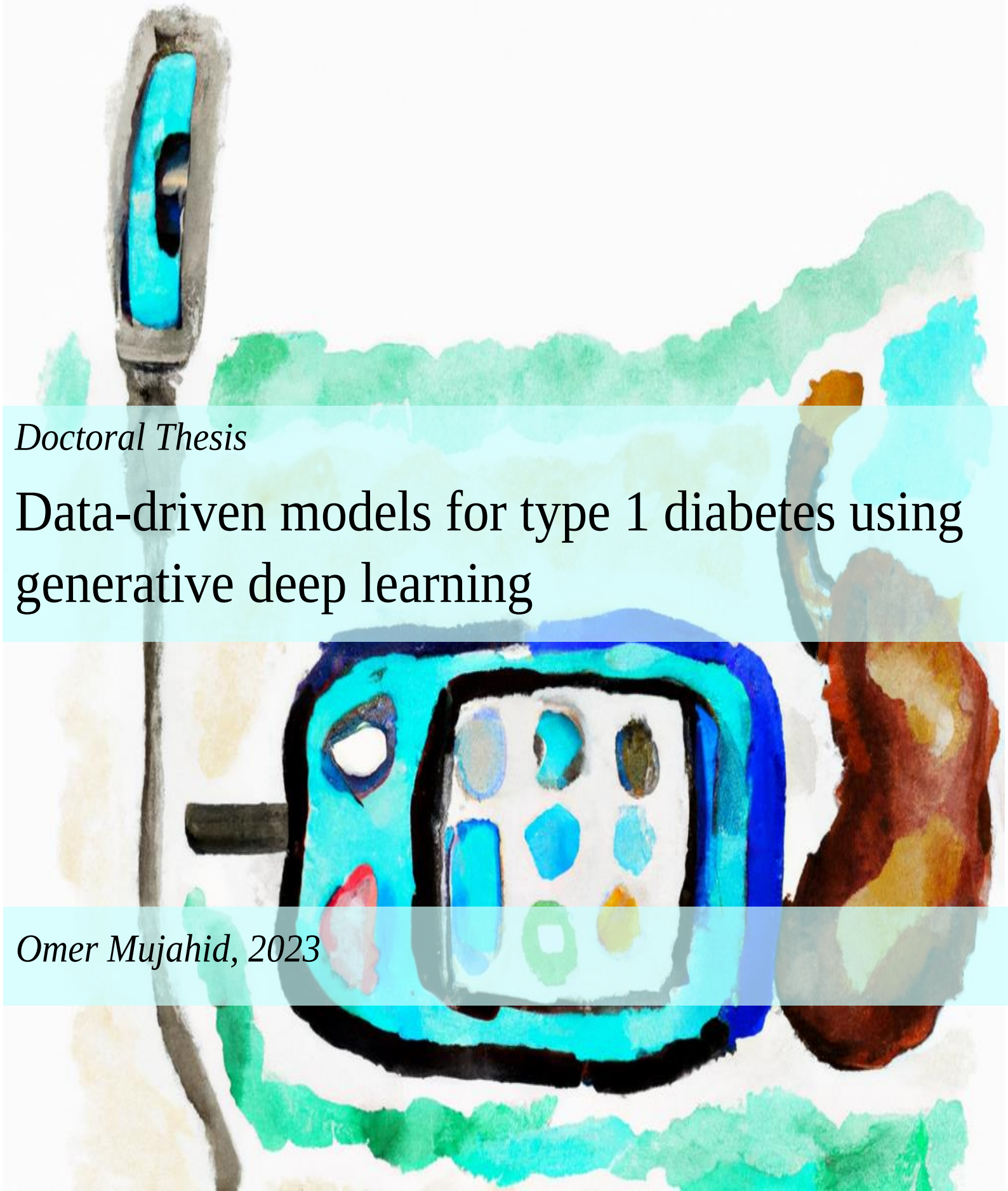
Universitat de Girona



Doctoral Thesis

Data-driven models for type 1 diabetes using generative deep learning

Omer Mujahid, 2023

DOCTORAL THESIS

DATA-DRIVEN MODELS FOR TYPE 1 DIABETES USING GENERATIVE DEEP LEARNING

OMER MUJAHID

2023

DOCTORAL PROGRAMME IN TECHNOLOGY

Supervised by:

Prof. Josep Vehí Casellas and Dr. Ivan Contreras

Tutored by:

Prof. Josep Vehí Casellas

Presented in partial fulfillment of the requirements for a doctoral degree from the University of Girona

DOCTORAL THESIS

DATA-DRIVEN MODELS FOR TYPE 1 DIABETES USING GENERATIVE DEEP LEARNING

A dissertation presented in partial fulfillment of the requirements for a doctoral degree from the University of Girona.

By:

Omer Mujahid

Supervisors:

Prof. Josep Vehí Casellas

Dr. Iván Contreras

Universitat de Girona

APPROVAL TERM

The Doctoral Thesis entitled **"DATA-DRIVEN MODELS FOR TYPE 1 DIABETES USING GENERATIVE DEEP LEARNING"**, presented by **Omer Mujahid**, on May 4, 2023, has been evaluated and approved by the examination committee as partial fulfillment of the requirements for a doctoral degree from the University of Girona.

Examination Committee:

_____

Dr.

_____

Dr.

_____

Dr.

Girona, May 4, 2023.

I dedicate this work to my parents, Muhammad Mujahid and Farzana Mujahid, whose unwavering support has been my guiding force throughout my academic journey. And to my wife, Amina Sundas, whose love and dedication inspire me to be my best self.

ACKNOWLEDGEMENTS

The completion of this thesis is not solely attributed to one individual, but rather to the collective efforts of many individuals. It is with great appreciation that I extend my acknowledgment to those who have played a significant role in this challenging journey.

I would like to extend my utmost appreciation to my supervisor, Prof. Josep Vehí, who envisioned the outcome of this thesis, provided valuable insights, and never shied away from sharing his knowledge. After that, I would like to thank my co-supervisor, Ivan Contreras, who helped me frame the problems, implement models, and write papers. I would also like to thank my colleague at MICE Lab, Aleix Beneyto, for his help in reviewing my write-ups and providing assistance and resources in the implementation of closed-loop therapies.

I am also indebted to my colleagues, Sayyar and Ernesto, for answering my numerous queries and providing resources that helped in the formation of this thesis. I would like to express my gratitude to my fellow PhD students, Alvita, Najib, and Ibrahim, for our technical discussions and their eagerness to provide help whenever I needed it. Furthermore, I am sincerely grateful to our project manager, Sara, who helped me carry out the formal procedures involved in my Ph.D.

My deepest gratitude goes to my parents, Mujahid and Farzana, who have been the greatest source of strength for me in this journey. I would like to thank my siblings, Sana, Maria, and Ali, for their support in taking care of my parents and managing affairs at home while I was away.

Last but not least, I am eternally grateful to my wife, Amina, who lifted me up every time I was down during this journey and stood by me through thick and thin.

Omer Mujahid
Girona, Spain
April, 2023

LIST OF PUBLICATIONS

This thesis is based on a compendium of the following publications:

1. **Omer Mujahid**, Ivan Contreras and Josep Vehí. Machine learning techniques for hypoglycemia prediction: trends and challenges, *Sensors*, 21(2):546, **2021**.
2. **Omer Mujahid**, Ivan Contreras, Aleix Beneyto, Ignacio Conget, Marga Giménez, and Josep Vehí. Conditional synthesis of blood glucose profiles for T1D patients using deep generative models, *Mathematics*, 10(20):3741, **2022**.
3. **Omer Mujahid**, Ivan Contreras, Aleix Beneyto and Josep Vehí. Generative deep learning for the development of a type 1 diabetes simulator, *Communications Medicine*, Under Review.

The research work leading to this thesis resulted in additional journal and conference publications, which are listed below and sorted by publication date.

Journals

4. Josep Noguer, Ivan Contreras, **Omer Mujahid**, Aleix Beneyto, and Josep Vehi. Generation of Individualized Synthetic Data for Augmentation of the Type 1 Diabetes Data Sets Using Deep Learning Models, *Sensors*, 22(13):4944, **2022**.

Conferences

1. **Omer Mujahid**, Ivan Contreras, and Josep Vehi. Generation of realistic virtual diabetes patients using a Pix2Pix GAN: in pursuit of an AI-based diabetes simulator. In Virtual Diabetes Technology Meeting 2021. Abstract

2. Josep Noguer, **Omer Mujahid**, Josep Vehi, and Ivan Contreras. Improvement of the predictive ability of nocturnal hypoglycemia models by using synthetic data created by generative adversarial networks. In 15th Advanced Technologies & Treatments for Diabetes (ATTD 2022), Barcelona - Spain. Poster

3. **Omer Mujahid**. Enhancing predictive models thanks to synthetic data. In the 44th International Engineering in Medicine and Biology Conference (EMBC 2022), Glasgow - United Kingdom. Minisymposium

Book Chapters

1. Josep Vehi, **Omer Mujahid**, and Ivan Contreras. "Aim and Diabetes." In Artificial Intelligence in Medicine, pp. 701-709. Cham: Springer International Publishing, 2022.

