). User-specific titration of insulin occurs using an adaptive learning postprandial hypoglycaemia avoidance (ALPHA) algorithm [190] which selects the optimal bolus insulin based on the prior glycaemic outcomes of the user (Figure 3.1c). To ensure recommendations conform to physician standards, we developed an expert-knowledge quality control algorithm (Figure 3.1d, Extended Data 3.E1-E7). The heuristic Quality Control algorithm is designed for user safety and operates independent of the machine-learning framework. The KNN-DSS system delivers one or more recommendations from a set of 12 unique recommendations for insulin adjustments and dosage behaviors with respect to long-acting basal insulin, fast-acting carbohydrate-to-insulin ratio (carb:ins), and correction insulin dosage (Table 3.1). The top three meal and basal insulin recommendations are selected by KNN classification, whereas recommendations for correction doses and compliance with care are supplied by the heuristic ALPHA and quality control algorithms.

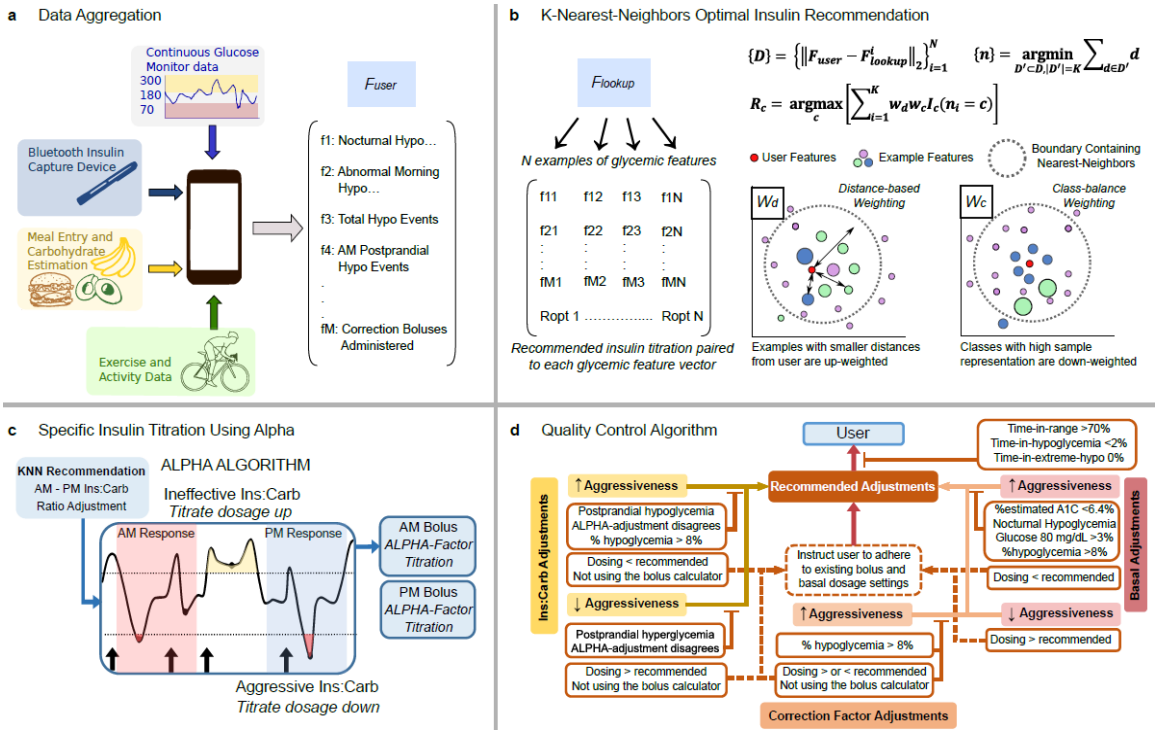


Figure 3.1: Decision support engine framework to identify user-specific insulin titrations.

a, The user data are aggregated and processed for extracting glucose, insulin, meal, and exercise features that may be used to optimally titrate insulin doses. **b**, The user features ***F***_user are matched to the closest examples in the look-up table for the K-nearest neighbours algorithm, ***F***_lookup. The distance between user features and the examples in the look-up table are calculated as {***D***}, and the *K* examples within minimum distance to user features, {***n***}, are weighted by distance, *w*_d, and class-size, *w*_c; the final insulin dosage recommendations, ***R***_c, are returned by the KNN algorithm. **c**, For those recommendations indicated by the KNN-DSS, the ALPHA algorithm assigns an aggressiveness factor that titrates carbohydrate ratios and correction factors to improve time in target range and reduce time in hypoglycaemia. **d**, A quality control algorithm is employed to ensure that KNN-DSS recommendations adhere to physician standards

We validated the accuracy and safety of KNN-DSS generated recommendations compared with those of board-certified endocrinologists using 687 days of real-world data collected from 25 adult participants on MDI therapy. We demonstrated efficacy of the KNN-DSS during two 52-week studies *in silico* and characterized the response of the

engine to dynamic disturbances in glycaemic patterns. Lastly, we report the results of a short, proof-of-concept, single-centre clinical study to evaluate the safety of the KNN-DSS in human participants (Supplementary Table 3.S1: Study 1).

Table 3.1: Recommendations delivered by the KNN-DSS engine.
Recommendations for insulin dosage are titrated to be higher or lower during different time windows using the specified titration method. [a] Behavioral recommendation.

Recommendation	Message to User	Adjustment Window	Titration Method
Basal Adjustment	"You may need (less/more) basal insulin. It is recommended that you (decrease/increase) your AAA insulin from BBB to CCC units.	AM PM	*Adjustment by +/- 10% from weekly baseline settings*
Carb:ins Ratio	"You may need (less/more) insulin before (breakfast/lunch/dinner/a meal). It is recommended that you change your carb ratio from AAA to BBB."	7:00 – 11:00 11:00 – 15:00 15:00 – 20:00 20:00 – 7:00	*ALPHA Algorithm: meal bolus glycaemic response and assignment of dosage titration.*
Correction Factor	"You may need (less/more) insulin in (morning / afternoon / evening / nighttime) to bring down high glucose levels. It is recommended that you change your correction factor from AAA to BBB."	7:00 – 11:00 11:00 – 15:00 15:00 – 20:00 20:00 – 7:00	*ALPHA Algorithm: correction bolus glycaemic response and assignment of dosage titration.*
Bolus Adherence[a]	"You may have taken (more/less) insulin than recommended by the bolus calculator during certain times of day. It is recommended that you use the amount recommended by the bolus calculator"	All Day	N/A
Basal Adherence [a]	"We have found that the amount of basal insulin that you are taking is different than the amount we recommend. Taking the recommended amount may improve your glucose levels."	All Day	N/A

To compare the recommendations of the KNN-DSS with endocrinologist recommendations, we used data collected from 25 adult participants with T1D during a

28-day outpatient study (Supplementary Table 3.S1: Study 1). One of three endocrinologists from Oregon Health and Science University (OHSU) medical center analysed the glucose and insulin dosing data from each participant and provided recommended adjustments to basal insulin, and fast-acting meal insulin and correction insulin during four different windows of time (7:00 – 11:00, 11:00 – 15:00, 15:00 – 20:00, or 20:00 – 7:00) (Table 3.1). Recommendations regarding daily bolus calculator use were also provided. The KNN-DSS recommendations were labelled as being in full agreement, partial agreement, full disagreement, or partial disagreement with the physician. The accuracy of recommendations delivered by the KNN-DSS as compared to those of board-certified endocrinologists was quantified using a modified Sorenson-Dice coefficient [191] (equation (3.6)). We measured a combined agreement of 67.9% between endocrinologist and KNN-DSS recommendations, whereas 6.4% of recommendations were in disagreement. In 16.7% of recommendations the engine identified an issue and physician did not indicate a recommendation, and 9% of recommendations were not comparable (Table 3.2). We performed additional analysis on a subset of these participants who exhibited consistent use of the bolus calculator and adherence to their insulin dosage settings (greater than 75%). For this subset of participants, the engine recommendations were in full agreement with physicians 50.8% of the time and exhibited an overall agreement of 67.5% with the recommendations of physicians. We observed that over 99% of recommendations delivered across the 100 weeks of data by the KNN-DSS passed a safety review in which the endocrinologists reviewed each recommendation for the potential to cause hypoglycaemia episodes and overnight events.

The measure of physician agreement with the KNN-DSS is similar to those found in published studies involving other decision support systems. In an international, multicenter study,[47] Nimri et al. compared physician-recommended adjustments of

insulin pump settings and found the expected full agreement of different physicians to be between 41% and 45%, and the expected full disagreement between 9% and 12%. The results demonstrated high variability among physicians both internationally and in practice at the same institution. Nimri et al. used multiple study centers and evaluated insulin pump settings, which differentiated their study from that of a single-center study on people using MDI. Additional studies that compared adjustments in MDI therapy reported that general practitioners and endocrinologists alike identified 67% of indicated changes to insulin,[192] and reported an agreement of 63% between physician-recommended and software-recommended titrations to insulin dosages.[193]

We measured inter-physician recommendation variability on a dataset collected previously from participants with T1D using sensor augmented pump therapy [194] (Supplementary Table 3.S1: Study 2). We found a value of 41.2% in relation to full agreement among endocrinology faculty at OHSU (Supplementary Table 3.S2). The KNN-DSS engine, trained with a virtual platform, demonstrates high full agreement (50.8%) and partial agreement (67.5%) with endocrinologist recommendations when validated on real-world data (Study 1), and exceeds inter-physician agreement found in Study 2.

Table 3.2: Agreement between KNN-DSS and endocrinologist recommendations using real-world human data.

Agreement is calculated using a Sorensen-Dice coefficient similarity comparing physician recommendations to engine recommendations for each week of observed data. Adherence refers to participant use of the bolus calculator.

Recommendation Comparison	% Agreement	% Disagreement	% Additional	% Not comparable
Assessing recommendations on all participant data (N = 78 weeks)	*Full* 27.8 *Partial* 40.1 Overall **67.9**	*Full* 0.4 *Partial* 6.0 Overall **6.4**	*Physician* 6.4 *Engine* 10.3 Overall **16.7**	**9.0**

Assessing recommendations on participant data with > 75% recommendation adherence (N = 39 weeks)	*Full* 50.8 *Partial* 16.7 Overall **67.5**	*Full* 2.6 *Partial* 8.8 Overall **11.4**	*Physician* 13.2 *Engine* 7.9 Overall **21.1**	**0.0**

We evaluated the ability of the KNN-DSS to improve glycaemic outcomes using two virtual patient simulators [57, 75]. Each virtual patient was given real-world meal scenarios previously recorded during a clinical trial of automated insulin delivery therapy. These meal scenarios are meant to rigorously challenge algorithm performance with realistic eating patterns.

Evaluation *in silico* demonstrated the ability of the KNN-DSS to identify problematic glycaemic patterns and to deliver effective insulin dosage recommendations. In the first *in silico* study using the OHSU T1D simulator [57], 29 virtual patients with varying adherence to insulin therapy, varying circadian insulin sensitivities and carbohydrate misestimation of ±30%, were evaluated in a 75-week study of weekly decision support. After 12 weeks of use, the KNN-DSS and supporting algorithms improved virtual patient outcomes considerably, increasing average percent time in target range from 59.5% to 79.8% (P = 2E-5), maintaining percent time in hypoglycaemia at target (<2%) and reducing inter-individual variability (Figure 3.2a, Supplementary Table 3.S3).

After the optimal insulin dosing settings were obtained at 52 weeks of simulation, we disturbed the system by imposing new insulin dosing errors and changes to patient settings (Supplementary Table 3.S4). The engine was able to correct the problems with the invalid dosing settings, gradually improve the time in target range, and reduce the time in hypoglycaemia (Figure 3.2b). Patient settings largely trended towards their pre-

disturbance values (Figure 3.2c), which indicates the engine's ability to respond to dynamic changes and retrieve the original settings. Since there are many possible combinations of long-acting and short-acting insulin therapy that may improve glycaemic control in a person with T1D, not all therapy settings returned to the original settings prior to the disturbance. We nonetheless found that all of the glycaemic outcomes of the virtual patients still improved following the disturbance (Figure 3.2b).

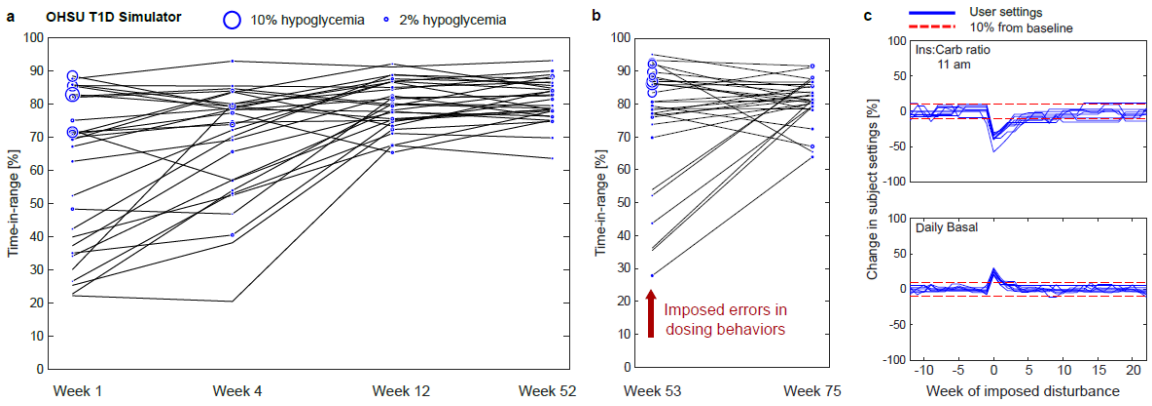


Figure 3.2: Engine performance in improving subject outcomes *in silico.* **a**, Outcomes of the *in silico* evaluation of the KNN-DSS over 52 weeks. Virtual patients from the OHSU T1D simulator undergo weekly use of the KNN-DSS engine. **b,** At 52 weeks, new insulin settings and dosing behaviors are imposed on patients, the effects of which are measured at 53 weeks. **c,** Evolution of patient settings 20 weeks from baseline following a disturbance at week 52. The red dashed line indicates 10% from baseline.

In the second 52-week *in silico* study, we used a benchmarking platform accepted by the United States Food and Drug Administration for the evaluation of artificial pancreas algorithms, the UVA-Padova simulator [75], to evaluate the performance of the KNN-DSS on 100 virtual adult patients, 100 virtual adolescent patients, and 100 virtual pediatric patients. In adult patients, percent time in target range improved from 75.1% at baseline to 81.8% (P=1E-4) at the study conclusion. Percent time in hypoglycaemia was reduced from 4.0% at baseline to 0.55% (P=2E-12) at the end of the study (Supplementary Table 3.S3, Extended Data 3.E8). *In silico* studies may

provide optimistic estimations regarding glycaemic outcomes because virtual study patients exhibit near perfect adherence to recommended insulin adjustments, which often does not happen in real-world studies. Nonetheless, the results shown here indicate that study participants who use the recommendations provided by the KNN-DSS may expect a reduction in hypoglycaemia and an improvement in time in target range after 12 weeks. Notably, at 4 weeks, both the UVA-Padova and the OHSU T1D simulators showed reductions in glycaemic variability and hypoglycaemia and a small average improvement in glycaemic time in target range, with many patients showing a reduction in time in target range. The KNN-DSS is designed to prioritize safety and reduce hypoglycaemia as well as optimize insulin dosages and percent time in target range; as a consequence, time in target range may initially be reduced as problematic and aggressive insulin doses are titrated down, and improves significantly after 12 weeks of continued use. Performance was also evaluated using the UVA-Padova simulator in an adolescent and pediatric population. For adolescents and children, we also observed significant improvements in glycaemic outcomes, as percent time in target range increased from 68.2% to 75.2% (P = 8E-9) for adolescents, and 65.7% to 69.1% (P = 0.02) for children after 12 weeks (Supplementary Table 3.S5).

We evaluated the safety of the KNN-DSS in a single-center feasibility study. Sixteen adults out of 25 adults with type 1 diabetes on MDI therapy (Supplementary Table 3.S1: Study 1) underwent weekly KNN-DSS augmented decision support. These 16 participants were given weekly recommendations on dosing and behavioral changes from a physician who reviewed the glucose history of participants and recommendations suggested by the KNN-DSS. Although this small pilot study was not powered to detect any impact from the intervention, and the study duration of 4 weeks was too short to observe a significant impact on time in target range, we did observe that participants exhibited a 25% decrease in hypoglycaemia events over a 24-hr period (from 0.86 to

0.64 events per day, P = 0.051), a 33% decrease in daytime hypoglycaemia events (from 0.43 to 0.29 events per day, P = 0.096), a 43% decrease in hypoglycaemia overnight (from 0.50 to 0.29 events per day, P = 0.04) and an 76% decrease in serious hypoglycaemia (<54 mg dL^-1) overnight (from 0.48% to 0.11%, P=0.03) during the final week compared with the first week in the study (Extended Data 3.E9) . Fifteen of the 16 participants completed all 4 weeks of the study whereas one participant chose to exit the study after two weeks. In a similar manner, we observed no substantial changes in daytime percent time in target range comparing the final week to week 1 (54.2% to 52.6%, P = 0.63), a decrease in time in target range overnight (56.7% to 45.7%, P = 0.034), a non-significant decrease in time in target range over a 24-hr period (55.3% to 49.1%, P = 0.06), and a non-significant increase in mean glucose (172.4 mg dL^-1 to 185.5 mg dL^-1, P = 0.08). As shown in Supplementary Table 3.S3, our simulation studies confirm the findings of the clinical study, which indicates that, although a reduction in hypoglycaemia can be expected after four weeks, a longer study is required before we can expect substantial improvements in time in target range.

Other groups have also reported on the impact of automated decision support systems on glycaemic outcomes, with results primarily indicating that hypoglycaemia can be reduced. Breton et al. recruited 24 participants with T1D (8 of whom were on MDI therapy) for a crossover study to evaluate automated decision support. Glycaemic outcomes were evaluated during a 48-hour in-patient session in which participants underwent standardized meals and exercise. They observed a statistically significant reduction of percent time in hypoglycaemia for the DSS vs. the control group (3.2% [1.3, 4.8] vs. 0.9% [0.4 2.3], P = 0.02) but no significant change in percent time in target range.[45] Herrero et al. described a case-based reasoning DSS (named ABC4D) and evaluated the system *in silico* using 20 representative adults and adolescents from the UVA-Padova simulator population. They found that after 4 weeks, percent time in target

range could be improved in adults from 75.2 ± 11.7 to 81.9 ±13.4 (P<0.05) and percent time in hypoglycaemia could also be reduced from 0.3±0.5 to 0 (P=0.17).[46] Although we also used the UVA-Padova population, a direct comparison of ABC4D to the KNN-DSS would be difficult as we evaluated our algorithm in 300 UVA patients. Reddy et al. further evaluated the ABC4D algorithm in a real-world, 6-week pilot study, but whereas both percent time in target range and percent time in hypoglycaemia improved slightly, the changes were not significant.[59] Other approaches described a reduction in hypoglycaemia after 12 weeks of basal and fast-acting insulin decision support, but no significance measures were reported [47]. A common theme of these studies is that it has not yet been shown that a DSS can improve percent time in target range in human studies. Our *in silico* results indicate that longer study durations (>12 weeks) will be necessary to demonstrate improvements in percent time in target range through use of the KNN-DSS.

The KNN-DSS was trained and validated for use with specific sensor technologies, insulin therapies, and target populations. The KNN-DSS system utilizes CGM devices that sample glucose at 5-min intervals. Flash glucose systems are compatible with the KNN-DSS algorithm; however these will require further testing to handle the asynchronous sample rates of flash glucose systems. Like flash glucose monitoring systems, fingerstick glucose is also measured at varying intervals (4-10 times per day), and more training and testing is needed before this could be incorporated into the KNN-DSS. The engine is also compatible with the large majority (~95%) of existing insulin therapies including fast-acting aspart or lispro and long-acting Lantus (glargine), Tresiba (degludec), and Toujeo (glargine U300) basal formulations. Intermediate-acting NPH insulin (which represents <5% of use cases) will require additional testing and evaluation. Although we report on adults, adolescents and children in this article, further work will be needed to assess performance in vulnerable and complex populations,

including the elderly and pregnant women. Moreover, virtual simulators have not been developed to fully represent these populations, which makes it challenging to incorporate these populations into the design. The KNN-DSS will need further training before it can be targeted to these groups.

We explored how specific glucose and insulin features were related to optimal recommendations calculated by the KNN-DSS algorithm (Supplementary Table 3.S6). We observed that the KNN-DSS mapping of specific glycaemic features to optimal recommendations matches intuition and, in general, is synonymous with physician opinions regarding titration of insulin dosing for people with T1D.

We have shown that an artificial intelligence decision support system with an expert-knowledge quality control algorithm can be used to help people with T1D identify problematic glycaemic patterns at the same level of accuracy as board-certified endocrinologists. Our unique *in silico* training platform enables us to generate training sets of diverse glycaemic profiles from the OHSU T1D simulator. The final engine design performs well on independent *in silico* virtual populations and, most notably, can identify insulin dosage issues in real-world human data. Further validation in longer clinical trials in humans is critical to understand how artificial intelligence-based decision support systems can improve glycaemic outcomes in people with T1D.

3.2 Methods

3.2.1 K-nearest-neighbours design

The K-nearest-neighbor classification algorithm (KNN) [70] is a supervised machine learning approach that matches input features with an outcome variable or class. We define the KNN input features as specific glycaemic outcomes, such as

percent time in target range (70-180 mg dL^-1), percent time spent in hypoglycaemia (<70 mg dL^-1), the number of meal- or correction-related hypoglycaemia episodes and so on (see Supplementary Table 3.S7 for complete list of features). The outcome variable or class that the KNN predicts is a recommended adjustment to insulin dosage that leads to an improved percent time in target range and a reduction in percent time in hypoglycaemia. This classification is accomplished using a look-up table that matches the weekly glycaemic features of a person with optimal dosing recommendations. The KNN approach identifies unique recommendations regarding long-acting basal insulin and fast-acting meal insulin. Recommendations regarding adherence and correction factors are accomplished separately by the ALPHA and Quality Control algorithms described in detail below (Figure 3.1).

3.2.2 Training Dataset generation

We generated the look-up table of optimal recommendations using an *in silico* virtual patient simulator consisting of 99 individuals with T1D exhibiting diverse glycaemic dynamics generated by variations in insulin sensitivity and daily insulin requirements, carbohydrate sensitivity, weight differences, circadian insulin sensitivity and variations in adherence to insulin dosing [57]. Circadian insulin sensitivity was achieved through continuous modulation of insulin-mediated peripheral glucose uptake, peripheral insulin uptake, and hepatic glucose production, varying within 20% of their original values [57]. CGM data and insulin data were obtained from 70 of the virtual patients over the course of a 15-week *in silico* study, while the other 29 patients were retained for an evaluation of algorithm performance in a separate *in silico* study. Real-world meal scenarios provided to the virtual test subjects were obtained from a previous clinical trial [28] in which we acquired 80 d of realistic meal patterns and carbohydrate content. Forty of these daily meal scenarios were used to train the KNN-DSS, while the

other 40 were used to validate the algorithm. These daily meal-scenarios were randomized and administered to virtual patients and the Pettus-Edelman [195] approach for CGM trend arrow adjustment and a bolus calculator was used to dose mealtime insulin. The optimal recommendations for an individual were identified by simulating the glycaemic outcomes from administering each dosing recommendation given in Table 3.1. The glycaemic outcomes of time in glucose target range (70-180 mg dL^-1), mean glucose, hypoglycaemia (<70 mg dL^-1), and serious hypoglycaemia (<54 mg dL^-1) were measured the week following each simulated dosing scenario. During the training of the KNN, these measured outcomes were used to define a heuristic objective function to select the optimal recommendations, {*R*_Opt}, to maximize the measured percent time in target range, and reduce percent time in hypoglycaemia (*X*_TIR, *X*_Hypo in equation (3.1), respectively). We provide this heuristic objective function in equation (3.1). In step 1, we identify a subset of recommendations, {*R*_Hypo}, that yields percent time in hypoglycaemia less than or equal to 2% and no serious hypoglycaemia. If the patients exhibit persistent hypoglycaemia, we identify the recommendations that minimize hypoglycaemia, {*R*_Hypo}. Recommendations that yield a percent time in hypoglycaemia greater than 2%, or any extreme hypoglycaemia, are excluded from the subset of optimal recommendations. In step 2, we identify from this subset the optimal recommendation, { *R*_Opt}, which yields the largest improvement in the time spent in target glucose range, or mean glucose if the percent time in target range is too small due to persistent hyperglycaemia.

$$R_{Opt} = \begin{cases} \textbf{\textit{Step 1a}}, & \{R_{Hypo}\} = \bigcup_{i=1}^{n} i * \delta(x_{hypo}) * \gamma(x_{serious\ hypo}) \quad \delta(a) = \begin{cases} 0, a > 2 \\ 1, a \leq 2 \end{cases} \quad \gamma(a) = \begin{cases} 0, a \neq 0 \\ 1, a = 0 \end{cases} \\ \textit{Step 1b}, & \{R_{Hypo}\} = \underset{X_{Hypo}}{\operatorname{argmin}} \sum_{i=1}^{n} x_i \quad \rightarrow perform\ if\ step\ 1a\ yeilds\ empty\ set \\ \textbf{\textit{Step 2a}}, & \{R_{Opt}\} = \underset{X_{TIR}}{\operatorname{argmax}} \sum_{i=R_{Hypo}} x_i \\ \textit{Step 2b}, & \{R_{Opt}\} = \underset{X_{MeanGlucose}}{\operatorname{argmin}} \sum_{i=R_{Hypo}} x_i \rightarrow perform\ if\ step\ 2a\ yeilds\ empty\ set \end{cases}$$

Eq. 3.1 Method of optimal recommendation identification during simulation

This optimal recommendation was stored with the glycaemic features from the prior week to form an observation in the look-up table. The observation of paired weekly glycaemic features and optimal insulin dosage were compiled into a look-up table used in the KNN algorithm. Additional real-world behavioral scenarios were imposed during the *in silico* study by programming the virtual patient to perform one or more of 13 insulin dosing errors (Supplementary Table 3.S4), to periodically accept or fail to accept the advice of the recommendation given each week and to use one of two different types of bolus calculators. The size of the final look-up table totaled 51,831 observations.

3.2.3 K-nearest neighbours parameter identification and feature selection

The features used in the KNN-DSS included features drawn from CGM, physical activity data, and long-acting and short-acting insulin data. We identified an optimal set of features through a feature selection technique called *'greedy'* sequential forward selection [196]. We determined the optimal number of neighbours by performing a grid search using the optimal set of features. The final KNN design used 30 neighbours and 25 glycaemic features to perform classification (see Supplementary Table 3.S7). The weighting scheme was decided by comparing classification accuracy across (1) no

weighting, (2) distance-based weighting, (3) class-based weighting, and (4) combined distance and class-based weighting (Supplementary Tables 3.S8-S9).

3.2.4 Feature Importance

We evaluated which features contributed most significantly to each recommendation by calculating the '*mutual information'* between each feature and the recommendation selected by the classifier. The mutual information between two random variables (for example, a feature *f* with a distribution F and a recommendation *r* with a distribution R) with a joint probability mass function P_FR) is defined according to equation (3.2).

$$I(F,R) = \sum_{f,r} P_{FR}(f,r)\log\left(\frac{P_{FR}(f,r)}{P_F(f)P_R(r)}\right)$$

Eq. 3.2 Mutual information criterion

The mutual information of features was calculated for each separate recommendation in the classifier. A one-vs-all approach was used, in which the mutual information considers a single recommendation class as positive and all other classes as negative. This was repeated for each recommendation class in order to generalize what features contribute most significantly to a given recommendation class. Relative feature importance determined for each recommendation is listed in Supplementary Table 3.S6.

3.2.5 Precise Insulin Titration

The KNN-DSS assumes that meal insulin doses are calculated using carb:ins ratio, and correction boluses are calculated correction factors and glucose trends [195, 197]. The also uses smart-bolus calculations that incorporate meal-time corrections and active insulin on board (IOB) (equation (3.3)).

$$\begin{cases} Recommended\ Insulin\ Bolus = Bolus_{meal} + Bolus_{correction} - IOB \\ Bolus_{meal} = \ Grams\ of\ Carbohydrate\ \ x\ \frac{1\ unit\ of\ insulin}{carb{:}ins} \\ Bolus_{correction} = (Current\ Glucose + Offset - \ Target\ Glucose)\ x\ \frac{1\ unit\ of\ insulin}{correction\ factor} \end{cases}$$

Eq. 3.3 Insulin bolus calculation

Precise titration of insulin is accomplished using heuristic approaches that adjust carb:ins ratios and correction factors. The ALPHA algorithm, described in our recent publication [190], retrospectively analyses the average glycaemic response of a person to insulin boluses and returns an aggressiveness factor (*A_f*). The framework of this algorithm, as well as the modifications to the original algorithm for the current paper implementation, is described as follows. Separate analysis is performed for both meal-related boluses and correction boluses, again across different windows of time (see Table 3.1). For each bolus entry, ALPHA adapts the aggressiveness factor if the glucose of the person is outside of target range following meals or corrections. The aggressiveness factors corresponding to each bolus are then used to calculate an average, smoothed aggressiveness factor (*A_f^Avg*) shown in equation (3.4). This smoothed aggressiveness factor is used to adjust carb:ins ratio and correction factor settings. ‘

$$A_f^{Avg}(k) = \frac{A_f(k) + A_f^{Avg}(k-1) + A_f^{Avg}(k-2)}{3}$$

Eq. 3.4 ALPHA aggressiveness factor

In the implementation discussed herein, the aggressiveness factor assigned to an individual bolus delivered at time k, *A_f* (k) is determined using a piece-wise linear adjustment that is a function of the minimum glucose (*G_min*) measured within 4 h of the last meal bolus or within 3 h after the last correction bolus (equation (3.5)).

$$A_f(k) = \begin{cases} 0.4, & 0 \le G_{min} \le G_{hypo} \\ \frac{\left(G_{min} - G_{hypo}\right) * \left(A_f^{Avg}(k-1) - 0.4\right)}{G_{eug-lower} - G_{hypo}} + 0.4, & G_{hypo} \le G_{min} \le G_{eug-lower} \\ A_f^{Avg}(k-1), & G_{eug-lower} \le G_{min} \le G_{eug-upper} \\ \frac{\left(G_{min} - G_{eug-upper}\right) * \left(1.3 - A_f^{Avg}(k-1)\right)}{G_{hyper} - G_{eug-upper}} + A_f^{Avg}, & G_{eug-upper} \le G_{min} \le G_{hyper} \\ 1.3, & G_{hyper} \le G_{min} \end{cases}$$

Eq. 3.5 Point estimate of ALPHA aggressiveness factor

If *G_min* is within the target range of *G_eug-lower* = 90 to *G_eug-upper* = 140 mg dL^-1 (where *G_eug-lower* and *G_eug-upper* are the lower and upper limits of a euglycaemic glucose target range, respectively), then the aggressiveness factor does not change and *A_f = A_f^Avg* . However, if *G_min* drops below *G_eug-lower*, the aggressiveness factor, *A_f* , is reduced proportionally down to a hypoglycaemia threshold of *G_hypo* (70 mg dL^-1). Below the hypoglycaemia threshold (*G_hypo*), *A_f* = 0.4 which means that the pre-meal insulin will be dosed at 40% of the original amount as shown in equation (3.5). In a similar manner, the aggressiveness factor is increased proportionally with respect to *G_min* if *G_min* is above the upper limit of target range (*G_eug-upper*) until *G_min* exceeds the hyperglycaemic threshold (*G_hyper*). Above *G_hyper*, the value of the insulin aggressiveness factor is 1.3. To ensure that the aggressiveness factor accurately reflects user glycaemic response, a minimum of 5 boluses must be observed within a specific window of time before a new aggressiveness factor is calculated.

Adjustment of carb:ins and correction factor occurs under two separate scenarios. The ALPHA algorithm is used to calculate the precise dosage adjustment to carb:ins only when indicated by the KNN classification procedure. In contrast, the ALPHA algorithm adjusts correction factors directly because the KNN classification does not address correction factors. The maximum adjustment to fast-acting insulin is constrained to ±15% per week. For adjustments to long-acting basal insulin, the dosage

is adjusted by ±10% per week when indicated by the KNN classification procedure. The KNN-DSS system accounts for the extended (> 5 d) pharmacologic steady state of ultra-long-acting Tresiba (degludec) and Toujeo (glargine U300) by constraining basal recommendations to one basal insulin recommendation every 2 weeks (see Extended Data 3.E2-E3). These constraints are a safety measure. Titrations to basal insulin are applied uniformly to all basal doses that may occur at different times of day. For example, for people who require twice-daily basal insulin injections, if a 10% reduction in basal insulin is recommended, both the morning and evening insulin will be reduced by 10%.

3.2.6 Quality Control Algorithm

A quality control algorithm was developed to ensure that recommendations delivered to the person adhere to physician standards. This algorithm incorporates expert knowledge based on physician input on titration of basal and bolus insulin regimens to ensure that engine recommendations are consistent with physician standards and are safe for a person with T1D. Quality control metrics for each recommendation delivered by the KNN-DSS are shown in Figure 3.1d and elaborated further in Extended Data 3.E1-E7.

3.2.7 Clinical Study Data and Physician Review

Data were obtained from 25 people with T1D who participated in a 4-week, outpatient study of CGM-augmented MDI therapy (Supplementary Table 3.S1: Study 1). After the data collection techniques were optimized on the first 9 participants, the remaining 16 participants received recommendations for dosing and behavioral changes on the basis of suggestions from the KNN-DSS system. At the end of each study week, an endocrinologist reviewed the data and identified one or more adjustments to insulin

therapy. Study participants were equipped with a Dexcom G6 sensor and Apple Watch to track glucose trends and physical activity. Participants used either long-acting Lantus or Tresiba insulin that was captured automatically using the Bluetooth enabled Gocap or Clipsulin insulin dose-capture devices. Participants used fast-acting Novolog (aspart) insulin captured during the study automatically with an InPen device. Participants were instructed to log meals and exercise using a custom food and exercise tracking app, and to dose insulin using the InPen app. Out of the 25 participants, 15 were female, the mean weight was 82.73 +/- 19.56 kg, and mean height 170.60 +/- 19.56 cm. The mean duration of diabetes was 15.52 +/- 6.92 years, mean age was 30.50 +/- 5.92 years, and mean A1c was 8.78 +/- 1.36 %. Additional information regarding study population characteristics, recruitment, and ethics oversight can be found in the Reporting Summary.

The study concluded with a total of 78 physician review sessions accounting for over 500 d of data from 25 participants. For each week of study data, physicians were instructed to identify one or more adjustments to insulin therapy from a set of 12 potential recommendations (Table 3.1). Data obtained during this clinical study were retrospectively analysed by the KNN-DSS to generate recommendations and calculate physician agreement metrics according to Supplementary Table 3.S10.

3.2.8 Safety Review

Recommendations generated by the KNN-DSS for human participants underwent safety evaluation by faculty at the Department of Endocrinology, Harold Schnitzer Diabetes Health Center. Safety of the KNN-DSS recommendations was assessed by having the physician determine whether the recommendation had the potential to cause hypoglycaemia or night-time hypoglycaemia events. A total of 100 safety reviews were performed.

3.2.9 Missing Data

At least four days of cumulative CGM data are required by the KNN-DSS framework to provide new recommendations. The KNN-DSS engine framework will refrain from providing a recommendation until sufficient CGM data are present. To address issues of missing data and data misclassification of insulin boluses that are common in real-world datasets, we developed an auxiliary insulin bolus estimation tool. Insulin boluses recorded by the Bluetooth-enabled insulin capture devices were first paired to announced meal entries. Boluses that were not within 20 min of an announced meal were counted as unlabeled boluses. For each unlabeled bolus, we evaluated the glucose level, glucose trend, and correction factor setting at the time of bolus administration. We then estimated the correction insulin dose that would have been called for by inputting this information into the Scheiner trend adjustment calculator [197]. Any remaining units of insulin are counted as a meal bolus. These estimated contributions are then combined with existing insulin data and are input to the algorithm.

3.2.10 Assessment of the Accuracy of KNN-DSS Recommendations as Compared to Endocrinologist Recommendations

Glycaemic outcomes and insulin dosing behaviors were analysed by one of three board-certified endocrinologists and by the KNN-DSS. Recommendations delivered by the KNN-DSS were compared to physician recommendations for each week of data collected during the clinical study. Similarity between KNN-DSS recommendations and physician recommendations were calculated using a modified Sorensen-Dice similarity coefficient[25].

$$DSC = \frac{\left|R_{Engine} \cap R_{Physician}\right|}{\left|R_{Engine}\right|}$$

Eq. 3.6 Modified Sorensen-Dice similarity

In equation (3.6), the similarity between physician and engine recommendations is calculated as the number of recommendations common to both sets, divided by the total number of recommendations delivered by the engine.

Recommendations were classified into one of three categories: '*agreement'*, '*disagreement'*, or '*additional treatment'* (Supplementary Table 3.S10). *Agreement* refers to an engine recommendation that was in full agreement with the physician recommendation (a perfect categorical match), or that was in partial agreement with the physician recommendation and titrates insulin in the same direction (for example, different categorical recommendations that both increase insulin). *Disagreement* refers to an engine recommendation that was in full disagreement with the physician (for example, one recommendation increases basal insulin and the other decreases basal insulin), or that partially disagrees with the physician and titrates overall insulin in a different direction (for example, one recommendation increases meal insulin and the other decreases basal insulin). *Additional treatment* refers to a scenario in which the engine recommended insulin dosage adjustments, but the physician indicated no change to the settings of the study participant, and vice-versa. Short-acting insulin bolus treatments reflect a 4 h pharmacokinetic activity; therefore insulin doses in adjacent treatment windows are highly correlated and are considered in partial agreement. A behavioral recommendation to be more adherent to a dosing regimen was counted as safe and in *agreement*. In some scenarios where the engine recommended to use the bolus calculator and the physician recommended to increase or decrease basal insulin, the recommendations are not comparable. The overall accuracy was calculated by averaging the similarity across all recommendations.

3.2.11 Inter-physician Recommendation Agreement

Using a dataset [194] collected during a 1-month outpatient clinical study of open-loop insulin therapy, three board-certified endocrinologists separately reviewed participant CGM data and dosing behaviors and recommended one or more adjustments to insulin therapy (Supplementary Table 3.S1: Study 2). The physicians then collectively reviewed participant data to reach a consensus on what recommendation should be given. The Sorensen-Dice coefficient (equation (3.6)) was then used to determine the agreement between individual physician recommendations, as well as the accuracy of physicians' recommendations, as compared to the consensus (Supplementary Table 3.S2). A summary of dataset description and usage is available in Supplementary Table 3.S1.

3.2.12 *In silico* evaluation

We evaluated the KNN-DSS during two *in silico* studies. In the first study, 29 virtual patients from the OHSU T1D simulator participated in a 75-week study in which the virtual patients used the decision support system weekly to adjust doses to basal insulin, mealtime insulin, and correction insulin. After 52 weeks, we changed certain insulin settings and dosing behaviors and monitored the ability of the engine to recover these settings. In the second study, 100 virtual patients from the UVA-Padova simulator participated in a similar 52-week study of engine usage. To simulate an MDI population using the UVA-Padova simulator, we replaced the default time-varying basal rate, which is characteristic of programmable insulin pumps, with a constant basal dosage that could be titrated weekly by the KNN-DSS.

Virtual patients from both studies exhibited inter-individual variations in weight, total daily insulin requirement, and circadian insulin sensitivities as described above. For

both the OHSU T1D and UVA-Padova simulators, patients were fed real-world meal scenarios (Supplementary Table 3.S1: Study 3) that ranged from 2-9 meals per day and occurred at varying intervals. In addition, we imposed errors in insulin dosing settings and adherence to dosing strategies (see Supplementary Table 3.S4), as well as statistical variation in estimation of meal amounts to reflect realistic use of bolus calculators, interstitial glucose CGM measurement noise and circadian variation in insulin sensitivity to reflect realistic glycaemic profiles. We evaluated study outcomes of percent time in target range and percent hypoglycaemia at time points of 1 week, 1 month, 3 months as well as at the study conclusion.

The virtual patients used for evaluation and a subset of the real-world meal scenarios were excluded from the training process. In this way, performance was analysed on virtual people with T1D that had not been observed before by the KNN-DSS.

3.2.13 Analysis and Statistical Power

The glycaemic outcomes of percent time in target range and percent time in hypoglycaemia were determined for the virtual patients at each time point of the study. The percent change was calculated across each week of the study compared with the first week of the study, before any recommendations were given. Results are reported by mean and s.d. for normally distributed outcomes, and median and interquartile range for non-parametric data. A students two-tailed, paired t-test of alpha = 0.05 was used to determine the significance in the change of glycaemic outcomes, and a two-tailed Wilcoxon signed-rank test was used to determine significance for non-parametric data. Cohen's d effect size for paired samples was calculated to account for the influence of the *in silico* framework [198] and large sample sizes on p-value statistics.

3.3 Article Information

3.3.1 Use of Human Subjects

All participants were adults enrolled under informed consent. The pilot study was approved by the Institutional Review Board at OHSU, and additional information can be found at clinicaltrials.gov under registration number NCT03443713.

3.3.2 Data Availability

The data generated *in silico* during this study and the code used for analysis is available from the corresponding author on request. Access to human participant data was granted for the current study, and further human data usage or sharing is subject to restrictions and is not publicly available. Requests for restricted, de-identified data on human participants can be submitted to the corresponding authors at OHSU. Requests will be assessed on a case-by-case basis, and are subject to a formal Repository Sharing Agreement. Additional reported outcomes of human participants can be found at clinicaltrials.gov under registration number NCT03443713.

3.3.3 Code Availability

The code used to generate *in silico* data for this study, the OHSU virtual patient population simulator code, is available at https://github.com/petejacobs/T1D_VPP. Access to the licensed software for the UVA-Padova virtual population was granted for the current study and it can be requested from the developers of this software directly at the University of Padova.

3.4 Extended and supplementary data

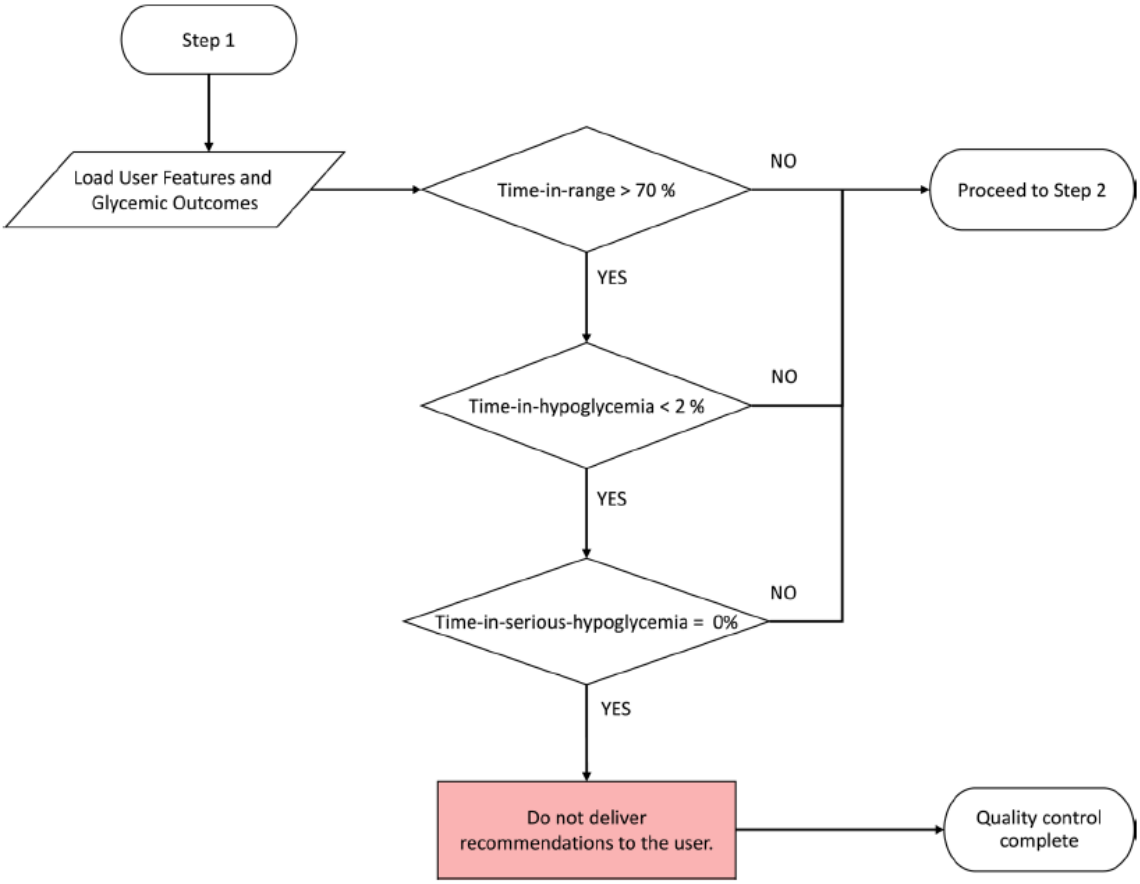


Figure 3.E1: Extended Data. Quality control algorithm to assess need for insulin titration.