2. Josep Vehi, **Omer Mujahid**, and Ivan Contreras. "Artificial Intelligence and Machine Learning for Diabetes Decision Support." In Advanced Bioscience and Biosystems for Detection and Management of Diabetes, pp. 259-272. Cham: Springer International Publishing, 2022.

LIST OF FIGURES

FIGURE **Page**

ACRONYMS AND ABBREVIATIONS

The following acronyms and abbreviations can be found in this thesis.

Acronyms and abbreviations

BG	Blood Glucose
CGM	Continuous Glucose Monitoring
CL	Closed-loop
CHO	Carbohydrate
CGAN	Conditional Generative Adversarial Network
CIBERDEM	*Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas*
DNN	Deep Neural Network
DGM	Deep Generative Model
DM	Diabetes Mellitus
GAN	Generative Adversarial Network
IDF	International Diabetes Federation
MICELAB	Modeling, Identification & Control Engineering Laboratory
OL	Open-Loop
P2P	Pixel-to-Pixel
S2S	Sequence-to-Sequence
T1D	Type 1 Diabetes
T2D	Type 2 Diabetes
UdG	University of Girona - *Universitat de Girona*

TABLE OF CONTENTS

Abstract

Modeling biological systems has always been challenging given the complexity of the processes involved in them. Experts have been employing physiological models to approximate the dynamics of biological systems; however, these models are constrained by the limitations of mathematical techniques that can only encompass part of the physical phenomena behind a biological system. Mathematical physiological models of the human glucose-insulin system are considered the gold standard of simulators in type 1 diabetes (T1D) healthcare. Though accurate to a certain degree, these models are not capable of simulating scenarios that could fully capture the real-life dynamics of a T1D patient. The underlying cause for this phenomenon could be attributed to the numerous hidden factors that are ignored during physiological modeling because of increasing model complexity or hurdles in their representation.

This work was carried out with a focus on accurate model approximation in T1D. The rationale is built on the hypothesis that generic function approximators such as deep neural networks (DNNs) have the ability to learn all that from data that cannot be modeled mathematically. Since deep generative models (DGMs) are implemented using DNNs, they are capable of learning the underlying probability distribution of a data set. This thesis presents several methodologies based on data-driven models using DGMs for improved model approximation in T1D. Firstly, a systematic review of data-driven models for predicting hypoglycemia is conducted. After that, a methodology for data augmentation in a hypoglycemia classifier using a generative adversarial network (GAN) is developed as part of this thesis. The next work in this series focuses on the conditional synthesis of realistic BG profiles of T1D patients. Finally, building on the work performed thus far, a T1D simulation environment is developed using a sequence-to-sequence GAN (S2S GAN) that is capable of synthesizing realistic patients with T1D.

The results obtained from these methods show the efficacy of DGMs for model formation in T1D. It has been demonstrated through these results that a highly precise approximation of the glucose-insulin system of patients with T1D can be obtained from data with the help of DGMs. Moreover, these models have been shown to generate novel data that is statistically similar to real data for all the standardized glycemic metrics. Furthermore, the causal synthesis of realistic T1D data has been shown in the work presented in this thesis.

RESUMEN

Modelar sistemas biológicos siempre ha sido un desafío dada la complejidad de los procesos involucrados en ellos. Actualmente, la implementación de modelos fisiológicos es el estandar para aproximarse a la dinámica de los sistemas biológicos; sin embargo, estos modelos están restringidos por las limitaciones de las técnicas matemáticas, que solo pueden abarcar una parte de los fenómenos físicos detrás de un sistema biológico. Los modelos fisiológicos matemáticos del sistema de glucosa-insulina humano se consideran el modelo de referencia de los simuladores en el ambito del cuidado de la salud en la diabetes tipo 1 (T1D). Aunque estos modelos son precisos hasta cierto punto, no son capaces de simular escenarios que puedan capturar completamente la dinámica de la vida real de un paciente con T1D. La causa subyacente de este fenómeno podría atribuirse a los numerosos factores ocultos que se ignoran durante el modelado fisiológico debido a la creciente complejidad del modelo o los obstáculos en su representación.

Este trabajo se ha llevado a cabo centrandose en una aproximación precisa del modelo en T1D. La justificación se basa en la hipótesis de que los aproximadores de funciones genéricas, como las redes neuronales profundas (DNN), tienen la capacidad de aprender todo aquello de los datos que no se puede modelar matemáticamente. Dado que los modelos generativos profundos (DGM) se basan en DNN, estos son capaces de aprender la distribución de probabilidad subyacente de un conjunto de datos. Esta tesis presenta varias metodologías basadas en modelos basados en datos que utilizan DGM para mejorar la aproximación del modelo en T1D. En primer lugar, se realiza una revisión sistemática de los modelos basados en datos para predecir la hipoglucemia. Posteriormente, como parte de esta tesis, se desarrolla una metodología para aumentar una base de datos en un clasificador de hipoglucemia utilizando una red generativa antagónica (GAN). A continuación, la tesis se centra en la síntesis condicional de perfiles realistas de glucosa en sangre de pacientes con T1D usando redes generativas condicionales. Finalmente, sobre la base del trabajo realizado hasta el momento, se desarrolla un entorno de simulación de pacientes con T1D utilizando una GAN de secuencia a secuencia (S2S GAN).

Los resultados obtenidos de estos métodos muestran la eficacia de los DGM para la generacion de modelos en la T1D. Esta tesis demuestra que las DGM son capaces de obtener una aproximación muy precisa del sistema glucosa-insulina de pacientes con T1D. Por otro lado, se ha demostrado que estos modelos generan datos ineditos que son estadísticamente similares a los datos reales para todas las métricas de glucemia estandar. Además, el trabajo presentado en esta tesis ha demostrado la síntesis causal de datos realistas en pacientes con

T1D.

RESUM

Modelar sistemes biològics sempre ha estat un desafiament atesa la complexitat dels processos involucrats en ells. Actualment, la implementació de models fisiològics és l'estàndard per aproximar-se a la dinàmica dels sistemes biològics; aquests models, però, estan restringits per les limitacions de les tècniques matemàtiques, que només poden abastar una part dels fenòmens físics darrere d'un sistema biològic. Els models fisiològics matemàtics del sistema de glucosa-insulina humà es consideren el model de referència dels simuladors en la cura de la salut de la diabetis tipus 1 (T1D). Tot i que són necessaris fins a cert punt, aquests models no són capaços de simular escenaris que puguin capturar completament la dinàmica de la vida real d'un pacient amb T1D. La causa subjacent d'aquest fenomen podria atribuir-se als nombrosos factors ocults que s'ignoren durant el modelatge fisiològic a causa de la complexitat creixent del model o dels obstacles en la seva representació.

Aquest treball s'ha dut a terme centrant-se en una aproximació precisa del model a T1D. La justificació es basa en la hipòtesi que els aproximadors de funcions genèriques, com les xarxes neuronals profundes (DNN), tenen la capacitat d'aprendre tot allò de les dades que no es pot modelar matemàticament. Com que els models generatius profunds (DGM) s'implementen mitjançant DNN, aquests són capaços d'aprendre la distribució de probabilitat subjacent d'un conjunt de dades. Aquesta tesi presenta diverses metodologies basades en models basats en dades que fan servir DGM per millorar l'aproximació del model a T1D. En primer lloc, es fa una revisió sistemàtica dels models basats en dades per predir la hipoglucèmia. Posteriorment, com a part d'aquesta tesi, es desenvolupa una metodologia per a l'augment de dades en un classificador d'hipoglucèmia utilitzant una xarxa generativa antagònica (GAN). El següent treball se centra en la síntesi condicional de perfils realistes de glucosa a la sang de pacients amb T1D usant xarxes generatives condicionals. Finalment, sobre la base del treball fet fins ara, es desenvolupa un entorn de simulació de pacients amb T1D utilitzant una GAN de seqüència a seqüència (S2S GAN).

Els resultats obtinguts d'aquests mètodes mostren l'eficàcia dels DGM per a la generació de models a la T1D. Aquesta tesi demostra que les DGM són capaces d'aconseguir una aproximació molt precisa del sistema glucosa-insulina de pacients amb T1D. D'altra banda, s'ha demostrat que aquests models generen dades inèdites que són estadísticament similars a les dades reals per a totes les mètriques de glucèmia estàndard. A més a més, el treball presentat en aquesta tesi ha demostrat la síntesi causal de dades realistes en pacients amb T1D.