Quality control algorithm to assess need for insulin titration. User data and glycaemic outcomes are loaded and compared against metrics for percent time in hypoglycaemia, percent time in target range, and percent time in serious hypoglycaemia. If users meet all metrics, recommendations for insulin titration are not required.

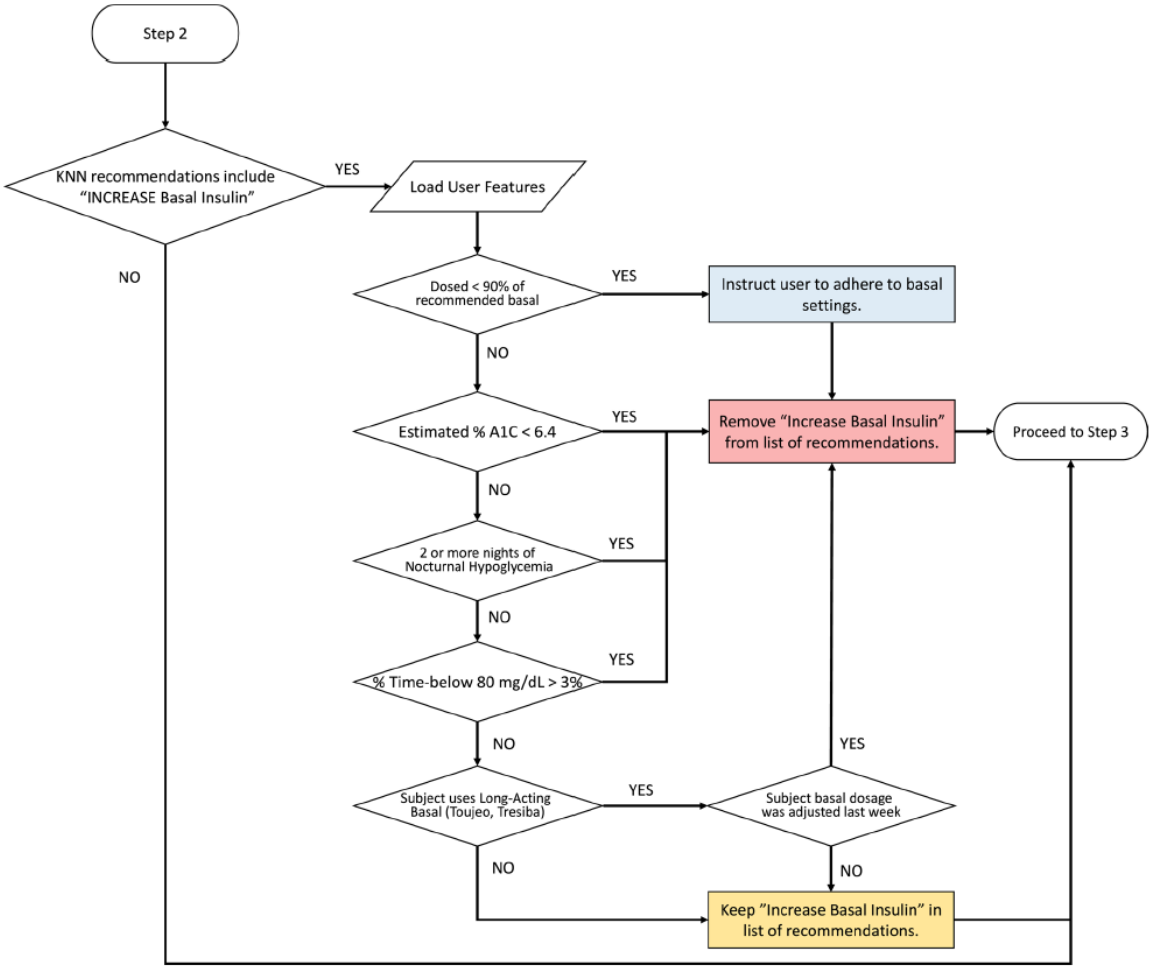


Figure 3.E2: Extended Data. Quality control algorithm to assess increasing basal insulin dosage

Quality control algorithm to assess increasing basal insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of nocturnal hypoglycaemia, near hypoglycaemia episodes, subject time in target range, subject adherence, and insulin formulation-dependent requirements.

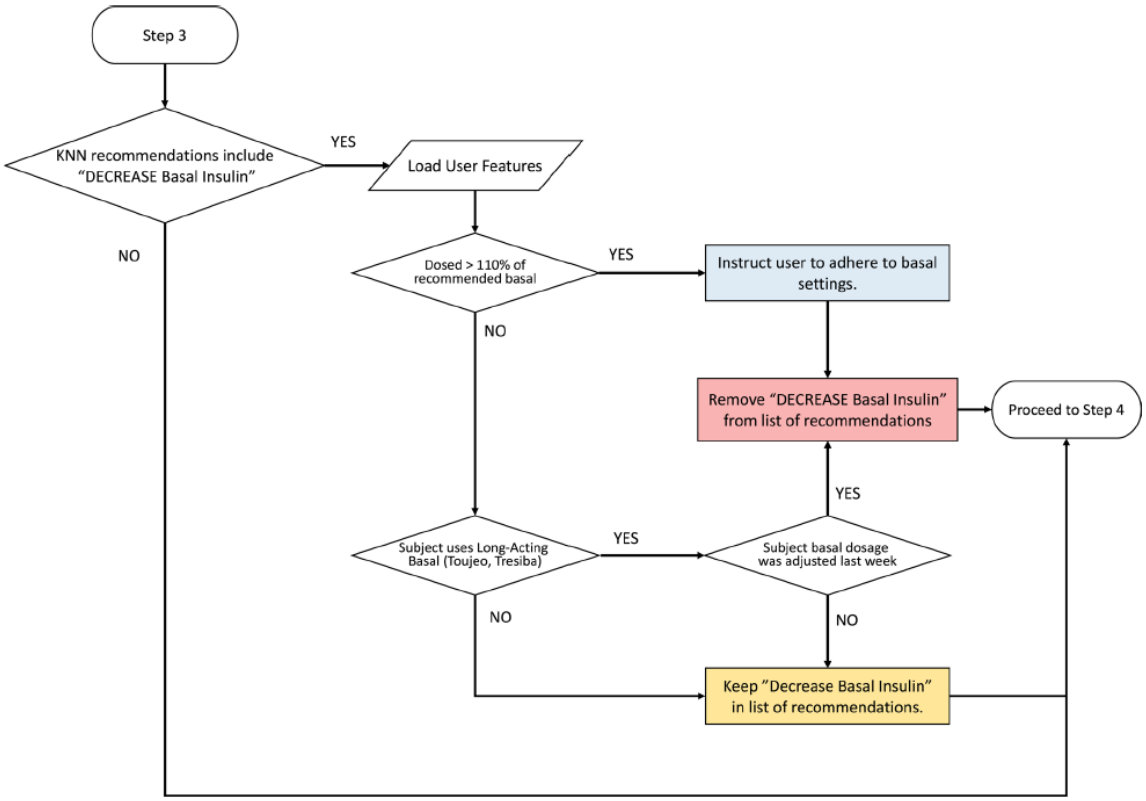


Figure 3.E3: Extended Data. Quality control algorithm to assess decreasing basal insulin dosage.

Quality control algorithm to assess decreasing basal insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for subject adherence, and insulin formulation-dependent requirements.

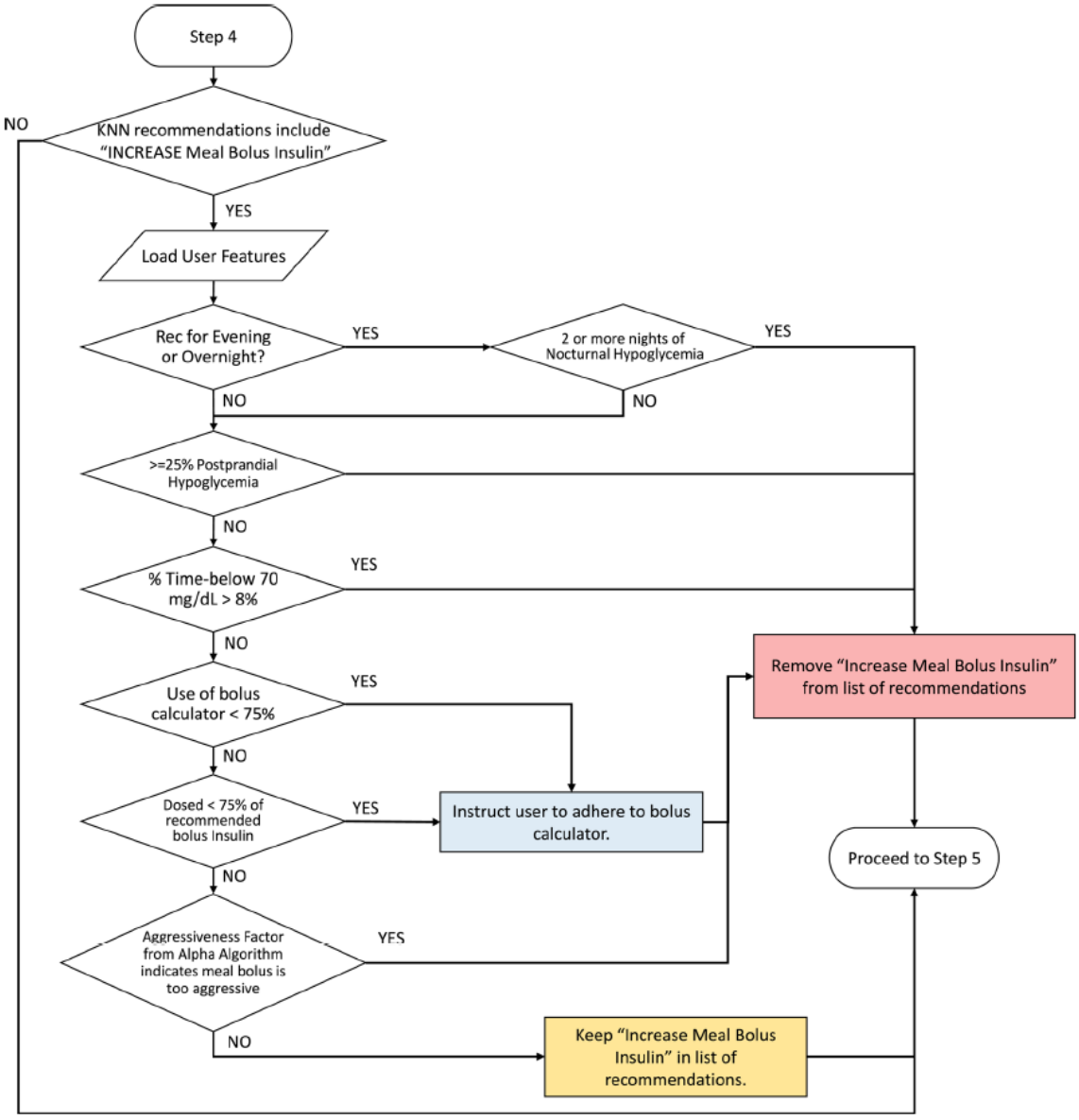


Figure 3.E4: Extended Data. Quality control algorithm to assess increasing meal bolus insulin dosage.

Quality control algorithm to assess increasing meal bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial hypoglycaemia, subject adherence, and factors returned by the ALPHA algorithm.

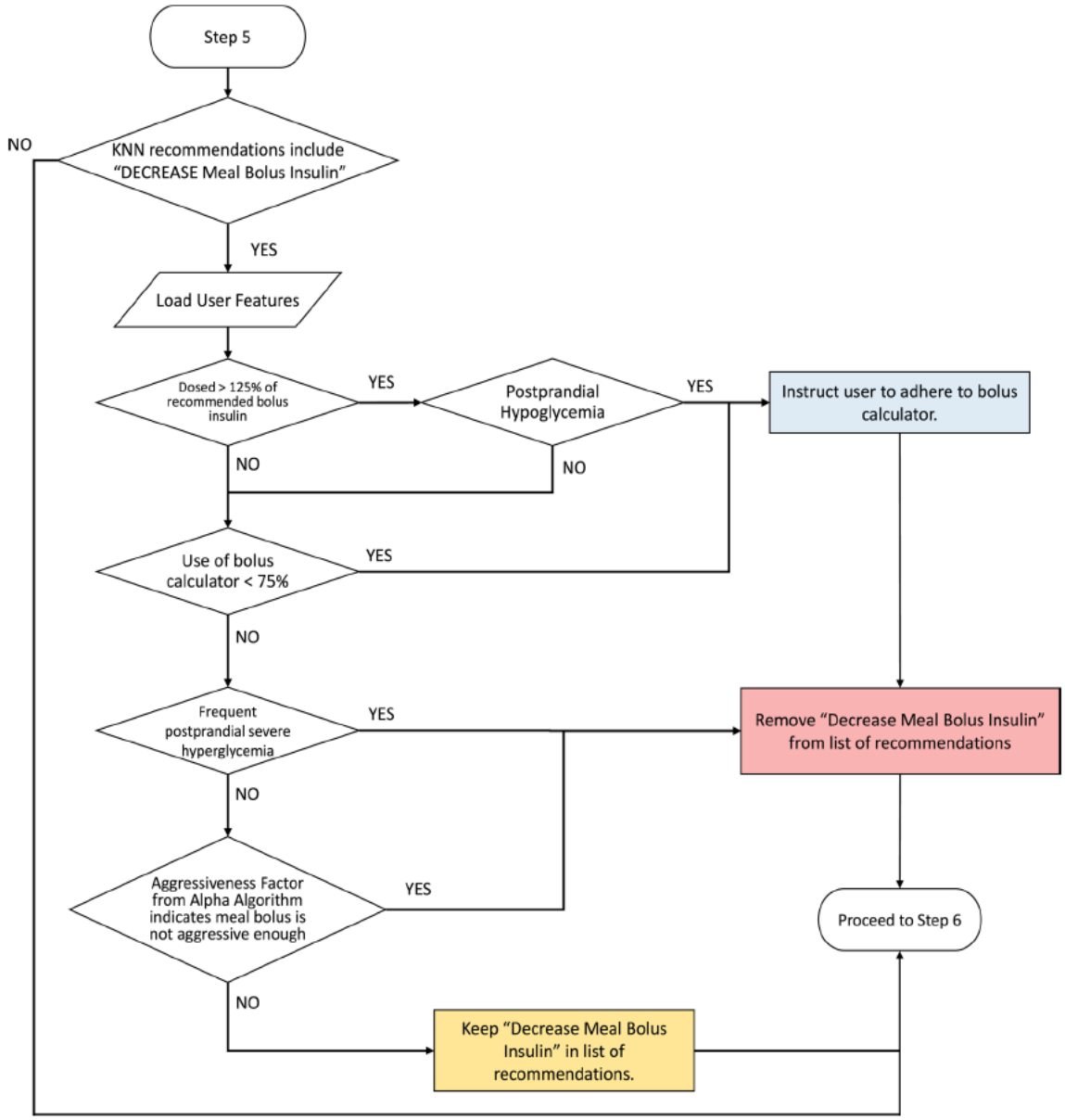


Figure 3.E5: Extended Data. Quality control algorithm to assess decreasing meal bolus insulin dosage.

Quality control algorithm to assess decreasing meal bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial severe hyperglycaemia, subject adherence, and factors returned by the ALPHA algorithm

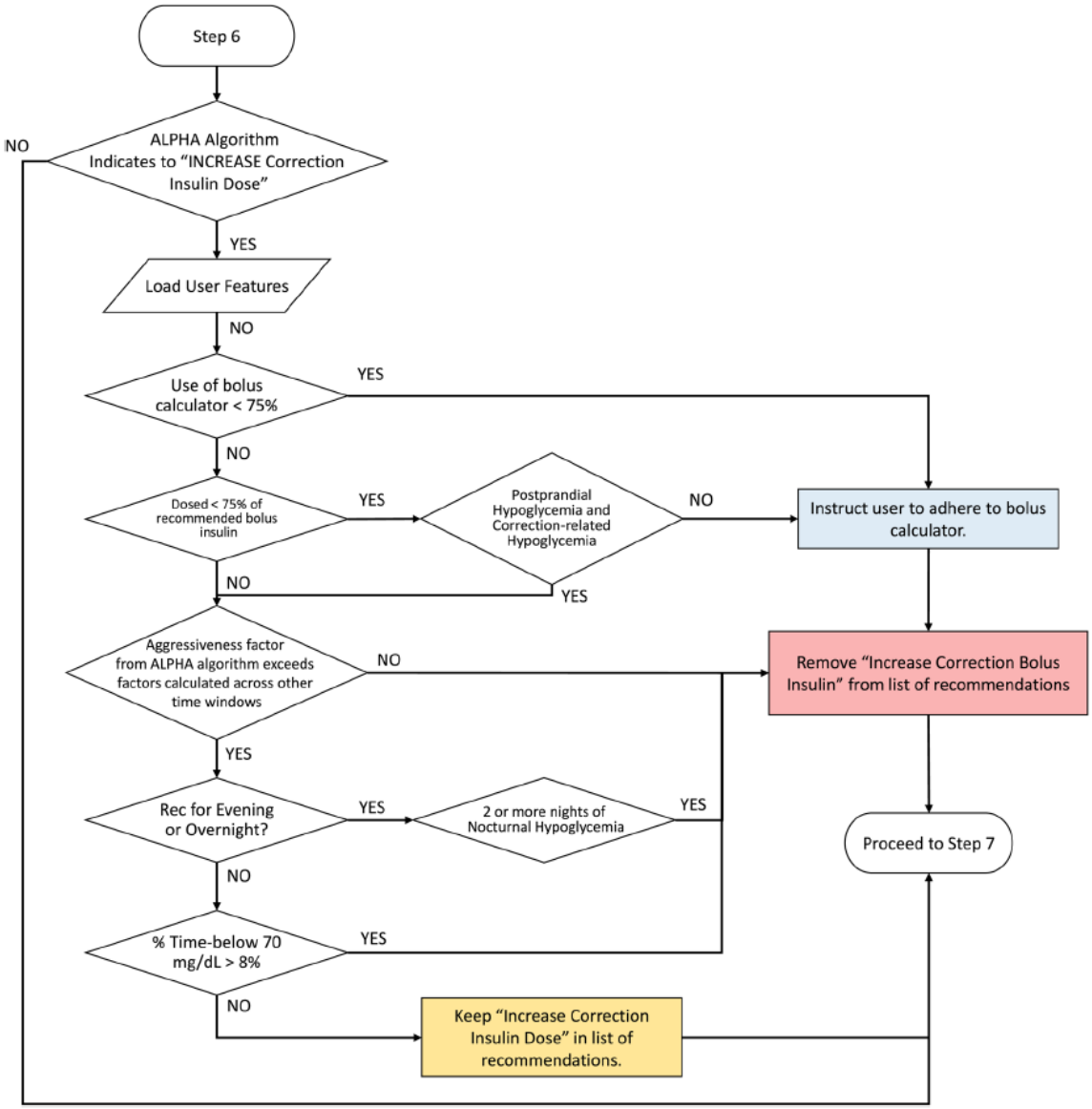


Figure 3.E6: Extended Data. Quality control algorithm to assess increasing correction bolus insulin dosage.

Quality control algorithm to assess increasing correction bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial and correction-related hypoglycaemia, subject adherence, and factors returned by the ALPHA algorithm.

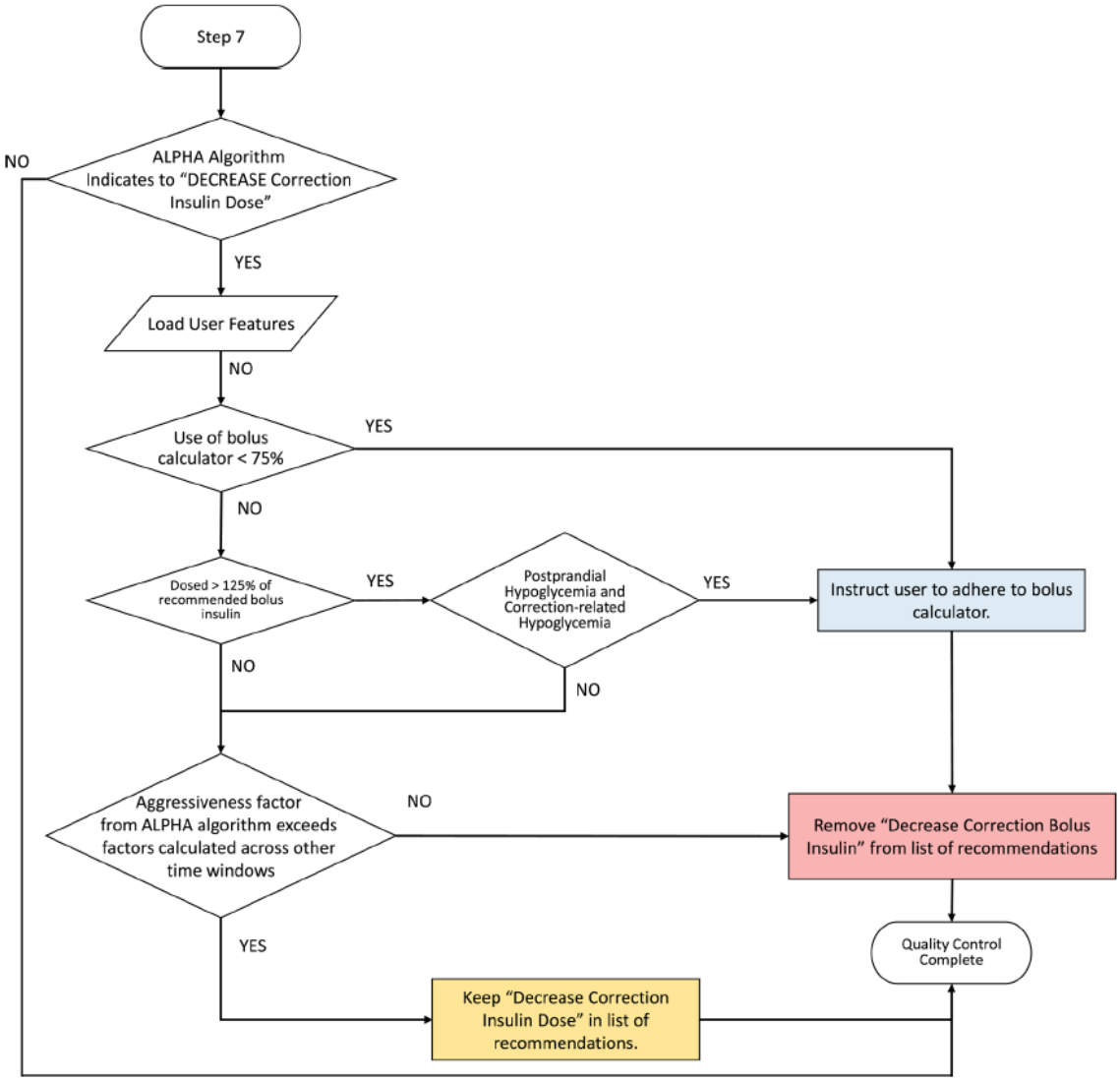


Figure 3.E7: Extended Data. Quality control algorithm to assess decreasing correction bolus insulin dosage.

Quality control algorithm to assess decreasing correction bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of subject adherence, postprandial and correction-related hypoglycaemia, and factors returned by the ALPHA algorithm.

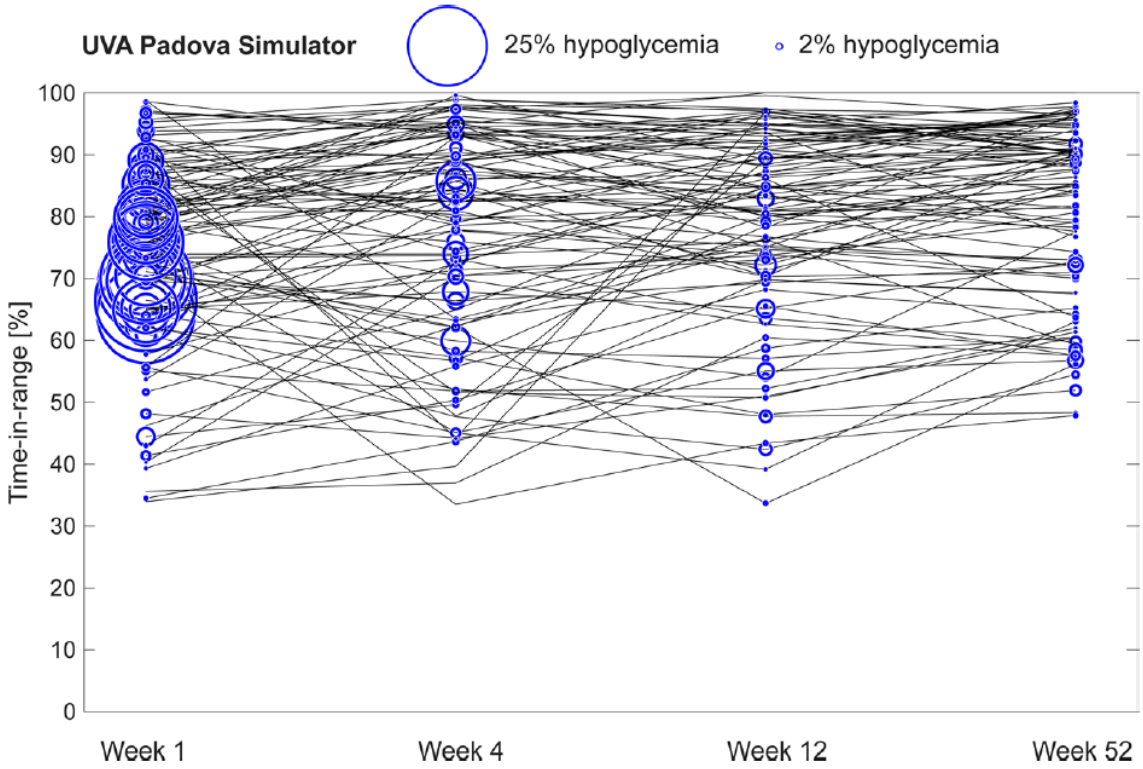


Figure 3.E8: Extended Data. KNN-DSS engine performance in improving subject outcomes in an independent virtual patient population.

KNN-DSS engine performance in improving subject outcomes in an independent virtual patient population. Glycaemic outcomes during a 52-week study of the FDA-approved UVA-Padova virtual patient simulator. Percent time in hypoglycaemia is indicated by the blue circular radius.

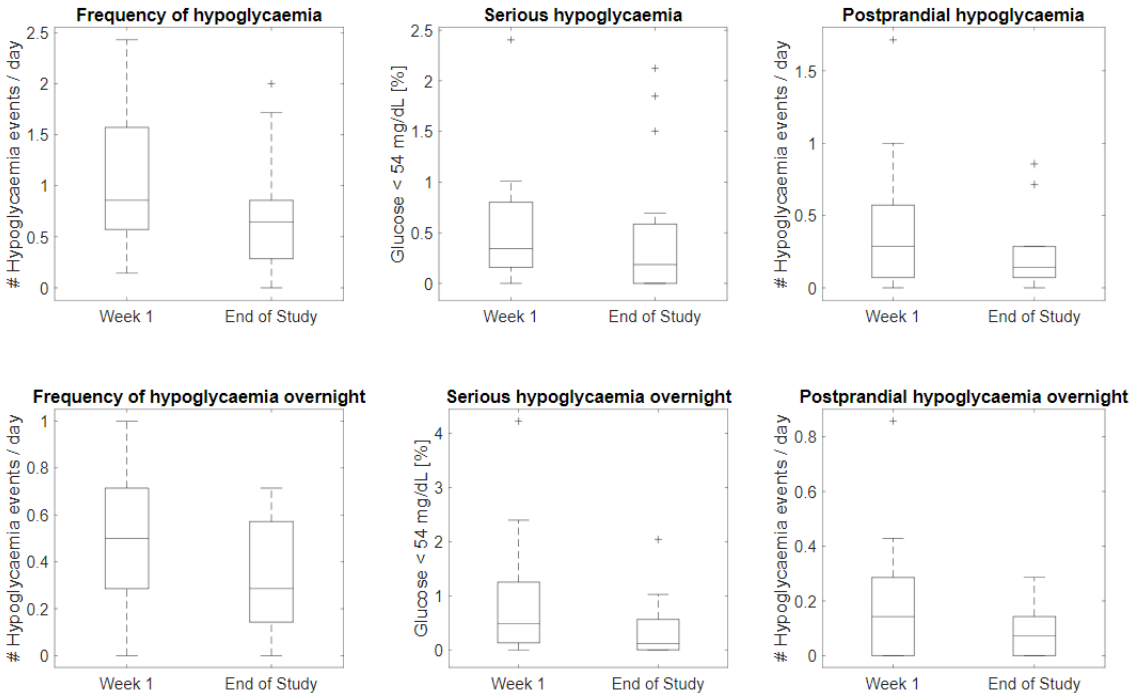


Figure 3.E9: Extended Data. Outcomes of a human pilot study evaluating KNN-DSS augmented decision support.

Outcomes of a human pilot study evaluating KNN-DSS augmented decision support over 4 weeks where the first recommendation is given at the start of week 2. For panels a-f, boxplot limits indicate the first and third quartiles, centerline indicates the median, and whiskers mark the last non-outlier data-point within 1.5xIQR. For panels a-f, participant data collected during week 1 and the final week of the study were compared using a two-tailed Wilcoxon signed-rank test, with significance level of 5%. a, Frequency of hypoglycaemia was nominally reduced on the final week compared to week 1 of the study (0.86 vs 0.64, P = 0.051, n = 16 independent subjects). b Serious hypoglycaemia was nominally reduced on the final week compared with week 1 of the study (0.34% vs. 0.19%, P = 0.56, n = 16 independent subjects). c Postprandial hypoglycaemia events were nominally reduced on the final week compared with week 1 (0.29 vs 0.14, P = 0.08, n = 16 independent subjects). d Frequency of overnight hypoglycaemia was significantly reduced on the final week compared to week 1 (0.50 to 0.29, P= 0.04, n = 16 independent subjects). e Serious hypoglycemia overnight was significantly reduced on the final week compared to week 1 (0.48% to 0.11%, P = 0.03, n = 16 independent subjects). f Postprandial hypoglycemia overnight was nominally reduced on the final week compared to week 1 (0.14 to 0.07, P = 0.06, n = 16 independent subjects).

Table 3.S1: Supplementary Table. Description of real-world human datasets used for KNN-DSS engine evaluation.

Dataset name	Purpose of dataset use and analysis	Description of data	Study details
Study 1	To compute agreement between KNN-DSS recommendations and endocrinologist recommendations	Continuous glucose monitor Bluetooth-enabled insulin pen Activity monitor device Exercise activity log Meal content and size User adherence to bolus calculations	*Participants:* 25 adults with type 1 diabetes who undergo multiple daily injection therapy. *Design:* Subjects participated in a 28-day study of multiple injection therapy to evaluate a new mobile app interface and collect data reflecting CGM-augmented MDI therapy. Physicians individually reviewed data from separate participants on a weekly basis and recommended changes to insulin therapy. Fifteen of the participants were given physician-reviewed recommendations provided by the KNN-DSS.
Study 2	To compute inter-physician variability in recommendations	Continuous glucose monitor Insulin pump infusion rate Insulin bolus calculations Activity monitor device Exercise announcements Meal content	*Cited:* Reddy et. al [194] *Participants:* 10 adults with type 1 diabetes undergoing open-loop insulin pump therapy. *Design:* Subjects participated in a 30-day study evaluating overnight glycemic response to aerobic and anaerobic exercise.

		Retrospective physician recommendations for insulin titration	Following completion of the study, a group of physicians individually and collectively reviewed the all participant data and identified recommended changes to insulin therapy.
Study 3	To design real-world meal scenarios to evaluate use of the KNN-DSS *in silico*	Meal scenarios including carbohydrate content and time of consumption	*Cited:* Castle et. al [28] *Participants:* 20 adults with type 1 diabetes previously utilizing open-loop insulin pump therapy. *Design:* Subjects participated a 4-arm study evaluating closed-loop automated insulin delivery algorithms. Each arm of the study lasted 4 days. Participant recorded meal scenarios and rescue carbs throughout the study.

Table 3.S2: Supplementary Table. Agreement of recommendations delivered by individual endocrinologists.

Agreement is calculated using the Sorenson-Dice coefficient similarity between physician recommendations for each segment of participant data (Equation 3.6).

Recommendation Comparison	% Agreement	% Disagreement	% Additional	% Not comparable
Calculated Agreement between Endocrinologists using Real-world Human Data				
Assessing agreement between individual physicians	*Full* 41.2 *Partial* 14.7 Overall **55.9**	*Full* 0.3 *Partial* 2.8 Overall **3.1**	**32.6**	**8.4**
Assessing agreement between physicians and a consensus recommendation	*Full* 63.3 *Partial* 6.8 Overall **70.1**	*Full* 0.1 *Partial* 0.6 Overall **0.7**	**24.1**	**5.1**

Table 3.S3: Supplementary Table. Evaluation of KNN-DSS *in silico*.
Simulation of engine use was performed using two virtual patient populations with outcomes determined at weeks 1, 4, 12, and the end of the virtual study. Target range is defined as glucose measured between 70-180 mg dL^-1. Effect sizes larger than d = 0.5 are considered moderate, while effect sizes greater than d = 0.8 are considered substantial. Degrees of freedom in statistical tests are df = 99 and df = 28 for the UVA simulator and OHSU T1D simulator, respectively. A students two-tailed, paired t-test of alpha = 0.05 was used to compare percent time-in-range from the start of the study to week 4, 12, and study completion. A two-tailed Wilcoxon signed-rank test was used to compare percent time-in-hypoglycemia from the start of the study to week 4, 12, and study completion.

***In silico* Evaluation**				
OHSU Simulator (N = 29)	**Baseline**	**Week 4**	**Week 12**	**End of Study**
% time-in-range ***(p-value, d = effect size, CI = confidence interval)***	59.5%± 22.8	68.6 ±17.2% *(p=2E-3, d = 0.64, CI = [3.7,14.5])*	79.8 ± 7.3% *(p=2E-5, d = 0.96, CI = [12.3,28.4])*	81.1 ± 6.8% *(p=9E-6, d = 1.01, CI = [13.4,29.8])*
% time-in-hypoglycemia	0.74 [0.04, 2.88] %	0.69 [0.50, 1.35]% *(p= 0.24, d = 0.41)*	1.14 [0.63, 1.41]% *(p= 0.79, d = 0.28)*	0.89 [0.66, 1.35]% *(p= 0.69, d = 0.32)*
UVA Simulator (N = 100)	**Baseline**	**Week 4**	**Week 12**	**End of Study**
% time-in-range	75.1% ± 16.1	77.2% ± 17.6 *(p= 2E-1, d = 0.14, CI = [-0.7,5.0])*	79.6% ± 15.4 *(p= 1E-2, d = 0.25, CI = [0.9,8.0])*	81.8% ± 13.9 *(p= 1E-4, d = 0.39, CI = [3.3,9.9])*
% time-in-hypoglycemia	4.0 [1.1, 13.9] %	1.22 [0, 2.2] % *(p=1E-12, d = 0.79)*	0.79 [0, 1.6] % *(p=2E-12, d = 0.78)*	0.55 [0, 1.2] % *(p=2E-12, d = 0.81)*

Table 3.S4: Supplementary Table. Imposed errors in insulin dosing.

* indicates that the imposed error was not used during training-set generation, but was used during KNN-DSS engine validation.

Insulin Dosing Scenario	Implementation
No Dosing Error	No change to user settings
Postprandial Hypoglycemia	Dosing 20% more insulin than recommended
Severe Postprandial Hypoglycemia	Dosing 40% more insulin than recommended
Postprandial Hyperglycemia	Dosing 20% less insulin than recommended
Severe Postprandial Hyperglycemia	Dosing 40% less insulin than recommended
Aggressive Basal	Dosing 40% more basal than recommended
Insufficient Basal	Dosing 40% less basal than recommended
Aggressive Basal and Insufficient Bolus	Dosing 40% more basal than recommended, and 40% less meal bolus than recommended
Aggressive Basal with Aggressive Bolus	Dosing 40% more basal than recommended, and 40% more meal bolus than recommended
Insufficient Basal with Insufficient Bolus	Dosing 40% less basal than recommended, and 40% less meal bolus than recommended
Insufficient Basal with Aggressive Bolus	Dosing 40% less basal than recommended, and 40% more meal bolus than recommended
Insufficient Basal with Aggressive Correction Bolus	Dosing 40% less basal than recommended, and 40% more correction bolus than recommended
Insufficient Basal with Insufficient Correction Bolus	Dosing 40% less basal than recommended, and 40% less correction bolus than recommended
Adherence to bolus calculators *	Dosing less than or more than indicated by the bolus calculator, by a value of 30%.

Table 3.S5: Supplementary Table. Results of *in silico* validation using the UVA-Padova simulator.

* indicates one subject was excluded due to computational errors in the simulation. Simulation of engine use was performed with outcomes determined at weeks 1, 4, 12, and the end of the virtual study. Target range is defined as glucose measured between 70-180 mg dL^-1. Effect sizes larger than d = 0.5 are considered moderate, while effect sizes greater than d = 0.8 are considered substantial. A students two-tailed, paired t-test of alpha = 0.05 was used to compare percent time-in-range from the start of the study to week 4, 12, and study completion. A two-tailed Wilcoxon signed-rank test was used to compare percent time-in-hypoglycemia from the start of the study to week 4, 12, and study completion.

***In silico* Evaluation**				
ADULTS UVA Simulator (N = 100)	**Baseline**	**Week 4**	**Week 12**	**End of Study**
% time-in-range ***(p-value, d = effect size, CI = confidence interval)***	75.1% ± 16.1	77.2% ± 17.6 *(p= 2E-1, d = 0.14, CI = [-0.7,5.0])*	79.6% ± 15.4 *(p= 1E-2, d = 0.25, CI = [0.9,8.0])*	81.8% ± 13.9 *(p= 1E-4, d = 0.39, CI = [3.3,9.9])*
% time-in-hypoglycemia	4.0 [1.1, 13.9]	1.22 [0,2.2] *(p=1E-12, d = 0.79)*	0.79 [0,1.6] *(p=2E-12, d = 0.78)*	0.55 [0,1.2] *(p=2E-12, d = 0.81)*
ADOLESCENTS UVA Simulator (N = 100)	**Baseline**	**Week 4**	**Week 12**	**End of Study**
% time-in-range ***(p-value, d = effect size, CI = confidence interval)***	68.2 ± 13.2	73.9 ± 12.7 *(p= 2E-12, d = .80, CI =[4.3, 7.1])*	75.2 ± 12.2 *(p= 8E-9, d = 0.63, CI = [4.8, 9.2])*	77.4 ± 12.9 *(p= 2E-10, d = 0.71, CI =[6.6,11.7])*
% time-in-hypoglycemia	6.27 [1.7,12.1]	2.31 [1.0,4.1] *(p= 3E-13, d= 0.85)*	2.03 [0.8,3.5] *(p= 4E-11, d =0.75)*	2.08 [1.1,3.4] *(p=9E-12, d = 0.76)*
PEDIATRIC UVA Simulator (N = 99*)	**Baseline**	**Week 4**	**Week 12**	**End of Study**
% time-in-range ***(p-value, d = effect size, CI = confidence interval)***	65.7% ± 16.9	69.4% ± 16.0 *(p= 2E-4, d = 0.38, CI = [1.8,5.5])*	69.1% ± 15.4 *(p= 2E-2, d = 0.24, CI = [0.6,6.2])*	71.6% ± 16.3 *(p= 5E-4, d = 0.36, CI = [2.6,9.1])*
% time-in-hypoglycemia	4.12 [1.5,9.7]	2.48 [1.1,4.2] *(p=1E-7, d =0.58)*	2.03 [1.0,3.4] *(p=1E-7, d =0.59)*	2.13 [1.1,3.4] *(p=7E-8, d =0.59)*

Table 3.S6: Supplementary Table. Top 5 features for each class of recommendations delivered by the K-nearest-neighbors algorithm, as ranked by mutual information criteria calculation.

The recommendation to decrease basal insulin was most closely related to frequency and duration of hypoglycemia as well as nighttime glycemic excursions. The recommendation to decrease meal bolus aggressiveness was most closely related to the person's postprandial hypoglycemia. The recommendation to increase insulin before meals was most closely related to serious hyperglycemic events and overall daytime glycemic excursions. The relationship of these specific glycemic features with their associated optimal recommendations matches intuition and, in general, is synonymous with physician opinions regarding titration of insulin dosing for people with T1D.

Recommendation	Top 5 Features
Decrease Basal Insulin	Percent Time-in-Hypoglycemia Frequency of Hypoglycemic Events Nocturnal Hypoglycemia Percent Time-in-Hyperglycemia Postprandial Hypoglycemia, Overnight
Increase Basal Insulin	Frequency of Hypoglycemic Events Percent Time-in-Hyperglycemia Percent Time-in-Hypoglycemia Nocturnal Hypoglycemia Postprandial Hypoglycemia, Overnight
Increase Meal Bolus Dosed 7AM-11AM	Percent Time-in-Hyperglycemia Percent Time-in-Serious-Hyperglycemia Postprandial Hyperglycemia, 3PM-8PM Percent Time-in-Hypoglycemia Postprandial Serious Hyperglycemia, 3PM-8PM
Increase Meal Bolus Dosed 11AM-3PM	Percent Time-in-Hyperglycemia Percent Time-in-Serious-Hyperglycemia Percent Time-in-Hypoglycemia Postprandial Hypoglycemia, 3PM-8PM Frequency of Hypoglycemic Events
Increase Meal Bolus Dosed 3PM-8PM	Percent Time-in-Hyperglycemia Postprandial Hyperglycemia, 3PM-8PM Postprandial Hypoglycemia, Overnight Frequency of Hypoglycemic Events Percent Time-in-Hypoglycemia

Increase Meal Bolus Dosed Overnight	Percent Time-in-Hyperglycemia Percent Time-in-Serious-Hyperglycemia Postprandial Serious Hyperglycemia, 3PM-8PM Correction Boluses Postprandial Serious Hyperglycemia, 3PM-8PM
Decrease Meal Bolus Dosed 7AM-11AM	Postprandial Hypoglycemia, 11AM-3PM Postprandial Hypoglycemia, 7AM-11AM Frequency of Serious Hypoglycemic Events Postprandial Serious Hypoglycemia, 11AM-3PM Postprandial Serious Hypoglycemia, 7AM-11AM
Decrease Meal Bolus Dosed 11AM-3PM	Postprandial Hypoglycemia, 3PM-8PM Percent Time-in-Hypoglycemia Frequency of Hypoglycemic Events Frequency of Serious Hypoglycemic Events Postprandial Hypoglycemia, 11AM-3PM
Decrease Meal Bolus Dosed 3PM-8PM	Postprandial Serious Hypoglycemia, Overnight Postprandial Hypoglycemia, Overnight Frequency of Serious Hypoglycemic Events Percent Time-in-Hypoglycemia Frequency of Hypoglycemic Events
Decrease Meal Bolus Dosed Overnight	Percent Time-in-Hypoglycemia Frequency of Hypoglycemic Events Nocturnal Hypoglycemia Postprandial Hypoglycemia, Overnight Frequency of Serious Hypoglycemic Events

Table 3.S7: Supplementary Table. Glycemic features selected as inputs into the KNN-DSS look-up table.