CHAPTER 1

INTRODUCTION

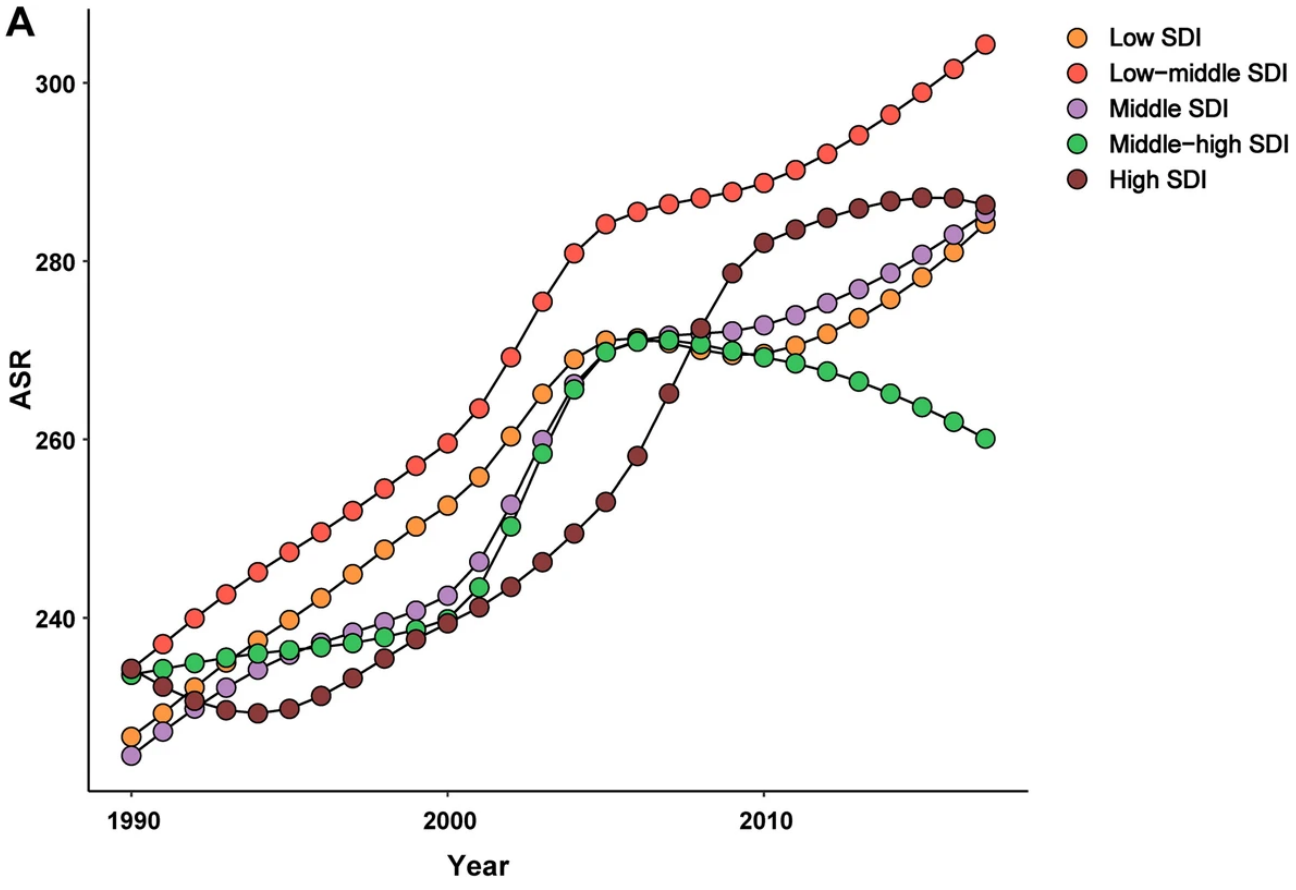
This chapter presents an overview of the importance of models in diabetes healthcare in section 1.1 followed by a discussion of conventional ways of modeling in type 1 diabetes (T1D) in Section 1.2. Section 1.3 provides a run-through of the various advantages data-driven models offer with a focus on deep generative models. Section 1.4 talks us through the main objectives of the thesis whereas section 1.5 gives us a structure of this thesis.

1.1 The Importance of Modeling in Diabetes Mellitus

Diabetes mellitus (DM) is a global epidemic with over 422 million people suffering from it globally and its prevalence increases with time (Lovic et al., 2020). The worldwide incidence rate of DM has increased by 102% from 1990 to 2017 (Liu et al., 2020). Moreover, according to the international diabetes federation, the number of people living with diabetes are predicted to rise from 537 million in 2021 to 643 million by 2030 and 783 million by 2045. Figure 1.1 shows the yearly increase in the incidence rate of diabetes. DM is a diverse set of diseases that

influences the use of blood glucose (BG) in the body and is characterized by hyperglycemia caused by the absolute or relative reduction in insulin action or production (Alam et al., 2014). The chronic conditions in diabetes are termed as type 1 diabetes (T1D) and type 2 diabetes (T2D), whereas those conditions that are potentially reversible are prediabetes and gestational diabetes. T2D is the insufficient production of insulin by the pancreas or the inept response of the cells to the insulin produced. T1D, on the other hand, is a disorder caused by the inability of the human pancreas to produce insulin. It is an autoimmune disease in which the human immune system attacks the islet cells in the pancreas leaving it permanently damaged. T1D is also referred to as insulin-dependent diabetes because of the fact that people suffering from it are forever dependent on insulin administration.

Figure 1.1: The ASIR of diabetes mellitus caused by SDI regions, from 1990 to 2017. (ASIR, age-standardized incidence rate; SDI, socio-demographic index) (Liu et al., 2020).



Even though every type of diabetes is dangerous, T1D may be considered the most fatal. If left unattended, T1D may lead to life-threatening comorbidities and serious complications. People living with T1D lead stressful lives because of the efforts required to manage this complex disease. It is estimated that a person living with T1D has to make about 180 more decisions per day regarding their health as compared to someone without diabetes (Digitale et al., 2019). Decisions like the choice of meal, physical activity, and insulin administration make T1D one of the most complex diseases to manage. The need for improved therapeutic regimens and treatments has always been there in T1D healthcare. Over the years, engineers have aspired to come up with cutting-edge technologies that could help patients with T1D manage their disease better. Technologies like automatic insulin delivery systems that harness the control action of closed-loop controllers (Dovc and Battelino, 2020), insulin pumps, smart pens, and mobile applications (Beck et al., 2019) have been making the lives of patients with T1D easier.

The advancements in technology for a certain disease has often been complemented by corresponding progress in the modeling of the disease (Nath et al., 2018). The reason for this is the numerous advantages models of biological systems offer for technological purposes. Of the many advantages they present, models help develop our understanding of the disease. Moreover, they can be used to test and evaluate new treatments, therapies, and technologies before they are tested on humans. Furthermore, they may aid the process of personalized treatment in healthcare and enhance our ability to predict the outcome of a disease. Models also help expose functions of complex biological systems (Harline et al., 2021). In addition to all these advantages, the major contribution of biological models can be seen in the assembly of simulation environments. Biological simulation environments are tools that strive to provide us with a glimpse of real patients' dynamics using a computer. All biological simulators employ some kind of biological model in integration with certain other mechanisms to emulate the behavior of a real-life biological phenomenon. The human glucose-insulin system is an example

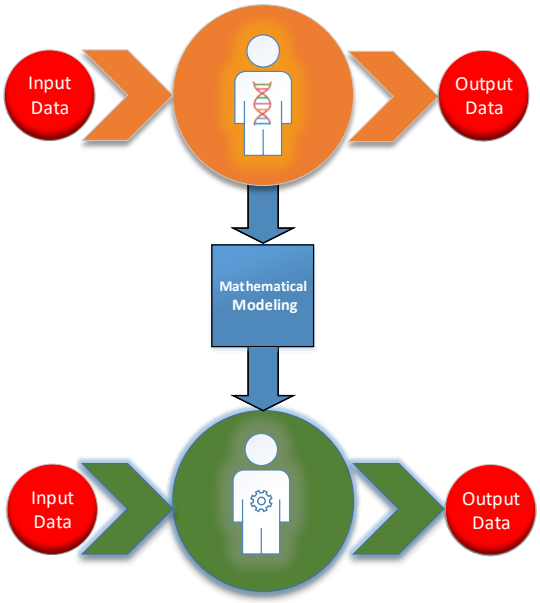
of a complex biological system with a nested hierarchy of complex patterns and processes. Over the years, scientists and engineers have developed several models of the glucose-insulin system with the aim of accurate approximation of the system. These models of the human glucose-insulin system have been utilized by several diabetes simulators (Man et al., 2014) (Hovorka et al., 2004) (Bergman and Urquhart, 1971).

1.2 Models in Type 1 Diabetes

Traditionally, the modeling of biological systems has been carried out using mathematical techniques. Mathematical physiological and pharmacokinetic models of biological systems are pretty much considered the standard way of modeling in biology. Figure 1.2 depicts the concept of mathematical modeling from the internal physiology of the biological system. While mathematical models provide advantages in terms of model transparency and interpretability, they are incapable of encompassing the entire physiology of biological systems. These models strive for accurate approximation but suffer from phenomena that are either hidden or too complex to model. In T1D, the aim of modeling is to emulate the glucose-insulin action in patients precisely. However, as discussed above, mathematical models in T1D suffer from inaccuracies because the domain knowledge that guides model development is in itself incomplete (Engelhardt et al., 2016). The physiology of a human glucose-insulin system is affected by numerous factors such as patients' life choices and habits and other life disturbances such as menstruation, depression, medication, etc. These factors define the trajectory of a patient's glycemic profile and are necessary to be taken into consideration in order to obtain realistic models. With mathematical modeling, taking these factors into consideration is extremely difficult as they increase the complexity of modeling immensely. Hence, these mathematical models suffer from inaccuracies that arise in them in the form of grey noise. This noise directly influences the simulations based on the mathematical models, often giving birth to wrongly simulated scenarios. The issues with mathematical models provide a justification for

the exploration of other modeling techniques that would improve the approximation of the glucose-insulin system in patients with T1D and enable the creation of simulation environments that are as closer to reality as possible. This research work has been conducted with this aim in mind. The alternative route taken for model approximation in T1D is with the data-driven models using generative deep learning.

Figure 1.2: The mathematical modeling approach that estimates the system's physiology with a set of equations and model parameters.



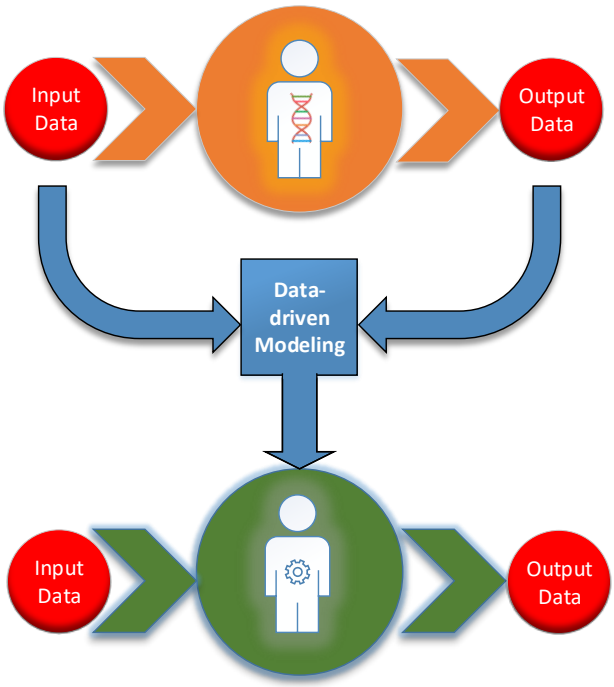
1.3 Data-driven Models

Data-driven modeling refers to a category of methodologies that involve the acquisition of models from input and output data without the explicit incorporation of the underlying physical principles governing the system. The idea of data-driven modeling could broadly be understood

from Figure 1.3. It can be observed from the figure that model approximation is done from the input and output data and no attempt at interpreting the internal physiology of the biological system is made. Hence, the most important resource in data-driven modeling is data. With the advent of new technologies in diabetes, acquiring diabetes-related data has become much easier. Continuous glucose monitors (CGM) are becoming part of the standard care for patients with T1D in the first world countries and their use is increasing with each passing day. CGM has made the acquisition of large quantities of BG data possible. Moreover, technologies like smart insulin pens and connected mobile applications have enabled users to store their insulin-related information in mobile applications and other connected devices. According to a population-based study conducted from 1995 to 2017, among 96,547 participants, the percentage of patients using insulin pump therapy increased from 1% in 1995 to 53% in 2017, while the percentage of patients using CGM rose from 3% in 2006 to 38% in 2017 (van den Boom et al., 2019). Furthermore, the use of CGM in clinical trials has increased as well, making the acquisition of large quantities of BG data through clinical trials possible (Fox et al., 2021). Moreover, carbohydrate information could be fed and stored by patients using mobile applications, whereas additional related data, e.g. physical activity information could be acquired using physical activity trackers like Fitbit, Apple watch, etc. With the availability of multidimensional diabetes-related data, the notion of data-driven models in diabetes healthcare has held pace over the last few years and its popularity is still on the rise.

Designers have utilized different statistical and machine learning techniques to harness the power of data into useful models. Moreover, these models have been employed for different purposes such as prediction, classification, regression, etc. Using generic function approximators like deep neural networks (DNNs) for model approximation is one such type of approach. DNNs have the ability to learn complex underlying probability distributions from data. According to the universal approximation theorem, an artificial neural network with a single hidden layer can learn any function given that it is sufficiently wide. Some studies have shown that

Figure 1.3: The data-driven modeling approach that involves learning a model from the input and output data of a system



DNNs have the capability of surpassing mathematical models in terms of system approximation provided that the training data is available in sufficient quantity (Nikzad et al., 2012) (Nalisnick et al., 2018). In addition to the various tasks mentioned above that the data-driven models could perform, if set up a particular way, they are also capable of generating new data instances. In probability theory, such models are referred to as generative models. These models include

the distribution of the data itself by learning a joint probability distribution P(X,Y), or P(X) if there are no labels. They are different than discriminative models, which learn a conditional probability distribution P(X|Y) instead.

When DNNs are used for the task of learning a joint probability distribution and then generating novel samples from the learned distribution, they are referred to as deep generative models (DGMs). Over the past few years, the effectiveness of DGMs in producing realistic data samples in different fields of life has been proven by various studies (Newton, 2019) (Coutinho-Almeida et al., 2021) (Gonzalez-Abril et al., 2022). There are several types of DGM architectures employed by designers for generating various types of data such as images, text, music, etc. These architectures include generative adversarial networks (GANs), variational autoencoders (VAEs), normalizing flows, diffusion models, etc. Each one of these models then has several variants that can be used for the generation or conditional generation of novel data samples. Since the research work presented in this thesis utilizes GANs and its variants, the focus of the discussion will be GANs.

1.4 Objectives

Building on the rationale presented above, the major goal of this thesis could be written down as:

"To enhance the current state-of-the-art models in T1D by endeavoring to attain a comprehensive estimation of the glucose-insulin dynamics."

The principal objective can be divided into specific research objectives that are pursued by this thesis.

- To assess the current state of data-driven models in T1D with a focus on hypoglycemia prediction by considering the effects of prediction horizon, type of data, and type of model.

- To demonstrate the efficacy of deep generative models by generating realistic BG data for patients with T1D conditioned on the plasma insulin approximation of the patient.
- To condition the generation of BG values on more than one variable (i.e. insulin and carbohydrates) and prove improved model approximation.
- To set up a T1D simulation environment with open-loop (OL) and closed-loop (CL) therapies using a deep generative model.

1.5 Thesis Structure

This thesis is organized into four main chapters. After this introductory chapter, Chapter 2 provides a comprehensive compilation of the articles on which this thesis is founded. Chapter 3 critically analyzes the key contributions, challenges, and constraints that are inherent to the research presented in these articles. In particular, this chapter offers an in-depth evaluation of the research questions that have been addressed, the methodologies employed, and the results obtained. Finally, Chapter 4 provides a conclusive discussion of the future research avenues that are emerging from the current work, and highlights their potential implications for the field.

CHAPTER 2

DATA-DRIVEN MODELS FOR TYPE 1 DIABETES USING GENERATIVE DEEP LEARNING

This chapter consists of four sections. Section 2.1 presents a review paper on the latest trends and challenges in using machine learning for hypoglycemia prediction. Section 2.2 consists of a paper wherein realistic blood glucose profiles are synthesized using conditional generative adversarial networks. Section 2.3 comprises a manuscript that lays out a framework for the development of T1D simulator using DGMs.

- **2.1 Machine learning techniques for hypoglycemia prediction: trends and challenges**
- **2.2 Conditional synthesis of blood glucose profiles for T1D patients using deep generative models**
- **2.3 Generative deep learning for the development of a type 1 diabetes simulator**

2.1 Machine learning techniques for hypoglycemia prediction: trends and challenges

In this publication, we have performed a systematic review of the state-of-the-art of machine learning models in hypoglycemia prediction.

Title: Machine learning techniques for hypoglycemia prediction: trends and challenges

Authors: **Omer Mujahid**, Ivan Contreras, and Josep Vehí

Journal: MDPI Sensors

Volume: 21, Number:2 Published: January 2021

DOI: https://doi.org/10.3390/s21020546

Quality index: JCR 2021 Engineering, Electrical & Electronic, Impact factor: 3.847, Q2 (95/276)

sensors

MDPI

Review

Machine Learning Techniques for Hypoglycemia Prediction: Trends and Challenges

Omer Mujahid [1], Ivan Contreras [1] and Josep Vehi [1,2,*]

[1] Model Identification and Control Laboratory, Institut d'Informatica i Applicacions, Universitat de Girona, 17003 Girona, Spain; omer.mujahid@udg.edu (O.M.); ivancontreras@udg.edu (I.C.)
[2] Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), 17003 Girona, Spain
* Correspondence: josep.vehi@udg.edu

Abstract: (1) Background: the use of machine learning techniques for the purpose of anticipating hypoglycemia has increased considerably in the past few years. Hypoglycemia is the drop in blood glucose below critical levels in diabetic patients. This may cause loss of cognitive ability, seizures, and in extreme cases, death. In almost half of all the severe cases, hypoglycemia arrives unannounced and is essentially asymptomatic. The inability of a diabetic patient to anticipate and intervene the occurrence of a hypoglycemic event often results in crisis. Hence, the prediction of hypoglycemia is a vital step in improving the life quality of a diabetic patient. The objective of this paper is to review work performed in the domain of hypoglycemia prediction by using machine learning and also to explore the latest trends and challenges that the researchers face in this area; (2) Methods: literature obtained from PubMed and Google Scholar was reviewed. Manuscripts from the last five years were searched for this purpose. A total of 903 papers were initially selected of which 57 papers were eventually shortlisted for detailed review; (3) Results: a thorough dissection of the shortlisted manuscripts provided an interesting split between the works based on two categories: hypoglycemia prediction and hypoglycemia detection. The entire review was carried out keeping this categorical distinction in perspective while providing a thorough overview of the machine learning approaches used to anticipate hypoglycemia, the type of training data, and the prediction horizon.