Hypoglycemia is defined as measured glucose < 70 mg dL^-1, serious hypoglycemia is defined as measured glucose < 54 mg dL^-1, hyperglycemia is defined as glucose > 180 mg dL^-1, and serious hyperglycemia is defined as glucose > 300 mg dL^-1. To count an individual hypoglycemic or hyperglycemic event, measured glucose must be within the specified range and preceded by 20 continuous minutes of euglycemia; this hysteresis approach is used to prevent over-counting glycemic events that occurred due to CGM measurement noise.

Feature	Description	Normalization
Frequency of Hypoglycemic events	Overall number of independent hypoglycemic events detected in the users CGM data. Data is converted to a frequency measurement (# hypoglycemic events per # days of data).	Normalized to the expected frequency of hypoglycemic events for people with T1D (1.16 hypoglycemic events per day, a value determined from a separate dataset [194]).
Nocturnal Hypoglycemia	Number of hypoglycemic events detected between 10PM-7AM, normalized to the overall number of independent hypoglycemic events detected in the users CGM data.	‘’
Abnormal Morning Hypoglycemia	Number of abnormal morning hypoglycemic events detected in the user CGM data. Data is normalized to overall number of hypoglycemic events detected in the users CGM data.	‘’
Postprandial Hypoglycemia 7AM – 11AM	Number of hypoglycemic events detected following a meal bolus dosed between 7AM-11AM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Postprandial Hypoglycemia 11AM – 3PM	Number of hypoglycemic events detected following a meal bolus dosed between 11AM – 3 PM. Data is then normalized to the overall number of hypoglycemic	‘’

	events detected in the users CGM data.	
Postprandial Hypoglycemia 3PM – 8PM	Number of hypoglycemic events detected following a meal bolus dosed between 3PM – 8PM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Postprandial Hypoglycemia Overnight	Number of hypoglycemic events detected following a meal bolus dosed between 8PM – 7AM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Correction-Related Hypoglycemia	Number of hypoglycemic events detected following correction boluses dosed. Data is normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Frequency of Serious Hypoglycemic events	Overall number of independent serious hypoglycemic events detected in the users CGM data. Data is converted to a frequency measurement (# serious hypoglycemic events / # days of data).	‘’
Serious Postprandial Hypoglycemia 7AM – 11AM	Number of serious hypoglycemic events detected following a meal bolus dosed between 7AM-11AM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Serious Postprandial Hypoglycemia 11AM – 3PM	Number of serious hypoglycemic events detected following a meal bolus dosed between 11AM-3PM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’

Serious Postprandial Hypoglycemia 3PM – 8PM	Number of serious hypoglycemic events detected following a meal bolus dosed between 3PM-8PM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Serious Postprandial Hypoglycemia Nighttime	Number of serious hypoglycemic events detected following a meal bolus dosed between 8PM-7AM. Data is then normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Correction-Related Serious Hypoglycemia	Number of serious hypoglycemic events detected following correction boluses dosed. Data is normalized to the overall number of hypoglycemic events detected in the users CGM data.	‘’
Percent Time-in-Hypoglycemia	The percent of time that the users CGM measured less than 70 mg/dL.	N/A
Postprandial Hyperglycemia 3PM-8PM	Number of hyperglycemic events detected following a meal bolus dosed between 3PM-8PM. Data is then normalized to the overall number of hyperglycemic events detected in the users CGM data.	Normalized to the expected frequency of hyperglycemic events for people with T1D (2.05 hyperglycemic events per day, a value determined from a separate dataset [194]).
Correction Related Hyper	Number of hyperglycemic events detected following correction boluses dosed. Data is normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’
Serious Postprandial Hyper 7AM – 11AM	Number of serious hyperglycemic events detected following a meal bolus dosed between 7AM-11AM. Data is then normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’

Serious Postprandial Hyper 11AM – 3PM	Number of serious hyperglycemic events detected following a meal bolus dosed between 11AM-3PM. Data is then normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’
Serious Postprandial Hyper 3PM – 8PM	Number of serious hyperglycemic events detected following a meal bolus dosed between 3PM-8PM. Data is then normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’
Serious Postprandial Hyper Overnight	Number of serious hyperglycemic events detected following a meal bolus dosed between 8PM-7AM. Data is then normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’
Correction Related Serious Hyperglycemia	Number of serious hyperglycemic events detected following correction boluses dosed. Data is normalized to the overall number of hyperglycemic events detected in the users CGM data.	‘’
Percent Time in Hyperglycemia	The percent of time that the users CGM measured greater than 180 mg/dL	N/A
Percent Time in Serious Hyperglycemia	The percent of time that the users CGM measured greater than 300 mg/dL	N/A
Correction Boluses	Number of detected correction boluses dosed.	Normalized to the expected maximum number of correction boluses based on treatment recommendations (8 doses per 24hr period if users dosed once / three hours)

Table 3.S8: Supplementary Table. Look-up table class representation.

Recommendation	Count	Percent
Increase Basal	3723	7.18%
Decrease Basal	12041	23.23%
Increase AM Bolus	2574	4.97%
Increase PM Bolus	2549	4.92%
Increase Evening Bolus	3185	6.14%
Increase Overnight Bolus	1046	2.02%
Decrease AM Bolus	3744	7.22%
Decrease PM Bolus	4678	9.03%
Decrease Evening Bolus	7064	13.63%
Decrease Overnight Bolus	1565	3.02%
No Adjustment Recommended	9662	18.64%

Table 3.S9: Supplementary Table. Classification accuracy for weighting schemes.

K-nearest-neighbors classification weighting	% Classification Accuracy
No weighting imposed	0.6886
Class and distance-based weighting	0.7385
Distance-based weighting	0.7221
Class-based weighting	0.6864

Table 3.S10: Supplementary Data. Criteria used for labeling a recommendation from the KNN-DSS as in full agreement, partial agreement, full disagreement, or partial disagreement with a physician.

Additional: No change to settings indicated. As an example, for each unique recommendation delivered by Party A, we define subsets of recommendations corresponding to the categories perfect agreement, partial agreement, disagreement, and partial disagreement. We then calculate the similarity between these subsets and the recommendations delivered by Party B. This determines the agreement between Party A and Party B with respect to perfect agreement, partial agreement, disagreement, and partial disagreement.

KNN-DSS Recommendation	**Physician Recommendation**			
	Full Agreement	*Partial Agreement*	*Full Disagreement*	*Partial Disagreement*
Increase Basal Insulin	Increase Basal Insulin	Increase short-acting insulin doses for meals or corrections	Decrease basal insulin	Decrease short-acting insulin doses for meals or corrections
Decrease Basal Insulin	Decrease basal insulin	Decrease short-acting insulin doses for meals or corrections	Increase Basal Insulin	Increase short-acting insulin doses for meals or corrections
Increase Meal Insulin, AM	Increase meal insulin AM	Increase meal or correction insulin (in adjacent time windows overnight and afternoon) Increase basal insulin Adhere to bolus calculator or basal dosage.	Decrease meal insulin AM.	Decrease meal or correction insulin (in adjacent time windows overnight and afternoon) Increase or decrease meal or correction insulin (non-adjacent time window evening) Decrease basal insulin
Increase Meal Insulin, Afternoon	Increase meal insulin afternoon	Increase meal or correction insulin (in adjacent time windows AM and evening) Increase basal insulin Adhere to bolus calculator or basal dosage	Decrease meal insulin afternoon.	Decrease meal or correction insulin (in adjacent time windows AM and evening) Increase or decrease meal or correction insulin (non-adjacent time window overnight) Decrease basal insulin
Increase Meal Insulin, Evening	Increase meal insulin evening	Increase meal or correction insulin (in adjacent time windows afternoon and overnight) Increase basal insulin Adhere to bolus calculator or basal dosage.	Decrease meal insulin evening	Decrease meal or correction insulin (in adjacent time windows afternoon and overnight) Increase or decrease meal or correction insulin (non-adjacent time window AM) Decrease basal insulin
Increase Meal Insulin, Overnight	Increase meal insulin overnight	Increase meal or correction insulin (in adjacent time windows evening and AM) Increase basal insulin	Decrease meal insulin overnight	Decrease meal or correction insulin (in adjacent time windows evening and AM) Increase or decrease meal or correction insulin (non-adjacent time window afternoon)

		Adhere to bolus calculator or basal dosage		Decrease basal insulin
Decrease Meal Insulin, AM	Decrease meal insulin AM	Decrease meal or correction insulin (in adjacent time windows overnight and afternoon) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase meal insulin AM Partial disagreement	Increase meal or correction insulin (in adjacent time windows overnight and afternoon) Increase or decrease meal or correction insulin (non-adjacent time window evening) Increase basal insulin
Decrease Meal Insulin, Afternoon	Decrease meal insulin afternoon	Decrease meal or correction insulin (in adjacent time windows AM and evening) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase meal insulin afternoon	Increase meal or correction insulin (in adjacent time windows AM and evening) Increase or decrease meal or correction insulin (non-adjacent time window overnight) Increase basal insulin
Decrease Meal Insulin, Evening	Decrease meal insulin evening	Decrease meal or correction insulin (in adjacent time windows afternoon and overnight) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase meal insulin evening	Increase meal or correction insulin (in adjacent time windows afternoon and overnight) Increase or decrease meal or correction insulin (non-adjacent time window AM) Increase basal insulin
Decrease Meal Insulin, Overnight	Decrease meal insulin overnight	Decrease meal or correction insulin (in adjacent time windows evening and AM) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase meal insulin overnight	Increase meal or correction insulin (in adjacent time windows evening and AM) Increase or decrease meal or correction insulin (non-adjacent time window afternoon) Increase basal insulin
Increase Correction Insulin, AM	Increase correction insulin AM	Increase meal or correction insulin (in adjacent time windows overnight and afternoon) Increase basal insulin Adhere to bolus calculator or basal dosage	Decrease correction insulin AM	Decrease meal or correction insulin (in adjacent time windows overnight and afternoon) Increase or decrease meal or correction insulin (non-adjacent time window evening) Decrease basal insulin
Increase Correction Insulin, Afternoon	Increase correction insulin afternoon	Increase meal or correction insulin (in adjacent time windows AM and evening) Increase basal insulin Adhere to bolus calculator or basal dosage	Decrease correction insulin afternoon	Decrease meal or correction insulin (in adjacent time windows AM and evening) Increase or decrease meal or correction insulin (non-adjacent time window overnight) Decrease basal insulin
Increase Correction Insulin, Evening	Increase correction insulin evening	Increase meal or correction insulin (in adjacent time windows afternoon and overnight)	Decrease correction insulin evening	Decrease meal or correction insulin (in adjacent time windows afternoon and overnight)

		Increase basal insulin Adhere to bolus calculator or basal dosage		Increase or decrease meal or correction insulin (non-adjacent time window AM) Decrease basal insulin
Increase Correction Insulin, Overnight	Increase correction insulin overnight	Increase meal or correction insulin (in adjacent time windows evening and AM) Increase basal insulin Adhere to bolus calculator or basal dosage	Decrease correction insulin overnight	Decrease meal or correction insulin (in adjacent time windows evening and AM) Increase or decrease meal or correction insulin (non-adjacent time window afternoon) Decrease basal insulin
Decrease Correction Insulin, AM	Decrease correction insulin AM	Decrease meal or correction insulin (in adjacent time windows overnight and afternoon) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase correction insulin AM	Increase meal or correction insulin (in adjacent time windows overnight and afternoon) Increase or decrease meal or correction insulin (non-adjacent time window evening) Increase basal insulin
Decrease Correction Insulin, Afternoon	Decrease correction insulin afternoon	Decrease meal or correction insulin (in adjacent time windows AM and evening) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase correction insulin afternoon	Increase meal or correction insulin (in adjacent time windows AM and evening) Increase or decrease meal or correction insulin (non-adjacent time window overnight) Increase basal insulin
Decrease Correction Insulin, Evening	Decrease correction insulin evening	Decrease meal or correction insulin (in adjacent time windows afternoon and overnight) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase correction insulin evening	Increase meal or correction insulin (in adjacent time windows afternoon and overnight) Increase or decrease meal or correction insulin (non-adjacent time window AM) Increase basal insulin
Decrease Correction Insulin, Overnight	Decrease correction insulin overnight	Decrease meal or correction insulin (in adjacent time windows evening and AM) Decrease basal insulin Adhere to bolus calculator or basal dosage	Increase correction insulin overnight	Increase meal or correction insulin (in adjacent time windows evening and AM) Increase or decrease meal or correction insulin (non-adjacent time window afternoon) Increase basal insulin
No change to settings	No change to settings	N/A	N/A	N/A
Use bolus calculator	Use the bolus calculator	Modify short-acting bolus	N/A	N/A
Adhere to basal insulin dosage	Adhere to basal insulin dosage	Modify basal dosage	N/A	N/A

4 Quantifying the impact of physical activity on future glucose trends using artificial intelligence

Summary:

- Adults with T1D who perform aerobic exercise exhibit considerable variability in glucose outcomes, even if someone performs the exact same exercise regimen multiple times.
- Changes in glucose will be experienced differently by people with different physical fitness levels. People with higher aerobic fitness will experience significantly steeper glucose trends and significantly lower minimum glucose during aerobic exercise than people with lower aerobic fitness.
- Adaptive predictive algorithms can predict the changes that one will experience during exercise, and can be personalized to better predict an individual's aerobic exercise glucose patterns. The performance of adaptation on predictive accuracy serves as an upper bound on predictive accuracy for real-world approaches.

This manuscript submitted in June and currently awaiting reviewer comments

Abstract:

Prevention of hypoglycemia (glucose <70 mg/dL) during aerobic exercise is a major challenge in type 1 diabetes. Providing predictions of glycemic changes during and following exercise can help people with type 1 diabetes avoid

hypoglycemia. A unique dataset representing 320 days, and 50,000+ time points of glycemic measurements was collected in adults with type 1 diabetes who participated in a 4-arm crossover study evaluating insulin-pump therapies, whereby each participant performed 8 identically designed in-clinic exercise studies. We demonstrate that even under highly controlled conditions, there is considerable intra- and interparticipant variability in glucose outcomes during and following exercise. Participants with higher aerobic fitness exhibited significantly lower minimum glucose and steeper glucose declines during exercise. Adaptive, personalized artificial intelligence (AI) algorithms were designed to predict exercise-related glucose changes. These algorithms achieved high accuracy in predicting the minimum glucose and hypoglycemia during and following exercise sessions, for all fitness levels.

4.1 Introduction

Physical activity has been shown to reduce cardiovascular risk factors in people with type 1 diabetes [31] and regular physical exercise has recently been shown to result in improved time in target glucose range (70-180 mg/dL)[32]. However, exercise is also known to cause substantial changes in glucose. These changes in glucose vary per exercise modality [162, 194, 199-202] and are most dramatic during steady aerobic exercise [161]. There is an increased risk of hypoglycemia during exercise that occurs due to altered muscular uptake of glucose during exercise, and delayed hypoglycemia that can occur on nights following exercise due to changes in insulin-sensitivity [50, 203-205]. These dynamic processes underlying glucose uptake are compounded by regular

bouts of exercise [206, 207]. While regular exercise can improve overall health, voiding hypoglycemia during exercise is a known challenge for people with type 1 diabetes[208].

Continuous glucose monitoring technologies (CGM) can provide real-time alerts to the occurrence of hypoglycemia (< 70 mg/dL) or hyperglycemia (> 180 mg/dL) during exercise. And while certain commercial CGM systems like the Dexcom CGM have recently been reported to achieve 13.3% mean absolute relative error (MARE) during aerobic activity [209], use of CGM alone is not sufficient to prevent hypoglycemia. Commercially available automated insulin delivery (AID) systems have been shown to improve time in glucose target range across real-world daily activities [99, 210], but the exercise modalities of these systems are limited to user-selected modifications to basal insulin and target glucose during announced physical activity [42, 43]. AID algorithms that incorporate real-time physical activity data to prevent hypoglycemia typically reduce automated insulin and in the case of dual-hormone systems increase glucagon in anticipation of glucose drops during aerobic exercise [28, 211, 212], but even these systems do not completely eliminate exercise-induced hypoglycemia. Consensus statement guidelines have been developed to help people with type 1 diabetes make decisions regarding modification of insulin dosages and carb intake prior to and during exercise [82, 213], but people with type 1 diabetes will oftentimes need to use trial-and-error approaches to learn how to avoid hypoglycemia during exercise. Both automated hormone delivery and decision support systems currently lack the ability to accurately predict exercise-induced changes in glucose. In addition, there can be significant inter and intra-person variability in glucose changes during exercise. Exercise-related glucose changes in people with type 1 diabetes have not yet been precisely quantified in individuals and across populations when considering different insulin therapies, or baseline fitness levels.

Artificial intelligence (machine learning) is a powerful tool whereby machines are designed to solve problems or perform sophisticated tasks, and can even provide help to make medical decisions, or decision support, for diabetes management. Machine learning approaches have been used in disease detection [214], insulin dose modification through decision support [215, 216], and can be expanded to provide exercise decision support directly to a person living with type 1 diabetes, or to AID systems in order to adjust insulin during physical activity[29, 50]. While algorithms that have been designed to predict future glucose exhibit relatively low root mean squared error (RMSE) during non-exercise periods (14.0 mg/dL-18.0 mg/dL) [217-219], recent studies have indicated that the accuracy of these algorithms is oftentimes far worse during exercise (46.16 mg/dL) [51]. Machine learning models have already been developed to predict changes in glucose immediately following aerobic exercise [50, 51, 65, 167], and, when integrated with a decision support system, increase the minimum glucose measured during in-clinic exercise sessions [45]. Still, these algorithms oftentimes have poor accuracy during real-world scenarios [51], demonstrate large variability in performance between individuals [220] and have not been evaluated across varying physical fitness levels.

Population machine learning models are trained on a group of people and are designed to generalize to provide predictions for all people. Whereas a *personalized* model learns an individual's unique physiology in order to improve prediction accuracy for an individual. Personalized models can be designed by training machine learning models specifically on an individual's data [65], by clustering a number of similar people into groups prior to model training and then training a model on that cluster [48, 221], or by adapting a model in real time using newly observed data [51]. It is not yet clear how personalization impacts the prediction accuracy of exercise-related changes in glucose.

Herein we characterize the impact of aerobic exercise on glucose changes using a unique dataset collected during highly controlled, aerobic exercise sessions in adults with type 1 diabetes. Glucose variations are characterized per participant, insulin therapy, and are further explored with respect to baseline physical fitness. Personalized machine learning models were then designed to estimate the minimum glucose during aerobic exercise and four hours following the start of exercise, and to quantify the impact of personalization on model accuracy. We considered three machine learning algorithms, including a multivariate adaptive regression spline (MARS) model [71], a previously described logistic regression model [45], and a previously described autoregressive model with exogenous inputs (ARX) [65]. The dataset used to train and benchmark the approach was collected in a previously published study whereby aerobic exercise was performed 8 times per study participant under identical exercise intensity and duration, meal content and timing conditions, and across multiple hormone therapies including single-hormone, dual-hormone, predictive low-glucose suspend, and standard of care insulin pump therapies [28]. The findings obtained from this unique dataset can serve as a benchmark for comparison with other adaptive prediction algorithms, since we anticipate that the repeatability of the changes in glucose will be substantially reduced under free-living exercise conditions compared with these controlled conditions.

4.2 Results

Variations in blood glucose dynamics during identically designed exercise scenarios

To evaluate the repeatability of exercise-related glucose changes, participant glucose outcomes were obtained from 20 adults with type 1 diabetes who each

performed 8 identically-designed aerobic exercise sessions at 70% VO_2max for 40 minutes (N = 160 observations). To control for additional variability in glucose trends that can impact exercise-related glucose changes, the in-clinic exercise sessions were designed such that participants consumed a self-selected breakfast at 8 am, daily activities at 10 am, lunch at 12 pm, and performed exercise at 2 pm. Meals of identical content were consumed at the same time, and aerobic treadmill exercise was performed at the same time for each of the 8 in-clinic visits. Figure 1 shows the variability in the changes in blood glucose during exercise for each participant across the entire study (Figure 1A) and also organized by insulin therapy (Figure 1B-E). The difference in exercise-related blood glucose changes measured during highly controlled exercise sessions (Figure 1B-E, connecting dashed and solid lines) are reported as the difference averaged across all study arms, per participant in Table 4.1. Blood glucose dropped during exercise for nearly every exercise session, and CGM dropped further the 4 hour period after exercise was concluded for some subjects (Figure 1B-E, circles). Despite highly repeatable exercise conditions, food intake, and glucose management strategies, there was still substantial intra-participant variability of the change in glucose during exercise across all 8 identical exercise scenarios, ranging across participants from 23.1 mg/dL (participant 13) to 56.4 mg/dL (participant 9) (Table 4.1). While variability is smaller for some participants when looking at the two exercise sessions performed under a given hormone therapy, substantial variability in glucose changes during exercise is still observable for other study participants (Figure 1B-E). The average change in blood glucose during exercise and variability in this change is reported per therapy arm and per participant in Table 4.1.

Physical fitness impacts changes in glucose observed during physical activity

Baseline aerobic fitness was assessed by VO_2max norms for men and women using a rating scale from the American College of Sports Medicine [222] that ranks individuals on a scale of very poor, poor, fair, good, excellent, and superior. We found that participants with higher aerobic fitness (rated as good, excellent, and superior VO_2max) exhibited significantly lower minimum glucose during aerobic exercise than those with lower aerobic fitness (rated as very poor, poor, and fair VO_2max) (75.9 mg/dL vs 103.1 mg/dL, $p < 0.001$). Participants with higher aerobic fitness also exhibited lower CGM-measured minimum glucose compared with participants with lower aerobic fitness in the 4-hours following the start of exercise (70.4 mg/dL vs 85.4 mg/dL, $p < 0.001$). And, the higher aerobic participants had significantly steeper glucose drops during exercise (-2.2 mg/dL/min vs -1.8 mg/dL/min, $p < 0.05$) (Figure 2A-C). Participants with higher aerobic fitness exhibited lower glucose values across the in-clinic study days (Figure 2D,E), with significantly lower glucose during activities of daily living when they were physically active ($p < 0.05$), during the aerobic exercise, and in the overnight period following in-clinic aerobic exercise.

Population model predictions achieve good prediction accuracy

Three types of population machine learning models were designed: a MARS model to predict minimum glucose following exercise, a logistic regression model to predict hypoglycemia following exercise, and an ARX model to predict CGM values at the end of exercise. Participant features used to create the models were extracted from the data collected during each of the in-clinic exercise sessions (N = 160 exercise sessions) and are defined in Supplementary Table 4.S1. Leave-out-one-subject cross-validation was used during algorithm training to develop generalizable predictive models

(Supplementary Figure S1). Predictive accuracy of the three machine learning models to predict minimum blood glucose at the end of exercise and also CGM-measured minimum glucose during the 4 hours following the start of exercise are reported in Table 4.2. The population MARS model estimated minimum glucose during exercise with an MAE of 20.0 mg/dL; a sensitivity of 63%, and an accuracy of 67% to predict hypoglycemia when cross-validated across all 20 participants with each participant left out during the training. The population logistic regression model achieved a sensitivity of 64% and accuracy of 61% in predicting hypoglycemia during exercise when cross-validated on all 20 participants. The population ARX model exhibited worse MAE than the MARS model, 23.8 mg/dL, and achieved the highest sensitivity (71%) and accuracy (81%) to predict CGM-measured glucose < 70 mg/dL 40 minutes after the start of exercise, when cross-validated across all 20 participants.

For longer prediction horizons of 4 hours after the start of exercise, the population MARS model exhibited a MAE of 20.1 mg/dL, and a sensitivity of 62% and an accuracy of 56% to detect CGM-measured hypoglycemia when cross-validated across all 20 participants. The results of the logistic regression model to predict hypoglycemia during exercise and 4 hours following the start of exercise were similar for both 45-minutes and 4-hour prediction horizons. The logistic regression model achieved a sensitivity of 63% and accuracy of 58% to detect CGM-measured hypoglycemia when cross-validated across all 20 participants. The ARX model was not designed for the 4-hour predictive window and therefore results are not shown.

Prior exercise-related glucose drops help predict future glucose drops

The benefit of personalization was evaluated by first considering whether the inclusion of participant exercise history, or data collected during previous exercise

sessions, can improve accuracy to predict glucose during exercise. To do this, a second MARS model was designed that also incorporates participant exercise history features as inputs to the MARS model (Supplementary Table 4.S2). Exercise data features that were found to be predictive of future glucose trends included (1) the participant's average metabolic expenditure measured during other aerobic exercise sessions, and (2) the average change in glucose measured during other aerobic exercise sessions by the participant. When evaluated on the holdout set, the MARS model that included exercise history reduced MAE by 39%, from 23.4 mg/dL to 14.3 mg/dL, improved sensitivity to predict hypoglycemia during exercise from 50% to 73%, and improved accuracy from 75% to 81% (Table 4.2). Cross-validation across all 20 participants showed that the inclusion of participants' exercise history into the MARS model reduced MAE from 20.0 mg/dL to 17.6 mg/dL, improved sensitivity from 63% to 66% to detect hypoglycemia, and improved accuracy from 67% to 70%.

For longer prediction horizons of 4 hours, the MARS model that included exercise history performed similarly to the MARS model that was designed without exercise history, when cross-validated across all 20 participants (Table 4.2).

Adaptive personalization improved the accuracy of predictive models

The benefit of personalization was also investigated through adaptation of the machine learning models to better predict individual participants' exercise-related glucose changes. Stochastic gradient descent [223] was used to incorporate the exercise information obtained from a participants exercise session (e.g., data collected during their first study visit) in order to update the population model parameters. The adapted model was then used to predict the same participant's outcomes for a separate, held-out exercise session (e.g., their second study visit). This adaptation procedure was repeated for each exercise session, enabling the machine learning model parameters to

adapt to an individual's data over time as more exercise sessions were observed. Personalization of the model coefficients through stochastic gradient descent adaptation improved the accuracy of all of the predictive algorithms (see Table 4.2) to estimate glucose during exercise and 4 hours after the start of exercise. The improvement from adaptation was not influenced by the order of the observed exercise sessions, and we report the results from the original order prior to shuffling. Gradient descent adaptation of model coefficients reduced the predictive error of the MARS model from an MAE of 20.0 mg/dL to an MAE of 18.1 mg/dL, reduced sensitivity from 63% to 61%, and significantly improved the 20-fold cross-validation accuracy of the MARS model in predicting hypoglycemia during exercise from 67% to 78% ($p<0.05$). The predictive error per-participant can be seen in Table 4.1. Adaptation of the logistic regression parameters improved the sensitivity to predict hypoglycemia during exercise from 64% to 68%, and significantly improved the accuracy from 61% to 70% ($p<0.05$), when cross-validated across all 20 participants. Adaptation of the ARX model improved the cross-validation MAE from 23.8 mg/dL to 22.0 mg/dL, and improved the sensitivity to detect hypoglycemia during exercise from 71% to 76% and accuracy from 81% to 83%.

For longer prediction horizons of 4 hours following the start of exercise, adaptation reduced the predictive error and improved the accuracy of all of the models. The personalization through adaptation of the MARS model coefficients significantly reduced the MAE from 20.1 mg/dL to 18.3 mg/dL, reduced sensitivity from 62% to 56%, and significantly increased the accuracy to predict CGM-measured hypoglycemia 4 hours following exercise from 56% to 68% ($p<0.05$). The adaptation of the MARS model designed to include prior exercise session metrics reduced the MAE from 21.1 mg/dL to 18.2 mg/dL, reduced sensitivity from 74% to 57% and increased the accuracy to detect CGM-measured hypoglycemia 4 hours following exercise from 57% to 69% when cross-validated across all 20 participants. Adaptation of the logistic regression model

increased sensitivity from 63% to 64%, and significantly improved the accuracy from 58% to 70% ($p<0.05$) to predict CGM-measured hypoglycemia in the 4 hours following exercise when cross-validated across all 20 participants.

Figure 3 shows the Parkes consensus grid of the MARS model cross-validation across all 20 participants in predicting glucose at the end of exercise. Personalization of the population MARS model increased the number of observations in the consensus error grid region A from 110 observations to 115 observations, with no changes in regions C, D, or E. When exercise history was included in the design of the MARS model, adaptation increased the values in region A to 118 observations, with no observations in regions D and E and 99% of observations in the combined A + B regions (Figure 3C).

Physical fitness impacts predictive performance

The MARS models performed equivalently for higher fitness vs. lower fitness study participants (Table 4.3). The ARX performed worse for the higher fitness participants than the lower fitness participants. The accuracy to detect hypoglycemia during exercise, and in the 4 hours following start of exercise, was nominally lower in all machine learning models when evaluated on participants with higher aerobic fitness. Adaptation improved the accuracy to predict hypoglycemia for participants with higher and lower aerobic fitness, and across both prediction horizons (Table 4.3).

Table 4.1: Changes in glucose during exercise and participant-specific predictive error of the AI models.

Results of AI models exhibiting the best performance to predict minimum glucose in the 4 hours following exercise. † indicates participants with excellent aerobic fitness. Participant 18 SMBG was not available for the standard-care arm, and is not reported.

Participant ID	**Mean Glucose Drop During Exercise [mg/dL]**	**Glucose Drop per Study Arm**				Average difference measured during identical exercise [mg/dL]	**Predicted minimum glucose**			
							End of exercise		***+ 4hr***	
		Standard Care	**Predictive Low-Glucose Suspend**	**Single-hormone AP**	**Dual-hormone AP**		***(Model = MARS + Exercise History + personalization)***		***(Model = MARS + personalization)***	
							MARE [%]	RMSE [mg/dL]	MARE [%]	RMSE [mg/dL]
1†	-92.4 ± 24.0	-74.5 ± 10.6	-79.5 ± 3.5	-122.0 ± 12.7	-93.5 ± 31.8	20.8	11.9	8.5	13.8	10.0
2†	-100.3 ± 23.4	-107.0 ± 1.4	-122.0 ± 1.4	-106.5 ± 12.0	-65.5 ± 12.0	9.5	17.6	13.8	24.9	18.5
3	-41.4 ± 37.2	-61.0 ± 39.6	-54.0 ± 25.5	11.0 ± 9.9	-61.5 ± 3.5	27.8	11.7	15.9	26.1	33.5
4†	-91.1 ± 47.6	-132.0 ± 58.0	-93.0 ± 35.4	-93.5 ± 57.3	-46.0 ± 22.6	61.3	19.1	16.9	17.9	13.1
5†	-104.1 ± 28.4	-74.0 ± 25.5	-105.0 ± 25.5	-110.0 ± 36.8	-127.5 ± 4.9	32.8	22.7	21.4	25.3	18.8
6	-118.5 ± 48.1	-119.5 ± 13.4	-127.5 ± 91.2	-96.5 ± 57.3	-130.5 ± 54.4	76.5	24.2	41.9	21.9	30.6
7†	-83.6 ± 54.4	-79.5 ± 46.0	-109.5 ±102.5	-43.5 ± 10.6	-101.8 ± 52.0	74.6	10.2	8.0	10.5	8.4
8†	-101.1 ± 39.1	-145.5 ± 48.8	-106.5 ± 20.5	-86.0 ± 5.7	-66.5 ± 31.8	37.8	23.5	20.1	18.3	15.2
9†	-86.6 ± 56.4	-37.3 ± 34.3	-65.0 ± 56.6	-111.5 ± 61.5	-132.5 ± 53.0	72.6	17.1	17.2	33.1	43.2
10†	-97.8 ± 50.0	-71.0 ± 2.8	-74.5 ± 101.1	-126.5 ± 29.0	-119.0 ± 36.8	60.0	28.9	35.0	15.7	28.9
11	-94.4 ± 33.9	-95.5 ± 14.8	-93.0 ± 63.6	-106.5 ± 54.4	-82.5 ± 14.8	52.3	23.5	24.6	22.8	23.0
12†	-55.4 ± 26.7	-20.8 ± 15.2	-66.0 ± 22.6	-83.0 ± 5.7	-52.0 ± 7.1	17.9	22.1	41.0	17.8	21.3
13	-36.0 ± 23.1	-9.0 ± 2.8	-48.0 ± 15.6	-27.0 ± 18.4	-60.0 ± 9.9	16.5	16.3	15.1	13.0	12.4
14	-112.6 ± 42.8	-98.0 ± 46.7	-102.0 ± 65.1	-97.8 ± 35.7	-152.5 ± 29.0	62.4	19.5	24.6	19.3	18.5
15	-88.1 ± 36.7	-104.5 ± 82.7	-102.5 ± 17.7	-73.0 ± 11.3	-72.5 ± 16.3	45.3	16.1	15.9	16.1	16.8
16†	-77.1 ± 33.9	-57.5 ± 17.7	-45.0 ± 25.5	-91.5 ± 27.6	-114.5 ± 16.3	30.8	23.4	20.4	19.8	21.0
17	-69.4 ± 38.0	-23.0 ± 19.8	-111.5 ± 0.7	-79.5 ± 17.7	-63.5 ± 36.1	26.3	22.0	28.1	22.8	32.0
18	-97.0 ± 48.7	N/A	-115.5 ± 2.1	-135.0 ± 4.2	-40.5 ± 43.1	23.3	19.4	23.5	39.0	53.5
19	-74.0 ± 23.6	-78.0 ± 25.5	-68.0 ± 12.7	-95.0 ± 15.6	-55.0 ± 33.9	31.0	20.0	30.1	25.3	29.0
20†	-63.0 ± 54.0	-140.0 ± 21.2	-54.5 ± 20.5	-13.0 ± 33.9	-44.5 ± 26.2	36.0	20.1	20.1	22.5	19.0
Mean ± Std	**-84.2 ± 43.26**	**-80.4 ± 46.6**	**-87.1 ± 42.9**	**-84.3 ± 44.2**	**-84.1 ± 40.7**	**40.8 ± 20.9**	**19.5 ± 4.7**	**22.1 ± 9.4**	**21.3 ± 6.8**	**23.3 ± 11.3**

Table 4.2: Comparing the effect of adaptation on the performance of models designed to predict exercise-related changes in glucose.

*indicates that the significance $p < 0.05$ determined Wilcoxon signed-rank test for paired, non-parametric data comparing the change in error or accuracy on a per-participant basis. † These models return predicted CGM, not SMBG values. The ARX model is only designed to predict glucose in 40 minute windows, and the results for a 4 hour prediction horizon are not shown

	Population Model		**Adapted Model Coefficient Personalization**		**Comparison between population and personalized**	
	RMSE *(MAE)* [mg/dL]	[Sensitivity, Specificity] *(Accuracy)* [%]	RMSE *(MAE)* [mg/dL]	[Sensitivity, Specificity] *(Accuracy)* [%]	Δ MAE [%]	Δ Accuracy [%]
Predicting minimum glucose at the end of exercise						
MARS Model						
Training, 16-fold CV	24.1 *(19.2)*	[73, 67] *(69)*	--	--	**--**	**--**
Validation, Holdout Set	26.5 *(23.4)*	[50, 86] *(75)*	23.1 *(19.6)*	[70, 86] *(81)*	-16.2	+ 8.3
Validation, 20-fold CV	24.6 *(20.0)*	[63, 63] *(67)*	23.0 *(18.1)*	[61, 78] *(78)*	**- 9.5**	**+ 16.1***
MARS Model + Exercise History Features						
Training, 16-fold CV	23.1 *(18.2)*	[75, 65] *(68)*	--	--	**--**	**--**
Validation, Holdout Set	18.7 *(14.3)*	[73, 86] *(81)*	19.7 *(15.8)*	[73, 95] *(88)*	+ 10.1	+ 7.7
Validation, 20-fold CV	22.6 *(17.6)*	[66, 69] *(70)*	22.1 *(17.5)*	[51, 83] *(77)*	**- 0.6**	**+ 10.1***
ARX Model: Population Model†						
Training, 16-fold CV	28.8 *(22.7)*	[71, 94] *(83)*	--	--	**--**	**--**
Validation, Holdout Set	32.8 *(28.6)*	[59, 87] *(72)*	27.6 *(23.3)*	[59, 87] *(72)*	-18.7	+ 0
Validation, 20-fold CV	29.6 *(23.8)*	[71, 91] *(81)*	27.7 *(22.0)*	[76, 90] *(83)*	**- 7.4**	**+ 3.1**
Logistic Regression						
Training, 16-fold CV	--	[66, 67] *(66)*	--	--	--	--
Validation, Holdout Set	--	[73, 76] *(75)*	--	[73, 90] *(84)*	--	+ 12.5
Validation, 20-fold CV	--	[64, 56] *(61)*	--	[68, 61] *(70)*	--	**+ 15.5***
Predicting Minimum Glucose 4 hours after exercise						
MARS Model†						
Training, 16-fold CV	25.8 *(19.7)*	[67, 68] *(68)*	--	--	**--**	**--**
Validation, Holdout Set	25.7 *(21.6)*	[18, 76] *(56)*	21.5 *(16.3)*	[33, 96] *(78)*	- 24.8	+ 38.9
Validation, 20-fold CV	25.1 *(20.1)*	[62, 51] *(56)*	23.3 *(18.3)*	[56, 70] *(68)*	**- 9.0 ***	**+ 21.4 ***
MARS Model + Exercise History Features†						
Training, 16-fold CV	24.8 *(18.6)*	[79, 61] *(69)*	--	--	**--**	**--**
Validation, Holdout Set	30.7 *(26.1)*	[29, 61] *(47)*	23.0 *(16.0)*	[56, 96] *(84)*	-38.8	+ 80.0
Validation, 20-fold CV	26.3 *(21.1)*	[74, 52] *(57)*	23.9 *(18.2)*	[57, 70] *(69)*	**- 13.8**	**+ 20.0**
Logistic Regression†						
Training, 16-fold CV	--	[57, 72] *(65)*	--	--	--	**--**
Validation, Holdout Set	--	[32, 77] *(50)*	--	[53, 92] *(69)*	--	**+** 37.5
Validation, 20-fold CV	--	[63, 50] *(58)*	--	[64, 74] *(70)*	--	**+ 20.4 ***

Table 4.3: Comparing the effect of aerobic fitness on the performance of models designed to predict exercise-related changes in glucose.

* indicates that the significance $p < 0.05$ as determined by Wilcoxon rank-sum test for non-parametric data, comparing algorithm performance on participants with higher aerobic and lower aerobic fitness rankings.

Model	Personalization	Accuracy [%]		MARE [%]	
		Lower VO_2max **n = 70 obs**	***Higher VO_2max*** **n = 88 obs**	***Lower VO_2max*** **n = 70 obs**	***Higher VO_2max*** **n = 88 obs**
Predicting minimum glucose during exercise					
MARS Model	Population	73	65	23	20
	Adaptation	*84*	*72*	*20*	*20*
MARS Model + Exercise History	Population	74	65	19	19
	Adaptation	*86*	*70*	*19*	*20*
ARX Model	Population	88	75 *	16	38*
	Adaptation	*85*	*82*	*16*	*34**
Logistic Regression	Population	65	55	--	--
	Adaptation	*72*	*65*	--	--
Predicting minimum glucose in the 4 hours following exercise					
MARS Model	Population	57	55	25	21
	Adaptation	*70*	*69*	*22*	*20*
MARS Model+ Exercise History	Population	63	52	25	24
	Adaptation	*69*	*68*	*23*	*20**
Logistic Regression	Population	58	56	--	--
	Adaptation	*72*	*63*	--	--

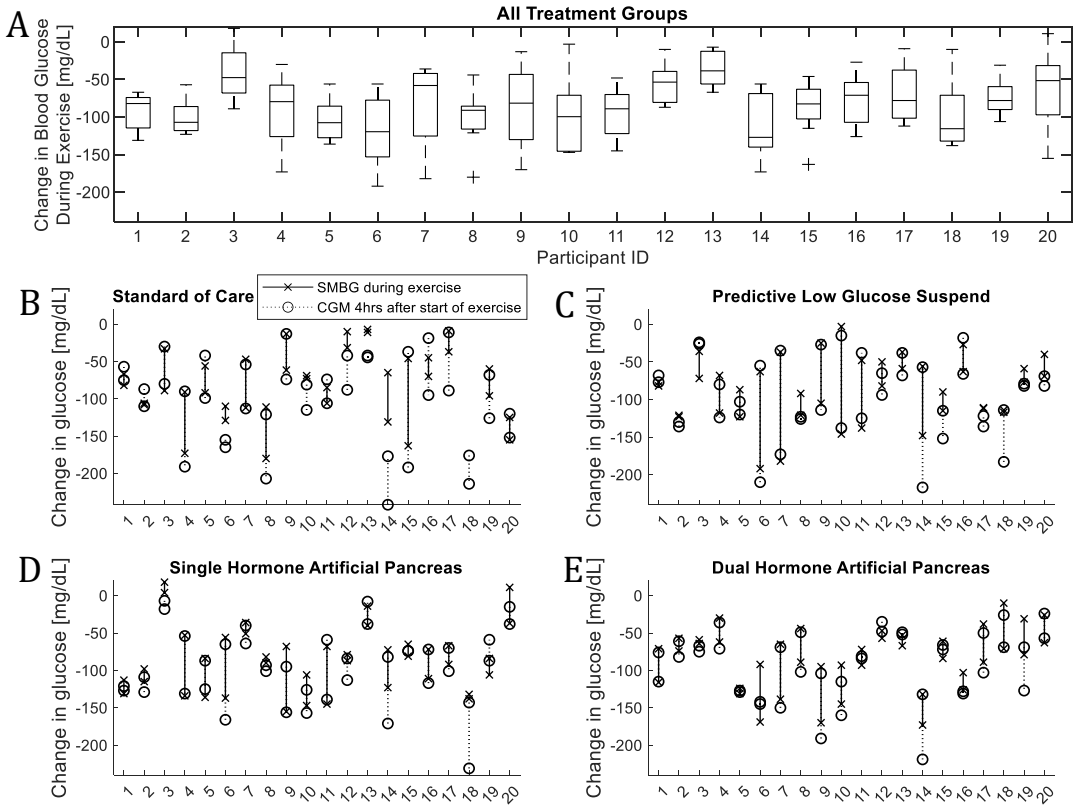


Figure 4.1: Change in blood glucose measured during identical aerobic exercise sessions.

(A) The change in glucose measured during 8 identical exercise sessions across a 4-arm clinical study. The box plot represents the median and interquartile range of the change in glucose measured during exercise, with cross symbols representing outlier values.

(B-E) the change in glucose measured during aerobic exercise within a given insulin therapy. The black x symbol represents the change in glucose measured during an exercise session, and there are two x symbols per participant per study arm. The line drawn between two black x symbols represents the difference in glucose outcomes measured between the two identically-designed exercise sessions. The open black circle represents the change glucose measured from the start of exercise, to the minimum glucose measured 4 hours after exercise, and these outcomes are connected by a dotted black line.