Keywords: hypoglycemia; machine learning; prediction; detection; artificial intelligence; decision support system (DSS)

Citation: Mujahid, O.; Contreras, I.; Vehi, J. Machine Learning Techniques for Hypoglycemia Prediction: Trends and Challenges. *Sensors* **2021**, *21*, 546. https://doi.org/10.3390/s21020546

Received: 17 December 2020
Accepted: 12 January 2021
Published: 14 January 2021

1. Introduction

Hypoglycemia is the drop in blood glucose (BG) below critical levels [1]. The BG level at which hypoglycemia occurs, however, has long been a topic of much debate in medical circles [2]. The most accepted definition is that when the BG level drops below 70 mg/dL or 3.9 mmol/L, hypoglycemia is diagnosed [3]. It is one of the most lethal conditions that may arise most commonly in type 1 diabetics (T1D) followed by type 2 diabetics (T2D). Hypoglycemia may lead to loss of consciousness, confusion, seizures, and in extreme cases, death [4]. The symptoms of hypoglycemia, however, may vary for different individuals based on several factors. For a symptom to be associated with hypoglycemia, it is important that it satisfies Whipple's triad [5]. This essentially means that the symptom is consistent with hypoglycemia, the blood glucose level is below the normal range, and the symptom is relieved when the plasma glucose level is increased to normal or above. The symptoms of hypoglycemia are eventually connected to neuronal glucose deprivation, which in layman terms means glucose deprivation of the human nervous system and brain [6]. These symptoms could then be categorized into neurogenic symptoms caused by glucose deprivation of the autonomic nervous system and the neuroglycopenic symptoms caused by the glucose deprivation of the central nervous system. Based on the level of BG, the symptoms of hypoglycemia can be categorized as mild, moderate, and extreme. It is in

the extreme form that hypoglycemia is most lethal. In most cases, however, hypoglycemia does not have any symptoms at all and occurs silently. The silent arrival of hypoglycemia is one of the causes of distress for its sufferers. Along with the physical discomforts that hypoglycemia brings, the mental torments are a major reason for the diabetics to despise its existence. The insecurity and fear of hypoglycemia causes the life quality of diabetics to degrade immensely [7]. The main cause of hypoglycemia is the reduction in blood glucose levels because of an overdose of insulin or a low intake of food/carbohydrates [8]. The reduction in BG that is caused by insulin or any other form of drugs is known as iatrogenic hypoglycemia [9]. Hypoglycemia may occur because of multiple other reasons, i.e., kidney failure, liver complications, hyperthyroidism, starvation, and the consumption of certain drugs including alcohol. Hypoglycemia may be classified into multiple groups based on factors, i.e., the agent of cause, the time of the day it occurs, the age of the individual, the severity of the glycemic event, and the connection to another condition in the body [10].

Based on the time of occurrence, hypoglycemia is commonly characterized into daytime hypoglycemia, postprandial hypoglycemia [11], and nocturnal hypoglycemia [10]. Daytime hypoglycemia typically means the hypoglycemic event that occurs during the day. Postprandial hypoglycemia refers to the hypoglycemic event after the patient has eaten. It could also be referred to as reactive hypoglycemia, whereas nocturnal hypoglycemia means the hypoglycemic event occurs during the night when the patient is sleeping. Each type of hypoglycemia has its own associated risks. Patients run the risk of postprandial hypoglycemia when they misestimate the amount of carbohydrates (CHO) consumed in each meal. Varying insulin sensitivity is also a major factor in misanalysing the amount of bolus insulin needed and might lead to postprandial hypoglycemia [12]. Nocturnal hypoglycemia, on the other hand, is a much bigger problem than any other form of hypoglycemia. The reason is that nocturnal hypoglycemia occurs when the patient is sleeping and is virtually incapable of defending him-/herself against the glycemic event. The fact that over half of all the extreme hypoglycemic episodes occur during sleep add to the severity of this type of hypoglycemia.

Since hypoglycemia is a combination of various symptoms when blood glucose drops below 70 mg/dL and sometimes it is entirely asymptomatic, diagnosing it is very hard and it is near impossible for a human to predict its occurrence in advance. In the case of a hypoglycemic event, the initial treatment could be consuming 15 to 20 g of fast acting carbohydrates [2], and even though the consumption of glucose seems like the only solution to overcome an ailment that is caused by the deficiency of glucose, it takes 10–15 min for the human body to process glucose [4]. This means that the patient has already experienced mental and physical trauma before returning to a normal glycemic state. Moreover, clinical evidence and observational data show that the recommended glycated haemoglobin (HbA)targets are not met in the majority of T1D patients [13]. A more appropriate approach is to manage the blood glucose in such way that hypoglycemia is prevented.

There has been an immense surge in the use of technologies for diabetes management. Glucose monitoring systems have been one of the trending topics in biomedicine [14]. Multiple glucose monitoring devices are available these days that provide periodic or flash updates of the patient's glucose levels. Some commercially available devices include Medtronic CGM, Abbott FreeStyle Libre, and Dexcom CGM systems [15,16]. These devices contain a continuous glucose monitoring (CGM) sensor along with a portable monitor that displays glucose levels and in some cases provides alarms of adverse glycemic events. The CGM sensors measure glucose dynamically and have a tiny filament inserted beneath the skin. These sensors remain in contact with the interstitial fluid with the help of an enzymatic electrode. Such electrodes use enzymes to cause reduction–oxidation reactions and then measure the amount of current or voltage produced by the movement of electrons, which is often concentration dependant [15]. The latest commercially available CGM sensors such as the FreeStyle Libre by Abbot give a BG value reading with a sampling time

of 1 s and has a lifespan of 14 days after its first use, during which it does not need to be calibrated. This makes the process of testing BG less painful.

These monitoring devices are sometimes used in coordination with an insulin pump to form the sensor-augmented pump (SaP) therapy [17]. The SaP forms an important component of a closed-loop artificial pancreas (AP) system. Such systems have been worked upon for many years [18]. A closed-loop AP system has three main components, a CGM, an insulin pump, and control algorithm that controls the insulin dose. In other cases, the glucose monitoring systems, when used in coordination with an artificially intelligent decision-making module that gives suggestions about insulin and carbohydrate intake to the patients, form a decision support system (DSS). The DSS has proven to be an apt therapy for multiple daily injections (MDI) users, which is the most common method of insulin treatment for diabetic patients.

Machine learning (ML) has emerged as one of the major fields of artificial intelligence (AI) in recent times, and its impact on healthcare has been huge [19,20]. The concept of ML has its roots in computer science, statistics, and optimization. With a focus on enabling the computer to train itself without being explicitly programmed, ML gives a computer the power to predict outcomes up to a certain level of accuracy. In many medical scenarios knowledge of an adverse event beforehand could prevent an emergency and in many cases save lives. The quality of ML to predict the future makes it a great tool to anticipate such events [21]. Hypoglycemia, being one of such events, may also be anticipated using ML. The uncertainty associated with the occurrence of a hypoglycemic event looms on the horizon for T1Ds, making their lives ever so miserable. Biomedical engineers, therefore, want to come up with efficient predicting models in order to reduce the uncertainty and improve the life quality of diabetics. This is the reason that there has been an exponential increase in research work focused on ML techniques to predict adverse glycemic events in general and hypoglycemia in particular [22]. It is still too early to say that most such works are truly ready to be made commercially available for the public use; however, encouraging results have been seen in several of these works. It is known that ML techniques feed on large amounts of data in order for their prediction to be accurate. Moreover, the data need to be diverse and free of any corruption and irregularities [23]. To have such data for any biomedical application is a hard task because of the involvement of many such constraints that affect the quality of the data being acquired. Medical data are renowned for being complex and disordered. The limitations associated with sensors, noncompliance of patients to the study protocols, faults in the study protocols, and unwillingness of patients to undergo the study are some of the factors that affect the quality of data available for the training of ML algorithms. It is for this reason that biomedical data require a lot of pre-processing and filtration before being ready to be fitted with an ML model.

The Aim of Hypoglycemia Prediction

Experts have tried to identify hypoglycemia based on different characteristics but most of the times, hypoglycemia is asymptomatic and is often unrecognized. This is one reason hypoglycemia can prove deadly. The absence of signs and prior indicators may cause the patients to act undesirably in the wake of a hypoglycemic event and consequently move themselves into disaster. Though the occurrence of hypoglycemia is hard to determine, it is often observed in patients who take insulin regularly [24]. Of the patients who take insulin, type 1 diabetics are three times more likely to experience hypoglycemia as compared to type 2 diabetics [25].

In many cases even when a hypoglycemic is recognized by the patient, it is often too late to prevent it. Hence, taking carbohydrates/glucose when a hypoglycemic event is taking place will not help the cause. It is therefore necessary to have a mechanism that could inform the patient in advance about the occurrence of a hypoglycemic event in the future. The aim of such a system should be to correctly forecast a hypoglycemic event in the future and then inform/warn the patient about it. A prediction system like this could be efficiently embedded in a decision support system (DSS). A DSS could then

guide the patient about the steps and measures to be taken to prevent the predicted event from happening.

This review focuses on the performance and potential of several such works. The works that are reviewed here are explicitly focused on ML techniques for hypoglycemia prediction/detection in T1Ds. Tables 1 and 2 shows the entire collection of manuscripts reviewed. It could be observed from these tables that the majority of the works done in the domain of hypoglycemia prediction/detection were published in 2019 and 2020. This is proof of a rising trend in the use of ML models for hypoglycemia prediction/detection. It is important to mention here that throughout this review, the ML frameworks are not discussed explicitly. No effort in establishing a ranking criterion has been made. The reason for this is that a large variety of ML frameworks are used in the literature and also, the factors defining the frameworks are diverse. Since no two studies used a common framework for ML modelling, comparing research works based on their frameworks was a hard task. Another reason of refraining from any sort of quantitative comparison was to keep the review as impartial as possible and let the readers establish an understanding of the work done in the field of ML-based prediction of hypoglycemia.

The methodology of the entire review process is discussed in the next section. Results obtained from the review are discussed in the section after that. A thorough analysis of the reviewed manuscripts is done in the results section based on a distinction between studies aimed at hypoglycemia detection and prediction, the data used to train the ML models, the type of ML models used, and the prediction/forecasting horizon. A discussion about the entire review is presented in the succeeding section followed by a conclusion of the presented work.

Table 1. Summary of reviewed manuscripts addressing hypoglycemia detection: ML model used to perform the detection, type of data (ToD) used to train the ML model, the size of the cohort, recording duration (RD), age of population (AoP), gender of the participants in study, treatment method (TM).

Ref	Year	ML Model	ToD	Cohort	RD	AoP	Gender	TM
[26]	2019	cTAKES	clinical notes	395 and 460 notes	-	-	-	-
[27]	2017	LDA [a]	BG, breath samples	56	1 bag each	CH [n], ADO [p], ADU [q]	F [r]: 55.36%, M [s]: 44.64%	-
[28]	2018	PLSR [b], ANN [d]	temp, IR [c], Z	20	2 days	ADU	-	-
[29]	2018	CNN [e]	EHR	500 records	95,246 sentences	-	-	-
[30]	2019	LSVM [f], LR [g], RF [h]	Secure messages	3000 messages	-	-	-	-
[31]	2020	XGBoost	EHR	17,658	4 years	ADU	F: 47% M: 53%	MDI
[32]	2016	SVM [i]	HR, temp, GSR	1	2 months	ADU	M: 100%	IP [u]
[33]	2018	RF, MLP [t]	BG, PA	93	4 months	CH, ADO, ADU	F: 46.2% M:53.7%	IP
[34]	2016	DT [j]	ECG, breath data, accelerometer	5	260 h	-	-	IP
[35]	2016	DL [k]	ECG	15	10 h	CH	-	MDI
[36]	2019	DL	EHR	500 records	-	-	-	-
[37]	2015	-	EEG	15	-	CH, ADO, ADU	-	-
[38]	2019	KNN [l]	camera, BG	14	850 samples/subjects	ADU	-	-

[a] LDA: Linear Discriminant Analysis; [b] PLSR: Partial Least Square Regression; [c] IR: Infra-Red; [d] ANN: Artificial Neural Network; [e] CNN: Convolutional Neural Network; [f] LSVM: Linear Support Vector Machine; [g] LR: Logistic Regression; [h] RF: Random Forest; [i] SVM: Support Vector Machine; [j] DT: Decision Tree; [k] DL: Deep Learning; [l] KNN: K-Nearest Neighbor; [n] CH: Children; [p] ADO: Adolescents; [q] ADU: Adults; [r] F: Female; [s] M: Male; [t] MLP: Multilayer Perceptron.

Table 2. Summary of reviewed manuscripts addressing hypoglycemia prediction: ML model used to perform the detection, prediction horizon (PH) in minutes, the size of the cohort, recording duration (RD), type of data (ToD) used to train the ML model, treatment method (TM), age of population (AoP), gender of the participants in study.

Ref	Year	ML Model	PH (min)	Cohort					
				Size	RD	ToD	TM	AoP	Gender
[39]	2019	LDA	35	463	4721 nights	BG, Insulin	IP	ADU	F: 58% M: 42%
[40]	2019	DT, RF	30	55	244 exercise sessions	HR, BG	IP	ADU	F: 60% M: 40%
[41]	2020	MLP, SVM	360	10	12 weeks	HR, BG, CHO	MDI	ADU	F:80% M:20%
[42]	2016	Extreme ML NN	360	16	4.09 days	ECG	-	CH	-
[43]	2019	RF, SVM, KNN, LR	30	104	113 days	BG	MDI	ADU	F: 60% M: 40%
[44]	2017	k-mean clustering	540	34	10 days	BG	-	ADU	-
[45]	2020	RMRF [a]	-	127	2525 nights	BG, PA, Insulin, CHO	-	-	-
[46]	2019	Ensemble of commonly used ML models	30	104	Between 2014 and 2015	1 (BG)	-	-	-
[47]	2020	RF	360	9800	1 mil nights	BG	IP	ADU	F: 51% M: 49%
[48]	2020	SVR [b]	360	124	22,804 nights	BG, Insulin	IP	ADO, ADU	F: 60% M: 40%
[49]	2016	stochastic models	60, 240, 360	34	150 days	BG	MDI	CH, ADU	-
				179	476 days				
[50]	2019	ANN	30	N/A	1 Week	BG	-	-	-
[51]	2018	ANN	30, 60	6	8 weeks	Insulin, BG, PA, CHO	IP	ADU	-
[52]	2020	LR, RF	0–15, 15–30, 30–45, 45–60	112	90 days	BG, Insulin, CHO	IP	ADO, ADU	F: 39.2% M: 60.7%
[53]	2019	ANN, SVM, AB [c], GNB [d]	240	10 10	Several months	BG, PA, Insulin, CHO	MDI	ADU	-
[54]	2017	CART	15	33	72 to 96 h	BG	-	-	-
[55]	2020	MLR, LASSO	420	100	162,000 traces	BG, CHO	-	ADU	-
				218	2 months				
[56]	2020	MDP [e]	210	NIDDK repository	6 Treatment points	BG, Insulin, CHO	-	ADU	-

Table 2. *Cont.*

Ref	Year	ML Model	PH (min)	Cohort					
				Size	RD	ToD	TM	AoP	Gender
[57]	2019	ARIMA, RF, SVM	15	25	14 days	BG	IP	ADU	F: 44% M: 56%
[58]	2020	KKR [f]	30	11	7–50 days	BG	IP	ADU	-
[59]	2019	GE [g], SVM, ANN	60	100	14 days	BG, Insulin, CHO, PA	IP	ADU	-
			240	10	6 weeks				
			360	6	8 weeks				
[60]	2019	RF, SVM, ANN	120	6	8 weeks	Insulin, BG, PA, CHO	IP	ADU	-
[61]	2020	KNN	10,080	70	15 weeks	BG, insulin, CHO	MDI	-	-
[62]	2019	GRU [h]	45	40	4 days	BG	-	-	-
[63]	2017	DL	30	25	N/A	BG	-	CH, ADO	-
[64]	2020	RNN [i]	30	10	360 days	Insulin, BG, PA, CHO	IP	-	-
				6	8 weeks				
[36]	2019	RNN	30	124	27,466 days	BG, Insulin	IP	ADU	-
[65]	2020	DRL [j]	The meal duration	10 10	6 months	BG, CHO	MDI	ADO, ADU	-
[66]	2019	XGBT [q]	The meal duration	100	2 months	BG, Insulin, CHO	-	ADU	-
[67]	2019	KNN	The meal duration	100	4 days	BG, Insulin, CHO	-	ADU	-
[68]	2019	SVM	The meal duration	10		BG, Insulin, CHO	IP	ADU	F:20%, M: 80%
[69]	2016	Combination of NH predictors	360	34	150 days	BG	MDI	CH, ADU	-
				179	476 days				
[70]	2016	ACL [k]	1440	28 100		BG, PA, Insulin, CHO	-	CH, ADO, ADU	-
[71]	2019	10 Different ML Methods	30	6	8 weeks	Insulin, BG, PA	IP	ADU	-