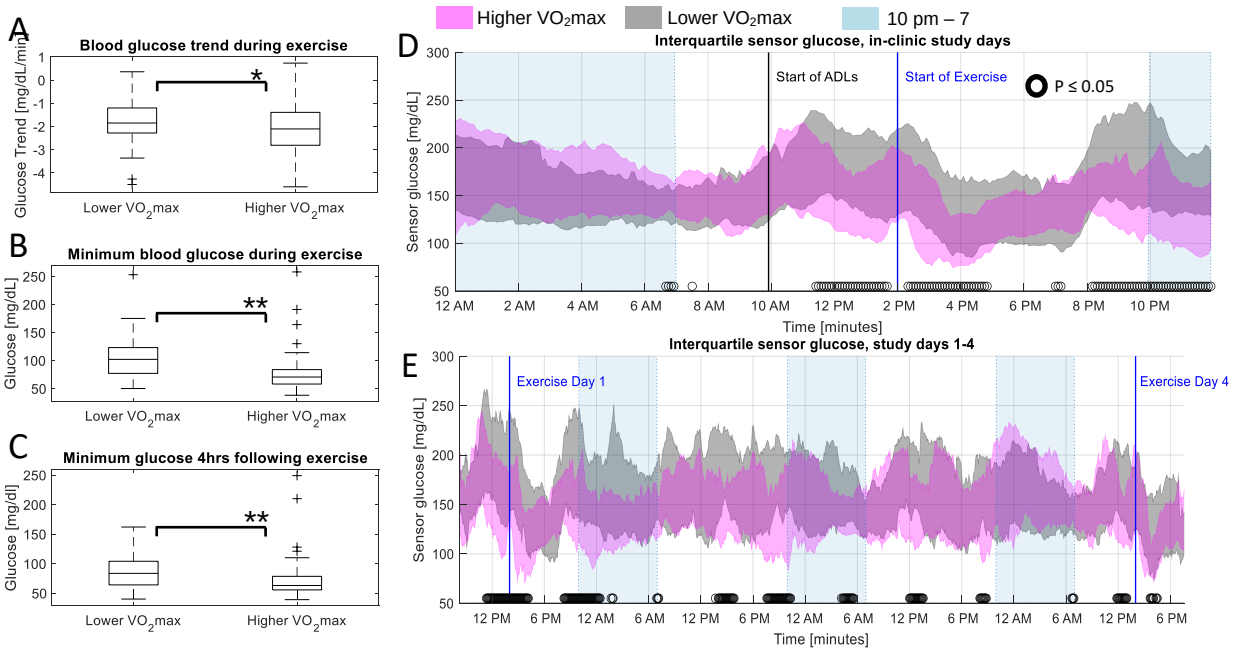
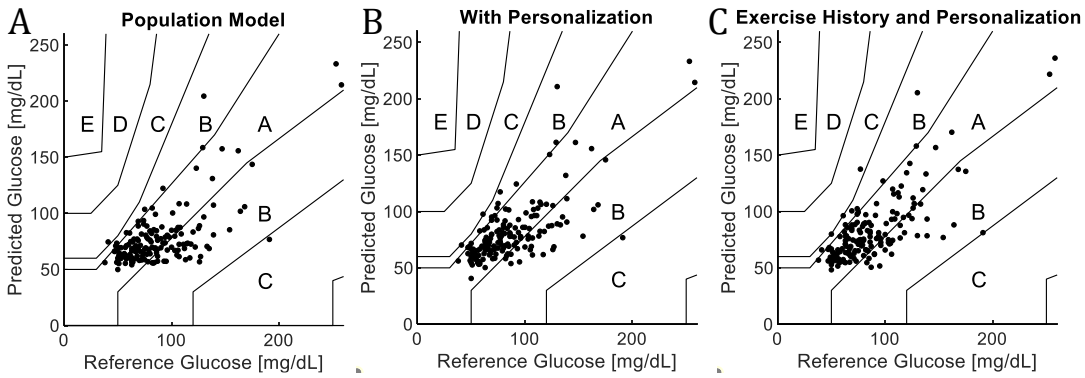


Figure 4.2: Differences in glycemic response across baseline physical fitness. Box plots represent the median and interquartile range of the data.

* represents significant differences p<0.05 between boxplot groups as determined by an independent t-test. ** represents significant differences p<0.05 between boxplot groups as determined by Wilcoxon rank-sum test. ○ represents significant differences p<0.05 between sensor glucose as determined by a Wilcoxon rank-sum test. (A) The slope of glucose during aerobic exercise is significantly steeper in participants with higher aerobic fitness (n = 88 observations collected from 11 participants) than participants with lower aerobic fitness (n = 70 observations collected from 9 participants) (-2.2 mg/dL/min vs -1.8 mg/dL/min, p = 0.03). (B) The minimum glucose measured during aerobic exercise is significantly lower in participants with higher aerobic fitness (n = 88 observations collected from 11 participants) than in participants with lower aerobic fitness (n = 70 observations collected from 9 participants) and (75.9 mg/dL vs 103.1 mg/dL, p = 4.7E-9) (C) The minimum glucose measured by CGM in the 4-hrs following the start of aerobic exercise is significantly lower in participants with higher aerobic fitness (n = 88 observations collected from 11 participants) than in participants with lower aerobic fitness (n = 70 observations collected from 9 participants) and (70.4 mg/dL vs 85.4 mg/dL, p = 3.3E-5) (D) Interquartile range of sensor glucose obtained from participants during in-clinic study days 1 and 4. Participants with higher aerobic fitness exhibit significantly lower glucose during activities of daily living and aerobic exercise, and in the nighttime following exercise (p < 0.05). The lower aerobic fitness group is represented by grey area (n = 72 sensor traces collected from 9 participants). The higher aerobic fitness group is represented by magenta area (n = 88 sensor traces collected from 11 participants). During the in-clinic exercise study visits, activities of daily living were performed starting at 10 am, and exercise at 70% VO_2max was performed at 2 pm. There are fewer sensor traces from 9 pm – 12 (lower fitness, n = 36, higher fitness, n = 45), as overnight data is only available from study day 1, whereas participants exited the clinical study on day 4 and overnight sensor data is therefore not available. (E) Interquartile range of sensor glucose across the entire 4-day study. The lower aerobic fitness group is represented by grey area (n = 36 sensor traces collected from 9 participants). The higher aerobic fitness group is represented by magenta area (n = 44 sensor traces collected from 11 participants).



A	B	C
A = 69.6%	A = 72.8%	A = 74.7%
B = 29.1%	B = 26.0%	B = 24.6%
C = 1.3 %	C = 1.3 %	C = 0.6 %
D = 0%	D = 0%	D = 0%
E = 0%	E = 0%	E = 0%
A + B = 98.7%	A + B = 98.7%	A + B = 99.4%

Figure 4.3: Consensus Error Grid for models predicting minimum glucose at the end of exercise.

The regions of the consensus error grid indicate the clinical impact of prediction errors. Observations that land in regions A and B indicate safe predictions. Observations that lay in regions C, D, and E may result in clinical errors such as missed hypoglycemia, or false positive hypoglycemia that results in excessive carbohydrate intake. The percentage of observations falling within each region is listed below each figure.

(A) Population MARS model validation (n = 158 observations of exercise data collected from 20 participants) without including prior exercise history.

(B) The MARS model predictions after personalization of population model coefficients (n = 158 observations of exercise data collected from 20 participants).

(C) The predictions of the MARS model that incorporates exercise history features, with additional personalization of the model coefficients (n = 158 observations of exercise data collected from 20 participants).

4.3 Conclusions

Herein we demonstrate that there is substantial variability in glucose changes during aerobic exercise in people with T1D even under highly repeatable food intake and

exercise conditions, and that these changes are impacted by baseline physical fitness levels. We also present new adaptive glucose forecasting algorithms and demonstrate how personalization and prior history can improve accuracy in predicting minimum glucose during and following aerobic exercise. To our knowledge, this is the first analysis of exercise-related glucose changes and prediction strategies using an ideal dataset of highly regimented, identical study exercise visits and across multiple insulin therapies. In the published clinical study data set used to train the proposed predictive algorithms [28], the specific variations in glucose during exercise were not presented and a demonstration of differences between individuals with varying aerobic fitness was not presented. The data demonstrate that individuals living with type 1 diabetes will experience considerable variability during exercise, even when exercise occurs in the context of identical meals, exercise intensity and duration, insulin therapy, and scheduled daily activities. For some participants, the magnitude of this variability was diminished when examined within the context of an individual insulin therapy. From a clinical perspective, this highlights the challenge and uncertainty that individuals face during aerobic exercise; even if someone could undertake the exact same daily activities, meals, and exercise practices, there will be differences in their glucose outcomes during exercise. Part of this variability is explained by insulin therapy, but there are many other factors such as activity level in the days preceding exercise, and stressors such as sleep quality, illness, or timing of menstrual cycle that affect glycaemia following exercise. And, baseline physical fitness can also have a significant impact on glycemic outcomes during exercise. The high intra- and interparticipant variability in glucose trends during exercise presents an opportunity for adaptive machine learning approaches to help people with type 1 diabetes avoid acute and long-term complications related to hypoglycemia.

The impact of exercise on glucose trends during exercise, and across participants with varying physical fitness levels, is still an open question [82, 224]. While an inverse relationship has previously been observed between the regularity of exercise and the rate of severe hypoglycemia[31], it has also been reported that participants with higher aerobic fitness exhibit a greater risk of hypoglycemia [225]. We contribute definitive findings that participants with higher aerobic fitness exhibit significantly steeper glucose trends during exercise, experienced significantly lower glucose at the end of exercise, and exhibit nominally lower variability in their glycemic outcomes. And although participants with varying aerobic fitness exhibited significantly different glucose outcomes following exercise, measures such as VO_2max and fitness ranking require in-clinic evaluation, and are not yet feasible for incorporation into the design of the accessible predictive algorithms. It was also observed that participants with higher aerobic fitness were shown to have significantly lower CGM across the entirety of the 4-arm clinical study; sensor readings for these participants were significantly lower during activities of daily living, exercise, and in the nighttime and 48-hrs following aerobic exercise. This precise knowledge can help to inform new strategies to help people of different fitness levels avoid exercise-related hypoglycemia.

Other groups have presented various methods at predicting glucose during exercise. Reddy et al. [50] developed a hypoglycemia prediction algorithm during exercise using a decision tree and random forest algorithm. This random forest model utilized data within first 10 minutes of aerobic exercise to form predictions, and achieved an 86% sensitivity and 87% specificity to hypoglycemia. This approach does not describe adaptation or personalization of models or utilize exercise history. It was also limited in that it required data during the first 10 minutes of exercise to estimate hypoglycemia which makes it impossible for the algorithm to provide automated hormone dosing or decision support prior to the start of exercise. The algorithms

proposed in this manuscript do not use data during the exercise event. The proposed algorithms were designed for use prior to the start of exercise for the purpose of modifying hormone doses and/or carbohydrate intake. The ARX model that we evaluated in this paper was presented originally in Romero-Ugalde et al., where model was designed to predict CGM values at 30 minutes following aerobic stair-step exercise, and achieved an RMSE of 7.75 mg/dL[65]. We repeated the methods described in Romero-Ugalde et al., and while we discovered this method achieves fair accuracy to predict CGM < 70 mg/dL, we were unable to achieve the performance that was previously reported and instead obtained an RMSE of 28.3 mg/dL. While the ARX model described by Romero-Ugalde et al. did not achieve the same predictive error as the MARS model, the adaptation methods presented herein improved the accuracy of the ARX model to predict CGM < 70 mg/dL and reduced the predictive RMSE. Breton et al. developed a hypoglycemia prediction algorithm utilizing the contextual physical activity predictors identified by Ben Brahim et al. [167]. The accuracy of this model was not reported and does not describe personalization [45]. In the current paper, we used identical features described by Breton et al. and demonstrated the performance of the model. We additionally showed that adaptation can significantly improve the performance in predicting hypoglycemia during exercise. Each of the prior publications as well as our findings identified the importance of CGM and SMBG measurements at the start of exercise as a critical predictive feature. The current manuscript extends the work done previously by emphasizing the importance of personalization and physical fitness considerations when designing glucose forecasting algorithms during exercise.

Personalization of the population-based machine learning models was shown to improve the accuracy in almost every model-framework, across both short-term and long-term prediction horizons, and across all validation scenarios. Adaptation of model parameters using stochastic gradient descent was shown to significantly improve the

accuracy of detecting hypoglycemia during exercise for the MARS and logistic regression models. And adaptation of the MARS and ARX models improved overall accuracy of predictions in terms of MAE. Personalization of the MARS framework that included exercise history as an input feature significantly improved predictive accuracy to detect hypoglycemia during exercise. Taken together across all of the models and validation strategies presented in Table 2, personalization resulted in an average reduction in minimum glucose error estimations by 12.9%, and an average increase in hypoglycemia prediction accuracy of 21.0%. A strength of the personalization methods presented in this manuscript is the simplicity of the gradient descent approach, which is computationally inexpensive and can be implemented easily in other predictive frameworks with just a few lines of code.

In summary, individuals on insulin pump therapy who perform aerobic exercise under highly regimented, nearly identical conditions and intensities will experience day-to-day variations in exercise-related glucose changes during and following exercise. Baseline physical fitness significantly impacts changes in glucose during exercise. Under these controlled conditions, glucose data at the start of exercise, as well as data from prior exercise sessions are informative of anticipated changes in glucose during future exercise sessions across participants of varying physical fitness levels. And while machine learning models can predict the expected changes in glucose during exercise and be personalized to provide more accurate predictions, further work is needed to accurately predict hypoglycemia in participants with higher baseline physical fitness. Further studies are forthcoming to determine the performance of our adaptation strategy on at-home exercise session data across participants with varying physical fitness. The scientific community is invited to apply this benchmarking dataset in their research by contacting the corresponding author for access to the data.

4.4 Research Design and Methods

This analysis was performed upon approval of OHSU Institutional Review Board, study number 00019659.

4.4.1 Study Population and Setting

This analysis utilized data obtained during a previous clinical study [28]. The data was collected from 20 adults with type 1 diabetes (N = 20, 14 F, Age 34.5 +/- 4.7y, duration diabetes 19.7 +/- 8.6 y, BMI 26 +/- 5.7, HbA1C 7.5 +/- 0.8, VO_2max 37.1 +/- 9.6) who participated in a 4-arm study. Each study arm consisted of 4 days of either (1) single-hormone automated insulin therapy, (2) dual-hormone (insulin and glucagon) automated therapy, (3) predictive low glucose suspend CGM-augmented pump therapy, or (4) standard of care CGM-augmented pump therapy. Participants visited the clinic on days 1 and 4 of each study arm. During in-clinic study visits, participants consumed a self-selected breakfast, lunch and dinner and performed aerobic exercise in the afternoon. Each participant consumed the same meals at the same time and performed the same physical activity at the same time for each of the 8 in-clinic visits (4 arms x 2 days). Aerobic exercise was performed at 70% VO_2max and lasted for 40 minutes. Participant accelerometer and heartrate data were obtained using ZephyrLife BioPatch devices (Zephyr, Annapolis, MD). The automated insulin and glucagon delivery systems were controlled using a custom exercise-aware algorithm [226] installed on a Google Nexus smart phone. This automated delivery system wirelessly communicated the t:slim pumps (Tandem, San Diego, CA) and G5 CGM sensors (Dexcom, San Diego, CA) via Bluetooth. During the control arm, participants used their own insulin pumps. The insulin pumps in this study were filled with aspart insulin (Novo Nordisk, Plainsboro, NJ). This secondary analysis utilized participant data obtained from G5 devices, t:slim devices,

ZephyrLife BioPatch devices and self-monitored blood glucose (SMBG) Contour Next devices (Bayer, Whippany, NJ).

4.4.2 Input features and outcome measures for model design

Participant features were extracted from the data collected by the study devices during each of the in-clinic exercise sessions (N = 160 exercise sessions). No observations were excluded from analysis on the basis of artifacts in the time series data, such as noise in CGM data due to calibration or movement, or signal dropout. The input features derived from the clinical data are defined in Supplementary Table S1. Additional features describing participant exercise history are defined in Supplementary Table S2. The final input features for each model were determined from *Greedy sequential variable selection* [196], or reproduced as described in previous publications [45, 65]. The algorithms were trained to predict (1) the minimum glucose from the start of exercise to the end of exercise as measured using self-monitored blood glucose (SMBG), and (2) minimum glucose 4 hours following the start of exercise as measured by continuous glucose monitor (CGM). SMBG measurements were measured by all participants at the start and end of exercise per study protocol, however SMBG was not always measured in the 4 hour period following exercise therefore CGM is used for the 4-hour prediction model. Participant age, sex, and VO_2max were used to classify each participant into categories of higher (including good, excellent, and superior VO_2max) aerobic fitness or lower (including very poor, poor and fair VO_2max) aerobic fitness, as defined by the American Society of Sports Medicine VO_2max aerobic fitness norms [222].

4.4.3 Development of the population prediction models

Three types of machine learning models were designed: a MARS model to predict minimum glucose following exercise, a logistic regression model to predict hypoglycemia following exercise, and an ARX model to predict CGM values at the end of exercise. A second personalized MARS model was designed that incorporates participant exercise history features as inputs to the model (Supplementary Table S2). Each population model was designed using a training set, which consisted of data from 16 participants. The population machine learning models were trained using leave-one-participant-out cross-validation, meaning the input features and model parameters were selected using fifteen of the participants in the training set, and then performance was evaluated on the sixteenth held-out participant. The machine learning models were then evaluated on data from a holdout set, which consisted of data from the 4 participants who were not used in the training set. These 4 holdout participants were sampled to ensure that they were representative of the population and had the same frequency of hypoglycemia and minimum glucose as the training set. The general predictive accuracy of the models were evaluated using a 20-fold leave-one-participant-out cross-validation, where the model parameters were retrained on 19 participants and the model performance evaluated on 1 held-out participant (Supplementary Figure S2).

4.4.3.1 MARS Model to predict minimum SMBG following exercise

A MARS model implements a linear regression framework that also considers the numerical range of the predictors. Each input feature (Supplementary Table S1) was processed into paired hinge-functions, representing the feature values above and below a specific hinge point (ie, SMBG values above and below a hinge point of 150 mg/dL are with separate model coefficients). Next, *Greedy sequential variable selection* [196] was

used to iteratively identify optimal hinge-functions to predict minimum glucose during exercise. The MARS model coefficients were designed using a weighted regression; this approach places a penalty on MARS model misestimation of observations < 70 mg/dL. This essentially minimizes predictive error as well as improves sensitivity and specificity of the algorithm to detect hypoglycemia. The final model structure used to predict the minimum glucose during aerobic exercise is shown in equation (1). The model coefficients β_0, β_1, β_2, and β_3 along with the hinge points are solved for each model separately during model training.

$$\begin{aligned} &Minimum\ glucose\ during\ exercise \\ &= \beta_0 + \beta_1 * \max\left(0, CGM_{Start\ of\ Exercise} - 254\right) \\ &+ \beta_2 \max\left(0, CBG_{Start\ of\ Exercise} - 124\right) \\ &+ \beta_3 \max\left(0, HR_{10\ minutes\ prior\ to\ exercise} - 97.15\right) \\ &+ \beta_4\ CGM\ Trend_{25\ minutes\ prior} \end{aligned}$$

Eq. 4.1 MARS model to predict minimum glucose during exercise

4.4.3.2 ARX model to predict CGM following exercise

Romero-Ugalde et al. developed ARX models to forecast CGM measurements during aerobic exercise[65]. We used the methods and features described by Romero-Ugalde to reproduce a population ARX model. The exercise sessions in our dataset lasted on average for 43.2 ±14 minutes, therefore the ARX model was designed to predict CGM 40 minutes following the start of exercise. The final model structure is shown below in equation (2) where the coefficients β_0, β_1, β_2, and β_3 are solved for during model training.

$$CGM\ 40\ minutes\ after\ the\ start\ of\ exercise$$
$$= \beta_0 + \beta_1 CGM_{Start\ of\ Exercise}$$
$$+ \beta_2 CGM_{10\ minutes\ preceding\ exercise}$$
$$+ \beta_3 CGM_{20\ minutes\ preceding\ exercise}$$

Eq. 4.2 ARX model to predict CGM during exercise

4.4.3.3 Logistic regression to predict hypoglycemia during and following exercise

Breton et al. published a logistic regression model to predict hypoglycemia during exercise. We used the identical variables described by Breton et al. [45] to train a population logistic regression model to predict the occurrence of hypoglycemia during aerobic exercise and in the 4 hours following exercise. The inputs to this model were the CGM at the start of exercise, the average CGM trend in the hour preceding exercise, and the ratio of the active insulin (IOB) at the start of exercise to the participant's total daily insulin requirement (TDIR). The participant TDIR is defined as the total insulin dosed per day on average. The model is shown in equation (3) where the coefficients β_0, β_1, β_2, and β_3 are solved for during model training.

$$Probability\ of\ Hypoglycemia$$
$$= logit\ (\beta_0 + \beta_1 CGM_{Start\ of\ Exercise}$$
$$+ \beta_2 Average\ CGM\ Trend_{Prior\ Hour}$$
$$+ \beta_3 \frac{IOB_{Start\ of\ Exercise}}{TDIR})$$

Eq. 4.3 Logistic regression model to predict hypoglycemia during exercise

4.4.3.4 MARS model designed with exercise history as a form of personalization

The methods described above were used create a second personalized MARS model that incorporates exercise history from a given participant. The model was designed by identifying the optimal features included in Supplementary Table 1, and also exercise history features included in Supplementary Table 2 that describe participants' glucose dynamics during prior exercise sessions. The population model to detect minimum glucose during exercise is shown below in equation (4) whereby the coefficients β_0-β_6 were solved for each model separately during training of the model.

$$
\begin{aligned}
&Minimum\ glucose\ during\ exercise \\
&= \beta_0 + \beta_1 \max\left(0, CGM_{Start\ of\ Exercise} - 254\right) \\
&+ \beta_2 \max\left(0, 254 - CGM_{Start\ of\ Exercise}\right) \\
&+ \beta_3 \max\left(0, HR_{10\ minutes\ prior\ to\ exercise} - 97.15\right) \\
&+ \beta_4 \max(0, Average\ \Delta CGM_{other\ exercise\ sessions} + 84.92) \\
&+ \beta_5 \max(0, -84.92 - Average\ \Delta CGM_{other\ exercise\ sessions}) \\
&+ \beta_6 \max(0, 5.97 - Average\ MET_{other\ exercise\ sessions})
\end{aligned}
$$

Eq. 4.4 MARS + Exercise history model to predict minimum glucose

4.4.4 Real-time Model Adaptation

To determine the impact of adaptation on prediction accuracy, the population model parameters were adapted to each participant left-out of model training using data from the participant's exercise observations. Stochastic gradient descent [223] was used to update the population model parameters using the participant's most recent observed exercise session, and the adapted model was then used to predict the same

participant's outcomes of the next exercise session. This adaptation procedure was repeated successively for each held-out exercise observation, updating the population model parameters over time to better reflect a held-out participant's glucose dynamics as each exercise session was observed. In order to determine if the order of the exercise sessions impacted prediction accuracy, the order of the 8 identical exercise sessions were shuffled four times and the adaptation procedure was repeated.

4.4.5 Statistical Analysis

Significance of model error before and after personalization was evaluated using a two-tailed Wilcoxon signed-rank test, significance level of alpha = 0.05. Significance testing was performed on a per-participant level, df = 19, to compare the change in error and accuracy to detect hypoglycemia before and after personalization. Significance of glucose outcomes for participants in different physical fitness categories was evaluated using a two-tailed students t-test for parametric data, and a two-tailed Wilcoxon rank-sum test for non-parametric data, significance level alpha = 0.05. Glucose outcomes measured during exercise for each participant was explored with a boxplot. The centerline of the boxplot indicates the median measurement and box edges represent the 25^{th} and 75^{th} percentiles. Model performance was assessed using root mean squared error (RMSE), mean absolute error (MAE), as well as the sensitivity, specificity and accuracy to detect observations with level 1 hypoglycemia (< 70 mg/dL). Leave-4-participant-out cross-validation was used to create a receiver operating curve for each algorithm to determine the optimal predictive threshold to detect hypoglycemia. The optimal threshold for each algorithm was then used to evaluate algorithm sensitivity, specificity, and accuracy to detect hypoglycemia for left-out participant data (Supplementary Figure S2). The Parkes consensus error grid analysis[101] was used to determine the clinical impact of the algorithm predictions. Model design and assessment

were performed in Matlab 2019b (MathWorks, Natick, MA). A power analysis was performed previously for the published clinical study; a study size of 20 participants was sufficient to detect a -3.3% change in % time-in-hypoglycemia and a 16.3% change in % time-in-target glucose (70-180 mg/dL), for >80% power and an alpha = 0.0125 [28].

4.5 Acknowledgements

Funding. Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under Award Number 1DP3DK101044-01, F31DK121436 and 1 R01DK120367-01. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Duality of Interest. N.S.T. has nothing to disclose. P.G.J. and J. R. C. have a financial interest in Pacific Diabetes Technologies, Inc. a company that may have a commercial interest in this type of research. No other potential conflicts of interest relevant to the article were reported.

Author Contributions. N.S.T and P.G.J. contributed to the writing, literature review, methodology, data analysis and interpretation, figures and tables. J.R.C. and J.E.Y. contributed to the writing and data interpretation. C.M.L and G.M.Y. contributed to the data analysis and methodology. N.S.T. is the guarantor of this work and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Data Availability. De-identified research data used in this analysis is available as a CSV file, upon completion and signature of a sharing agreement, as defined by the Oregon Health & Science University Institutional Review Board (OHSU IRB).

Prior Publication. Parts of this study were presented as a poster at the American Diabetes Association 78th Scientific Sessions, Orlando, FL, June 22-26 2018

4.6 Supplementary Information

4.6.1 Supplementary Data

Table 4.S1: Supplementary Table. Data features obtained from participant wearables during a 4-arm artificial pancreas study.

Name	Description
CBG Start of Exercise	The glucose value as measured by SMBG at the start of aerobic exercise.
CGM Start of Exercise	The glucose at the start of exercise as measured by CGM. The CGM directly following SMBG calibration is used to avoid calibration artifact.
Average CGM Prior 25 minutes	The average glucose in the 25 minutes preceding exercise. If a calibration point is detected, the average CGM following the most recent calibration point is used.
CGM Trend (25, 20, 15, 10) Minutes Prior to Exercise	The glucose trends in the last 25 min preceding exercise, as calculated by CGM data. The trend is determined using data preceding SMBG calibration points, if SMBG is present at the start of exercise.
Average CGM Trend Prior Hour	The average glucose trend in the hour preceding exercise, calculated by CGM data.
Coefficient of Variation of Glucose	The coefficient of variation of CGM data in the 4 hours directly preceding aerobic exercise.
Insulin On Board	The current active insulin directly preceding the start of exercise. Measured in Units of insulin, as calculated using a linear decay formulation.
Meal On Board	The current active meal carbohydrate directly preceding the start of exercise. Measured in grams of carbohydrate, as calculated using a linear decay formulation.
Recent Meal	Binary variable indicating if a meal was consumed in the 2 hrs preceding exercise.
HR 10 minutes prior to exercise	The heartrate of the participant directly preceding the start of aerobic exercise.
Average HR prior 25 minutes	The average heartrate of the participant in the 25 min directly preceding exercise.
HR Trend (25, 20, 15, 10) Minutes Prior to Exercise	The trend in the heartrate of the participant in the 25 minutes directly preceding exercise.

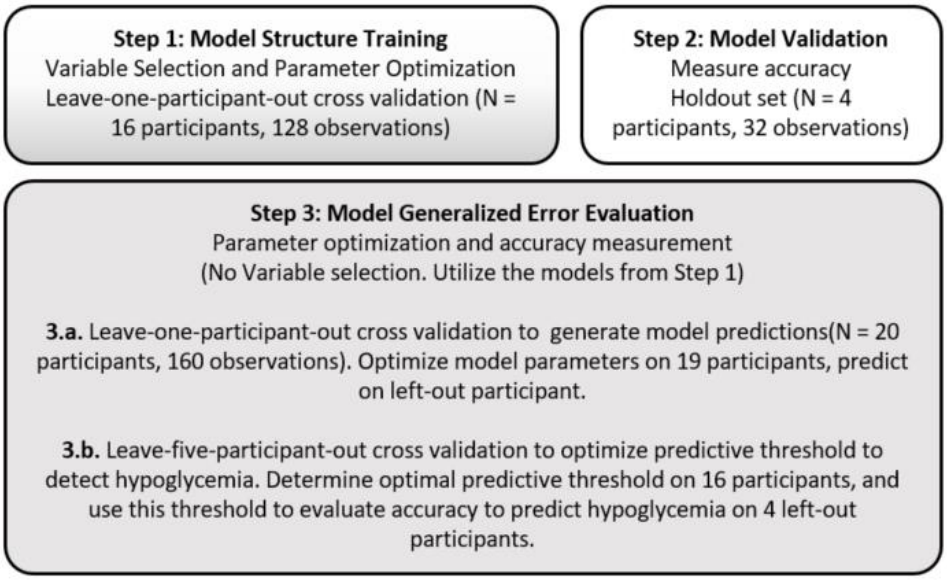
MET $_{\text{10 minutes prior to exercise}}$	The estimated metabolic expenditure of the participant directly preceding the start of exercise. Calculated using the Zakeri algorithm.
MET $_{\text{prior 25 minutes}}$	The average energy expenditure of the participant in the 25 min preceding aerobic exercise
Trend MET $_{\text{(25, 20, 15, 10) Minutes Prior to Exercise}}$	The trend in the energy expenditure of the participant in the 25 min preceding the start of aerobic exercise.
TDIR	Total daily insulin requirement of the participant in Units of insulin
TDIR $_{\text{estimated}}$	Total daily insulin requirement as the average insulin dosed over the preceding study days.
Weight	Weight in Kilograms
Height	Height in centimeters
Age	Age in years

Table 4.S2: Supplementary Table. Data features representing exercise history.

For example, on a given day of exercise, this exercise history data would represent the participant's data collected during other exercise sessions.

Features from prior exercise sessions	
Δ(CGM/CBG) $_{\text{Identical Exercise}}$	The change in glucose measured during a separate in-clinic exercise session.
Average Δ(CGM/CBG) $_{\text{in-clinic exercise sessions}}$	The average change in glucose measured during all other identical in-clinic exercise sessions recorded by the user.
Average Δ(CGM/CBG) $_{\text{other exercise sessions}}$	The average change in glucose measured during all other exercise sessions recorded by the user during the entire 4-week study, both in-clinic at home.
HR $_{\text{Identical Exercise}}$	The heartrate measured 10 minute after the start of exercise, recorded during a separate in-clinic exercise session.
Average HR $_{\text{in-clinic exercise sessions}}$	The heartrate measured 10 minutes after the start of exercise, collected during all other identical in-clinic exercise sessions recorded by the user, which is then averaged.
Average HR $_{\text{other exercise sessions}}$	The heartrate measured 10 minutes after the start of exercise, collected during all other exercise sessions recorded by the user during the entire 4-week study, both in-clinic at home, which is then.
MET $_{\text{Identical Exercise}}$	The metabolic expenditure measured 10 minutes after the start of exercise, recorded during an identical in-clinic exercise session.
Average MET $_{\text{in-clinic exercise sessions}}$	The metabolic expenditure measured 10 minutes after the start of exercise, averaged across all other identical in-clinic exercise sessions recorded by the user.
Average MET $_{\text{other exercise sessions}}$	The metabolic expenditure measured 10 minutes after the start of exercise, averaged across all other exercise sessions recorded by the user during the entire 4-week study, both in-clinic at home.

Figure 4.S3: Supplementary Figure. Model Training and Validation



5 New physiology models that utilize metabolic expenditure data from activity sensors to forecast changes in glucose during aerobic exercise

Summary:

- This chapter proposes new ordinary differential equation models that can be used to better simulate glucose trends of people living with type 1 diabetes during aerobic exercise.
- When incorporated into a virtual patient population, these models can be used to evaluate the performance of automated insulin delivery algorithms during exercise. These models can also be used to develop new methods to avoid hypoglycemia during aerobic exercise.
- The proposed models improve upon the existing OHSU T1D exercise simulator by better representing insulin and non-insulin mediated glucose uptake and by requiring only physical activity metrics from wearable devices to forecast glucose trends *in silico*.

Abstract

Background: People with type 1 diabetes can experience dramatic changes in glucose during aerobic exercise leading to hypoglycemia. Current ordinary differential equation (ODE) physiological models of glucose dynamics do not correctly estimate glucose changes during exercise. This project compares multiple candidate mathematical model structures of glucose metabolism during exercise. These candidate

models include metabolic expenditure as an input derived from body-worn physical activity sensors.

Methods: Models were identified and evaluated using a secondary analysis of data collected from adults with T1D who participated in a glucose-tracer infusion study (n=17, 11 F, weight 78.2 ± 11.0, TDIR 56.2 ± 12.7). After randomization into moderate intensity (40-45% VO_{2max}) vs. high intensity (60-65% $V0_{2max}$) groups, participants performed in-clinic exercise on three separate occasions clamped at infusions of low (basal), medium (1.5x basal), or high (3x basal) insulin infusion rates, and performed 45 minutes of aerobic treadmill exercise. Five models were designed to demonstrate how aerobic exercise impacts endogenous glucose production (EGP), glucose transport and disposal from plasma. These models were fit using a Hamiltonian Monte-Carlo sampling scheme in the RSTAN Bayesian inference software.

Results: Model 5 was the most accurate during model training (root-mean-squared error 0.84±0.32 mmol/L). Model 5 exhibited a lower RMSE than the existing OHSU T1D exercise simulator when evaluated on a separate validation dataset (RMSE 3.17±1.23 mmol/L vs 3.23±1.54 mmol/L).

Conclusions: New models are propose to simulate the impact of aerobic exercise on insulin- and non-insulin mediated glucose disposal, and endogenous glucose production. The proposed models improve upon the existing OHSU T1D exercise simulator accuracy and are feasible for incorporation into existing virtual patient populations and model predictive control closed-loop algorithms.

5.1 Introduction

Type 1 diabetes is caused by pancreatic β-cell dysregulation with subsequent insulin depletion. People who live with type 1 diabetes must maintain glucose within a

target range (3.9 -10 mmol/L) through injections of exogenous insulin. However, errors in insulin dosing, and variations in insulin sensitivity, often complicate glycemic control. Exercise has been shown to cause short-term increases in insulin sensitivity [33], reduce the risk of developing cardiovascular disease [31], and, more recently has been shown to improve glucose control in the 24-hour period following physical activity [32]. However sustained aerobic exercise also causes drastic changes in glucose and carries an increases the risk of hypoglycemia (< 3.9 mmol/L), which can lead to disorientation, blackouts, and even coma if untreated. Fear of hypoglycemia is a barrier to the incorporation of exercise into daily habits [35], and low confidence in controlling glucose during exercise has been reported by many people with type 1 diabetes [36]. There is great interest in mobile decision support tools that can help to manage glucose during exercise and avoid hypoglycemia [36].

Recent advancements in automated insulin delivery (AID) account for physical activity announcements in order to modify basal insulin rate or glucagon delivery prior to exercise [28, 212]. These approaches include manual entry of intended physical activity [42, 43], automatic detection through accelerometers [28], and also the incorporation of predictive models [51, 169]. However, this does not entirely prevent exercise-related hypoglycemia, even when glucagon is included as an additional hormone delivery of the system to help prevent hypoglycemia during and following exercise [28]. Many AID developers utilize population physiology models of insulin-glucose dynamics, also known as virtual patient populations [75, 103, 107] , in order to evaluate how AID algorithms will perform under conditions of exercise. These virtual patient populations have the advantage of forecasting or simulating glucose outcomes for given insulin dosing strategies or automated insulin delivery (AID) systems [39, 40] , or to replay an individual’s glycemic outcomes with different insulin doses [45, 56] . On a population level, our group has shown that the OHSU *in silico* virtual patient populations reflects

similar glucose outcomes to people with T1D from real-world clinical trials [57]. Other simulators have been published which are accepted by the FDA to validate insulin pump AID algorithms [75]. While there are many generalized models for insulin-glucose dynamics and controls, there are not many models that describe the impact of exercise on glucose [44]. Realistic virtual patient populations can better reflect the performance of AID systems in human populations. AIDs will perform better during exercise if they can incorporate models that can realistically describe how different types of physical activity impact glucose disposal and endogenous glucose production[227, 228].

Physical activity has been mathematically represented in glucose compartment models in various ways. Published models incorporate the relationship between active muscle mass and glucose disposal [84], the impact of VO_2 max on insulin secretion and hepatic glucose production [86], and the relationship between heartrate and insulin-independent and insulin-mediated glucose disposal [85]. In this chapter, we introduce new models of aerobic exercise that account for insulin-mediated and non-insulin mediated impacts of exercise, and incorporate these mechanisms into the design of the existing OHSU T1D virtual patient simulator [57]. The models were evaluated using a previously published dataset whereby people living with type 1 diabetes performed aerobic exercise under an insulin clamp to three insulin delivery rates and under moderate and intense exercise intensities [229]. The models were designed to utilize metabolic expenditure data collected from smartwatches or chest-strap activity monitors to estimate exercise intensity. Candidate model structures were optimized to match the clinical study dataset using Bayesian inference and Hamiltonian Monte Carlo sampling [77] . This powerful method is well-suited to the task of performing system identification to identify the parameters of a system of ordinary differential equations and ill-posed problems. The new models that describe the impact of aerobic exercise on glucose production and disposal are expected to ultimately be used to provide more accurate *in*

silico evaluation and design of AID systems and decision support systems under the conditions of aerobic exercise.

5.2 Research methods and design

5.2.1 Study Data

This secondary analysis utilized a unique dataset collected from adults living with type 1 diabetes who participated in a glucose tracer infusion study that investigated glucose disposal during exercise (n = 17, 11 F, weight 78.2 ± 11.0, TDIR 56.2 ± 12.7). [229]. Participants were randomized into either a moderate intensity (40-45% VO2~~max~~) exercise group or a high intensity (60-65% VO_{2max}) exercise group. Participants then each performed three in-clinic study days whereby they were clamped at an insulin rate of low (basal), medium (1.5x basal), or high (3x basal) insulin rates, and performed aerobic exercise at the intensity assigned to them. Enriched glucose tracer was infused during the study to mimic endogenous glucose appearance, and dextrose was infused in order to maintain glucose within target range (70-180 mg/dL) during the 3-hour run-in preceding exercise, and in the resting period following exercise to study closeout. Each participant was fitted with a ZephyrLife Biopatch (Zephyr, Annapolis, MD) chest monitor to collect activity and metabolic expenditure data. Plasma venous glucose concentration as measured by Yellow Springs Instrument 3200 (YSI Inc., Yellow Springs, OH) was measured in regular intervals during the study.

A second validation set was used to compare performance of the proposed exercise models with the existing virtual patient population exercise model. The validation set was obtained during a previous clinical study [28], from 20 adults with type 1 diabetes (N = 20, 14 F, Age 34.5 +/-4.7y, duration diabetes 19.7 +/- 8.6 y, BMI 26 +/-

5.7) participating in a clinical evaluating insulin pump therapies. Participants performed 2 in-clinic exercise studies, whereby participants consumed a self-selected breakfast in the morning, a self-selected lunch at 12 pm, and a self-selected dinner. Participants performed aerobic exercise at 2 pm in the afternoon, consisting of treadmill exercise at 70% VO_{2max} that lasted for 40 minutes. This validation analysis utilized participant data obtained from G5 CGM devices (Dexcom, San Diego, CA), t:slim insulin pump devices (Tandem, San Diego, CA), ZephyrLife BioPatch devices, and self-monitored blood glucose (SMBG) Contour Next devices (Bayer, Whippany, NJ). The exercise models proposed in this study were validated on study data collected during one arm of this previously published study, the single-hormone automated insulin therapy arm of this previously published study.

5.2.2 Data Processing

Data used in this analysis was obtained directly from insulin infusion rate [u/hr], dextrose infusion [mmol/kg/min], enriched glucose tracer [%], and blood glucose [mmol/L] that was collected for each participant study visit. Three-axis accelerometer and heartrate measurements were obtained from ZephyrLife BioPatch devices worn by study participants during exercise, and were used to calculate estimated metabolic expenditure (MET). All data was interpolated to a 5-minute sampling interval starting from minute 0 to the last measured glucose data point during the study prior to Bayesian inference procedures and validation procedures.

5.2.3 Model Structure

Five candidate models are proposed to forecast glucose dynamics during aerobic exercise. The candidate models are modified from an existing virtual patient model. This previously published virtual patient population consists of 8 mathematical compartments

that describe subcutaneous insulin transport and insulin action, as well as glucose uptake and disposal [57, 103, 106]. The virtual patient model that serves as a base structure for the candidate models is shown below.

5.2.3.1 Virtual patient model

Insulin absorption system. Insulin injections are represented by the value u_i, and insulin is shown to move subcutaneously through two compartments, S_1 and S_2. Insulin doses that diffuse in the plasma compartment, I, are converted to plasma concentration through the volume of distribution parameter V_I, and undergoes baseline clearance at a rate of k_e [57, 103, 106].

$$\dot{S_1} = u_I - \frac{S_1}{t_{maxI}}$$

$$\dot{S_2} = \frac{S_1}{t_{maxI}} - \frac{S_2}{t_{maxI}}$$

$$\dot{I} = \frac{S_2}{t_{maxI} V_I} - k_e I$$

Eq. 5.1 Insulin absorption system, virtual patient population

Insulin activity system. The plasma insulin, I, is represented to act on glucose homeostasis through three major functions: glucose transport (X_1), intracellular glucose disposal (X_2), and suppression of hepatic glucose production (X_3). These functional representations are activated from plasma insulin at a rate of Sf_1*k_{a1}, Sf_2*k_{a2}, and Sf_3*k_{a3}, respectively, and deactivated at a rate of k_{a1}, k_{a2}, and k_{a3} [57, 103, 106].

$$\dot{X_1} = -k_{a1} X_1 + S_{f1} k_{a1} I$$

$$\dot{X_2} = -k_{a2} X_2 + S_{f2} k_{a2} I$$

$$\dot{X_3} = -k_{a3} X_3 + S_{f3} k_{a3} I$$

Eq. 5.2 Insulin action system, virtual patient population

Glucose system. Glucose is represented as existing in two compartments, the observable plasma concentration, Q_1, and an unobservable interstitial and intracellular

compartment, Q_2. Glucose flows between these two compartments; glucose flows from plasma to tissues due to facilitation of insulin X_1, and back into plasma at a rate of k_{12}. Plasma glucose undergoes baseline clearance Fc_{01}, and renal excretion at a rate F_R, and further intracellular/interstitical clearance as facilitated by insulin X_2. Hepatic glucose enters plasma at a rate of EGP_0, and this process is suppressed by insulin by a factor of $(1-X_3)$. Glucose absorbed from consumption of carbohydrates is represented by U_g [57, 103, 106].

$$\dot{Q}_1 = -Q_1(X_1 + F_{01}^c) - F_R + k_{12}Q_2 + EGP_0(1 - X_3) + U_g$$

$$\dot{Q}_2 = X_1Q_1 - (k_{12} + X_2)Q_2$$

Eq. 5.3 Glucose system of virtual patient population

The models presented in this approach were personalized to each study participant in order to better estimate the impact of physical activity on glucose dynamics. To do this, the parameters Sf_1, Sf_2, and Sf_3, were statistically sampled to reflect the weight and total daily insulin requirement of each unique study participant, as described by Resalat et al. [57]. Participant data obtained from the exercise physiology study data [229] is inputted into the personalized physiologic models to mathematically simulate glucose trends during parameter optimization. The candidate model topologies introduced in this chapter (Figure 5.1 – Figure 5.5) include additional compartments and fluxes that are added to the base model.

5.2.3.2 Proposed Model 1

Model 1 hypothesizes physical activity to impact non-insulin-mediated glucose disposal directly from plasma. Model 1 includes two additional compartments to the virtual patient model, 1) an endogenous glucose flux compartment represented by EGP, and 2) an activity effect compartment represented by A_1. The final structure is a 10-compartment model that details the impact of physical activity on glucose clearance and

endogenous glucose production (Figure 5.1). Model 1 includes the virtual patient population structure described in equation (5.1)-(5.3), with modifications listed below in differential equation form.

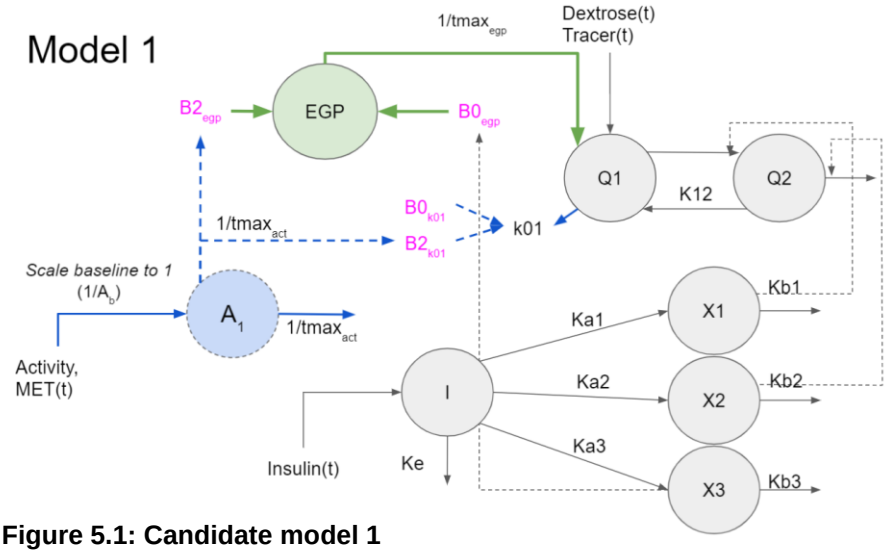


Figure 5.1: Candidate model 1

Physical activity is modeled to dispose of glucose directly from plasma, Q1, and increase endogenous glucose production. Endogenous glucose flux is represented by the green compartment, EGP, and the fast-acting impact of physical activity is represented by the blue compartment A_1. The grey compartments represent the existing virtual patient population. The solid lines connecting compartments represent transfer into and out of compartments, and the dotted lines represent the impact of compartments. Estimable parameters are represented by pink text, and fixed parameters are represented in black text.

Physical activity model. A new activity compartment, A_1, is proposed to represent the effect of physical activity on glucose physiology. Estimated metabolic expenditure data obtained from wearable devices¸ U_{MET}, is input into the A_1 compartment. Metabolic expenditure inputs are normalized to baseline activity data measured at rest, A_b. The effect of exercise on physiology is modeled to return to pre-activity baseline at a rate of $1/t_{maxAct}$. T_{maxAct} is fixed at a value of 10 minutes.

$$\dot{A_1} = \frac{1}{A_b} U_{MET} - \frac{1}{t_{maxAct}} A_1$$

Eq. 5.4 Activity system of proposed Model 1

Insulin absorption model. In the clinical exercise physiology study, participants were infused with IV insulin clamp, and did not undergo subcutaneous insulin infusion as

shown in the virtual patient insulin model Equation 5.1. The insulin absorption model is altered to represent insulin infusion rates of $U_{I.V.}$.

$$\dot{S_1} = U_{Injection} - \frac{S_1}{t_{maxI}}$$

$$\dot{S_2} = \frac{S_1}{t_{maxI}} - \frac{S_2}{t_{maxI}}$$

$$\dot{I} = U_{I.V.} + \frac{S_2}{t_{max} V_I} - k_e I$$

Eq. 5.5 Insulin absorption of proposed Model 1

Hepatic glucose flux model. This model proposes an additional endogenous glucose compartment¸ EGP, that is replenished by baseline endogenous glucose production β_{0EGP}. Baseline endogenous glucose production is suppressed by insulin by a factor of $(1-X_3)$. Hepatic glucose is modeled to increase due to the effect of physical activity A_1. A_1 is has a baseline of zero such that baseline physical activity does not impact endogenous glucose production. Delayed flux of endogenous glucose from the liver into plasma is represented by the component t_{maxEGP}, which is fixed at a value of 10 minutes. Baseline appearance of hepatic glucose β_{0EGP}, and the impact of physical activity on hepatic glucose flux, β_{2EGP}¸ are parameters that are identified using Bayesian methods further described below.

$$E\dot{G}P = \beta_{0EGP}(1 - X_3) + \beta_{2EGP}\left(\frac{A_1}{t_{maxAct}} - 1\right) - \frac{EGP}{t_{maxEGP}}$$

Eq. 5.6 Hepatic glucose flux system of proposed Model 1

Glucose model. Model 1 hypothesizes physical activity impacts the disposal of glucose directly out of plasma. The impact of physical activity on glucose clearance from plasma is represented by parameter k_{01}. This clearance parameter, k_{01}, is composed of an estimable baseline clearance, β_{0k01}, and the clearance due to activity measured above baseline, β_{2k01}. Activity effect from A_1 is centered at zero such that baseline physical activity does not impact glucose disposal. In the clinical study, participants were

infused with D_{10} glucose and D_2 tracer glucose directly into the plasma, and so absorption of glucose from the gut is not represented in equation 5.3. The Q_3 compartment was used to represent a delay in the mixing of infused glucose into the body from Q_3 into Q_1. The infused glucose is transported from Q_3 to Q_1 at a rate of k_{10}.

$$k_{01} = \beta_{0k01} + \beta_{2k01}\left(\frac{A_1}{t_{maxAct}} - 1\right)$$

Eq. 5.7 Glucose disposal of proposed Model 1

$$\dot{Q}_1 = -Q_1\left(X_1 + \frac{F_R}{GV_G} + k_{01} + F_{01}^c\right) + k_{12}Q_2 + \frac{1}{t_{maxEGP}}EGP + k_{10}Q_3$$

$$\dot{Q}_2 = X_1Q_1 - Q_2(k_{12} + X_2)$$

$$\dot{Q}_3 = U_{D2} + U_{D10} - k_{10}Q_3$$

Eq. 5.8 Glucose flux of proposed Model 1

Estimable parameters. The estimable parameters of proposed model 1 include (1) β_{0EGP}, baseline EGP generation, (2), β_{2EGP}, the impact of physical activity on EGP (3) β_{0k01}, baseline glucose clearance from plasma and (4) β_{2k01}, the impact of physical activity on non-insulin mediated glucose disposal from plasma.

5.2.3.3 Proposed Model 2

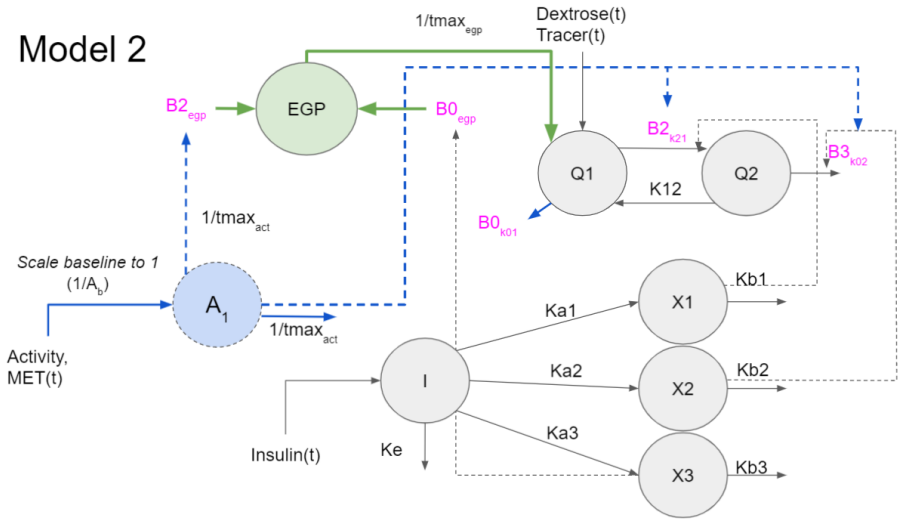


Figure 5.2: Candidate Model 2

Physical activity is modeled to impact non-insulin-mediated glucose transfer from plasma compartment Q_1 to unobserved compartment Q_2, and to impact insulin-mediated clearance of glucose from compartment Q_2. Note the blue dotted arrow leading to β_{3k02} impacts insulin - mediated clearance.

Model 2 hypothesizes physical activity to impact non-insulin-mediated transfer of glucose from plasma to unobserved tissue compartments, and to impact insulin-mediated disposal from the unobserved compartment (Figure 5.2). Model 2 retains the structure and modifications of proposed Model 1 (Equations 5.4-5.8) with additional modifications to the glucose model shown below.

Glucose model. The effect of exercise on glucose physiology from compartment A_1 is modeled to impact non-insulin-mediated transport of glucose from plasma Q_1 to the unobserved tissue space Q_2. A_1 is shown to modify the transfer rate k_{21} as determined by estimable parameter β_{2k21}. Physical activity is additionally defined to impact insulin-mediated disposal from the compartment Q2; A_1 is shown to modify the insulin-mediated disposal by X_2 via the estimable parameter β_{3k02}. Additional clearance of glucose directly from plasma is modeled by a parameter k_{01}, which is defined as the estimable parameter β_{0k01}.

$$k_{21} = \beta_{2k21}\left(\frac{A_1}{t_{maxAct}} - 1\right)$$

$$k_{02} = \beta_{3k02}\left(\frac{A_1}{t_{maxAct}} - 1\right)$$

$$k_{01} = \beta_{0k01}$$

Eq. 5.9 Glucose disposal of proposed Model 2

$$\dot{Q}_1 = -Q_1\left(X_1 + F_{01}^c + \frac{F_R}{GV_G} + k_{01} + k_{21}\right) + k_{12}Q_2 + \frac{1}{t_{maxEGP}}EGP + k_{10}Q_3$$

$$\dot{Q}_2 = Q_1(X_1 + k_{21}) - Q_2\left(k_{12} + X_2(1 + k_{02})\right)$$

$$\dot{Q}_3 = U_{D2} + U_{D10} - k_{10}Q_3$$

Eq. 5.10 Glucose flux of proposed Model 2

Estimable parameters. The estimable parameters of proposed model 2 include (1) β_{0EGP}, baseline EGP generation, (2) β_{2EGP}, the impact of physical activity on EGP, (3) β_{0k01}, baseline glucose clearance, (4) β_{2k21}, the impact of physical activity on non-insulin-mediated glucose transport out of plasma, and (5) β_{3k02}, the impact of physical activity on insulin-mediated glucose clearance.

5.2.3.4 Proposed Model 3

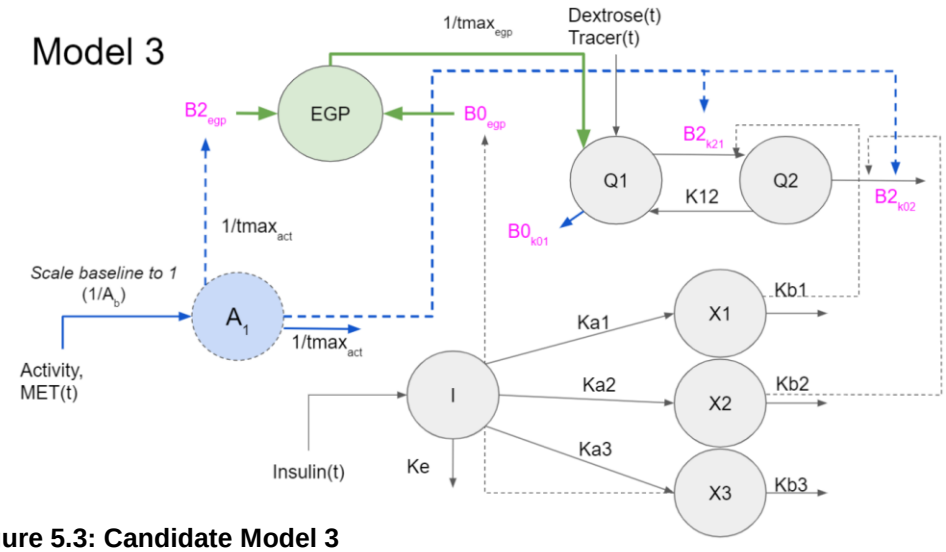


Figure 5.3: Candidate Model 3

Physical activity is modeled to impact non-insulin-mediated glucose transfer from plasma compartment Q_1 to unobserved compartment Q_2, and cause non-insulin-mediated clearance of glucose from the unobserved compartment Q_2. Note the blue dotted arrow leading to β_{2k02} differs from model 2, and is non-insulin-mediated.

Model 3 hypothesizes physical activity to impact non-insulin-mediated transfer and also non-insulin-mediated disposal of glucose (Figure 5.3). Model 3 retains the structure and modifications of Model 2, with additional modifications to the glucose system shown below.

Glucose model. The effect of exercise on glucose physiology from the compartment A_1 is modeled to impact non-insulin-mediated intracellular glucose disposal from the compartment Q2, via the estimable parameter β_{2k02}.

$$k_{21} = \beta_{2k21}\left(\frac{A_1}{t_{maxAct}} - 1\right)$$

$$k_{02} = \beta_{2k02}\left(\frac{A_1}{t_{maxAct}} - 1\right)$$

$$k_{01} = \beta_{0k01}$$

Eq. 5.11 Glucose disposal of proposed Model 3

$$\dot{Q}_1 = -Q_1\left(X_1 + F_{01}^c + \frac{F_R}{GV_G} + k_{01} + k_{21}\right) + k_{12}Q_2 + \frac{1}{t_{maxEGP}}EGP + k_{10}Q_3$$

$$\dot{Q}_2 = Q_1(X_1 + k_{21}) - Q_2(k_{12} + X_2 + k_{02})$$

$$\dot{Q}_3 = U_{D2} + U_{D10} - k_{10}Q_3$$

Eq. 5.12 Glucose flux of proposed Model 3

Estimable parameters. The estimable parameters of proposed Model 3 include (1) β_{0EGP}, baseline EGP generation, (2) β_{2EGP}, the impact of physical activity on EGP, (3) β_{0k01}, baseline glucose clearance, (4) β_{2k21}, the impact of physical activity on non-insulin-mediated glucose transport out of plasma, and (5) β_{2k02}, the impact of physical activity on non-insulin-mediated glucose clearance.

5.2.3.5 Proposed Model 4

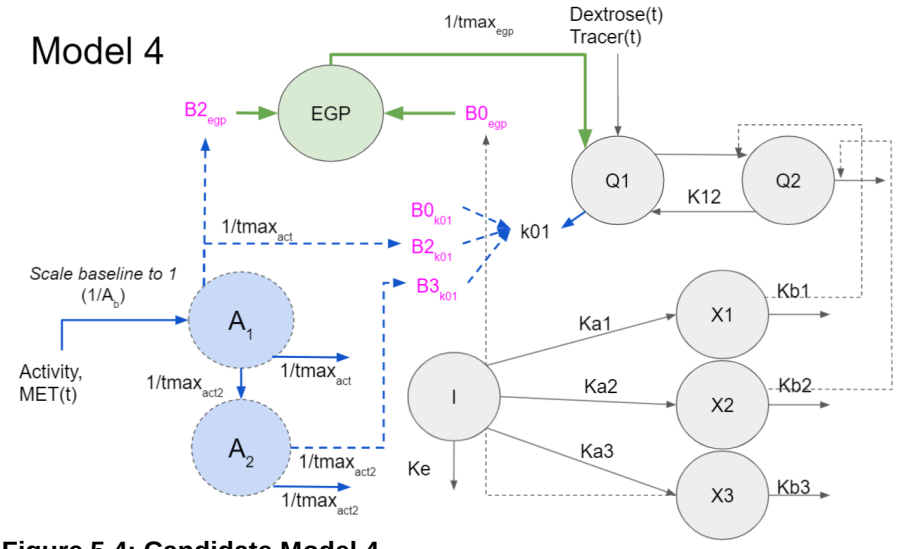


Figure 5.4: Candidate Model 4

Physical activity is modeled to dispose of glucose directly from plasma compartment Q_1. The impact of physical activity is represented by two compartments, A_1 and A_2, which represent the fast-acting and slow-acting impacts of physical activity on physiology, respectively. Both fast-acting and slow-acting processes contribute to glucose disposal from plasma. The fast-acting compartment A_1 is modeled to impact endogenous glucose production.

Model 4 hypothesizes physical activity to impact glucose disposal directly from plasma in two phases, a fast-acting disposal and a slow-acting disposal (Figure 5.4). Model 4 retains the structure and modifications of Model 1, including the hepatic glucose

flux system that is by the fast-acting activity component A_1. Additional modifications to the physical activity model and glucose model are shown below.

Physical activity model. The effect of physical activity on glucose physiology is modeled by two compartments, A_1 and A_2, which represent the fast-acting and slow-acting impacts of exercise, respectively. Estimated metabolic expenditure data obtained from wearable devices¸ U_{MET}, is input into the A_1 compartment. Metabolic expenditure inputs are normalized to baseline activity data measured at rest, A_b, and are centered around zero such that baseline physical activity does not impact glucose disposal. The fast-acting effect of exercise on physiology is modeled to return to pre-activity baseline at a rate of $1/t_{maxAct}$. T_{maxAct} is fixed at value 10 minutes. The slow-acting effect of exercise on physiology is modeled to return to pre-activity baseline at a rate of $1/t_{maxAct2}$, and $t_{maxAct2}$ is fixed at 400 minutes.

$$\dot{A_1} = \frac{1}{A_b} U_{MET} - \frac{1}{t_{maxAct}} A_1$$

$$\dot{A_2} = \frac{1}{t_{maxAct2}} A_1 - \frac{1}{t_{maxAct2}} A_2$$

Eq. 5.13 Activity system of proposed model 4

Glucose model. Model 4 hypothesizes the fast-acting and long-acting effect of exercise to dispose of glucose directly from plasma. The impact of physical activity on glucose clearance from plasma is represented by parameter k_{01}. This clearance parameter, k_{01}, is composed of an estimable baseline clearance, β_{0k01}, the clearance due to fast-acting activity, β_{2k01}, and the clearance due to slow-acting activity, β_{3k01}.

$$k_{01} = \beta_{0k01} + \beta_{2k01}\left(\frac{A_1}{t_{maxAct}} - 1\right) + \beta_{3k01}\left(\frac{A_2}{t_{maxAct}} - 1\right)$$

Eq. 5.14 Glucose disposal of proposed Model 4

$$\dot{Q}_1 = -Q_1\left(X_1 + \frac{F_R}{GV_G} + k_{01} + F_{01}^c\right) + k_{12}Q_2 + \frac{1}{t_{maxEGP}}EGP + k_{10}Q_3$$

$$\dot{Q}_2 = X_1Q_1 - Q_2(k_{12} + X_2)$$

$$\dot{Q}_3 = U_{D2} + U_{D10} - k_{10}Q_3$$

Eq. 5.15 Glucose model of proposed Model 4

Estimable parameters. The estimable parameters of proposed Model 4 include (1) β_{0EGP}, baseline EGP generation, (2) β_{2EGP}, the impact of physical activity on EGP, (3) β_{0k01}, baseline glucose clearance, (4) β_{2k01}, the fast-acting effect of physical activity on non-insulin-mediated glucose disposal from plasma, and (5) β_{3k01}, the slow-acting effect of physical activity on non-insulin-mediated glucose disposal from plasma.

5.2.3.6 Proposed model 5

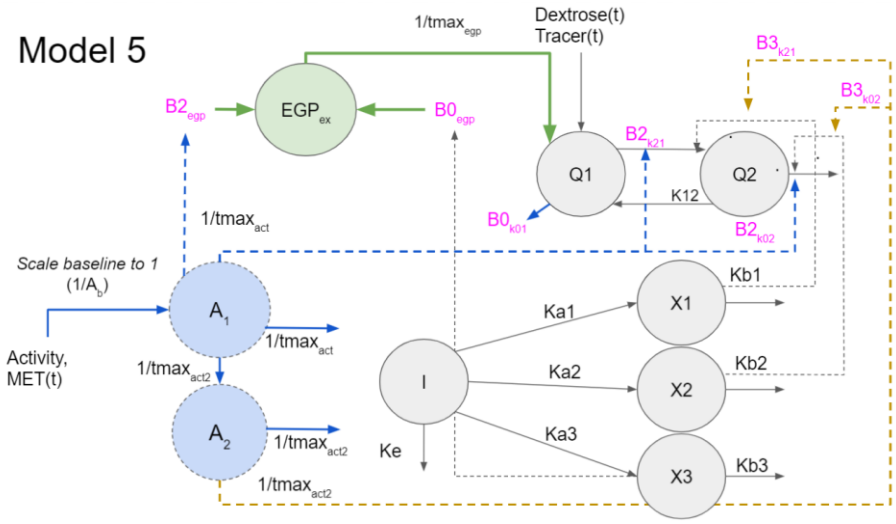


Figure 5.5: Candidate model 5

Physical activity is modeled to impact both insulin- and non-insulin-mediated glucose clearance through two timescales. The fast-acting activity compartment A_1 is modeled to impact non-insulin-mediated glucose transfer from Q_1 to Q_2, to increase non-insulin-mediated clearance of glucose from the unobserved compartment Q_2, and to increase endogenous glucose production. The slow-acting compartment A_2 is modeled to impact insulin-mediated glucose transfer from compartment Q_1 to compartment Q_2, and to impact insulin-mediated clearance of glucose from compartment Q_2.

Model 5 hypothesizes physical activity to impact glucose disposal in two phases, a fast-acting disposal and a slow-acting disposal. The fast-acting effect of physical

activity is modeled to impact non-insulin-mediated transfer and disposal of glucose. The slow-acting effect of physical activity is modeled to impact insulin-mediated transfer and disposal of glucose (Figure 5.5). Model 5 retains the structure and modifications of proposed Model 4, with additional modifications to the glucose system shown below.

Glucose model. The fast-acting effect of exercise represented by compartment A_1 is modeled to impact non-insulin-mediated transport of glucose from plasma Q_1 to the unobserved tissue space Q_2, via the transfer rate k_{21} and estimable parameter β_{2k21}. A_1 also impacts non-insulin-mediated glucose disposal from the compartment Q2 via the rate estimable parameter β_{2k02}. The slow-acting exercise effect A_2 impacts insulin-mediated transport of glucose and disposal of glucose. A_2 also increases transport of glucose from Q_1 to Q_2, by modifying X_1 activity via the rate k_{21_i} and estimable parameter β_{3k21}. A_2 also increases insulin-mediated disposal from compartment Q_2, by modifying X_2 activity via the rate k_{02_i} and estimable parameter β_{3k02}.

$$k_{21} = \beta_{2k21}\left(\frac{A_1}{t_{maxA}} - 1\right)$$

$$k_{02} = \beta_{2k02}\left(\frac{A_1}{t_{maxA}} - 1\right)$$

$$k_{21\_i} = \beta_{3k21}\left(\frac{A_2}{t_{maxA}} - 1\right)$$

$$k_{02\_i} = \beta_{3k02}\left(\frac{A_2}{t_{maxA}} - 1\right)$$

$$k_{01} = \beta_{0k01}$$

Eq. 5.16 Glucose disposal of proposed Model 5

$$\dot{Q}_1 = -Q_1\left(X_1 + k_{21} + F_{01}^c + \frac{F_R}{GV_G} + k_{01}\right) + k_{12}Q_2 + \frac{1}{t_{maxEGP}}EGP + k_{10}Q_3$$

$$\dot{Q}_2 = Q_1\left(X_1(1 + k_{21\_i}) + k_{21}\right) - Q_2\left(k_{12} + X_2\left(1 + k_{02\_i}\right) + k_{02}\right)$$

$$\dot{Q}_3 = U_{D2} + U_{D10} - k_{10}Q_3$$

Eq. 5.17 Glucose system of proposed Model 5

Estimable parameters. The estimable parameters of proposed Model 5 include (1) β_{0EGP}, baseline EGP generation, (2) β_{2EGP}, the impact of physical activity on EGP, (3) β_{0k01}, baseline glucose clearance, (4) β_{2k21}, the fast-acting effect of physical activity on non-insulin-mediated glucose transfer from plasma, Q_1, (5) β_{2k02}, the fast-acting effect of exercise on glucose disposal from Q_2, (6) β_{3k21}, the slow-acting effect of exercise on insulin-mediated glucose transfer from plasma, Q_1 and (7) the slow-acting effect of exercise on insulin-mediated glucose disposal from Q_2.

The complete ODE systems implemented in STAN for each model are shown in appendix 1-5.

5.2.4 Parameter optimization

Bayesian inference was used to optimize the unknown parameters of each candidate model structure in order to accurately simulate the participant data. Bayesian inference is used to define a theoretical joint probability distribution, p(y,x,θ), which represents the model system, the observed study data, x and y, and the unknown model parameters, θ. Bayes theorem is used decompose the theoretical joint probability distribution into two general expressions, show in Equation 5.18.

$$p(y, x, \theta) = p(y, x|\theta) * p(\theta) = p(\theta|y, x) * p(y, x)$$

Eq. 5.18 Bayes conditional probability theorem

The first term of Equation 5.18 describes the likelihood of the experimental data and model parameters, p(y,x|θ), and includes prior assumptions of the parameter values, p(θ). The second term of Equation 5.18 describes the probability of the inferred parameters given the study data, p(θ | y,x), and is termed the posterior probability. To solve for the posterior probability, equation 5.18 is rearranged and modified by the normalization factor p(y,x), seen in equation 5.19.

$$p(\theta|y,x) = \frac{p(y,x|\theta) * p(\theta)}{p(y,x)}$$

Eq. 5.19 Bayes conditional probability estimation

While an explicit formulation for each component of equation 5.19 cannot be defined except when the likelihood and the prior have conjugate forms, Hamiltonian Monte Carlo sampling [77] is a technique that is used to map the posterior probability distribution of each unknown parameter. The RSTAN package (R-Studio) was used to interface with STAN sampling software. The likelihood function is defined by the real-world data and is assumed to be sampled from a Gaussian distribution that is centered at the mean of the generative model estimates. The STAN implementation of the ODE models is shared in Appendices 1-5. The Gaussian likelihood function was implemented for two different datasets per exercise physiology study: 1) the enriched glucose tracer data was used to solve for candidate model parameters regarding glucose disposal, and 2) the overall glucose data was used to solve for candidate model parameters regarding glucose disposal and endogenous glucose production. The candidate ODE models were initialized at steady-state and were iteratively solved using Runge-kutta discretization within the STAN program. Each parameter assumed a uniform distribution, with specific constraints shown in Table 5.1.

Table 5.1: Boundaries imposed on parameter estimates.

MODEL	PARAMETER	(DESCRIPTION)	PRIORI CONSTRAINTS
ALL MODELS	β_{0k01}	(Baseline clearance of plasma glucose)	B ~ U(0,1)
	β_{0egp}	(Baseline hepatic glucose production)	B ~ U(0,0.2)
	$\beta 2_{egp}$	(Short-term impact of physical activity on hepatic glucose production)	B ~ U(0,1)
MODEL 1	β_{2k01}	(Short-term impact of physical activity on non-insulin mediated clearance of plasma glucose)	B ~ U(0,1)
MODEL 2	β_{2k21}	(Short-term impact of physical activity on non-insulin mediated glucose transport)	B ~ U(0,2)

	β_{3k02}	(Short-term impact of physical activity on insulin-mediated glucose clearance)	B ~ U(0,2)
MODEL 3	β_{2k21}	(Short-term impact of physical activity on non-insulin mediated glucose transport)	B ~ U(0,1)
	β_{2k02}	(Short-term impact of physical activity on non-insulin-mediated glucose clearance)	B ~ U(0,1)
MODEL 4	β_{2k01}	(Short-term impact of physical activity on non-insulin mediated clearance of plasma glucose)	B ~ U(0,1)
	β_{3k01}	(Long-term impact of physical activity on non-insulin mediated clearance of plasma glucose)	B ~ U(0,1)
MODEL 5	β_{2k21}	(Short-term impact of physical activity on non-insulin mediated glucose transport)	B ~ U(0,1)
	$\beta 2_{k02}$	(Short-term impact of physical activity on non-insulin-mediated glucose clearance)	B ~ U(0,1)
	$\beta 3_{k21}$	(Long-term impact of physical activity on insulin mediated glucose transport)	B ~ U(0,1)
	$\beta 3_{k02}$	(Long-term impact of physical activity on insulin mediated glucose clearance)	B ~ U(0,1)

The RSTAN model performs Hamiltonian Monte Carlo sampling of the estimated parameters for 3000 iterations. Each model was solved in RSTAN using 4 sampling chains, with 1000 sampling iterations used for a burn-in period and 2000 iterations used for parameter inference. Parameter convergence was evaluated by a defined parameter R_{hat}, which compares the standard deviation of sampled parameter values within a chain, and across all chains. If the parameter $R_{hat} = 1$, then the intrachain and interchain variability is equivalent and the sampling procedure has converged. The parameter values returned by the sampling procedure are then used to construct a posterior parameter distribution. The median value of the posterior distribution is interpreted as the model parameter estimate.

The sampling procedure was performed independently on each participant, and reflect the participant-specific glucose dynamics. The model parameters were estimated

by simultaneously comparing the data from the three study visits (1x basal insulin, 1.5x basal infusion, and 3x basal infusion) completed by each participant.

The candidate structures were compared for accuracy following Bayesian inference of model parameters. The root mean squared error (RMSE) was used to calculate the difference between model forecasts and the empirical study data. RMSE was calculated across the entire study, and also during the exercise period. The change in glucose during exercise was measured across the range of glucose values from the start of exercise to the end of the exercise period. RMSE was used to compare changes in glucose during exercise that were estimated by the model and the empirical data.

5.2.5 Model Validation

The candidate models were validated on a separate dataset not used for Bayesian inference and parameter optimization. Model topologies of the proposed candidate models were written in MATLAB script and imported into the OHSU virtual patient simulator. The median estimate was calculated across participant-specific parameter estimates for each model parameter, and used to create population candidate models. Virtual patients were matched to the clinical study participants in the validation dataset based on participant weight and TDIR. The population candidate models were initialized at steady-state 2 hours prior to the start of exercise in the validation set. Two validation scenarios were used to compare model performance. In the first validation scenario, virtual patients were treated with the clinical study data (IIR, Meal announcements, Rescue carbs, activity MET data). In the second validation scenario, virtual patients were treated with simulated AID pump therapy with OHSU's FMPD algorithm, and also clinical study data from meals and activity MET data. Model accuracy was assessed through RMSE calculated across the entire 4-hr simulation period, and during the exercise period.

5.3 Results

5.3.1 Model fit

Glucose forecasts were generated by the STAN sampling procedure during model optimization (Figure 5.6A), and are presented as an interquartile range across all participant study sessions for the candidate models (Figure 5.6B-F). This graph includes data from all insulin infusion rates, and exercise intensities. The root mean squared error (RMSE) between model forecasts and the empirical glucose collected from participants in the physiology study was calculated across the entire study and surrounding the exercise periods (minutes 180 to 225 from the start of the study) are shown in Table 5.2.

Model 4 exhibited the lowest RMSE overall, 0.83 mmol/L, and Model 5 exhibited the lowest RMSE during the exercise period, 0.98 mmol/L. Model 2 exhibited the highest RMSE overall, 1.10 mmol/L, and Model 1 exhibited the highest RMSE during the exercise period, 1.17 mmol/L.

Table 5.2: Training error measured between model forecasts and study data, averaged across all study participants.

MODEL	RMSE OVERALL [MEAN ± STDEV]	RMSE EXERCISE PERIOD [MEAN ± STDEV]
MODEL 1	0.94 ± 0.30	1.17 ± 0.57
MODEL 2	1.10 ± 0.35	1.06 ± 0.54
MODEL 3	0.85 ± 0.32	1.00 ± 0.50
MODEL 4	0.83 ± 0.32	0.99 ± 0.45
MODEL 5	0.84 ± 0.32	0.98 ± 0.50

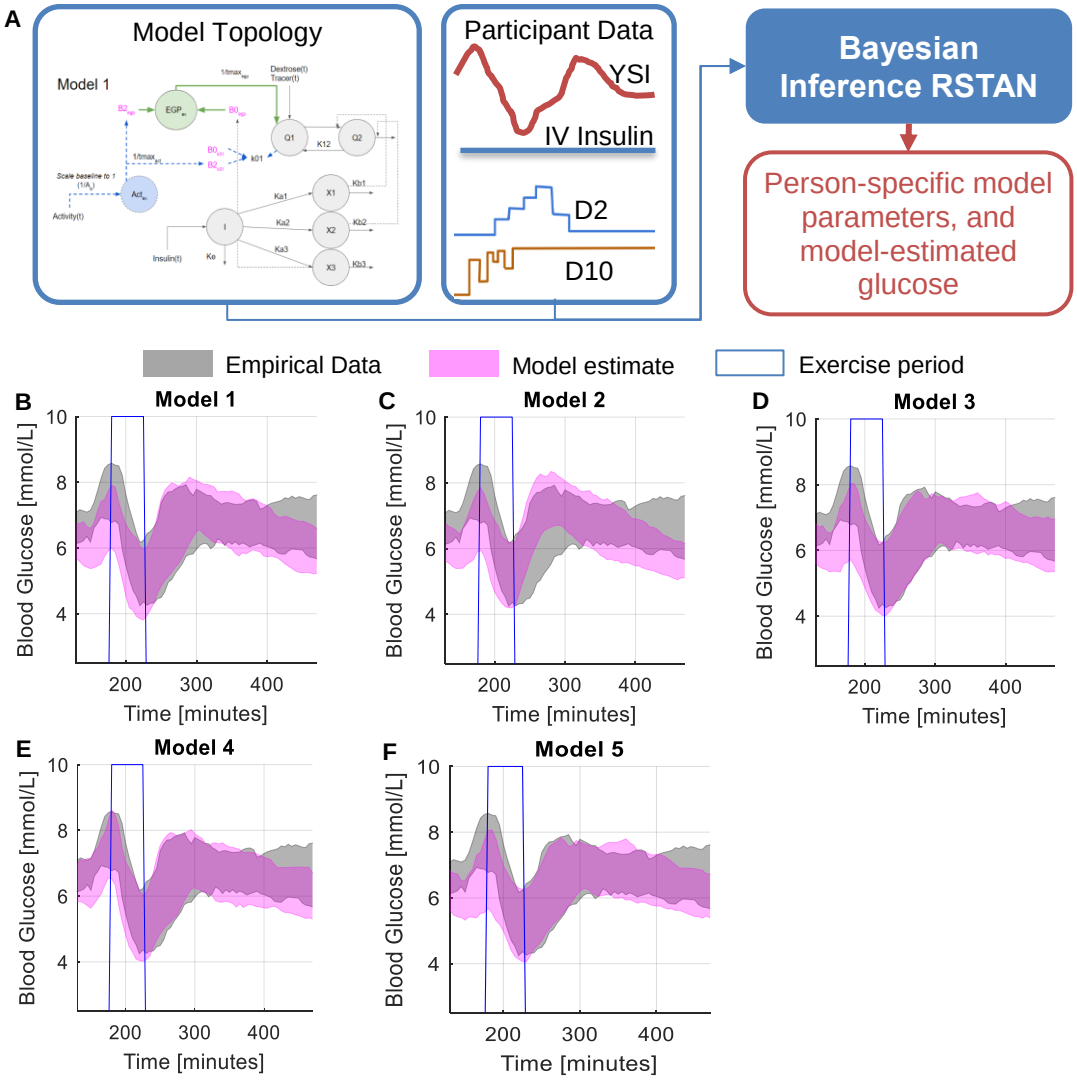


Figure 5.6: Model estimation results.

A. Optimization procedure for proposed candidate models parameters. **B-F.** Measured plasma glucose during an in-clinic exercise study, presented as an interquartile range across study visits. Aerobic exercise occurred between minutes 180-225.

5.3.2 Model validation

Validation of the proposed exercise models using a separate dataset is shown in Figures 5.7 and 5.8. The proposed models were also compared to the OHSU T1D exercise model [57]. Validation scenario 1 is displayed in Figure 5.7A, whereby the virtual patient models were initialized 2 hours prior to the start of exercise, and empirical

insulin-pump data was inputted into the virtual patient models. The RMSE was calculated between the virtual patient glucose trends and the empirical study glucose sensor data, and is reported in table 5.3. The exercise period occurred in simulation minutes 120-180. Model 5 exhibited the lowest RMSE across the entire simulation period, 3.17 mmol/L, as well the lowest RMSE during the exercise period, 3.16 mmol/L. Model 4 exhibited the highest RMSE across the entire simulation period, 3.32 mmol/L, and Model 1 exhibited the highest RMSE during the exercise period, 3.38 mmol/L (Table 5.3).

The OHSU T1D exercise model exhibited RMSE similar to the proposed models 1-5, however the OHSU T1D simulator exhibited considerably delayed glucose clearance in comparison to the empirical study data (Figure 5.7B). The proposed Models 1-5 did not exhibit a delay in glucose clearance during exercise, and generally estimated lower glucose ranges during and following the exercise period (Figure 5.7C-G).

Table 5.3: Validation error measured between study data and virtual patient glucose trends, averaged across all study participants.

MODEL	RMSE OVERALL [MEAN ± STDEV]	RMSE EXERCISE PERIOD [MEAN ± STDEV]
OHSU T1D EXERCISE SIMULATOR	3.23 ± 1.54	3.35 ± 2.25
MODEL 1	3.28 ± 1.20	3.38 ± 1.36
MODEL 2	3.30 ± 1.31	3.33 ± 1.57
MODEL 3	3.20 ± 1.28	3.22 ± 1.51
MODEL 4	3.32 ± 1.22	3.35 ± 1.35
MODEL 5	3.17 ± 1.23	3.16 ± 1.47

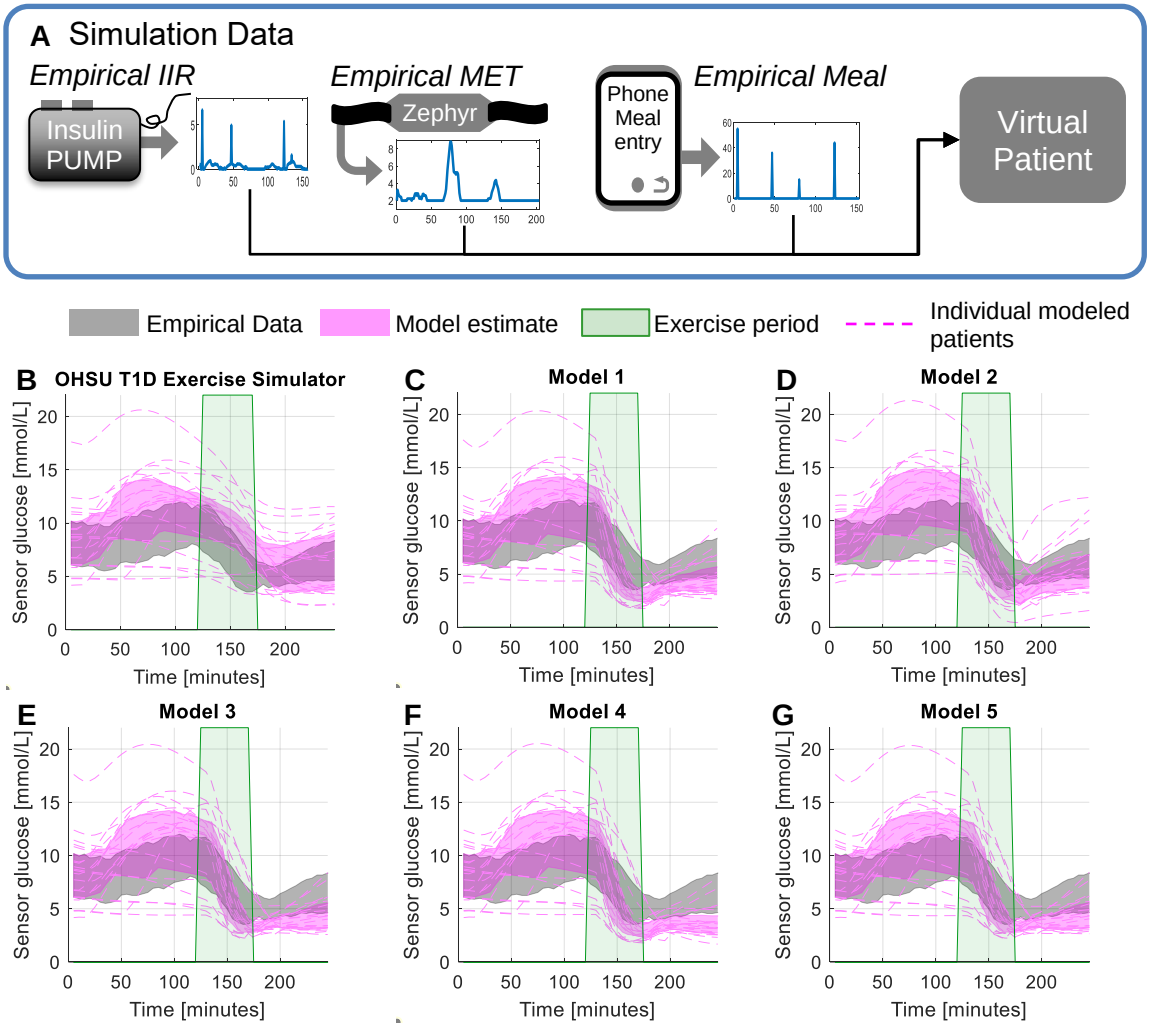


Figure 5.7: Model validation results of a virtual patient population treated with empirical insulin, meal, and exercise study data inputs

A. Input data sources for model validation in the virtual patient population, with empirical insulin inputs. **B-G.** Measured sensor glucose during an in-clinic exercise study, presented as an interquartile range across study visits. Aerobic exercise occurred between minutes 120-180.

The second validation scenario is displayed in Figure 5.8A, whereby the virtual patients were initialized 2 hours prior to the start of exercise, and underwent simulated AID insulin pump therapy by the OHSU FMPD algorithm. The RMSE calculated between the virtual patient glucose trends and the empirical study glucose sensor data is reported in Table 5.4.

The RMSE decreased for all models in validation scenario 2, whereby virtual patients underwent AID insulin pump therapy. Proposed Model 5 exhibited the lowest RMSE during the entire simulation period, 2.87 mmol/L. The OHSU T1D simulator exhibited the lowest RMSE during the exercise period, 2.87 mmol/L. Proposed Model 2 exhibited the highest RMSE across the entire simulation period, 2.98 mmol/L. Proposed Model 1 exhibited the highest RMSE during the exercise period, 3.15 mmol/L.