Table 2. *Cont.*

Ref	Year	ML Model	PH (min)	Cohort					
				Size	RD	ToD	TM	AoP	Gender
[72]	2020	RF, GBT	The meal duration	100	162 meal conditions	BG, Insulin, CHO	-	-	-
[73]	2019	DRL	N/A	10 10	30 days	CGM, CHO	-	ADO, ADU	-
[74]	2019	DL	30, 60	10	6 months	BG, Insulin, CHO	IP	ADO, ADU	-
				6	8 weeks				
				10	180 days				
[75]	2020	ARM [r], RF, LGBM [l], FCNNs [m],GCNN [n]	30, 60	141	9083 days	BG	-	-	-
				30	30 days				
[76]	2018	ANN	30	12	1 year	BG	IP	ADU	F: 50%, M: 50%
[77]	2017	GP [o], RF, KNN, GE	30	10	N/A	BG, insulin, CHO	-	-	-
[78]	2020	DL	30, 60	10 10	6 months	BG, Insulin, CHO, PA		ADU	-
[79]	2019	RF, SVM	15	25	14 days	BG	MDI	ADU	F: 44%, M: 56%
[80]	2017	LSTM [p]	30, 60, 90	106	7 days	BG	Both	-	-

[a] RMRF: Repeated Measures Random Forest; [b] SVR: Support Vector Regression; [c] AB: Adaboost; [d] GNB: Gaussian Naïve Bayes; [e] MDP: Markov Decision Process; [f] KKR: Kernel Ridge Regression; [g] GE: Grammatical Evolution; [h] GRU: Gradient Recurrent Unit; [i] RNN: Recurrent Neural Network; [q] XGBT: Extreme Gradient Boosted Tree; [k] ACL: Actor Critic Learning; [l] LGBM: Light Gradient Boosting; [m] FCNN: Fully-convolutional Neural Networks; [n] GCNN: Gradually Connected Neural Networks; [o] GP: Genetic Programming; [p] LSTM: Long Short-term Memory; [q] ARM: Autoregressive Model.

2. Materials and Methods

This paper looks at the broader horizon of the work done in the domain of ML for hypoglycemia prediction. The searched manuscripts were obtained by combining the results of multiple individual searches to form a pool of 900 manuscripts. PubMed and Google Scholar were used for the selection of manuscripts. PubMed was selected because it is the premier source of published research in biomedicine and life sciences available on the internet. Google scholar was selected for manuscript searching to enlarge the search area. English articles of the past five years were considered in this review. We excluded studies that involved type 2 diabetes or were review articles.

Manuscripts were searched through the advanced searching options in PubMed. The search was carried out by first combining the keywords 'machine learning' and 'hypoglycemia' with the help of an 'AND' logical operator to search all the fields provided in PubMed advanced search option. This search yielded a total of 41 manuscripts. Later, keywords 'artificial intelligence' and 'hypoglycemia' were searched together for all the fields, which yielded a total of 47 manuscripts. 'Machine learning' was also searched together with 'blood glucose prediction', yielding a total of 119 manuscripts. The keywords 'hypoglycemia' and 'machine learning' were then searched together with a series of other keywords by using the same logical operator to obtain the following results: prediction (23), detection (10), hypoglycemic event (15), and adverse glycemic event (4). The keyword 'hypoglycemia' was then solely searched with other keywords using the logical operator yielding the following results: support vector machine (9), random forest (15), deep learning (8), ANN (6), supervised learning (21), and clustering (82). In google scholar, keywords 'machine learning', and 'hypoglycemia' were searched. This search was carried out to expand the pool of the total shortlisted manuscripts. A total of 500 manuscripts were searched using google scholar. All these individual searches were then combined together to form a grand pool of 900 manuscripts. Here, it is important to understand that the majority of the manuscript searched in both google scholar and PubMed were similar because both platforms provide distinct methods of article searching. A thorough review of the selected manuscript pool was performed. Moreover, the bibliographies of the selected manuscripts were looked into for a detailed analysis of the manuscripts cited in these works. The shortlisted manuscripts were then scrutinized to obtained the final collection of 57 papers by using the methodology given in Figure 1.

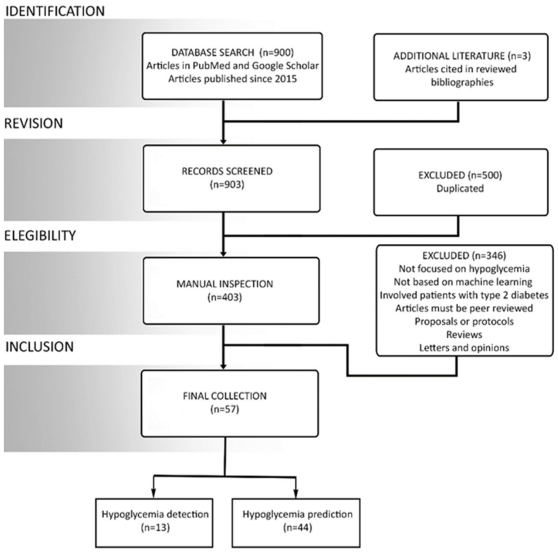


Figure 1. Overview of the review process and classification of literature.

3. Results

Results were obtained from the 57 shortlisted papers after a comprehensive analysis of the attributes that were found to be most impactful in deciding the quality of the work. The details of all the reviewed manuscripts are given in Tables 1 and 2. The results section is based on the following categories:

- The prediction and detection of hypoglycemia
- Type of data
- ML models
- Prediction horizon (PH)

3.1. The Prediction and Detection of Hypoglycemia

The first major classification of the manuscripts reviewed was done on the basis of where in time the ML models look for the occurrence of a hypoglycemic event. Hypoglycemia prediction essentially means forecasting the future hypoglycemic events. On the other hand, the detection of hypoglycemia only means detecting whether a hypoglycemic event has occurred at the present time or not. Many ML-based systems are just detection models. They do not look into the future to forecast the occurrence of an event. Even though this review primarily focuses on discussing the prediction models, the importance of detection models cannot be undermined. Automatic real-time detection of hypoglycemia may be crucial in many scenarios. In this section, works whose aim was to recognize or estimate the occurrence of a hypoglycemic event in the present have been identified. The purpose of doing so was to narrow down the review towards the works that were only focused on hypoglycemia prediction in the sections ahead. It is important to take into consideration that in order to have an ML algorithm that forecasts future events we must have time series data. In the context of this review, this is equivalent to saying that in order to predict the occurrence of a hypoglycemic event at a specific time in the future, the data used to train the ML models need to contain the BG, insulin, CHO or some other form of time series data.

From this, it can be deduced that works that do not use time series data do not try to predict the occurrence of hypoglycemia in the future but most often than not try to detect hypoglycemia in the present. The details of such works are given in Table 1. Physiological parameters of an electrocardiogram (ECG) were used to detect hypoglycemia by Ling et al. [42], Ranvier et al. [34], and San et al. [35]. Multiple systems used text and language processing for the detection of hypoglycemia. For instance, hypoglycemia was detected from electronic health records (EHRs) in the investigations proposed by Jin et al. [29], Ruan et al. [31], and Jin Li et al. [34]. Chen et al. [30] employed patient secure messages for automatic detection of hypoglycemia while Zhou et al. [26] aimed at detecting hypoglycemia by processing the text of clinical notes of patients. Temperature, near infra-red, and bio impedance sensors were employed in their system for the detection of BG trends during the occurrence of a hypoglycemic event by Tronstad et al. [28]. Marling et al. [32] used heart rate (HR), temperature, and galvanic skin response (GSR), and Juhl et al. [37] utilized electroencephalogram (EEG) data to perform hypoglycemia detection.

3.2. Type of Data: What Are the Current Models Trained on?

ML engineering primarily involves fitting ML models to a large amount of data in order to locate patterns and classify them into different label groups. For an ML model to work efficiently, a large quantity of good relevant data is required, which means that the data used to train ML models should be accurate, complete, and valid. However, in biomedical applications, the availability of good data for ML designers is rare, the reason being different natural and technical constraints involved in the process of data collection.

ML models for hypoglycemia prediction/detection may be trained on several types of data. The manuscripts we have reviewed used 12 different types of data to train ML models. These data include BG, insulin, carbohydrates (CHO), ECG, EHRs, HR, breath samples, temperature, clinical notes, secret messages, GSR, and EEG. Figure 2 shows the

distribution of the number of manuscripts for each type of data. It must be kept in mind that by data we mean the acquired data in their original form and not the extracted features.

Figure 2. Different types of data used to train ML models.

3.2.1. Blood Glucose (BG) Data

As one might think, the most relevant data while predicting events based on BG levels would be the BG values itself. Based on the profile of an individual's BG levels an ML system could be trained to predict the future BG values. This approach was used by the majority of the works that have been reviewed. Approximately 77% of the total manuscripts reviewed use BG data to train ML models for hypoglycemia prediction. It is, however, important to mention that not all of these models were trained on actual clinical data. Actual clinical data come from clinical trials. These trials are overseen by a clinical trial protocol that describes the terms and conditions under which the study is ought to be conducted. Some of these works use BG data from diabetes patient simulators. Diabetes simulators are platforms that are used to emulate certain physiological characteristics of a diabetic patient and allow the user to perform experiments by controlling different parameters related to insulin dosing strategies for diabetes patients. Diabetes simulators are often preferred in pre-clinical trials to evaluate the performance of new diabetes management systems/strategies. Some of the famous diabetes simulators include the UVA/PADOVA simulator and Hovorka model, etc. Of the works based on BG data reported in this review, 13.63% use simulated BG data from different diabetes patient simulators, while 68.18% use actual clinical data, whereas 18% of the manuscripts use both real and simulated patient data.