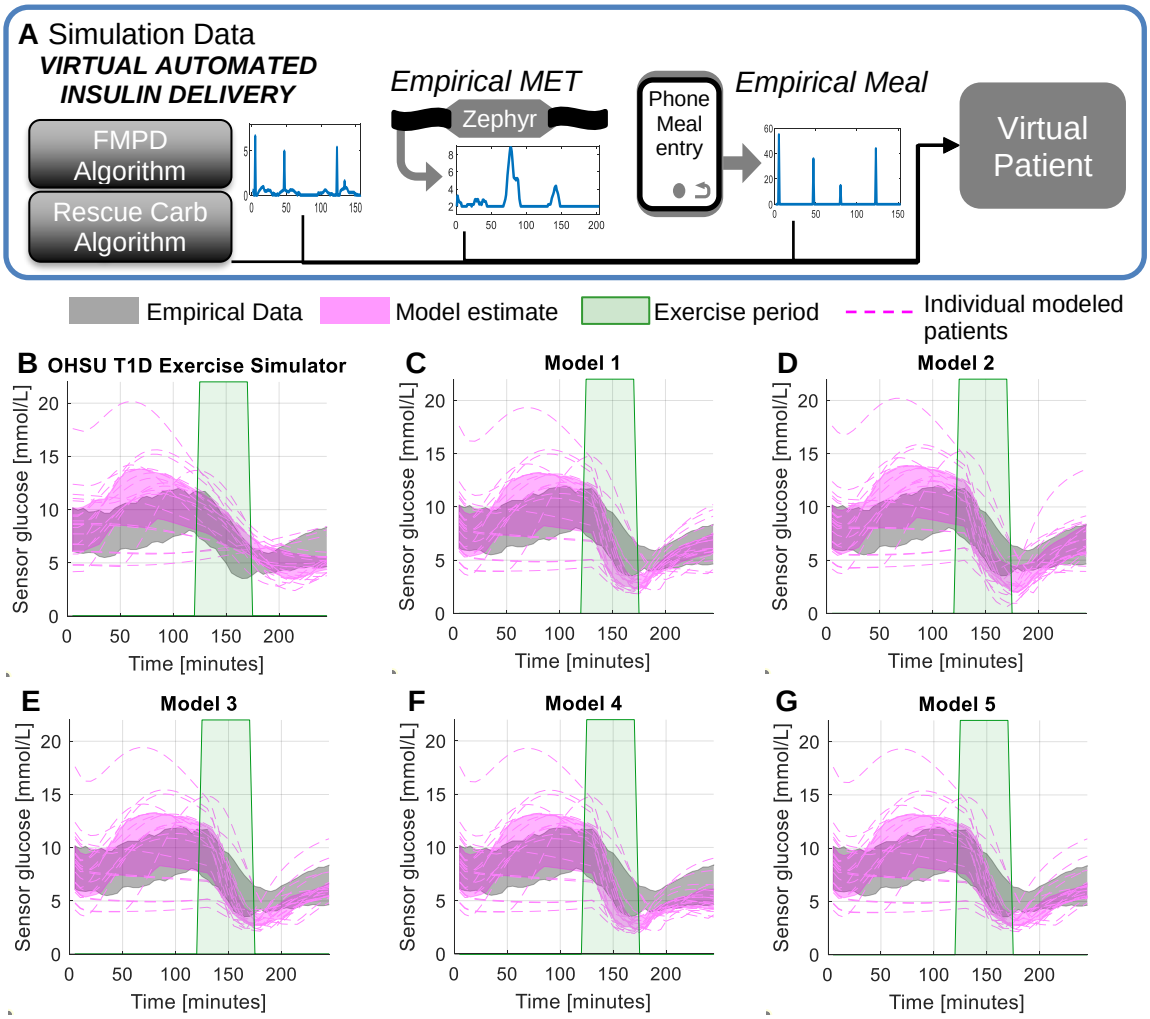


Figure 5.8: Model validation results of a virtual patient population treated with an automated insulin delivery algorithm, empirical meals and exercise study data inputs.

A. Input data sources for model validation in the virtual patient population, with empirical insulin inputs.

B-G. Measured sensor glucose during an in-clinic exercise study, presented as an interquartile range across study visits. Aerobic exercise occurred between minutes 120-180.

Table 5.4: Validation error measured between study data and virtual patient glucose trends, averaged across all study participants.

MODEL	RMSE OVERALL [MEAN ± STDEV]	RMSE EXERCISE PERIOD [MEAN ± STDEV]
OHSU T1D EXERCISE SIMULATOR	2.94 ± 1.24	2.87 ± 1.56
MODEL 1	2.97 ± 1.03	3.15 ± 1.40
MODEL 2	2.98 ± 1.14	3.13 ± 1.52
MODEL 3	2.89 ± 1.06	2.99 ± 1.43
MODEL 4	2.96 ± 1.04	3.13 ± 1.40
MODEL 5	2.87 ± 1.03	2.94 ± 1.39

5.3.3 Parameter Estimates

The RSTAN Bayesian inference procedure returned the median value of the sampled posterior distributions for each parameter. The median point-estimates for each parameter are presented as the average and interquartile range across all participants in Table 5.5. Further grouping of parameters into participants with moderate and intense exercise intensity are also presented in Table 5.5. The baseline glucose clearance parameters, B_{0k01}, were implemented for simulation purposes in RSTAN, and were not intended for use in the final virtual patient aerobic exercise models. As a result, the point estimates for baseline glucose clearance are not reported.

The point estimates for each proposed model parameter did not differ significantly when compared between participants who performed moderate exercise or intense exercise. For proposed Model 2, the point estimates for the parameter B_{3k02} settled at the pre-defined boundary value of 2. All other parameters for the candidate models converged within the defined boundary values and did not push against the assigned constraints.

Table 5.5: Estimated posterior distribution of model parameters across all participants and study dates.

Data is represented as mean, and 25th and 75th quartile. Point estimates of the range are the median of the estimated posterior distribution for each participant's simulation.

MODEL	PARAMETER	ALL PARTICIPANTS, N = 17			MODERATE EXERCISE, N = 10			INTENSE EXERCISE, N = 7		
		MEAN,	MEDIAN,	IQR	MEAN,	MEDIAN,	IQR	MEAN,	MEDIAN	IQR
MODEL 1	β0egp	1.38E-02	1.36E-02	[0.011,0.016]	1.23E-02	1.18E-02	[0.009,0.016]	1.59E-02	1.50E-02	[0.012,0.020]
	β2egp	1.01E-02	8.23E-03	[0.007,0.016]	9.17E-03	7.97E-03	[0.006,0.015]	1.14E-02	1.18E-02	[0.007,0.016]
	β2k01	3.86E-02	3.80E-02	[0.026,0.048]	3.84E-02	3.37E-02	[0.026,0.052]	3.89E-02	3.88E-02	[0.032,0.047]
MODEL 2	β0egp	1.58E-02	1.46E-02	[0.012,0.018]	1.38E-02	1.37E-02	[0.01,0.018]	1.87E-02	1.65E-02	[0.015,0.023]
	β2egp	9.06E-04	1.05E-04	[6E-05,4E-04]	1.41E-03	1.03E-04	[6.2E-05,0.004]	1.81E-04	1.11E-04	[6.7E-05,0.0001]
	β2k21	4.21E-02	3.67E-02	[0.027,0.052]	4.14E-02	3.38E-02	[0.026,0.048]	4.32E-02	3.89E-02	[0.029,0.056]
	B3k02	1.89E+00	2.00E+00	[1.99,2.0]	1.82E+00	2.00E+00	[2.00, 2.00]	2.00E+00	2.00E+00	[2.0,2.0]
MODEL 3	β0egp	1.39E-02	1.42E-02	[0.010,0.016]	1.24E-02	1.30E-02	[0.009,0.016]	1.59E-02	1.50E-02	[0.012,0.021]
	β2egp	1.02E-02	9.81E-03	[0.007,0.013]	9.29E-03	1.04E-02	[0.007,0.011]	1.13E-02	9.22E-03	[0.007,0.017]
	β2k21	1.73E-02	1.68E-02	[0.006,0.026]	1.47E-02	1.27E-02	[0.003,0.022]	2.06E-02	1.87E-02	[0.012,0.026]
	β2k02	1.10E-01	1.03E-01	[0.068,0.147]	1.05E-01	1.14E-01	[0.037,0.177]	1.17E-01	9.38E-02	[0.087,0.118]
MODEL 4	β0egp	1.37E-02	1.35E-02	[0.011,0.016]	1.23E-02	1.21E-02	[0.009,0.016]	1.57E-02	1.45E-02	[0.012,0.02]
	β2egp	1.19E-02	1.11E-02	[0.008,0.018]	1.07E-02	1.07E-02	[0.008,0.016]	1.37E-02	1.22E-02	[0.010,0.019]
	β2k01	3.99E-02	3.90E-02	[0.027,0.051]	3.92E-02	3.45E-02	[0.026,0.052]	4.08E-02	4.20E-02	[0.032,0.049]
	β3k01	4.48E-02	4.91E-02	[0.012,0.073]	4.50E-02	3.81E-02	[0.01,0.086]	4.44E-02	4.91E-02	[0.017,0.067]
MODEL 5	β0egp	1.37E-02	1.35E-02	[0.010,0.016]	1.22E-02	1.19E-02	[0.009,0.015]	1.59E-02	1.50E-02	[0.012,0.02]
	β2egp	9.73E-03	8.97E-03	[0.007,0.012]	8.74E-03	9.54E-03	[0.007,0.011]	1.11E-02	8.97E-03	[0.007,0.016]
	β2k21	1.59E-02	1.30E-02	[0.005,0.022]	1.34E-02	9.14E-03	[0.002,0.021]	1.95E-02	1.92E-02	[0.011,0.025]
	β2k02	1.15E-01	1.11E-01	[0.078,0.173]	1.13E-01	1.16E-01	[0.042,0.173]	1.18E-01	9.42E-02	[0.089,0.118]
	β3k21	1.67E-01	1.02E-02	[0.005,0.207]	1.62E-01	7.23E-03	[0.005,0.216]	1.75E-01	1.17E-02	[0.0055,0.16]
	β3k02	6.32E-01	7.73E-01	[0.269,0.952]	6.72E-01	8.38E-01	[0.37,0.95]	5.76E-01	7.01E-01	[0.209,0.933]

5.4 Discussion

New models are proposed that modify virtual patient simulators with mathematical representations of glucose disposal during exercise. The models provide a mathematical representation of how aerobic exercise impacts insulin-mediated and non-insulin mediated glucose disposal.

In Model 1, parameters are introduced to represent non-insulin-mediated glucose disposal directly from plasma; β_{2k01} describes the impact of exercise on non-insulin mediated glucose clearance, and endogenous glucose production is represented by β_{0EGP} and β_{2EGP}. While the parameter estimates for Model 1 converged within the specified boundary conditions, the reported training error and validation error is higher than other models.

Models 2 and 3 are designed to reflect a more physiologically accurate topology, whereby exercise stimulates non-insulin-mediated glucose transport from the plasma compartment (Q_1), into an unobserved glucose compartment (Q_2), by the parameter β_{2k21}. In Model 2, the additional parameter β_{3k02} represents the impact of physical activity on insulin-mediated glucose clearance. However, the distribution of β_{3k02} was estimated to max-out at the specified upper bound, 2; this indicates that the Model 2 topology cannot adequately describe the study data and should be disqualified as a candidate model. In Model 3, the parameter β_{2k02} is designed to reflect the impact of physical activity on non-insulin-mediated intracellular glucose disposal from Q_2. The parameters for Model 3 converged within the defined constraints, and Model 3 exhibited fair performance during model training and validation.

Models 4 and 5 explored the two-phase impact of exercise on glucose disposal as defined by a fast-acting activity compartment (A_1) that acts on the order of 10

minutes, and a slow-acting activity compartment (A_2) that acts on the order of 400 minutes. Model 4 introduced two parameters, β_{2k02} and $\beta_{3k02,}$ to describe the fast-acting and slow-acting impact of physical activity on non-insulin-mediated plasma glucose disposal. While the parameters of Model 4 converged within the specified constraints and the training error for Model 4 was low, the validation error for Model 4 was higher than other proposed models. Model 5 adopted a more complex topology that better reflects exercise physiology; two parameters, β_{2k21} and β_{2k02}, are introduced to model the fast-acting impact exercise on non-insulin-mediated glucose transport and intracellular glucose disposal; and two more parameters, β_{3k21} and β_{3k02}, reflect the slow-acting impact of exercise on insulin-mediated glucose changes. Model 5 achieved a low RMSE during model training and model validation, and exhibited convergence of all parameters within the defined constraints. And, the model parameters did not differ significantly between participants who performed varying intensities of exercise, indicating that the Model 5 structure can appropriately describe both moderate and intense exercise.

The OHSU T1D virtual patient population [57] currently implements a model of aerobic exercise that utilizes percent active muscle mass and estimated PVO_{2max} to impact insulin-mediated glucose uptake [84]. While this model can accurately capture aerobic exercise dynamics, and has been used to develop and evaluate automated insulin delivery systems, there are limitations to consider. A drawback of this approach is that the percent active muscle mass input is difficult to measure, and is typically qualitatively estimated across exercise types. Another drawback is the requirement of PVO_{2max} as an input, which is a person-specific measure that requires in-clinic evaluation and is not directly measureable during daily activities. In the OHSU T1D simulator, PVO_{2max} is estimated by calculating the proportion of immediate metabolic expenditure, $U_{MET}(t)$, divided by an individual's maximum metabolic expenditure. Yet another drawback is that maximum metabolic expenditure is a person-specific metric that

requires intensive in-clinic exercise sessions. A major strength of the proposed Model 5 is that it requires only metabolic expenditure as an input. Metabolic expenditure can be approximated from wearable activity monitors using wrist-worn heartrate and activity graph data.

The OHSU T1D exercise model exhibited varying performance during model validation. While the OHSU T1D exercise model achieved the lowest RMSE when validated in virtual patients undergoing simulated AID therapy, these findings were not reproduced when same virtual patients dosed insulin according to empirical insulin data. In addition, the OHSU T1D exercise model exhibited a noticeable delay in exercise-mediated glucose clearance as compared to the empirical data and the proposed Models 1-5. Additional points for consideration include 1) the existing OHSU T1D model simulates insulin-mediated glucose disposal, requiring insulin-on-board to simulate exercise; 2) endogenous glucose production is not modeled to increase during exercise as observed in clinical studies [230]. While proposed Model 5 exhibits comparable performance to the existing OHSU T1D simulator, it also models insulin- and non-insulin-mediated changes in glucose during exercise, and the impact of exercise on endogenous glucose production.

Existing physiology studies have suggested that two major mechanisms induce changes in glucose during exercise; one is a fast-onset process that mediates glucose uptake, and the other is a slow and steady change that impacts insulin-mediated glucose disposal [33, 34]. Proposed Model 5 was designed to reflect these insulin- and non-insulin-mediated changes in glucose clearance according to a fast-acting and long-acting timescale. While these mechanisms were assumed to operate within fixed timescales of 10 minutes and 400 minutes, one future direction is to precisely estimate the time-scales associated with the fast-acting and long-acting effects of aerobic exercise.

In conclusion, we contribute a new framework to model the impact of aerobic exercise on glucose disposal. These approaches describe exercise-induced glucose transfer and disposal, and endogenous glucose production, using metabolic data obtained directly from wearable activity monitors. Incorporation of new aerobic exercise models in to the OHSU T1D simulator will have many applications, including the design of new AID systems that can better predict glucose trends during physical activity and the design of decision support systems to prevent hypoglycemia during aerobic exercise. Additional model topologies that capture the impact of various exercise modalities on glucose disposal are forthcoming.

5.5 Appendix

Model 1 RSTAN code

```
// Model for Tracer glucose ------------------------------------------------------------------------
functions {
real[] G_D2_ODE(real t, real[] Q, real[] parms, real[] rdummy, int[] idummy) {
   G = G_T_mmolkg/Vdg;
   if (G<4.5)     // Calculate Fc01 baseline disappearance of glucose from the first compartment
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;

   if (G < 9)     // Calculate Renal disappearance of glucose from the first compartment
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // mmol/kg/min

   k01 = B0k01 + B2k01*(Q[10]/tmaxk01 - 1);  // Dk01/dt (1/min)
   dQdt[1] = (- Fc01 - Fr/G/Vdg - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3];  // DQ1/dt, (mmol/kg/min)
   dQdt[2] = Q[7]*Q[1] - (k12 + Q[8])*Q[2];                                 // DQ2/dt, (mmol/kg/min)
   dQdt[3] = U_D2 - k10*Q[3];                                               // DQ3/dt, (mmol/kg/min)
   dQdt[4] = - (1/tmaxI)*Q[4];                                              // DS1/dt, (mU/kg/min)
   dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                               // DS2/dt, (mU/kg/min)
   dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                        // DI/dt,  (mU/L/min)
   dQdt[7] = kb1*Q[6] - ka1*Q[7];                                           // DX1/dt  (1/min/min)
   dQdt[8] = kb2*Q[6] - ka2*Q[8];                                           // DX2/dt  (1/min/min)
   dQdt[9] = kb3*Q[6] - ka3*Q[9];                                           // DX3/dt  (1/min)
   dQdt[10]= U_A/Ab - Q[10]/tmaxk01;                                        // DMET/dt (1/min)
   return dQdt;
 }

// Model for Overall Glucose -----------------------------------------------------------------------
real[] G_T_ODE(real t, real[] Q, real[] parms, real[] rdummy, int[] idummy) {

   G = G_T_mmolkg/Vdg; //(mmol/L)
   if (G<4.5)       // Calculate Fc01 baseline disappearance of glucose from the first compartment
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;
```

```
   if (G < 9) // Calculate Renal disappearance of glucose from the first compartment
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // (mmol/kg/min)

   // Calculate the exercise and insulin mediated disappearance from the first glucose compartment
   k01 = B0k01 + B2k01*(Q[10]/tmaxk01-1);

   dQdt[1] = (- Fc01 - Fr/G/Vdg - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3] +  (1/tmaxegp)*Q[11];// DQ1/dt (mmol/kg/min)
   dQdt[2] = Q[7]*Q[1] - (k12 + Q[8])*Q[2];                                    // DQ2/dt (mmol/kg/min)
   dQdt[3] = U_D2 + U_D10 - k10*Q[3];                                          // DQ3/dt (mmol/kg/min)
   dQdt[4] = - (1/tmaxI)*Q[4];                                                 // DS1/dt (mU/kg/min)
   dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                  // DS2/dt (mU/kg/min)
   dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                           // DI/dt  (mU/L/min)
   dQdt[7] = kb1*Q[6] - ka1*Q[7];                                              // DX1/dt (1/min/min)
   dQdt[8] = kb2*Q[6] - ka2*Q[8];                                              // DX2/dt (1/min/min)
   dQdt[9] = kb3*Q[6] - ka3*Q[9];                                              // DX3/dt (1/min)
   dQdt[10]= U_A/Ab - Q[10]/tmaxk01;                                           // Dk01/dt (1/min)
   dQdt[11]= B0egp*(1-Q[9]) +  B2egp*(Q[10]/tmaxk01-1) - (1/tmaxegp)*Q[11];    // DEGP/dt (mmol/kg/min)
   return dQdt;
 }

model { // reference the true data. The data is scaled.
  G_D2_mmolkg[1:nt] ~ normal(G_hat[1:nt, 1], 1.0E-3);//Tracer
  G_T_mmolkg[1:nt] ~ normal(G_hat[1:nt, 11], 1.0E-1);//Overall glucose
}
```

Model 2 RSTAN Code

```
// Tracer glucose model
real[] G_D2_ODE(real t, real[] Q, real[] parms, real[] rdummy, int[] idummy) {
   G = G_T_mmolkg/Vdg;
   if (G<4.5)     // Calculate Fc01 baseline disappearance of glucose from the first compartment
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;

   if (G < 9)     // Calculate Renal disappearance of glucose from the first compartment
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // mmol/kg/min

   k21 = B2k21*(Q[10]/tmaxact - 1);  // 1/min
   k02 = B3k02*(Q[10]/tmaxact - 1);  // 1/min
   k01 = B0k01;
   dQdt[1] = (- Fc01 - Fr/G/Vdg - k21 - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3];  // DQ1/dt, (mmol/kg/min)
   dQdt[2] = (Q[7]+k21)*Q[1] - (k12 + Q[8]*(1+k02))*Q[2];       // DQ2/dt, (mmol/kg/min)
   dQdt[3] = U_D2 - k10*Q[3];                                   // DQ3/dt, (mmol/kg/min)
   dQdt[4] = - (1/tmaxI)*Q[4];                                  // DS1/dt, (mU/kg/min)
   dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                   // DS2/dt, (mU/kg/min)
   dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];            // DI/dt,  (mU/L/min)
   dQdt[7] = kb1*Q[6] - ka1*Q[7];                               // DX1/dt  (1/min/min)
   dQdt[8] = kb2*Q[6] - ka2*Q[8];                               // DX2/dt  (1/min/min)
   dQdt[9] = kb3*Q[6] - ka3*Q[9];                               // DX3/dt  (1/min)
   dQdt[10]= U_A/Ab - Q[10]/tmaxact;                            // DMET/dt (1/min)
   return dQdt;
 }


// Overall Glucose Model
real[] G_T_ODE(real t, real[] Q, real[] parms, real[] rdummy, int[] idummy) {
   G = G_T_mmolkg/Vdg;          //(mmol/L)
   if (G<4.5)     // Calculate Fc01 baseline disappearance of glucose from the first compartment
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;

   if (G < 9)     // Calculate Renal disappearance of glucose from the first compartment
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // (mmol/kg/min)
```

```
    // Calculate the exercise and insulin mediated disappearance from the first glucose compartment
    k21 = B2k21*(Q[10]/tmaxact - 1);  // 1/min
    k02 = B3k02*(Q[10]/tmaxact - 1);  // 1/min
    k01 = B0k01;

    dQdt[1] = (- Fc01 - Fr/G/Vdg - k21 - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3] +  (1/tmaxegp)*Q[11];// DQ1/dt
(mmol/kg/min)
    dQdt[2] = (Q[7]+k21)*Q[1] - (k12 + Q[8]*(1+k02))*Q[2];                    // DQ2/dt (mmol/kg/min)
    dQdt[3] = U_D2 + U_D10 - k10*Q[3];                                        // DQ3/dt (mmol/kg/min)
    dQdt[4] = - (1/tmaxI)*Q[4];                                               // DS1/dt (mU/kg/min)
    dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                // DS2/dt (mU/kg/min)
    dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                         // DI/dt  (mU/L/min)
    dQdt[7] = kb1*Q[6] - ka1*Q[7];                                            // DX1/dt (1/min/min)
    dQdt[8] = kb2*Q[6] - ka2*Q[8];                                            // DX2/dt (1/min/min)
    dQdt[9] = kb3*Q[6] - ka3*Q[9];                                            // DX3/dt (1/min)
    dQdt[10]= U_A/Ab - Q[10]/tmaxact;                                         // Dk01/dt (1/min)
    dQdt[11]= B0egp*(1-Q[9]) +  B2egp*(Q[10]/tmaxact-1) - (1/tmaxegp)*Q[11];  // DEGP/dt (mmol/kg/min)
    return dQdt;
  }

model { // reference the true data. The data is scaled.
  G_D2_mmolkg[1:nt_L] ~ normal(G_hat[1:nt_L, 1], 1.0E-3); // Tracer
  G_T_mmolkg[1:nt_L] ~ normal(G_hat[1:nt_L, 11], 1.0E-1); // Overall glucose
}
```

Model 3 RSTAN Code

```
// Tracer glucose model
    G = G_T_mmolkg/Vdg;
    if (G<4.5) // Calculate Fc01 baseline disappearance of glucose from the first compartment
        Fc01 = F01/4.5/Vdg; //(1/min)
    else
        Fc01 = F01/G/Vdg;

    // Calculate Renal disappearance of glucose from the first compartment
    if (G < 9)
        Fr = 0;
    else
        Fr = 0.003*(G-9)*Vdg; // mmol/kg/min

// Calculate the non-insulin mediated transfer from the first glucose compartment, and non-insulin-mediated disposal from
the unobserved glucose compartment.
    k01 = B0k01;
    k21 = B2k21*(Q[10]/tmaxAct - 1);  // Dk01/dt (1/min)
    k02 = B2k02*(Q[10]/tmaxAct - 1);  // Dk01/dt (1/min)
    // This function will be used to model the D2 infusion, the glucose tracer that mimics endogenous glucose production.
          // Disappearance                    //Appearance Q2, Infusion,
    dQdt[1] = (- Fc01 - Fr/G/Vdg - k21 - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3];   // DQ1/dt, (mmol/kg/min)
    dQdt[2] = (k21 + Q[7])*Q[1] - (k02 + k12 + Q[8])*Q[2];                          // DQ2/dt, (mmol/kg/min)
    dQdt[3] = U_D2 - k10*Q[3];                                                      // DQ3/dt, (mmol/kg/min)
    dQdt[4] = - (1/tmaxI)*Q[4];                                                     // DS1/dt, (mU/kg/min)
    dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                      // DS2/dt, (mU/kg/min)
    dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                               // DI/dt,  (mU/L/min)
    dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                  // DX1/dt  (1/min/min)
    dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                  // DX2/dt  (1/min/min)
    dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                  // DX3/dt  (1/min)
    dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                              // DMET/dt (1/min)
    return dQdt;


// Overall glucose model
    G = G_T_mmolkg/Vdg;          //(mmol/L)
    if (G<4.5)     // Calculate Fc01 baseline disappearance of glucose from the first compartment
        Fc01 = F01/4.5/Vdg; //(1/min)
    else
        Fc01 = F01/G/Vdg;

    // Calculate Renal disappearance of glucose from the first compartment
    if (G < 9)
```

```
    Fr = 0;
  else
    Fr = 0.003*(G-9)*Vdg; // (mmol/kg/min)

  // Calculate the non-insulin mediated transfer from the first glucose compartment, and non-insulin-mediated disposal
from the unobserved glucose compartment.
  k01 = B0k01;
  k21 = B2k21*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)
  k02 = B2k02*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)

            // Disappearance                              //Appearance Q2,  Infusion,  EGP compartment
  dQdt[1] = (- Fc01 - Fr/G/Vdg-k21-k01-Q[7])*Q[1] + k12*Q[2] + k10*Q[3] +(1/tmaxegp)*Q[11];// DQ1/dt (mmol/kg/min)
  dQdt[2] = (Q[7] + k21)*Q[1] - (k12 + k02 + Q[8])*Q[2];                                    // DQ2/dt (mmol/kg/min)
  dQdt[3] = U_D2 + U_D10 - k10*Q[3];                                                        // DQ3/dt (mmol/kg/min)
  dQdt[4] = - (1/tmaxI)*Q[4];                                                               // DS1/dt (mU/kg/min)
  dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                                // DS2/dt (mU/kg/min)
  dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                                         // DI/dt  (mU/L/min)
  dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                            // DX1/dt (1/min/min)
  dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                            // DX2/dt (1/min/min)
  dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                            // DX3/dt (1/min)
  dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                                        // DMET/dt (1/min)
  dQdt[11]= B0egp*(1-Q[9]) +  B2egp*(Q[10]/tmaxAct1-1) - (1/tmaxegp)*Q[11];                 // DEGP/dt (mmol/kg/min)

  return dQdt;
```

Model 4 RSTAN Code

```
// Tracer glucose model
  // Calculate Fc01 baseline disappearance of glucose from the first compartment
  G = G_T_mmolkg/Vdg;
  if (G<4.5)
    Fc01 = F01/4.5/Vdg; //(1/min)
  else
    Fc01 = F01/G/Vdg;

  // Calculate Renal disappearance of glucose from the first compartment
  if (G < 9)
    Fr = 0;
  else
    Fr = 0.003*(G-9)*Vdg; // mmol/kg/min

  k01 = B0k01 + B2k01*(Q[10]/tmaxAct1 - 1) + B3k01*(Q[11]/tmaxAct1 - 1);  // Dk01/dt (1/min)
  // This function will be used to model the D2 infusion, the glucose tracer that mimics endogenous glucose production.
       // Disappearance                        //Appearance Q2,  Infusion,
  dQdt[1] = (- Fc01 - Fr/G/Vdg - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3];          // DQ1/dt, (mmol/kg/min)
  dQdt[2] = Q[7]*Q[1] - (k12 + Q[8])*Q[2];                                         // DQ2/dt, (mmol/kg/min)
  dQdt[3] = U_D2 - k10*Q[3];                                                       // DQ3/dt, (mmol/kg/min)
  dQdt[4] = - (1/tmaxI)*Q[4];                                                      // DS1/dt, (mU/kg/min)
  dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                       // DS2/dt, (mU/kg/min)
  dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                                // DI/dt,  (mU/L/min)
  dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                   // DX1/dt  (1/min/min)
  dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                   // DX2/dt  (1/min/min)
  dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                   // DX3/dt  (1/min)
  dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                               // DMET/dt (1/min)
  dQdt[11]= Q[10]/tmaxAct2 - Q[11]/tmaxAct2;                                       // DMET2/dt (1/min)
  return dQdt;

// Overall glucose model
  // Calculate Fc01 baseline disappearance of glucose from the first compartment
  G = G_T_mmolkg/Vdg;          //(mmol/L)
  if (G<4.5)
    Fc01 = F01/4.5/Vdg; //(1/min)
  else
    Fc01 = F01/G/Vdg;

  // Calculate Renal disappearance of glucose from the first compartment
  if (G < 9)
    Fr = 0;
  else
    Fr = 0.003*(G-9)*Vdg; // (mmol/kg/min)
```

```
// Calculate the exercise and insulin mediated disappearance from the first glucose compartment
k01 = B0k01 + B2k01*(Q[10]/tmaxAct1-1)+B3k01*(Q[11]/tmaxAct1 - 1);

// This model mimics the total glucose in the compartments.
        // Disappearance                    //Appearance Q2,  Infusion,  EGP compartment
dQdt[1] = (- Fc01 - Fr/G/Vdg - k01 - Q[7])*Q[1] + k12*Q[2] +k10*Q[3]+(1/tmaxegp)*Q[12]; // DQ1/dt (mmol/kg/min)
dQdt[2] = Q[7]*Q[1] - (k12 + Q[8])*Q[2];                                                // DQ2/dt (mmol/kg/min)
dQdt[3] = U_D2 + U_D10 - k10*Q[3];                                                      // DQ3/dt (mmol/kg/min)
dQdt[4] = - (1/tmaxI)*Q[4];                                                             // DS1/dt (mU/kg/min)
dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                              // DS2/dt (mU/kg/min)
dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                                       // DI/dt  (mU/L/min)
dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                          // DX1/dt (1/min/min)
dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                          // DX2/dt (1/min/min)
dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                          // DX3/dt (1/min)
dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                                      // DMET/dt (1/min)
dQdt[11]= Q[10]/tmaxAct2 - Q[11]/tmaxAct2;                                              // DMET2/dt (1/min)
dQdt[12]= B0egp*(1-Q[9]) +  B2egp*(Q[10]/tmaxAct1-1) - (1/tmaxegp)*Q[12];               // DEGP/dt (mmol/kg/min)
```

Model 5 RSTAN Code

```
// Tracer glucose model
  G = G_T_mmolkg/Vdg;
   if (G<4.5)     // Calculate Fc01 baseline disappearance of glucose from the first compartment
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;

   // Calculate Renal disappearance of glucose from the first compartment
   if (G < 9)
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // mmol/kg/min

   k01 = B0k01;
   k21 = B2k21*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)
   k02 = B2k02*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)
   k21_i = B3k21*(Q[11]/tmaxAct1 - 1);
   k02_i = B3k02*(Q[11]/tmaxAct1 - 1);
   // This function will be used to model the D2 infusion, the glucose tracer that mimics endogenous glucose production.
        // Disappearance                                  //Appearance Q2,  Infusion,
   dQdt[1] = (- Fc01 - Fr/G/Vdg - k21 - k01 - Q[7])*Q[1] + k12*Q[2] +  k10*Q[3];    // DQ1/dt, (mmol/kg/min)
   dQdt[2] = (k21 + (1+k21_i)*Q[7])*Q[1] - (k02 + k12 + (1+k02_i)*Q[8])*Q[2];       // DQ2/dt, (mmol/kg/min)
   dQdt[3] = U_D2 - k10*Q[3];                                                       // DQ3/dt, (mmol/kg/min)
   dQdt[4] = - (1/tmaxI)*Q[4];                                                      // DS1/dt, (mU/kg/min)
   dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                       // DS2/dt, (mU/kg/min)
   dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                                // DI/dt,  (mU/L/min)
   dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                   // DX1/dt  (1/min/min)
   dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                   // DX2/dt  (1/min/min)
   dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                   // DX3/dt  (1/min)
   dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                               // DMET/dt (1/min)
   dQdt[11]= Q[10]/tmaxAct2 - Q[11]/tmaxAct2;                                       // DMET2/dt (1/min)
   return dQdt;


// Overall Glucose Model
   G = G_T_mmolkg/Vdg;   //(mmol/L)
   if (G<4.5)
      Fc01 = F01/4.5/Vdg; //(1/min)
   else
      Fc01 = F01/G/Vdg;

   // Calculate Renal disappearance of glucose from the first compartment
   if (G < 9)
      Fr = 0;
   else
      Fr = 0.003*(G-9)*Vdg; // (mmol/kg/min)
```

```
// Calculate the exercise and insulin mediated disappearance from the first glucose compartment
k01 = B0k01;
k21 = B2k21*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)
k02 = B2k02*(Q[10]/tmaxAct1 - 1);  // Dk01/dt (1/min)
k21_i = B3k21*(Q[11]/tmaxAct1 - 1);
k02_i = B3k02*(Q[11]/tmaxAct1 - 1);
// This model mimics the total glucose in the compartments.
          // Disappearance              //Appearance Q2,  Infusion,  EGP compartment
dQdt[1] = (- Fc01 - Fr/G/Vdg - k21 - k01 - Q[7])*Q[1] + k12*Q[2] +k10*Q[3] +(1/tmaxegp)*Q[12];// DQ1/dt (mmol/kg/min)
dQdt[2] = (k21 + (1+k21_i)*Q[7])*Q[1] - (k12 + k02 + (1+k02_i)*Q[8])*Q[2];                    // DQ2/dt (mmol/kg/min)
dQdt[3] = U_D2 + U_D10 - k10*Q[3];                                                            // DQ3/dt (mmol/kg/min)
dQdt[4] = - (1/tmaxI)*Q[4];                                                                   // DS1/dt (mU/kg/min)
dQdt[5] = (1/tmaxI)*Q[4] - (1/tmaxI)*Q[5];                                                    // DS2/dt (mU/kg/min)
dQdt[6] = U_I + (1/tmaxI)*Q[5]*(1/VdI) - ke*Q[6];                                             // DI/dt  (mU/L/min)
dQdt[7] = kb1*Q[6] - ka1*Q[7];                                                                // DX1/dt (1/min/min)
dQdt[8] = kb2*Q[6] - ka2*Q[8];                                                                // DX2/dt (1/min/min)
dQdt[9] = kb3*Q[6] - ka3*Q[9];                                                                // DX3/dt (1/min)
dQdt[10]= U_A/Ab - Q[10]/tmaxAct1;                                                            // DMET/dt (1/min)
dQdt[11]= Q[10]/tmaxAct2 - Q[11]/tmaxAct2;                                                    // DMET2/dt (1/min)
dQdt[12]= B0egp*(1-Q[9]) +  B2egp*(Q[10]/tmaxAct1-1) - (1/tmaxegp)*Q[12];                     // DEGP/dt (mmol/kg/min)
return dQdt;
```

6 Conclusion and future directions

Type 1 diabetes is a medical condition that requires daily management in order to maintain glucose within an optimal range [231], avoid short-term medical emergencies, and prevent long-term health complications. Exogenous insulin can be administered to enable glucose uptake and utilization, but the sole presence of exogenous insulin cannot compensate for the causative pancreatic dysfunction. Only precisely calculated, well-timed insulin doses that are titrated in response to dynamic changes in insulin sensitivity, can help to maintain a normoglycemic balance.

For people living with type 1 diabetes, this type of intensive insulin regimen requires life-long trial and error, with dangerous consequences in cases of therapeutic mistakes or mismanagement. A growing community of physicians and engineers aim to address these complications, and have developed technologies that provide real-time calculations for insulin doses, prediction of problematic glycemic patterns, insulin dose adjustment algorithms, and automated insulin and glucagon delivery algorithms in order to reduce user burden. Broadly, these algorithms provide decision support to people with type 1 diabetes in order to improve glucose control and avoid medical complications. The study of the diverse mechanisms which dictate changes in insulin sensitivity, with a particular emphasis on physical activity, occurs in parallel to technology development. These complementary fields of algorithm and exercise physiology research have increased the capacity of decision support and automated insulin delivery algorithms to modify insulin therapy in response to physical activity or other factors, and provide new *in silico* clinical trial frameworks to streamline the evaluation of these algorithms. The impact of these approaches is clear; evidence of improved glucose control with

automated insulin delivery algorithms, and a reduction, but not elimination, of hypoglycemia during daily activities and exercise [28, 29, 212].

This dissertation introduces new frameworks of decision support algorithms for the management of type 1 diabetes. The original work presented here contributes a new design strategy that utilizes virtual patient simulators to generate training data for the design artificial intelligence decision support algorithms. Data availability is a huge bottleneck in the design of both supervised and unsupervised machine learning algorithms. Oftentimes, data scientists face constraints in their analyses due to delays in data collection during clinical studies, lack of access to relevant existing datasets, sparsity of available data sets, or inherent limitations in the datasets due to the methods of data collection. The work presented in this dissertation overcomes these issues, demonstrating that a virtual patient population can be used to simulate patient data representing thousands of days of glucose trends following insulin therapy adjustments. We show that an artificial algorithm designed from this *in silico* data can recommend adjustments to insulin therapy that agree with recommendations of endocrinologists [62].

Another contribution of this dissertation is the creation of a new physiology model that describes glucose disposal during aerobic exercise in order to improve the realistic behavior of virtual patient simulators. The main limitation of *in silico* data generation is that the *in silico* data is only as realistic as the virtual patient simulator. The exercise model proposed in this dissertation builds on previous approaches of mathematical models of exercise physiology by utilizing metabolic expenditure data representing wearables such as chest harness or smart watch activity monitors, and modeling physiological changes in insulin and non-insulin mediated mechanisms of glucose disposal during exercise. Improvements in the realistic behavior of virtual patient populations provide many advantages to the realm of decision support and AID design.

New models of exercise dynamics are needed to better evaluate engineered algorithms designed to assist people with T1D avoid hypoglycemia during exercise. Incorporation of a new physiology model of aerobic exercise can provide more realistic evaluation of AID algorithms during *in silico* clinical trials. In addition, the updated virtual patient populations can be used to evaluate existing clinical guidelines that recommend insulin dose adjustment or carbohydrate intake preceding exercise, which often requires trial and error before demonstrating optimal performance in real-world settings. In addition, a model-predictive control closed-loop algorithm can use new models of aerobic exercise to accurately forecast glucose during exercise, and calculate changes in insulin dosing required to prevent hypoglycemia in the 24 hours following exercise. Our contributions in exercise physiology models add to the wealth of approaches designed to expand the modeling capabilities of virtual patient simulators; new models representing broad formulations of insulin, glucagon, and other therapeutics including pramlintide and insulin-sensitivity modifying agents will soon be incorporated into virtual patient simulators.

The incorporation of new models of aerobic exercise into existing virtual patient populations will also contribute to the design of new artificial intelligence algorithms. In chapter 2 of this dissertation, we reviewed the state of artificial intelligence in decision support and found many opportunities for the development of new recommender algorithms to assist people with diabetes to manage their glucose during exercise. The prediction of hypoglycemia in real-time is an ongoing effort for clinicians and engineers who work in the field of diabetes decision support. While many predictive algorithms of hypoglycemia have been described in the literature, the accuracy of these approaches are largely constrained to situations with low-to-moderate physical activity. Characterization of glucose trends during exercise is a complex task with varying clinical efficacy [28, 45, 51]. Chapter 4 reviews the glycemic response of adults who underwent

8 identical exercise scenarios in order to characterize the variability in glucose changes during exercise. This data is used to develop artificial intelligence algorithms that accurately estimate the occurrence of hypoglycemia during exercise and the 4-hour window following exercise. The importance of exercise history is further explored in this chapter, and is shown to impact the accuracy of the predictive algorithms to estimate changes in glucose during future exercise sessions. Lastly, it is shown that, in an ideal scenario, personalization of algorithms can significantly improve the accuracy to predict an individual's changes in glucose during exercise.

Taken together, this work provides new insights into the design of decision support algorithms for diabetes. The promising findings in chapter 3 can obviate common issues in data availability, in chapter 4 demonstrate the impact of personalization in predicting exercise-related hypoglycemia, and in chapter 5 provide new models to simulate glucose trends during aerobic exercise. The incorporation of these approaches into mobile decision support applications and insulin delivery systems are forthcoming and will be evaluated in clinical studies.

6.1 Future Directions

There are many opportunities to expand this work to address unsolved issues in decision support, as well as new applications and trends in diabetes technology.

6.1.1 Clinical targets for decision support systems

Complex artificial intelligence systems that exhibit stellar performance during *in silico* evaluation may only yield modest or insignificant improvements in outcomes during human trials [44]. There are many possible explanations for this, including (1) *in silico* evaluation strategies; *In silico* virtual patients are also assumed to adhere perfectly to

recommended therapies, however many factors affect adherence to therapy in real-world participants. As a result, virtual populations reflect optimistic algorithm performance; (2) inadequate design goals for decision support systems. Existing automated systems are largely focused on analyzing glycemic trends and identifying appropriate changes to insulin therapy, however these designs are not inclusive of complex social and behavioral factors that affect diabetes management in real-world populations; (3) length of real-world clinical evaluation. The therapeutic time-to-effect for weekly decision support is 12-weeks in perfectly adherent virtual patients, and will likely be delayed in real-world clinical trials where there is varying adherence to therapy. However, many clinical trials of automated decision support systems are less than 12 weeks in length.

In Figure 6.1, a multi-step process is outlined for clinical decision support systems. First, decision support systems are designed deliver appropriate recommendations to individuals living with diabetes, or their healthcare providers. Depending on the quality of the system, the recommendations can be assumed to be safe, effective, and easy to implement. Second, the recommendations must be reviewed and applied in daily practice—the automated system cannot undergo adequate evaluation if the system recommendations are not implemented in practice. Lastly, improvements in clinical outcomes may be measured after continued and prolonged use of the decision support therapy.

Automated decision support systems are designed to find optimal insulin therapy adjustments that can reduce hypoglycemia (glucose < 70 mg/dL) and improve percent time in range (70-180 mg/dL). These two glycemic targets are easily defined and quantified in virtual patient populations, and enable the use of *in silico* frameworks to estimate the clinical effect of adhering to proposed insulin adjustment therapy algorithms. However, virtual patient populations are approximations of glucose

dynamics, and can be improved to better reflect T1D physiology. Further work is needed to better reflect the diverse activities that impact insulin sensitivity and additional metabolic processes, and to model additional pharmaceutical therapies that affect glycemic control. In addition, virtual patient populations are limited in the ability to model realistic insulin dosing behaviors and schedules. As a result, engineering approaches that utilize virtual patient populations are largely constrained to basic insulin dosing scenarios, and do not quantitate or optimize social and psychological factors that affect glycemic control and adherence to prescribed therapies in real-world populations. In Figure 6.1, a red 'X' is shown to disrupt the decision support process between receiving recommended modifications to therapy, and incorporating those recommendations into daily practice. Limitations in system design that overlook psychosocial factors in clinical therapy may contribute to poor adherence to recommendations during real-world evaluation, and ultimately reduce the system performance in human studies.

In chapter 2, our comprehensive review of decision support systems for diabetes found that the most effective decision support systems were designed to provide recommendations that are accompanied by support from a certified diabetes educator or yearly review from physicians [93]. Certified diabetes educators or health coaches may provide support in order to resolve issues in system usability, education, and concerns about therapy, all of which are complex factors that are not easily addressed by engineered algorithms. Many smartphone applications are commercially available to provide decision support to people regarding health and weight loss, and have demonstrated consistent app-engagement and usability, user adherence, and improvements in health outcomes [232, 233]. A reasonable direction for automated diabetes decision support systems is to partner with existing frameworks that address exhibit targeted therapy utilization and coaching, thereby jointly addressing disruptions in the process outlined in Figure 6.1.

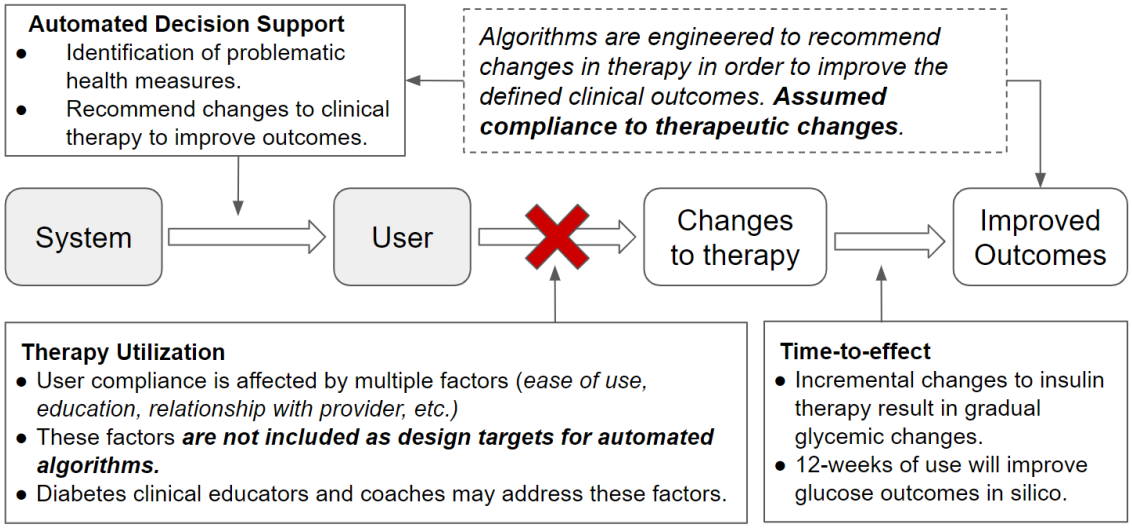


Figure 6.1: Process flowchart for clinical decision support systems.

Though a system may recommend impactful changes to insulin therapy, clinical outcomes may exhibit gradual changes that are not apparent until many months of use. In chapter 2 of this dissertation, it was observed that many proof-of-concept clinical trials of automated decision support systems were less than 12 weeks in length. In chapter 3 of this dissertation it was observed that *in silico* clinical evaluation of automated decision support did not measure statistically significant improvements in percent time in glucose target range until at least 12 weeks of use [62]. Recent clinical trials lasting at least 6 months in length have exhibited significant improvements in % A1C [61, 93]. It is possible that clinical trials may require at least 3 months of decision support therapy in order to properly evaluate decision support technology.

Taken together, many factors that affect glycemic control, including the accuracy of automated adjustments to insulin, adequate clinical trial length, and regular support from coaches or educators, must be addressed in order to significantly impact patient outcomes.

6.1.2 Exercise Decision Support

Another goal that stems from this dissertation is the creation of an exercise decision support system. New models that describe the impact of aerobic exercise and resistance exercise on glucose trends will be incorporated into a virtual patient population. Similar to the methods outlined in chapter 3, *in silico* trials can be used to simulate exercise of varying intensities, and modalities, across one week of multiple daily injection therapy. Resimulation procedures can be used to evaluate the impact of various treatment decisions on the clinical outcomes during exercise (i.e. consumption of carbohydrates, reduction of bolus insulin preceding exercise or following exercise, modifications to nighttime or daytime basal dose, etc.). This resimulation strategy will generate a training dataset that can be used to design an automated decision support system for exercise. Additional modeling in resistance exercise and interval training is forthcoming and will enrich ongoing approaches to prevent hypoglycemia.

6.1.3 Decision support for type 2 diabetes

Within the field of decision support there are many approaches that include type 2 diabetes [6, 93]. Many of these approaches utilize heuristic algorithms to improve glycemic control or to reduce cardiovascular risk factors [6]. New approaches in T2D decision support include glucose forecasting specific to individuals with type 2 diabetes [234], adjustment of insulin therapy [235], or detection of missed insulin basal doses [236]. The design of the automated decision support system introduced in chapter 3 may prove effective for individuals living with insulin-dependent type 2 diabetes, and is undergoing validation.

Additional strategies to manage type 2 diabetes include modeling pharmacotherapies used to manage type 2 diabetes (such as sulfonylureas, biguanides,

DPP4 inhibitors, etc.) on glucose physiology [237]. Recent development of virtual patient populations have been expanded to include glucose physiology of individuals with type 2 diabetes. These include modeling of endogenous insulin production [238], glucose absorption, and utilization and hepatic flux [239]. These new approaches may be incorporated into virtual patient populations of type 2 diabetes in order to perform *in silico* clinical evaluation, and, as demonstrated in chapter 3 may also be used to design new decision support systems for diabetes management.

6.1.4 Data considerations for predictive model design

Decision support algorithms and predictive algorithms can encounter design limitations due to inadequate data availability. These algorithms can also undergo real-world-use limitations if the data inputs to the algorithm differ from the data inputs present in the training set. The data used to train and operate algorithms is discussed in this section; this includes considerations for data acquisition, how to handle missing data, and the incorporation of new data types and sensor technologies.

6.1.4.1 Acquisition of real-world data to train predictive exercise algorithms

Chapter 4 of this dissertation introduced an artificial intelligence algorithm designed to estimate changes in glucose in the 4 hours following exercise. While it was shown that algorithms can be personalized in order to better predict an individual's exercise outcomes and can serve as a benchmark for algorithm performance, these specific algorithms are not optimized for use in real-world scenarios. Artificial intelligence algorithms are limited by the training dataset; the algorithms described in chapter 4 are designed using in-clinic data, which is considerably different than data metrics measured at-home. Future steps will require the creation of algorithms that are designed using data

collected across multiple exercise types (aerobic, resistance, high-intensity interval, etc.) in a real-world settings (at-home, parks, gyms, etc.). The newest T1-Dexi study funded by the Leona M. and Harry B. Helmsley Charitable trust will soon be publicly available, and provides CGM, insulin data, and exercise activity collected from adults with type 1 diabetes who performed broad types of exercise. Algorithms designed with this dataset are under development, and will serve as adaptive predictive models for use in real-world exercise sessions.

6.1.4.2 Algorithm handing of missing data inputs

Another consideration in the design of glucose predictive models during exercise is the handling of missing data. Data flow is not a continuous process; devices need to be calibrated, and devices experience hardware issues, networking errors, device failures, and misuse. However algorithms are typically designed and evaluated assuming data continuity. Missing data is often handled by interpolation for short stretches of missing data, or by halting the algorithm predictions until sufficient data is present. While there is great interest in algorithms that can provide on-demand and real-time predictions of glucose trends, there is no guarantee of data continuity or even the presence of specific devices. In real-world scenarios, the CGM transmitter may have expired, or an individual may have left their fitness watch at home. There is need for the development of a flexible algorithm that can account for missing devices or data gaps. Predictions will be available regardless of device and data availability through the use of ensemble predictors and classifiers that are trained on limited data availability.

6.1.4.3 Inclusion of personalized health data metrics

Algorithm personalization is an approach whereby predictive algorithms can be designed to better predict an individual user’s outcomes. Personalization is often

accomplished by training an algorithm using an individual person or group-of-similar-persons data, or by adapting a model's coefficients to an individual person over time. Adaptation of model parameters was shown in chapter 4 to better predict glycemic trends during exercise, however another personalization method to consider is the inclusion of other data metrics that are unique to individual people, such as demographic and genetic information. While the artificial intelligence MARS model proposed in chapter 4 did not identify demographic information to be predictive of glucose trends among 20 participants, other groups have noted sex-related differences in glucose uptake, disposal, and circulating hormonal profiles during exercise [240]. Age has also been implicated in changes in insulin sensitivity, and exercise-mediated changes in cellular glucose uptake [241]. Baseline genetic factors can also impact glucose disposal, and have not been incorporated into predictive algorithms. Polymorphisms and mutations of the regulatory protein TBC1D4 and have been found to impact insulin resistance and postprandial hyperglycemia in humans [242, 243], and postprandial hyperinsulinemia [244]. The proteins TBC1D1 and TBC1D4 have also been shown to undergo exercise-induced phosphorylation and regulation [245, 246], and specific mutations of these proteins impact exercise-related glucose uptake in rodent models [247]. Many other molecular regulators of exercise-stimulated glucose uptake have also been identified in rodent and human tissue [248], however further research is needed to determine how genetic polymorphisms can impact exercise-related glucose disposal in humans [243]. To date, there are no algorithms that utilize genetic testing or factors to predict glucose trends in real-time during daily activities or during exercise.

Glycemic control is also impacted by behavioral factors. Regular exercise is known to remodel muscle fiber content [249], and improve insulin sensitivity [249, 250], which varies between exercise modalities and intensity. A single bout of exercise can also prime skeletal muscle to uptake more glucose during subsequent exercise sessions

[251], and the underlying processes of this physical adaptation may be mathematically modeled in future approaches. In chapter 4, we demonstrated that an individual's aerobic fitness can significantly impact glucose trends. New techniques are needed to approximate changes in an individual's physical fitness for use in predictive algorithms or physiologic compartment models [251]. Taken together, these many personalized factors of demographics, genetics, exercise behaviors and physical fitness may be used in future algorithms to improve predicted glucose trends during exercise.

6.1.4.4 New and proposed sensor technologies for diabetes management

Automated decision support therapy systems, artificial intelligence algorithms, and automated insulin delivery systems are all designed with assumed data inputs from specific technologies (continuous glucose monitor, insulin pump or insulin pen, activity data, etc). Many new sensor technologies are being developed in the field of diabetes technologies. In this section, we describe our developments in glucose sensing catheter devices and also propose new sensor technologies that can improve advanced algorithms for diabetes management.

A new insulin catheter that measures glucose at the site of insulin delivery

Data presented in this section was published Oct 2020: Jacobs PG, Tyler NS, Vanderwerf SM, Mosquera-Lopez C, Seidl T, Cargill R, Branigan D, Ramsey K, Morris K, Benware S, Ward WK, Castle JR. Measuring glucose at the site of insulin delivery with a redox-mediated sensor. Biosens Bioelectron. 2020 Oct 1;165:112221. doi: 10.1016/j.bios.2020.112221. Epub 2020 Apr 29. PMID: 32729464.

Future technologies will enable the integration of glucose sensing with insulin delivery such that people with T1D will only need to insert a single device in their body to simultaneously sense glucose and deliver hormones. Glucose sensing and insulin delivery technology is evolving to become more integrated. In order to reduce the device burden on individuals living with T1D, a collaborator of our lab, Pacific Diabetes Technologies, developed an insulin cannula that is micropatterned with glucose sensing electrodes [252]. Previous to this work, it was not clear if insulin delivered close to the site of a glucose measurement impacted the accuracy of the glucose sensor [253]. Signal processing was performed to characterize sensor behavior during delivery of insulin solutions and saline solutions via the sensing cannula. These findings were used to modify sensor sampling techniques in order to improve the accuracy of glucose measurements.

In Figure 6.2, the average change in electrode measurement is studied following small basal doses, and large boluses of insulin and saline. While negligible changes in glucose sensor readings are observed following small basal deliveries, the delivery of large boluses cause sensor artifacts (6.2 A and B). These sensor artifacts can result in erroneous glucose measurements if not accounted for during sensor calibration. The sensor readings following basal and bolus delivery are observed to return to baseline and follow clinical YSI trends.

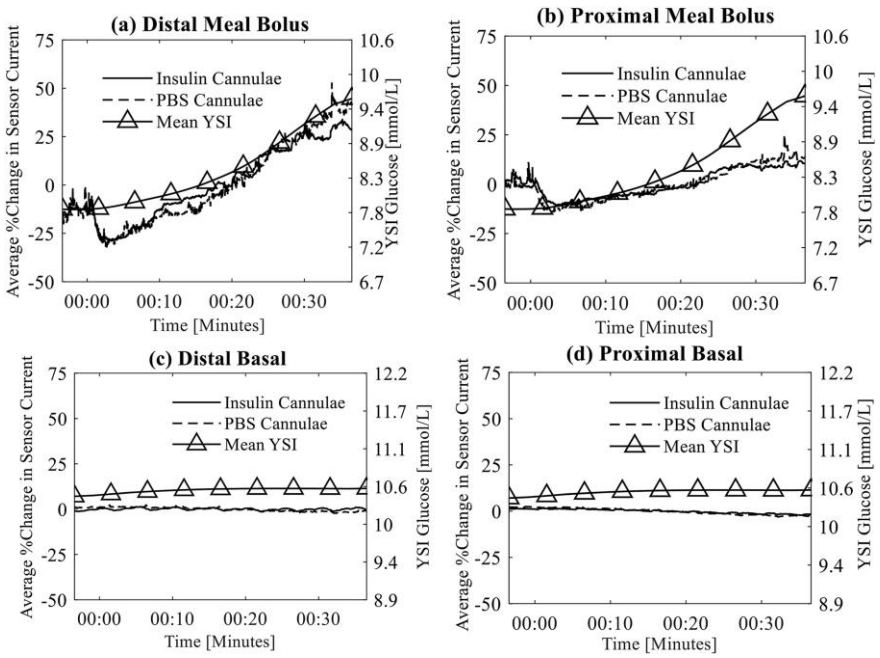


Figure 6.2: Sensor artifact for glucose-sensing catheter.

Top row shows the mean artifact of sensors from cannulae that are delivering insulin and PBS meal bolus amounts as measured by the distal (a) and proximal (b) sensing cannulae. The fluid delivery occurs at time 00:00. Bottom row shows mean artifact of sensors from cannulae that are delivering insulin and PBS basal amounts as measured by distal (c) and proximal (d) sensing cannulae. Blood glucose as measured by the YSI is also shown in each of these plots using the right y-axis. Notice that the sensors on the cannuae delivering insulin are nearly identical to those delivery PBS and therefore a difference between those lines is not apparent. *Figure and caption originally published in Jacobs PG, Tyler NS, Vanderwerf SM, Mosquera-Lopez C, Seidl T, Cargill R, Branigan D, Ramsey K, Morris K, Benware S, Ward WK, Castle JR. Measuring glucose at the site of insulin delivery with a redox-mediated sensor. Biosens Bioelectron. 2020 Oct 1;165:112221. PMID: 32729464.*

We proposed a new smart sampling algorithm that can be used to account for sensor artifacts following basal and bolus insulin delivery. This sampling method pauses sensor calibration following basal and bolus deliveries until the sensor artifact returns to baseline. The time required for the sensors to return to baseline following basal and bolus deliveries are shown in the survival curve in Figure 6.3. This analysis was used to impose a ~10 minute gap in sensor calibration following large boluses, and a ~3 minute gap in sensor calibration following small boluses. The smart sampling scheme was

shown to improve sensor accuracy as compared to clinical YSI data. This work contributes simple approach to improvement of sensor accuracy and will inform the design of future dual-sensing catheter technologies, and decision support algorithms. Smart sampling also provides an opportunity to incorporate predictive algorithms into glucose forecasting. The delivery of large boluses and the subsequent smart-sampling procedures will pause sensor calibration, and enable the use of predictive algorithms to forecast glucose values during periods of dilution artifacts. These approaches are forthcoming and human studies are planned.

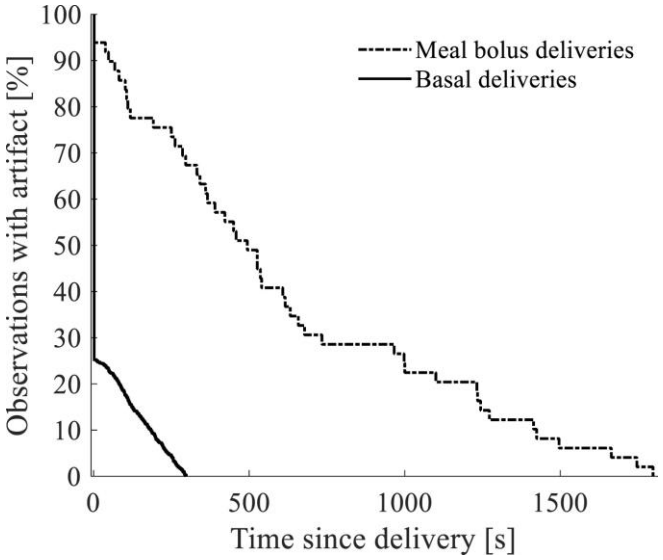


Figure 6.3: Survival curve of sensor artifacts.

Survival curve showing the percentage of observations after fluid delivery (insulin or PBS solution) that are free of the artifact vs. the time since the fluid was delivered. Immediately after fluid is delivered, most sensor observations exhibit some form of artifact, but the artifact is not present as time passes. Sensor recovery time is shown to be dependent on bolus volume whereby 70–80% of the smaller basal deliveries are free of artifact after 100 s while 70–80% of the sensor observations are free of artifact after 10 min for larger meal boluses. *Figure and caption originally published in Jacobs PG, Tyler NS, Vanderwerf SM, Mosquera-Lopez C, Seidl T, Cargill R, Branigan D, Ramsey K, Morris K, Benware S, Ward WK, Castle JR. Measuring glucose at the site of insulin delivery with a redox-mediated sensor. Biosens Bioelectron. 2020 Oct 1;165:112221. PMID: 32729464.*

Non-invasive glucose sensing technology

There are many new approaches to measure and estimate plasma and subcutaneous glucose content. While this dissertation develops models for use with commercial and experimental CGM systems that use sensing catheters, there are drawbacks to these approaches. Site-specific reactions can occur, and include infection and also fibrotic encapsulation which can impede sensor accuracy and calibration [254]. Cost is another consideration; in the US 30-40% of adults with type 1 diabetes use glucose sensing technology[255], but world-wide rates of use are much smaller, due to cost and availability [256, 257]. Limited access to technology and therapy prevents use of automated insulin delivery systems and the associated health benefits. However there are new approaches in non-invasive, low-cost glucose sensors that are currently under development. Infrared glucose sensors are now being developed by multiple tech startups and may be available in the near futures. Low-cost wearable microwave sensors [258], wearable contact-lens based polymers [259] are also under development as accessible glucose sensing technologies. And, paper-based sensors are under development for glucose and lactate sensing [260, 261]. Incorporation of these approaches into continuous glucose monitor systems is ongoing.

Additional data inputs

Existing physiology models and predictive models of glucose dynamics are designed to use readily available technologies; self-measured blood glucose, CGM-measured glucose, insulin doses captured by insulin pumps and smart pens, and also activity data captured in the form of accelerometer measurements or estimated metabolic expenditure. However many additional factors can impact glucose control; stress, hormonal profiles, sleep schedule [194, 262], illness, and even preceding

hypoglycemia are all known to cause variations in glucose dynamics. As such, there are many opportunities to identify new data inputs for use in predictive modeling.

In this dissertation, it was demonstrated that baseline aerobic fitness can significantly impact glucose trends during exercise. While physical fitness is typically assessed by specialists during in-clinic exams, new-generation fitness watches may enable estimation of percent VO_2 and baseline physical fitness, which can be inputted into predictive models. Estimation of sleep cycle is another consideration for predicting glucose trends during daily activities and during exercise. Numerous algorithms have been developed to estimate sleep cycle and wakefulness from existing activity wearables [263, 264], and may impact glucose estimation in real-time. Stress metrics have also been incorporated into many commercial fitness watches, and this information has been shown to improve the prediction of glucose trends in real time [265]. These new approaches hold promise in feasibly providing many additional data inputs obtained from one wearable device.

Another consideration is the development of new sensor technologies to estimate additional circulating biomarkers. Existing physiology models and predictive models are designed to incorporate insulin data obtained from insulin-pump or insulin pen logs. And while insulin doses have been described in physiologic models to diffuse through subcutaneous tissue compartments before reaching plasma, this physiologic process can be considerably different between individuals and contexts (eg, physical activity vs rest, varying injection sites, lipohypertrophy). As a result, physiologic models and predictive algorithms may experience considerable error due to mis-estimation of plasma insulin levels. One consideration is to develop new sensors that can estimate plasma insulin, in order to better estimate insulin-on-board at rest and during exercise. And this information can be used as inputs to predictive and physiology models to better forecast glucose trends in real-time. Other exercise-related biomarkers are another consideration

for use in predictive models. Existing algorithms can estimate exercise modalities from wrist-worn accelerometer data, however these approaches often mis-estimate precise metabolic expenditure and do not account for exercise training [266]. Circulating lactic acid levels have been reported to coincide with increased cellular metabolism during exercise [267], and can be used in conjunction with VO_2 data to estimate exercise intensity and muscular fatigue [268]. Future approaches in predictive modeling and physiologic models may benefit from plasma lactic acid measurements and provide new ways to model the impact of muscular fatigue and exercise endurance on glucose trends.

6.2 Preliminary findings for other decision support techniques

There were many approaches in data assimilation and decision support that were evaluated during the course of this dissertation work. In this section, I highlight some of these procedures and describe the caveats behind their usage.

6.2.1 Collaborative filtering engine for decision support

Many existing decision support tools are modeled after recommender systems and e-commerce systems [269-272]. Content-based collaborative filtering is a specific algorithm that is widely used in e-commerce commercial systems [270]. E-commerce provides product suggestions based on purchase and viewing history, and has recently been implemented in healthcare decision support [272]. The E-commerce framework is illustrated in Figure 6.4A. An E-commerce approach compares user preferences or interests, u_1-u_N, to the attributes of various movies or commercial products, m_1-m_M, and

finds the commercial product that is most similar to the user preferences. This comparison is accomplished through calculation of the dot product or cosine distance.

An E-commerce approach was initially proposed for the design of the automated decision support algorithm for type 1 diabetes presented in chapter 3. In Figure 6.4 B, the E-commerce system was designed to process user data into features, f_1-f_N, representing glucose during meals, specific behaviors and exercise; this user glycemic feature vector, F, functioned similarly to the user preference vector, U. To determine the appropriate recommendation, the glycemic features were related to specific recommendations, R_1-R_M, using a set of weights in a matrix, W. The cosine distance calculation was used to determine the recommendation with the highest importance given the users glycemic features (Figure 6.4B).

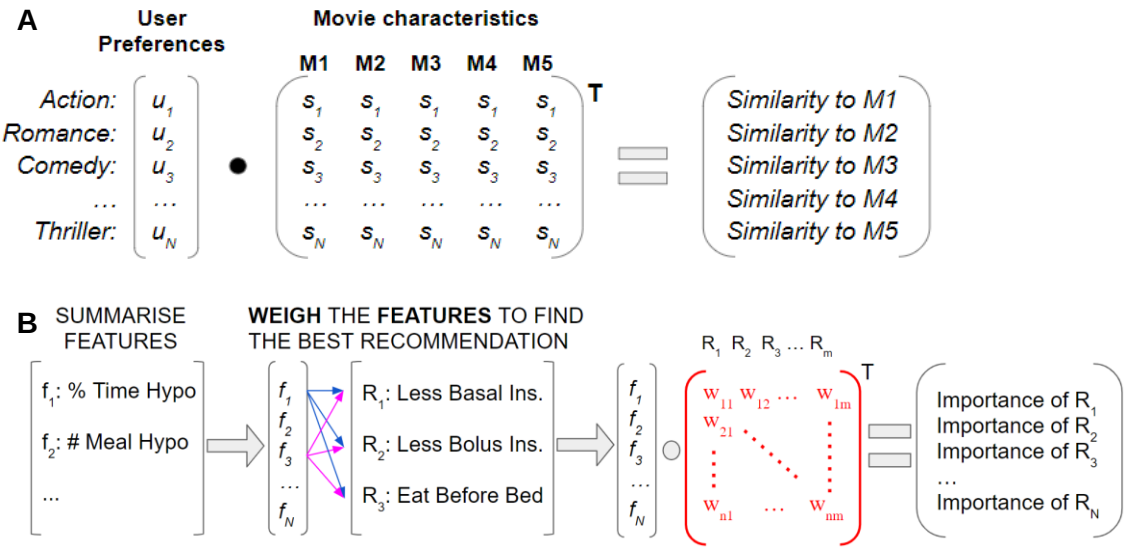


Figure 6.4: Collaborative filter method used to generate recommendations.

A. Recommendations for movies given preferred user genres, u_1-u_N, and the genre present in each movie, s_1-s_N. The dot product of user preference to movie genre will produce an importance or similarity measure of each movie to the user. B. A collaborative filter based decision support system. The user features, f_1-f_N, are related to recommended changes in insulin, R_1-R_m, by use of the weight matrix W. The entries of matrix W must be optimized.

The parameters of the e-commerce decision support system that require optimization are the parameters in matrix W that correlate glycemic problems (e.g. morning hypoglycemia) to recommended treatments (e.g. adjust morning ins:carb). Two different optimization strategies were evaluated to select specific weights that improve the identification and treatment of issues in glycemic control, (1) subgradient optimization and (2) simulated annealing.

Subgradient descent was not utilized in the final collaborative filter-based design because the parameter space was found to be non-smooth and the procedure was often trapped in local minima. The simulated annealing procedure was not influenced by the shape or smoothness of the parameter space, and was more successful in tuning the matrix W. However the results were still sub-optimal; *in silico* trials evaluating the matrix W did not result in major improvements in percent-time-in-range or reduction in hypoglycemia (Figure 6.5).

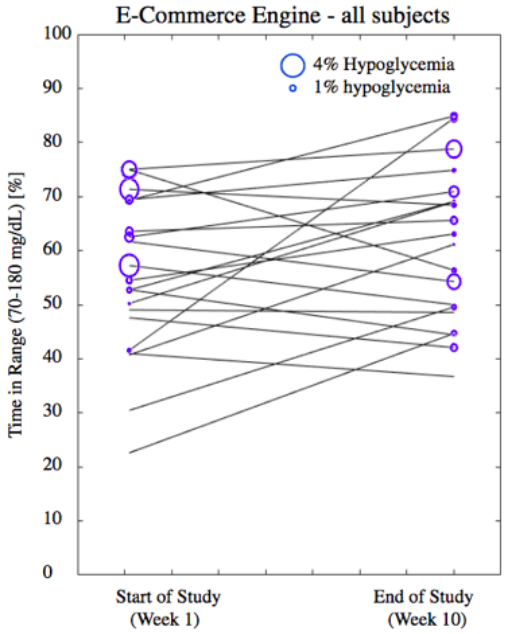


Figure 6.5: Measured glycemic outcomes after 10 weeks of weekly-use of the e-commerce decision support engine.

The y-coordinate of the data points represent the percent time in glucose target range, and each line represents the outcomes for an individual virtual subject. The radius of the blue circle represents percent time in hypoglycemia.

It was discovered upon further investigation that the cosine distance calculation, which is central to the e-commerce based classification, can be problematic when calculating importance of each recommendation. An example of cosine distance is illustrated in Figure 6.6.

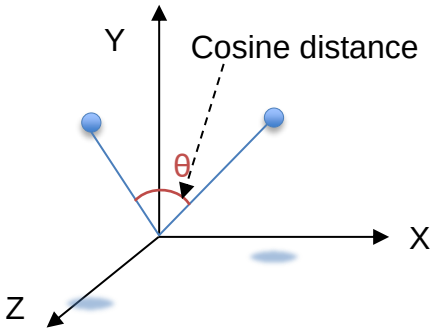


Figure 6.6: Cosine Distance
Two observations are represented along the features X, Y, and Z, and shown with shaded projections onto the X-Z axis. The cosine distance is the angular difference between the two observations in the feature space.

For any classification problem, the outcome classes must be separable along a multidimensional feature-space; a simplified 2D example is shown in Figure 6.7. Specifically, the cosine distance classification requires classes to be separable within angular sections, however, this was not possible with the glycemic features defined in the feature vector. In Figure 6.7, groups of observations belonging to a specific class label can be identified by identical numerical labels and color. Many classes overlap in multidimensional space and are confused by the E-commerce recommender engine, reducing system efficacy during *in silico* evaluation.

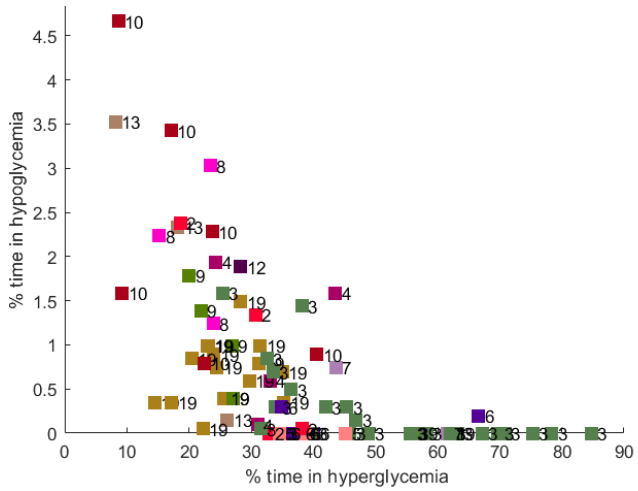


Figure 6.7: 2D representation of training observation feature space.
The features of percent time in hypoglycemia and percent time in hyperglycemia are plotted for a set of sample observations, across all classes. Each square represents a unique observation, and the number and color of the square indicates the class label. It can be seen that groups of observations belonging to the same class cannot be easily distinguished or separated from other class groups.

6.2.2 Net effect replay

One approach that was explored to minimize hypoglycemia during exercise was retrospective simulation of empirical glucose data. In this approach, a linear physiology model is used to perform realistic simulations of an individual's glucose across different insulin therapy parameters [56]. We utilized a dataset collected previously from adults with type 1 diabetes who participated in a 4-arm crossover study, and underwent 2 identical exercise study visits per study arm (8 visits total) [28]. First, a linear physiology model was optimized to an individual's study data using the idgrey() function in Matlab (Matworks, Inc., Natick NJ). Next, the net-effect approach [56] was used better forecast glucose measured during the first exercise study visit for a unique participant (Figure 6.8 A,B). The glucose of the first exercise study visit was then resimulated using varying levels of basal insulin in order to determine the optimal reduction in insulin to avoid hypoglycemia in the 4 hours following exercise (Figure 6.8C). Next, the net-effect method was applied to exercise study visit #2 from the same participant. The glucose

data in study visit #2 was resimulated using the optimal insulin reduction determined from resimulation of study visit #1. Reductions in hypoglycemia were measured and are shown in Table 6.1. This two-step resimulation procedure was performed across all participant data.

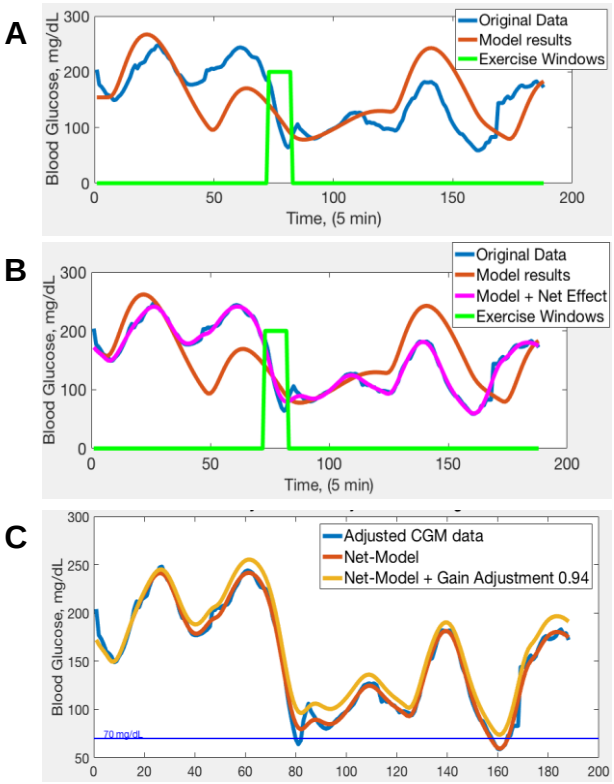


Figure 6.8: Glucose net-effect replay method.

A. Empirical glucose data and forecasted glucose. **B**. Model forecasts after incorporating net-effect optimization. **C**. Resimulation of glucose trends to determine the optimal reduction in basal insulin required to avoid hypoglycemia following exercise.

In Table 6.1, it can be seen that optimizing basal reduction on study visit #1, and applying this basal reduction on a separate exercise study visit could potentially reduce hypoglycemia by a relative 63.7%. However there are issues with this method that precluded real-world use. The net-effect procedure solves for an optimal forcing input that reduces simulation error, but this procedure does not optimize the model parameters. As a result, the parameters of the base physiology model that estimate the

impact of insulin therapy or carbohydrate consumption on glucose dynamics, may be suboptimal. The resimulation procedure shown in Figure 6.8 B-C will use the base physiology model structure to study the impact of varying basal or carbohydrate dosages, and these changes may be erroneous due to the model parameters. If the model parameters are optimized, this procedure can be used to make inferences about modifications in insulin therapy, but likely cannot be used to determine precise changes to insulin therapy.

Table 6.1: Predicted reduction in hypoglycemia following net-effect simulation of basal insulin reduction.

Predicted glucose outcomes on exercise study-visit #2 following net-effect replay and resimulation of glucose. Data is averaged across all participants and study arms.

GLUCOSE RANGE [MG/DL]	EMPIRICAL DATA	MODEL FORECAST	OPTIMAL BASAL REDUCTION	% CHANGE
70-180	73.2 %	74.1 %	73.2 %	- 1.2 %
< 70	2.2 %	2.0 %	0.7 %	- 63.7 %
> 180	23.6 %	23.9 %	26.1 %	9.1 %

It is hoped that the contributions presented in this dissertation, and the alternative approaches described in this section, can benefit the ongoing work of researchers and clinicians to design effective and safe decision support systems for diabetes.

Bibliography

1. Diabetes Atlas, *International diabetes federation.* IDF Diabetes Atlas, 9th edn. Brussels, Belgium: International Diabetes Federation, 2020.
2. American Diabetes Association, *9. Pharmacologic Approaches to Glycemic Treatment: Standards of Medical Care in Diabetes—2021.* Diabetes Care, 2021. **44**(Supplement 1): p. S111-S124.
3. Nathan, D.M. and D.E.R. Group, *The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview.* Diabetes care, 2014. **37**(1): p. 9-16.
4. Group, T.D.C.a.C.T.R., *Hypoglycemia in the Diabetes Control and Complications Trial.* Diabetes, 1997. **46**(2): p. 271.
5. Barnard, K., et al., *Use of an automated bolus calculator reduces fear of hypoglycemia and improves confidence in dosage accuracy in patients with type 1 diabetes mellitus treated with multiple daily insulin injections.* J Diabetes Sci Technol, 2012. **6**(1): p. 144-9.
6. Veazie, S., et al., *Rapid Evidence Review of Mobile Applications for Self-management of Diabetes.* Journal of General Internal Medicine, 2018. **33**(7): p. 1167-1176.
7. Health, M., *Mellitus Health Case Studies. Available online: https://www.mellitushealth.com/case-studies/.*
8. Rx, A. *Avhana Health. Available online: https://www.avhana.com/*.
9. Technologies, M.M. *EndoTool Glucose Management. Available online: https://monarchmedtech.com/endotool-glucose-management/*.
10. Glytec. *eGlycemic Management System Glucommaander. Available online: https://glytecsystems.com/solutions-overview/solutions-eglycemic-management-system/*. [cited 2021 March 2].
11. Wintergerst, M.W.M., et al., *Diabetic Retinopathy Screening Using Smartphone-Based Fundus Imaging in India.* Ophthalmology, 2020. **127**(11): p. 1529-1538.
12. Duffy, S., et al., *Using Community Health Workers and a Smartphone Application to Improve Diabetes Control in Rural Guatemala.* Global Health: Science and Practice, 2020. **8**(4): p. 699.
13. Nathan, D.M., *Long-Term Complications of Diabetes Mellitus.* New England Journal of Medicine, 1993. **328**(23): p. 1676-1685.
14. Carey, I.M., et al., *Risk of Infection in Type 1 and Type 2 Diabetes Compared With the General Population: A Matched Cohort Study.* Diabetes Care, 2018. **41**(3): p. 513-521.

15. Giovannucci, E., et al., *Diabetes and Cancer: A Consensus Report.* CA: A Cancer Journal for Clinicians, 2010. **60**(4): p. 207-221.

16. Wilcox, G., *Insulin and insulin resistance.* The Clinical biochemist. Reviews, 2005. **26**(2): p. 19-39.

17. Marín-Peñalver, J.J., et al., *Update on the treatment of type 2 diabetes mellitus.* World journal of diabetes, 2016. **7**(17): p. 354-395.

18. American Diabetes Association, *14. Management of Diabetes in Pregnancy: Standards of Medical Care in Diabetes—2021.* Diabetes Care, 2021. **44**(Supplement 1): p. S200-S210.

19. Yosten, G.L.C., *Alpha cell dysfunction in type 1 diabetes.* Peptides, 2018. **100**: p. 54-60.

20. Cooke, D.W. and L. Plotnick, *Type 1 diabetes mellitus in pediatrics.* pediatr Rev, 2008. **29**(11): p. 374-84.

21. Filippi, C.M. and M.G. von Herrath, *Viral trigger for type 1 diabetes: pros and cons.* Diabetes, 2008. **57**(11): p. 2863-2871.

22. Eizirik, D.L., L. Pasquali, and M. Cnop, *Pancreatic β-cells in type 1 and type 2 diabetes mellitus: different pathways to failure.* Nature Reviews Endocrinology, 2020. **16**(7): p. 349-362.

23. Puchalska, P. and P.A. Crawford, *Multi-dimensional Roles of Ketone Bodies in Fuel Metabolism, Signaling, and Therapeutics.* Cell metabolism, 2017. **25**(2): p. 262-284.

24. Control, C.f.D. and Prevention, *National diabetes statistics report, 2020.* Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services, 2020: p. 12-15.

25. Wilson, L. and J.R. Castle, *Recent Advances in Insulin Therapy.* Diabetes Technology & Therapeutics, 2020.

26. Cavanaugh, K., et al., *Addressing literacy and numeracy to improve diabetes care: two randomized controlled trials.* Diabetes Care, 2009. **32**(12): p. 2149-55.

27. Tatò, F., et al., *Circadian variation of basal and postprandial insulin sensitivity in healthy individuals and patients with type-1 diabetes.* Diabetes research (Edinburgh, Scotland), 1991. **17**(1): p. 13-24.

28. Castle, J.R., et al., *Randomized Outpatient Trial of Single- and Dual-Hormone Closed-Loop Systems That Adapt to Exercise Using Wearable Sensors.* Diabetes Care, 2018. **41**(7): p. 1471-1477.

29. Wilson, L.M., et al., *Dual-Hormone Closed-Loop System Using a Liquid Stable Glucagon Formulation Versus Insulin-Only Closed-Loop System Compared With a Predictive Low Glucose Suspend System: An Open-Label, Outpatient, Single-Center, Crossover, Randomized Controlled Trial.* Diabetes Care, 2020. **43**(11): p. 2721.

30. Ertl, A.C. and S.N. Davis, *Evidence for a vicious cycle of exercise and hypoglycemia in type 1 diabetes mellitus.* Diabetes/Metabolism Research and Reviews, 2004. **20**(2): p. 124-130.

31. Bohn, B., et al., *Impact of Physical Activity on Glycemic Control and Prevalence of Cardiovascular Risk Factors in Adults With Type 1 Diabetes: A Cross-sectional Multicenter Study of 18,028 Patients.* Diabetes Care, 2015. **38**(8): p. 1536-43.

32. Riddell, M., et al., *More time in glucose range during exercise days than sedentary days in adults living with type 1 diabetes.* Diabetes Technology & Therapeutics, 2020.

33. Sylow, L., et al., *Exercise-stimulated glucose uptake — regulation and implications for glycaemic control.* Nature Reviews Endocrinology, 2017. **13**(3): p. 133-148.

34. McMahon, S.K., et al., *Glucose Requirements to Maintain Euglycemia after Moderate-Intensity Afternoon Exercise in Adolescents with Type 1 Diabetes Are Increased in a Biphasic Manner.* The Journal of Clinical Endocrinology & Metabolism, 2007. **92**(3): p. 963-968.

35. Brazeau, A.-S., et al., *Barriers to physical activity among patients with type 1 diabetes.* Diabetes care, 2008. **31**(11): p. 2108-2109.

36. Wilson, L.M., et al., *Patient Input for Design of a Decision Support Smartphone Application for Type 1 Diabetes.* Journal of Diabetes Science and Technology, 2019. **14**(6): p. 1081-1087.

37. Beck, R.W., et al., *Effect of continuous glucose monitoring on glycemic control in adults with type 1 diabetes using insulin injections: the DIAMOND randomized clinical trial.* Jama, 2017. **317**(4): p. 371-378.

38. Bergenstal, R.M., et al., *Effectiveness of Sensor-Augmented Insulin-Pump Therapy in Type 1 Diabetes.* New England Journal of Medicine, 2010. **363**(4): p. 311-320.

39. Bequette, B.W., *A Critical Assessment of Algorithms and Challenges in the Development of a Closed-Loop Artificial Pancreas.* Diabetes Technology & Therapeutics, 2005. **7**(1): p. 28-47.

40. Haidar, A., *The Artificial Pancreas: How Closed-Loop Control Is Revolutionizing Diabetes.* IEEE Control Systems Magazine, 2016. **36**(5): p. 28-47.

41. Brown, S.A., et al., *Six-Month Randomized, Multicenter Trial of Closed-Loop Control in Type 1 Diabetes.* New England Journal of Medicine, 2019. **381**(18): p. 1707-1717.

42. Tandem Diabetes Care, I. *t:slim X2 Insulin Pump with Control-IQ Technology User Guide*. [.pdf] 2020 [cited 2020 14 October]; Available from: https://www.tandemdiabetes.com/docs/default-source/product-documents/t-slim-x2-insulin-pump/k%c3%a4ytt%c3%b6opas-(t-slim-x2--pumppu-control-iq-

teknologialla)cf14759775426a79a519ff1100a9fd393f6fb39775426a79a519ff1200a9fd39.pdf?sfvrsn=18a507d7_66.

43. Medtronic Minimed, I. *MiniMed 670G System User Guide*. [.pdf] 2017 [cited 2020 14 October]; Available from: https://www.medtronicdiabetes.com/sites/default/files/library/download-library/user-guides/MiniMed-670G-System-User-Guide.pdf.
44. Tyler, N.S. and P.G. Jacobs, *Artificial intelligence in decision support systems for type 1 diabetes.* Sensors, 2020. **20**(11): p. 3214.
45. Breton, M.D., et al., *Continuous Glucose Monitoring and Insulin Informed Advisory System with Automated Titration and Dosing of Insulin Reduces Glucose Variability in Type 1 Diabetes Mellitus.* Diabetes Technol Ther, 2018. **20**(8): p. 531-540.
46. Herrero, P., et al., *Advanced Insulin Bolus Advisor Based on Run-To-Run Control and Case-Based Reasoning.* IEEE J Biomed Health Inform, 2015. **19**(3): p. 1087-96.
47. Nimri, R., et al., *Adjusting insulin doses in patients with type 1 diabetes who use insulin pump and continuous glucose monitoring: Variations among countries and physicians.* Diabetes, Obesity and Metabolism, 2018. **20**(10): p. 2458-2466.
48. Montaser, E., et al., *Seasonal Local Models for Glucose Prediction in Type 1 Diabetes.* IEEE Journal of Biomedical and Health Informatics, 2019: p. 1-1.
49. Mosquera-Lopez, C., et al., *Predicting and Preventing Nocturnal Hypoglycemia in Type 1 Diabetes Using Big Data Analytics and Decision Theoretic Analysis.* Diabetes Technol Ther, 2020.
50. Reddy, R., et al., *Prediction of Hypoglycemia During Aerobic Exercise in Adults With Type 1 Diabetes.* J Diabetes Sci Technol, 2019. **13**(5): p. 919-927.
51. Hobbs, N., et al., *Improving Glucose Prediction Accuracy in Physically Active Adolescents With Type 1 Diabetes.* J Diabetes Sci Technol, 2019. **13**(4): p. 718-727.
52. Facchinetti, A., et al., *An online failure detection method of the glucose sensor-insulin pump system: improved overnight safety of type-1 diabetic subjects.* IEEE Trans Biomed Eng, 2013. **60**(2): p. 406-16.
53. Samadi, S., et al., *Automatic Detection and Estimation of Unannounced Meals for Multivariable Artificial Pancreas System.* Diabetes Technol Ther, 2018. **20**(3): p. 235-246.
54. Vasiloglou, M.F., et al., *A Comparative Study on Carbohydrate Estimation: GoCARB vs. Dietitians.* Nutrients, 2018. **10**(6): p. 741.
55. Lehmann, E.D., et al., *AIDA: an interactive diabetes advisor.* Comput Methods Programs Biomed, 1994. **41**(3-4): p. 183-203.

56. Patek, S.D., et al., *Empirical representation of blood glucose variability in a compartmental model*, in *Prediction Methods for Blood Glucose Concentration*. 2016, Springer. p. 133-157.

57. Resalat, N., et al., *A statistical virtual patient population for the glucoregulatory system in type 1 diabetes with integrated exercise model.* PLoS One, 2019. **14**(7): p. e0217301.

58. Miller, K.M., et al., *Current State of Type 1 Diabetes Treatment in the U.S.: Updated Data From the T1D Exchange Clinic Registry.* Diabetes Care, 2015. **38**(6): p. 971.

59. Reddy, M., et al., *Clinical Safety and Feasibility of the Advanced Bolus Calculator for Type 1 Diabetes Based on Case-Based Reasoning: A 6-Week Nonrandomized Single-Arm Pilot Study.* Diabetes Technol Ther, 2016. **18**(8): p. 487-93.

60. Liu, C., et al., *A Modular Safety System for an Insulin Dose Recommender: A Feasibility Study.* J Diabetes Sci Technol, 2020. **14**(1): p. 87-96.

61. Nimri, R., et al., *Insulin dose optimization using an automated artificial intelligence-based decision support system in youths with type 1 diabetes.* Nature Medicine, 2020. **26**(9): p. 1380-1384.

62. Tyler, N.S., et al., *An artificial intelligence decision support system for the management of type 1 diabetes.* Nature metabolism, 2020. **2**(7): p. 612-619.

63. Riddell, M.C., et al., *Exercise management in type 1 diabetes: a consensus statement.* The lancet Diabetes & endocrinology, 2017. **5**(5): p. 377-390.

64. McGaugh, S.M., et al., *The Development of an Exercise Advisor App for Type 1 Diabetes: Digitization Facilitates More Individualized Guidance.* Journal of Diabetes Science and Technology, 2020: p. 1932296820979811.

65. Romero-Ugalde, H.M., et al., *ARX model for interstitial glucose prediction during and after physical activities.* Control Engineering Practice, 2019. **90**: p. 321-330.

66. Szegedy, C., et al. *Going deeper with convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

67. Xie, J. and Q. Wang, *Benchmarking machine learning algorithms on blood glucose prediction for Type 1 Diabetes in comparison with classical time-series models.* IEEE Transactions on Biomedical Engineering, 2020: p. 1-1.

68. Tidepool. 5/28/2020]; Tidepool Big Data Donation Project]. Available from: https://www.tidepool.org/bigdata.

69. Trust, T.L.M.a.H.B.C. [cited 2020 5/28/2020]; Available from: https://helmsleytrust.org/.

70. Cover, T. and P. Hart, *Nearest neighbor pattern classification.* IEEE transactions on information theory, 1967. **13**(1): p. 21-27.

71. Friedman, J.H., *Multivariate Adaptive Regression Splines.* The Annals of Statistics, 1991. **19**(1): p. 1-67.

72. Hechenbichler, K. and K. Schliep, *Weighted k-Nearest-Neighbor Techniques and Ordinal Classification*. 2004.

73. Bicego, M. and M. Loog. *Weighted K-Nearest Neighbor revisited*. in *2016 23rd International Conference on Pattern Recognition (ICPR)*. 2016.

74. Tan, S., *Neighbor-weighted K-nearest neighbor for unbalanced text corpus.* Expert Systems with Applications, 2005. **28**(4): p. 667-671.

75. Man, C.D., et al., *The UVA/PADOVA Type 1 Diabetes Simulator: New Features.* Journal of Diabetes Science and Technology, 2014. **8**(1): p. 26-34.

76. Kruschke, J., *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* 2014.

77. Monnahan, C.C., J.T. Thorson, and T.A. Branch, *Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo.* Methods in Ecology and Evolution, 2017. **8**(3): p. 339-348.

78. Zisser, H., et al., *Clinical update on optimal prandial insulin dosing using a refined run-to-run control algorithm.* J Diabetes Sci Technol, 2009. **3**(3): p. 487-91.

79. Palerm, C.C., et al., *A Run-to-Run Control Strategy to Adjust Basal Insulin Infusion Rates in Type 1 Diabetes.* J Process Control, 2008. **18**(3-4): p. 258-265.

80. Herrero, P., et al., *Automatic Adaptation of Basal Insulin Using Sensor-Augmented Pump Therapy.* J Diabetes Sci Technol, 2018. **12**(2): p. 282-294.

81. Toffanin, C., et al., *Automatic adaptation of basal therapy for Type 1 diabetic patients: A Run-to-Run approach.* Biomedical Signal Processing and Control, 2017. **31**: p. 539-549.

82. Moser, O., et al., *Glucose management for exercise using continuous glucose monitoring (CGM) and intermittently scanned CGM (isCGM) systems in type 1 diabetes: position statement of the European Association for the Study of Diabetes (EASD) and of the International Society for Pediatric and Adolescent Diabetes (ISPAD) endorsed by JDRF and supported by the American Diabetes Association (ADA).* Diabetologia, 2020. **63**(12): p. 2501-2520.

83. Fabris, C., et al., *The Use of a Smart Bolus Calculator Informed by Real-time Insulin Sensitivity Assessments Reduces Postprandial Hypoglycemia Following an Aerobic Exercise Session in Individuals With Type 1 Diabetes.* Diabetes Care, 2020. **43**(4): p. 799-805.

84. Hernandez-Ordonez, M. and D.U. Campos-Delgado, *An extension to the compartmental model of type 1 diabetic patients to reproduce exercise periods with glycogen depletion and replenishment.* J Biomech, 2008. **41**(4): p. 744-52.

85. Man, C.D., M.D. Breton, and C. Cobelli, *Physical Activity into the Meal Glucose–Insulin Model of Type 1 Diabetes: In Silico Studies.* J Diabetes Sci Technol, 2009. **3**(1): p. 56-67.

86. Roy, A. and R.S. Parker, *Dynamic modeling of exercise effects on plasma glucose and insulin levels.* J Diabetes Sci Technol, 2007. **1**(3): p. 338-47.

87. Bergenstal, R.M., et al., *Safety of a Hybrid Closed-Loop Insulin Delivery System in Patients With Type 1 Diabetes.* JAMA, 2016. **316**(13): p. 1407-1408.

88. Sangave, N.A., T.D. Aungst, and D.K. Patel, *Smart Connected Insulin Pens, Caps, and Attachments: A Review of the Future of Diabetes Technology.* Diabetes Spectr, 2019. **32**(4): p. 378-384.

89. Cavanaugh, K., et al., *Association of numeracy and diabetes control.* Ann Intern Med, 2008. **148**(10): p. 737-46.

90. Ahola, A.J., et al., *Many patients with Type 1 diabetes estimate their prandial insulin need inappropriately.* J Diabetes, 2010. **2**(3): p. 194-202.

91. Foster, N.C., et al., *State of Type 1 Diabetes Management and Outcomes from the T1D Exchange in 2016–2018.* Diabetes Technology & Therapeutics, 2019. **21**(2): p. 66-72.

92. Wu, Y., et al., *Mobile App-Based Interventions to Support Diabetes Self-Management: A Systematic Review of Randomized Controlled Trials to Identify Functions Associated with Glycemic Efficacy.* JMIR Mhealth Uhealth, 2017. **5**(3): p. e35.

93. Charpentier, G., et al., *The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study).* Diabetes Care, 2011. **34**(3): p. 533-9.

94. Kirwan, M., et al., *Diabetes Self-Management Smartphone Application for Adults With Type 1 Diabetes: Randomized Controlled Trial.* J Med Internet Res, 2013. **15**(11): p. e235.

95. Drion, I., et al. *The Effects of a Mobile Phone Application on Quality of Life in Patients With Type 1 Diabetes Mellitus: A Randomized Controlled Trial*. Journal of diabetes science and technology, 2015. **9**, 1086-1091 DOI: 10.1177/1932296815585871.

96. Skrovseth, S.O., et al., *Data-Driven Personalized Feedback to Patients with Type 1 Diabetes: A Randomized Trial.* Diabetes Technol Ther, 2015. **17**(7): p. 482-9.

97. Perez-Gandia, C., et al., *Decision Support in Diabetes Care: The Challenge of Supporting Patients in Their Daily Living Using a Mobile Glucose Predictor.* J Diabetes Sci Technol, 2018. **12**(2): p. 243-250.

98. Kovatchev, B.P., et al., *Symmetrization of the Blood Glucose Measurement Scale and Its Applications.* Diabetes Care, 1997. **20**(11): p. 1655.

99. Garg, S.K., et al., *Glucose Outcomes with the In-Home Use of a Hybrid Closed-Loop Insulin Delivery System in Adolescents and Adults with Type 1 Diabetes.* Diabetes technology & therapeutics, 2017. **19**(3): p. 155-163.

100. Clarke, W.L., et al., *Evaluating Clinical Accuracy of Systems for Self-Monitoring of Blood Glucose.* Diabetes Care, 1987. **10**(5): p. 622.

101. Parkes, J.L., et al., *A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose.* Diabetes Care, 2000. **23**(8): p. 1143-8.

102. United States Food and Drug Administration. *United States Food and Drug Administration. Available online: https://www.fda.gov/* 2020; Available from: https://www.fda.gov/

103. Hovorka, R., et al., *Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes.* Physiol Meas, 2004. **25**(4): p. 905-20.

104. Goodwin, G.C., et al., *A systematic stochastic design strategy achieving an optimal tradeoff between peak BGL and probability of hypoglycaemic events for individuals having type 1 diabetes mellitus.* Biomedical Signal Processing and Control, 2020. **57**: p. 101813.

105. Cameron, F., et al., *Inpatient studies of a Kalman-filter-based predictive pump shutoff algorithm.* Journal of diabetes science and technology, 2012. **6**(5): p. 1142-1147.

106. Wilinska, M.E., et al., *Insulin kinetics in type-I diabetes: continuous and bolus delivery of rapid acting insulin.* IEEE Trans Biomed Eng, 2005. **52**(1): p. 3-12.

107. Wilinska, M.E., et al., *Simulation environment to evaluate closed-loop insulin delivery systems in type 1 diabetes.* J Diabetes Sci Technol, 2010. **4**(1): p. 132-44.

108. Resalat, N., et al. *OHSU Virtual Patient Population Simulator: Artificial Intelligence for Medical Systems (AIMS) Lab*. 2019 6/4/2020]; Available from: https://github.com/petejacobs/T1D_VPP.

109. Xie, J. and Q. Wang, *A Data-Driven Personalized Model of Glucose Dynamics Taking Account of the Effects of Physical Activity for Type 1 Diabetes: An In Silico Study.* J Biomech Eng, 2019. **141**(1).

110. He, M., et al., *CausalBG: Causal Recurrent Neural Network for the Blood Glucose Inference With IoT Platform.* IEEE Internet of Things Journal, 2020. **7**(1): p. 598-610.

111. Kadish, A.H., *Automation control of blood sugar a servomechanism for glucose monitoring and control.* Trans Am Soc Artif Intern Organs, 1963. **9**: p. 363-7.

112. Spencer, W.J., *A Review of Programmed Insulin Delivery Systems.* IEEE Transactions on Biomedical Engineering, 1981. **BME-28**(3): p. 237-251.

113. Peterson, C.M., L. Jovanovic, and L.H. Chanoch, *Randomized trial of computer-assisted insulin delivery in patients with type I diabetes beginning pump therapy.* Am J Med, 1986. **81**(1): p. 69-72.

114. Chanoch, L.H., L. Jovanovic, and C.M. Peterson, *The Evaluation of a Pocket Computer as an Aid to Insulin Dose Determination by Patients.* Diabetes Care, 1985. **8**(2): p. 172.

115. Skyler, J.S., et al., *Algorithms for adjustment of insulin dosage by patients who monitor blood glucose.* Diabetes Care, 1981. **4**(2): p. 311-8.

116. Bellomo, G., et al., *A simple computer program for insulin dose adjustment in diabetic patients.* Comput Methods Programs Biomed, 1988. **26**(3): p. 257-8.

117. Chiarelli, F., et al., *Controlled study in diabetic children comparing insulin-dosage adjustment by manual and computer algorithms.* Diabetes Care, 1990. **13**(10): p. 1080-4.

118. Berger, M.P. and D. Rodbard, *A pharmacodynamic approach to optimizing insulin therapy.* Comput Methods Programs Biomed, 1991. **34**(4): p. 241-53.

119. Hauser, T., et al., *Application of physicians' predictions of meal and exercise effects on blood glucose control to a computer simulation.* Diabet Med, 1993. **10**(8): p. 744-50.

120. Rutscher, A., et al., *KADIS--a computer-aided decision support system for improving the management of type-I diabetes.* Exp Clin Endocrinol, 1990. **95**(1): p. 137-47.

121. Salzsieder, E., et al., *Computer-aided systems in the management of type I diabetes: the application of a model-based strategy.* Comput Methods Programs Biomed, 1990. **32**(3-4): p. 215-24.

122. Rutscher, A., E. Salzsieder, and U. Fischer, *KADIS: model-aided education in type I diabetes. Karlsburg Diabetes Management System.* Comput Methods Programs Biomed, 1994. **41**(3-4): p. 205-15.

123. Augstein, P., et al., *Translation of personalized decision support into routine diabetes care.* J Diabetes Sci Technol, 2010. **4**(6): p. 1532-9.

124. Stadelmann, A., et al., *DIABETEX decision module 2--calculation of insulin dose proposals and situation recognition by means of classifiers.* Comput Methods Programs Biomed, 1990. **32**(3-4): p. 333-7.

125. Zahlmann, G., et al., *DIABETEX--a decision support system for therapy of type I diabetic patients.* Comput Methods Programs Biomed, 1990. **32**(3-4): p. 297-301.

126. Ambrosiadou, B.V., et al., *Decision support for insulin regime prescription based on a neural-network approach.* Med Inform (Lond), 1996. **21**(1): p. 23-34.

127. Gogou, G., et al., *A neural network approach in diabetes management by insulin administration.* J Med Syst, 2001. **25**(2): p. 119-31.

128. Andreassen, S., et al., *A probabilistic approach to glucose prediction and insulin dose adjustment: description of metabolic model and pilot evaluation study.* Comput Methods Programs Biomed, 1994. **41**(3-4): p. 153-65.

129. Cavan, D.A., et al., *Preliminary experience of the DIAS computer model in providing insulin dose advice to patients with insulin dependent diabetes.* Comput Methods Programs Biomed, 1998. **56**(2): p. 157-64.

130. Wong, X.W., et al., *In silico simulation of long-term type 1 diabetes glycemic control treatment outcomes.* J Diabetes Sci Technol, 2008. **2**(3): p. 436-49.

131. Rosales, N., et al., *Open-loop glucose control: Automatic IOB-based super-bolus feature for commercial insulin pumps.* Computer Methods and Programs in Biomedicine, 2018. **159**: p. 145-158.

132. Revert, A., et al., *Calculation of the Best Basal–Bolus Combination for Postprandial Glucose Control in Insulin Pump Therapy.* IEEE Transactions on Biomedical Engineering, 2011. **58**(2): p. 274-281.

133. Rossetti, P., et al., *Evaluation of a novel continuous glucose monitoring-based method for mealtime insulin dosing--the iBolus--in subjects with type 1 diabetes using continuous subcutaneous insulin infusion therapy: a randomized controlled trial.* Diabetes technology & therapeutics, 2012. **14**(11): p. 1043-1052.

134. Bellazzi, R., et al., *A telemedicine support for diabetes management: the T-IDDM project.* Comput Methods Programs Biomed, 2002. **69**(2): p. 147-61.

135. Schwartz, F.L., J.H. Shubrook, and C.R. Marling, *Use of case-based reasoning to enhance intensive management of patients on insulin pump therapy.* J Diabetes Sci Technol, 2008. **2**(4): p. 603-11.

136. Schwartz, F.L., et al., *Evaluating the automated blood glucose pattern detection and case-retrieval modules of the 4 Diabetes Support System.* Journal of diabetes science and technology, 2010. **4**(6): p. 1563-1569.

137. Torrent-Fontbona, F. and B. López, *Personalized Adaptive CBR Bolus Recommender System for Type 1 Diabetes.* IEEE Journal of Biomedical and Health Informatics, 2019. **23**(1): p. 387-394.

138. Biagi, L., et al., *Individual categorisation of glucose profiles using compositional data analysis.* Stat Methods Med Res, 2019. **28**(12): p. 3550-3567.

139. Pankowska, E., et al., *A Randomized Controlled Study of an Insulin Dosing Application That Uses Recognition and Meal Bolus Estimations.* J Diabetes Sci Technol, 2017. **11**(1): p. 43-49.

140. Sun, Q., M.V. Jankovic, and S.G. Mougiakakou, *Reinforcement Learning-Based Adaptive Insulin Advisor for Individuals with Type 1 Diabetes Patients under Multiple Daily Injections Therapy(.).* Conf Proc IEEE Eng Med Biol Soc, 2019. **2019**: p. 3609-3612.

141. Srinivasan, A., et al., *Novel insulin delivery profiles for mixed meals for sensor-augmented pump and closed-loop artificial pancreas therapy for type 1 diabetes mellitus.* Journal of diabetes science and technology, 2014. **8**(5): p. 957-968.

142. Anthimopoulos, M., et al., *Computer vision-based carbohydrate estimation for type 1 patients with diabetes using smartphones.* Journal of diabetes science and technology, 2015. **9**(3): p. 507-515.

143. Mahmoudi, Z., et al., *Sensor-based detection and estimation of meal carbohydrates for people with diabetes.* Biomedical Signal Processing and Control, 2019. **48**: p. 12-25.

144. Ensling, M., W. Steinmann, and A. Whaley-Connell, *Hypoglycemia: A Possible Link between Insulin Resistance, Metabolic Dyslipidemia, and Heart and Kidney Disease (the Cardiorenal Syndrome).* Cardiorenal Med, 2011. **1**(1): p. 67-74.

145. Buckingham, B.A., et al., *Outpatient safety assessment of an in-home predictive low-glucose suspend system with type 1 diabetes subjects at elevated risk of nocturnal hypoglycemia.* Diabetes Technol Ther, 2013. **15**(8): p. 622-7.

146. Calhoun, P.M., et al., *Efficacy of an Overnight Predictive Low-Glucose Suspend System in Relation to Hypoglycemia Risk Factors in Youth and Adults With Type 1 Diabetes.* J Diabetes Sci Technol, 2016. **10**(6): p. 1216-1221.

147. Biester, T., et al., *"Let the Algorithm Do the Work": Reduction of Hypoglycemia Using Sensor-Augmented Pump Therapy with Predictive Insulin Suspension (SmartGuard) in Pediatric Type 1 Diabetes Patients.* Diabetes technology & therapeutics, 2017. **19**(3): p. 173-182.

148. Sparacino, G., et al., *Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series.* IEEE Trans Biomed Eng, 2007. **54**(5): p. 931-7.

149. Perez-Gandia, C., et al., *Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring.* Diabetes Technol Ther, 2010. **12**(1): p. 81-8.

150. Zecchin, C., et al., *Reduction of number and duration of hypoglycemic events by glucose prediction methods: a proof-of-concept in silico study.* Diabetes Technol Ther, 2013. **15**(1): p. 66-77.

151. Daskalaki, E., et al., *An early warning system for hypoglycemic/hyperglycemic events based on fusion of adaptive prediction models.* Journal of diabetes science and technology, 2013. **7**(3): p. 689-698.

152. Contreras, I., et al. *A Hybrid Clustering Prediction for Type 1 Diabetes Aid: Towards Decision Support Systems Based upon Scenario Profile Analysis*. in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. 2017.

153. Toffanin, C., et al., *Multiple models for artificial pancreas predictions identified from free-living condition data: A proof of concept study.* Journal of Process Control, 2019. **77**: p. 29-37.

154. Oviedo, S., et al., *Risk-based postprandial hypoglycemia forecasting using supervised learning.* International Journal of Medical Informatics, 2019. **126**: p. 1-8.

155. Oviedo, S., et al., *Minimizing postprandial hypoglycemia in Type 1 diabetes patients using multiple insulin injections and capillary blood glucose self-monitoring with machine learning techniques.* Computer Methods and Programs in Biomedicine, 2019. **178**: p. 175-180.

156. Cappon, G., et al., *Classification of Postprandial Glycemic Status with Application to Insulin Dosing in Type 1 Diabetes-An In Silico Proof-of-Concept.* Sensors (Basel), 2019. **19**(14).

157. Schiffrin, A. and S. Suissa, *Predicting nocturnal hypoglycemia in patients with type I diabetes treated with continuous subcutaneous insulin infusion.* The American Journal of Medicine, 1987. **82**(6): p. 1127-1132.

158. Guemes, A., et al., *Predicting Quality of Overnight Glycaemic Control in Type 1 Diabetes using Binary Classifiers.* IEEE Journal of Biomedical and Health Informatics, 2019: p. 1-1.

159. Vehi, J., et al., *Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning.* Health Informatics J, 2019: p. 1460458219850682.

160. Bertachi, A., et al., *Prediction of Nocturnal Hypoglycemia in Adults with Type 1 Diabetes under Multiple Daily Injections Using Continuous Glucose Monitoring and Physical Activity Monitor.* Sensors (Basel), 2020. **20**(6).

161. Riddell, M.C., et al., *Diabetes Technology and Exercise.* Endocrinology and Metabolism Clinics of North America, 2020. **49**(1): p. 109-125.

162. Yardley, J.E., et al., *Resistance versus aerobic exercise: acute effects on glycemia in type 1 diabetes.* Diabetes Care, 2013. **36**(3): p. 537-42.

163. Lee, A.S., et al., *High-intensity interval exercise and hypoglycaemia minimisation in adults with type 1 diabetes: A randomised cross-over trial.* J Diabetes Complications, 2020. **34**(3): p. 107514.

164. Zaharieva, D.P., et al., *Improved Open-Loop Glucose Control With Basal Insulin Reduction 90 Minutes Before Aerobic Exercise in Patients With Type 1 Diabetes on Continuous Subcutaneous Insulin Infusion.* Diabetes Care, 2019. **42**(5): p. 824-831.

165. Moser, O., et al., *Reduction in insulin degludec dosing for multiple exercise sessions improves time spent in euglycaemia in people with type 1 diabetes: A randomized crossover trial.* Diabetes Obes Metab, 2019. **21**(2): p. 349-356.

166. Piotrowicz, A.K., et al., *An on-line support tool to reduce exercise-related hypoglycaemia and improve confidence to exercise in type 1 diabetes.* J Diabetes Complications, 2019. **33**(9): p. 682-689.

167. Ben Brahim, N., et al., *Identification of Main Factors Explaining Glucose Dynamics During and Immediately After Moderate Exercise in Patients With Type 1 Diabetes.* J Diabetes Sci Technol, 2015. **9**(6): p. 1185-91.

168. WILSON, L.M., et al., *1038-P: Results of Interim Analysis of a Randomized Crossover Study in Type 1 Diabetes (T1D) of a Dual-Hormone Closed-Loop System with Xerisol™ Glucagon vs. Insulin-Only Closed-Loop System vs. a Predictive Low Glucose Suspend System.* Diabetes, 2019. **68**(Supplement 1): p. 1038-P.

169. Hajizadeh, I., et al., *Incorporating Unannounced Meals and Exercise in Adaptive Learning of Personalized Models for Multivariable Artificial Pancreas Systems.* J Diabetes Sci Technol, 2018. **12**(5): p. 953-966.

170. Fabris, C., B. Ozaslan, and M.D. Breton, *Continuous Glucose Monitors and Activity Trackers to Inform Insulin Dosing in Type 1 Diabetes: The University of Virginia Contribution.* Sensors (Basel), 2019. **19**(24).

171. Ramkissoon, C.M., et al., *Detection and Control of Unannounced Exercise in the Artificial Pancreas Without Additional Physiological Signals.* IEEE Journal of Biomedical and Health Informatics, 2020. **24**(1): p. 259-267.

172. Beneyto, A., et al., *A New Blood Glucose Control Scheme for Unannounced Exercise in Type 1 Diabetic Subjects.* IEEE Transactions on Control Systems Technology, 2020. **28**(2): p. 593-600.

173. Garcia-Tirado, J., et al., *Closed-loop control with unannounced exercise for adults with type 1 diabetes using the Ensemble Model Predictive Control.* Journal of Process Control, 2019. **80**: p. 202-210.

174. Zhang, Y., K. Chanana, and C. Dunne, *IDMVis: Temporal Event Sequence Visualization for Type 1 Diabetes Treatment Decision Support.* IEEE Transactions on Visualization and Computer Graphics, 2019. **25**(1): p. 512-522.

175. Feigerlova, E., et al., *Effects of e-health educational interventions with patients as active participants on HbA1c level in type 1 diabetes on intensive insulin therapy: A systematic review and meta-analysis of randomized controlled trials.* Diabetes Metab Res Rev, 2020.

176. Franc, S., et al., *Real-life application and validation of flexible intensive insulin-therapy algorithms in type 1 diabetes patients.* Diabetes Metab, 2009. **35**(6): p. 463-8.

177. Xie, J. and Q. Wang, *A personalized diet and exercise recommender system for type 1 diabetes self-management: An in silico study.* Smart Health, 2019. **13**: p. 100069.

178. Tucholski, K., et al., *Assessment of optimal insulin administration timing for standard carbohydrates-rich meals using continuous glucose monitoring in children with type 1 diabetes: A cross-over randomized study.* J Diabetes Investig, 2019. **10**(5): p. 1237-1245.

179. Eissa, M.R., et al., *Intelligent Data-Driven Model for Diabetes Diurnal Patterns Analysis.* IEEE Journal of Biomedical and Health Informatics, 2020: p. 1-1.

180. Stewart, Z.A., et al., *Closed-Loop Insulin Delivery during Pregnancy in Women with Type 1 Diabetes.* N Engl J Med, 2016. **375**(7): p. 644-54.

181. Nosova, E.V., et al., *Leveraging technology for the treatment of type 1 diabetes in pregnancy: A review of past, current, and future therapeutic tools.* J Diabetes, 2020.

182. Kovatchev, B.P., et al., *Evening and overnight closed-loop control versus 24/7 continuous closed-loop control for type 1 diabetes: a randomised crossover trial.* The Lancet Digital Health, 2020. **2**(2): p. e64-e73.

183. Paris, C.A., et al., *Predictors of insulin regimens and impact on outcomes in youth with type 1 diabetes: the SEARCH for Diabetes in Youth study.* J Pediatr, 2009. **155**(2): p. 183-9.e1.

184. *Intensive Diabetes Treatment and Cardiovascular Outcomes in Type 1 Diabetes: The DCCT/EDIC Study 30-Year Follow-up.* Diabetes Care, 2016. **39**(5): p. 686-93.

185. Schwartz, F.L., et al., *Analysis of use of an automated bolus calculator reduces fear of hypoglycemia and improves confidence in dosage accuracy in type 1 diabetes mellitus patients treated with multiple daily insulin injections.* J Diabetes Sci Technol, 2012. **6**(1): p. 150-2.

186. Roze, S., et al., *Cost-effectiveness of continuous subcutaneous insulin infusion versus multiple daily injections of insulin in Type 1 diabetes: a systematic review.* Diabet Med, 2015. **32**(11): p. 1415-24.

187. McNally, K., et al., *Executive Functioning, Treatment Adherence, and Glycemic Control in Children With Type 1 Diabetes.* Diabetes Care, 2010. **33**(6): p. 1159-1162.

188. Sarbacker, G.B. and E.M. Urteaga, *Adherence to Insulin Therapy.* Diabetes Spectr, 2016. **29**(3): p. 166-70.

189. Steil, G.M., et al., *Use of Automated Clinical Decision Support (CDS) to Effect Glycemic Control in Elderly Patients with T1D.* Diabetes, 2018. **67**(Supplement 1): p. 921-P.

190. Resalat, N., et al., *Adaptive tuning of basal and bolus insulin to reduce postprandial hypoglycemia in a hybrid artificial pancreas.* Journal of Process Control, 2019. **80**: p. 247-254.

191. Sørensen, T.J., *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. 1948: I kommission hos E. Munksgaard.

192. Davidson, M.B., et al., *Comparison of Insulin Dose Adjustments by Primary Care Physicians and Endocrinologists.* Clinical Diabetes, 2018. **36**(1): p. 39.

193. Bashan, E. and I. Hodish, *Frequent insulin dosage adjustments based on glucose readings alone are sufficient for a safe and effective therapy.* Journal of Diabetes and its Complications, 2012. **26**(3): p. 230-236.

194. Reddy, R., et al., *The effect of exercise on sleep in adults with type 1 diabetes.* Diabetes Obes Metab, 2018. **20**(2): p. 443-447.

195. Pettus, J. and S.V. Edelman, *Recommendations for Using Real-Time Continuous Glucose Monitoring (rtCGM) Data for Insulin Adjustments in Type 1 Diabetes.* Journal of Diabetes Science and Technology, 2017. **11**(1): p. 138-147.

196. Whitney, A.W., *A Direct Method of Nonparametric Measurement Selection.* IEEE Trans. Comput., 1971. **20**(9): p. 1100-1103.

197. Scheiner, G., *Practical CGM: improving patient outcomes through continuous glucose monitoring*. 2015: American Diabetes Association.

198. White, J.W., et al., *Ecologists should not use statistical significance tests to interpret simulation model results.* Oikos, 2014. **123**(4): p. 385-388.

199. Colberg, S.R., et al., *Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association.* Diabetes Care, 2016. **39**(11): p. 2065-2079.

200. Moniotte, S., et al., *Outcomes of algorithm-based modifications of insulinotherapy during exercise in MDI vs insulin pump-treated children with type 1 diabetes: Results from the TREAD-DIAB study.* Pediatr Diabetes, 2017. **18**(8): p. 925-933.

201. Lascar, N., et al., *Attitudes and barriers to exercise in adults with type 1 diabetes (T1DM) and how best to address them: a qualitative study.* PLoS One, 2014. **9**(9): p. e108019.

202. Bussau, V.A., et al., *The 10-s maximal sprint: a novel approach to counter an exercise-mediated fall in glycemia in individuals with type 1 diabetes.* Diabetes Care, 2006. **29**(3): p. 601-6.

203. Wahren, J., *Glucose turnover during exercise in man.* Ann N Y Acad Sci, 1977. **301**: p. 45-55.

204. Man, C.D., M.D. Breton, and C. Cobelli, *Physical activity into the meal glucose-insulin model of type 1 diabetes: in silico studies.* J Diabetes Sci Technol, 2009. **3**(1): p. 56-67.

205. McMahon, S.K., et al., *Glucose requirements to maintain euglycemia after moderate-intensity afternoon exercise in adolescents with type 1 diabetes are increased in a biphasic manner.* J Clin Endocrinol Metab, 2007. **92**(3): p. 963-8.

206. Steenberg, D.E., et al., *Exercise training reduces the insulin-sensitizing effect of a single bout of exercise in human skeletal muscle.* The Journal of physiology, 2019. **597**(1): p. 89-103.

207. Boulé, N.G., et al., *Effects of Exercise Training on Glucose Homeostasis.* Diabetes Care, 2005. **28**(1): p. 108.

208. Wilson, L.M., et al., *Patient Input for Design of a Decision Support Smartphone Application for Type 1 Diabetes.* Journal of diabetes science and technology, 2020. **14**(6): p. 1081-1087.

209. Guillot, F.H., et al., *Accuracy of the Dexcom G6 Glucose Sensor during Aerobic, Resistance, and Interval Exercise in Adults with Type 1 Diabetes.* Biosensors (Basel), 2020. **10**(10).

210. Brown, S.A., et al., *Six-Month Randomized, Multicenter Trial of Closed-Loop Control in Type 1 Diabetes.* The New England journal of medicine, 2019. **381**(18): p. 1707-1717.

211. Wilson, L.M., et al., *Dual-Hormone Closed-Loop System Using a Liquid Stable Glucagon Formulation Versus Insulin-Only Closed-Loop System Compared With a Predictive Low Glucose Suspend System: An Open-Label, Outpatient, Single-Center, Crossover, Randomized Controlled Trial.* Diabetes Care, 2020.

212. Jacobs, P.G., et al., *Randomized trial of a dual-hormone artificial pancreas with dosing adjustment during exercise compared with no adjustment and sensor-augmented pump therapy.* Diabetes, Obesity and Metabolism, 2016. **18**(11): p. 1110-1119.

213. Riddell, M.C., et al., *Exercise management in type 1 diabetes: a consensus statement.* Lancet Diabetes Endocrinol, 2017. **5**(5): p. 377-390.

214. Li, Z., et al., *An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs.* Diabetes Care, 2018. **41**(12): p. 2509.

215. Tyler, N.S., et al., *An artificial intelligence decision support system for the management of type 1 diabetes.* Nature Metabolism, 2020. **2**(7): p. 612-619.

216. Tyler, N.S. and P.G. Jacobs, *Artificial Intelligence in Decision Support Systems for Type 1 Diabetes.* Sensors (Basel, Switzerland), 2020. **20**(11): p. 3214.

217. Pérez-Gandía, C., et al., *Artificial Neural Network Algorithm for Online Glucose Prediction from Continuous Glucose Monitoring.* Diabetes Technology & Therapeutics, 2010. **12**(1): p. 81-88.

218. Zecchin, C., et al., *Neural Network Incorporating Meal Information Improves Accuracy of Short-Time Prediction of Glucose Concentration.* IEEE Transactions on Biomedical Engineering, 2012. **59**(6): p. 1550-1560.

219. Zhu, T., et al., *Dilated Recurrent Neural Networks for Glucose Forecasting in Type 1 Diabetes.* Journal of Healthcare Informatics Research, 2020. **4**(3): p. 308-324.

220. Xie, J. and Q. Wang, *Benchmarking Machine Learning Algorithms on Blood Glucose Prediction for Type I Diabetes in Comparison With Classical Time-Series Models.* IEEE Transactions on Biomedical Engineering, 2020. **67**(11): p. 3101-3124.

221. Contreras, I., et al., *Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models.* PLOS ONE, 2017. **12**(11): p. e0187754.

222. American College of Sports, M. and B.A. Bushman, *ACSM's Complete Guide to Fitness & Health*. Vol. Second edition. 2017, Champaign, IL: Human Kinetics.

223. Ruder, S., *An overview of gradient descent optimization algorithms.* arXiv preprint arXiv:1609.04747, 2016.

224. Yardley, J.E. and R.J. Sigal, *Glucose management for exercise using continuous glucose monitoring: should sex and prandial state be additional considerations?* Diabetologia, 2021. **64**(4): p. 932-934.

225. Al Khalifah, R.A., et al., *Association of aerobic fitness level with exercise-induced hypoglycaemia in Type 1 diabetes.* Diabetic Medicine, 2016. **33**(12): p. 1686-1690.

226. Jacobs, P.G., et al., *Incorporating an Exercise Detection, Grading, and Hormone Dosing Algorithm Into the Artificial Pancreas Using Accelerometry and Heart Rate.* J Diabetes Sci Technol, 2015. **9**(6): p. 1175-84.

227. Resalat, N., et al. *Design of a dual-hormone model predictive control for artificial pancreas with exercise model*. in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016.

228. Resalat, N., et al., *Evaluation of model complexity in model predictive control within an exercise-enabled artificial pancreas*. Vol. 50. 2017. 8022-8028.

229. Nguyen, T.-T.P., et al., *Separating Insulin-Mediated and Non-Insulin-Mediated Glucose Uptake during Aerobic Exercise in People with Type 1 Diabetes.* American Journal of Physiology-Endocrinology and Metabolism, 2020.

230. Nguyen, T.-T.P., et al., *Separating insulin-mediated and non-insulin-mediated glucose uptake during and after aerobic exercise in type 1 diabetes.* American Journal of Physiology-Endocrinology and Metabolism, 2020. **320**(3): p. E425-E437.

231. American Diabetes Association, *6. Glycemic Targets: Standards of Medical Care in Diabetes—2021.* Diabetes Care, 2021. **44**(Supplement 1): p. S73.

232. Michaelides, A., et al., *Weight loss efficacy of a novel mobile Diabetes Prevention Program delivery platform with human coaching.* BMJ Open Diabetes Research & Care, 2016. **4**(1): p. e000264.

233. Ku, E.J., et al., *Clinical efficacy and plausibility of a smartphone-based integrated online real-time diabetes care system via glucose and diet data management: a pilot study.* Internal Medicine Journal, 2020. **50**(12): p. 1524-1532.

234. Faruqui, S.H.A., et al., *Development of a Deep Learning Model for Dynamic Forecasting of Blood Glucose Level for Type 2 Diabetes Mellitus: Secondary Analysis of a Randomized Controlled Trial.* JMIR Mhealth Uhealth, 2019. **7**(11): p. e14452.

235. Krishnamoorthy, D., et al., *A Model-free Approach to Automatic Dose Guidance in Long Acting Insulin Treatment of Type 2 Diabetes.* IEEE Control Systems Letters, 2020: p. 1-1.

236. Thyde, D.N., et al., *Machine Learning-Based Adherence Detection of Type 2 Diabetes Patients on Once-Daily Basal Insulin Injections.* Journal of Diabetes Science and Technology, 2020. **15**(1): p. 98-108.

237. Brazg, R., et al., *Effect of adding sitagliptin, a dipeptidyl peptidase-4 inhibitor, to metformin on 24-h glycaemic control and β-cell function in patients with type 2 diabetes.* Diabetes, Obesity and Metabolism, 2007. **9**(2): p. 186-193.

238. Ruan, Y., et al., *Modelling endogenous insulin concentration in type 2 diabetes during closed-loop insulin delivery.* BioMedical Engineering OnLine, 2015. **14**(1): p. 19.

239. Visentin, R., C. Cobelli, and C. Dalla Man, *The Padova Type 2 Diabetes Simulator from Triple-Tracer Single-Meal Studies: In Silico Trials Also Possible in Rare but Not-So-Rare Individuals.* Diabetes Technology & Therapeutics, 2020. **22**(12): p. 892-903.

240. Horton, T.J., et al., *Glucose kinetics differ between women and men, during and after exercise.* Journal of Applied Physiology, 2006. **100**(6): p. 1883-1894.

241. Vieira, R.F.L., et al., *Exercise activates AMPK signaling: Impact on glucose uptake in the skeletal muscle in aging.* Journal of Rehabilitation Therapy, 2020. **2**(2).

242. Manousaki, D., et al., *Toward Precision Medicine: TBC1D4 Disruption Is Common Among the Inuit and Leads to Underdiagnosis of Type 2 Diabetes.* Diabetes Care, 2016. **39**: p. dc160769.

243. Moltke, I., et al., *A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes.* Nature, 2014. **512**(7513): p. 190-193.

244. Dash, S., et al., *A truncation mutation in TBC1D4 in a family with acanthosis nigricans and postprandial hyperinsulinemia.* Proceedings of the National Academy of Sciences, 2009. **106**(23): p. 9350.

245. Treebak, J.T., et al., *Acute exercise and physiological insulin induce distinct phosphorylation signatures on TBC1D1 and TBC1D4 proteins in human skeletal muscle.* The Journal of Physiology, 2014. **592**(2): p. 351-375.

246. Jessen, N., et al., *Exercise increases TBC1D1 phosphorylation in human skeletal muscle.* American journal of physiology. Endocrinology and metabolism, 2011. **301**(1): p. E164-E171.

247. Stöckli, J., et al., *The RabGAP TBC1D1 Plays a Central Role in Exercise-Regulated Glucose Metabolism in Skeletal Muscle.* Diabetes, 2015. **64**(6): p. 1914.

248. Sylow, L., et al., *Exercise-stimulated glucose uptake - regulation and implications for glycaemic control.* Nat Rev Endocrinol, 2017. **13**(3): p. 133-148.

249. Yan, Z., et al., *Regulation of exercise-induced fiber type transformation, mitochondrial biogenesis, and angiogenesis in skeletal muscle.* Journal of Applied Physiology, 2010. **110**(1): p. 264-274.

250. Bird, S.R. and J.A. Hawley, *Update on the effects of physical activity on insulin sensitivity in humans.* BMJ open sport & exercise medicine, 2017. **2**(1): p. e000143-e000143.

251. Steenberg, D.E., et al., *Exercise training reduces the insulin-sensitizing effect of a single bout of exercise in human skeletal muscle.* J Physiol, 2019. **597**(1): p. 89-103.

252. Jacobs, P.G., et al., *Measuring glucose at the site of insulin delivery with a redox-mediated sensor.* Biosensors and Bioelectronics, 2020. **165**: p. 112221.

253. Ward, W.K., et al., *Can glucose be monitored accurately at the site of subcutaneous insulin delivery?* Journal of diabetes science and technology, 2014. **8**(3): p. 568-574.

254. McClatchey, P.M., et al., *Fibrotic Encapsulation Is the Dominant Source of Continuous Glucose Monitor Delays.* Diabetes, 2019. **68**(10): p. 1892.

255. Miller, K.M., et al., *Longitudinal Changes in Continuous Glucose Monitoring Use Among Individuals With Type 1 Diabetes: International Comparison in the German and Austrian DPV and U.S. T1D Exchange Registries.* Diabetes Care, 2020. **43**(1): p. e1.

256. Onisie, O., H. Crocket, and M. de Bock, *The CGM grey market: a reflection of global access inequity.* The Lancet Diabetes & Endocrinology, 2019. **7**(11): p. 823-825.

257. Graham, C., *Continuous Glucose Monitoring and Global Reimbursement: An Update.* Diabetes technology & therapeutics, 2017. **19**(S3): p. S60-S66.

258. Omer, A.E., et al., *Low-cost portable microwave sensor for non-invasive monitoring of blood glucose level: novel design utilizing a four-cell CSRR hexagonal configuration.* Scientific Reports, 2020. **10**(1): p. 15200.

259. Elsherif, M., et al., *Wearable Contact Lens Biosensors for Continuous Glucose Monitoring Using Smartphones.* ACS Nano, 2018. **12**(6): p. 5452-5462.

260. Li, M., et al., *A highly integrated sensing paper for wearable electrochemical sweat analysis.* Biosensors and Bioelectronics, 2021. **174**: p. 112828.

261. de Castro, L.F., et al., *Salivary diagnostics on paper microfluidic devices and their use as wearable sensors for glucose monitoring.* Analytical and Bioanalytical Chemistry, 2019. **411**(19): p. 4919-4928.

262. Jacobs, P. and R. Reddy, *Exercise, Sleep, and Type 1 Diabetes*. 2020. p. 145-157.

263. van Hees, V.T., et al., *Estimating sleep parameters using an accelerometer without sleep diary.* Scientific Reports, 2018. **8**(1): p. 12975.

264. de Arriba-Pérez, F., M. Caeiro-Rodríguez, and J.M. Santos-Gago, *How do you sleep? Using off the shelf wrist wearables to estimate sleep quality, sleepiness level, chronotype and sleep regularity indicators.* Journal of Ambient Intelligence and Humanized Computing, 2018. **9**(4): p. 897-917.

265. Sevil, M., et al., *Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management.* IEEE Transactions on Biomedical Engineering, 2021: p. 1-1.

266. Reddy, R.K., et al., *Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: evaluation study.* JMIR mHealth and uHealth, 2018. **6**(12): p. e10338.

267. Robergs, R.A., F. Ghiasvand, and D. Parker, *Biochemistry of exercise-induced metabolic acidosis.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2004. **287**(3): p. R502-R516.

268. Ražanskas, P., et al., *Predicting Blood Lactate Concentration and Oxygen Uptake from sEMG Data during Fatiguing Cycling Exercise.* Sensors, 2015. **15**(8).

269. Resnick, P., et al., *GroupLens: an open architecture for collaborative filtering of netnews*, in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. 1994, ACM: Chapel Hill, North Carolina, USA. p. 175-186.

270. Balabanovi, M., #263, and Y. Shoham, *Fab: content-based, collaborative recommendation.* Commun. ACM, 1997. **40**(3): p. 66-72.

271. Campos, P., F. Díez, and I. Cantador, *Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols.* User Modeling & User-Adapted Interaction, 2014. **24**(1/2): p. 67-119.

272. Gräßer, F., et al., *Therapy Decision Support Based on Recommender System Methods*. Vol. 2017. 2017. 1-11.