Moreover, of the works that use BG for hypoglycemia detection, some make use of sole BG data while others use BG plus a combination of different types of data such

as insulin, CHO, and PA data, etc. The distribution of different BG data combinations may be observed in Figure 2. About 36.36% of the manuscripts used BG data alone for hypoglycemia prediction; 6.82% of manuscripts predicted hypoglycemia by using BG combined with insulin data, and an equal amount of research work used BG combined with CHO. A total of 22.73% works used BG combined with insulin, CHO, and PA data. A total of 15.91% studies used BG, insulin, CHO, and PA data in conjunction while 11.36% studies made use of BG in combination with other sources of data, e.g., HR and ECG, etc.

3.2.2. Only BG

The details of the works that trained ML models only on BG data are given in Table 2. Most of these models are time series forecasting models and involve BG data that have certain timestamps associated with the actual BG values. Nocturnal hypoglycemia prediction was targeted in studies such as Kriukova et al. [44], Vu et al. [47], Sampath et al. [69], and Tkachenko et al. [49]. Seo et al. [43] proposed the prediction of postprandial hypoglycemia by training ML models with BG data while Jung et al. [54] predicted day-time hypoglycemia using similar data. Quan et al. [50], Dong et al. [62], and Mhaskar et al. [63] used neural networks trained on BG data for hypoglycemia prediction. On the other hand, Rodriguez et al. [57] used three different ML models trained on BG data from 25 patients. A KRR-based system was presented by Marcus et al. [58] while Seo et al. [46] proposed another model to predict hypoglycemia that used BG values.

3.2.3. BG Combined with Other Types of Data

From Figure 2 it is evident that of the manuscripts which use BG data for training, a portion use the combination of BG and insulin data. Jensen et al. [39] and Mosquera-Lopez et al. [48] proposed models that predicted nocturnal hypoglycemia from BG and insulin together. These systems showed moderate performance in terms of sensitivity and specificity. A recurrent neural network (RNN) was trained by Mosquera et al. [81] using BG and insulin data for adverse glycemic event prediction. This system was reported to be more than 90% accurate in predicting hypoglycemic events.

Some studies trained ML models on BG along with CHO values. The quantity of such studies is very low since the CHO data are often very inconsistent and not a lot of ML designers like to work with them. CHO, however, is an important feature to consider in insulin prediction models. Insulin bolus calculation was performed by Zhu et al. [65] and Giulia Noaro et al. [55]. Both of these works employed BG and CHO data. Zhu et al. [73] presented a system based on DRL. This study displayed an improved control of single hormone and dual hormone insulin delivery.

Other works such as Dave et al. [52], Shifrin et al. [56], and Cappon et al. [66] used BG along with insulin and CHO for prediction purposes. Aiello et al. [67] and Oviedo et al. [53] both aimed at postprandial hypoglycemia prediction by utilizing BG data combined with insulin and CHO data. Noaro et al. [72] proposed an insulin bolus calculator while Vehi et al. [59] proposed a hypoglycemia prediction and prevention system that employed BG, insulin, and CHO data for ML model training. A DSS that provides weekly insulin dosage recommendations for type1 diabetics was proposed by Tyler et al. [61].

There are certain works that along with BG, insulin and CHO made use of additional data e.g., physical activity (PA) data and HR etc. Nocturnal hypoglycemia was predicted from BG, insulin, CHO and PA data by Calhoun et al. [45], Bertachi et al. [51], Bertachi et al. [41] and Güemes et al. [60]. Glucose value forecasting is performed by Li et al. [78], Mayo et al. [71], Zhu et al. [64] and Daskalaki et al. [70] through the utilization of such a combination of data.

Certain works have also used BG data in combination with other data such as breath samples and camera samples, etc. Reddy et al. [40] predicted hypoglycemia at the start of an aerobic exercise. Hypoglycemia was predicted from breath samples using ML techniques by Siegel et al. [27]. Vahedi et al. [33] predicted BG levels from BG and PA while BG levels

were estimated from PPG signals using the mobile phone camera of a patient and the BG data in a study proposed by Zhang et al. [38].

3.2.4. Other Types of Data

We are aware that the beauty of ML lies in formalizing non-linear relationships between different data(s) and outcomes. Researchers have hence tried to predict hypoglycemia by training ML models using multiple types of other data. Of the works that we have reviewed, 22.81% are based on data other than BG as shown in Figure 2. These data include the EHR, ECG, GSR, EEG, clinical notes, secret messages, breath samples, and body temperature. The individual percentages of works based on these data are as follows: EHR 5%, ECG 5%, HR 5%, breath samples 3.5%, body temperature 3.5%, clinical notes 2%, secret messages 2%, GSR 2%, EEG, 2%.

3.3. Machine Learning Models

ML designers have a variety of ML algorithms at their disposal while implementing new designs. The choice of an ML algorithm is guided by multiple factors, i.e., the type of data used to train the model, the number of features, and most importantly, the quantity of data available [82]. The literature reviewed here shows that a total of 34 unique ML algorithms have been used as can be observed in Tables 1 and 2. These algorithms have been categorized into six major families of ML algorithms as shown in Figure 3. The most common of these families is the ANNs followed by the DTs, kernels, and others. If we talk about the most famous individual ML models, RF has been the choice of designers followed by SVM.

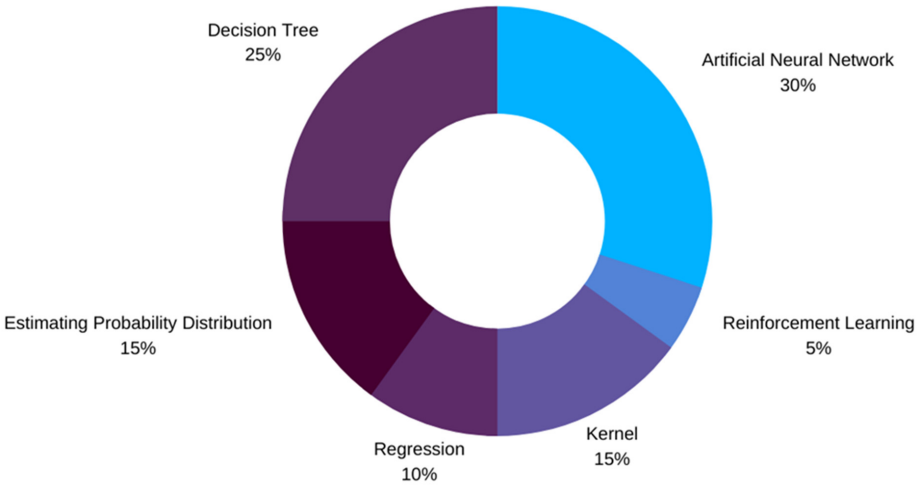


Figure 3. ML models used by studies, grouped into families based on similarity.

The majority of the studies have utilized the self-learning capabilities of ANN. These studies have employed multiple variants of ANN such as RNN, DL, CNN, MLPs, etc. ANNs were used by Bertachi et al. [41], Vahedi et al. [33], Zhu et al. [64], Mosquera-Lopez et al. [81], San et al. [35], Jin et al. [36], Mhaskar et al. [63], Li et al. [74], Li et al. [78], Bertachi et al. [51], Güemes et al. [60], Oviedo et al. [53], Vehi et al. [59], Quan et al. [50], and Amar et al. [75]. Unlike other ML models, ANNs extract their own features from the inputs based on their hidden parameters. ANNs were used together with a reinforcement learning algorithm in studies presented by Zhu, Li, Kuang, et al. [65], and Zhu, Li, Herrero, et al. [73].

DTs are predictive models that predict the outcome for a set of input features after testing the features through several tree branches. DTs too have multiple variants that were utilized in the studies cited in this review. The most famous variant of DT is RF. Because of

characteristics such as robustness to noise, handling of missing values and robustness to outliers, RF has been chosen by many ML designers for this application. RF creates a large number of decision trees and then outputs the mode of all the decision trees. This approach fixes the over-fitting problem of decision trees. Seo et al. [43], Güemes et al. [60], Vahedi et al. [33], G Noaro et al. [72], Vu et al. [47], Reddy et al. [40], Chen et al. [30], Dave et al. [52], Calhoun et al. [45], Amar et al. [75], Hidalgo et al. [77], and Rodriguez et al. [79] have all used RF for predicting/detecting hypoglycemia. Ruan et al. [31] and Cappon et al. [66] used the XGboost algorithm. XGboost is the gradient-boosted variant of DT and is aimed at enhancing the performance of decisions trees. Common DTs were employed by Ranvier et al. [34] and Reddy et al. [40].

Kernel-based SVM is the second most common choice of ML algorithms for designers working towards the goal of hypoglycemia detection/prediction. This is an indicator of the fact that SVM works well for such problems where the data-sets are relatively small. SVM is a binary linear classifier that maps feature points in space, creating different categories [83]. These categories are separated by a gap as wide as possible. When a test point is brought to the model, SVM maps it to one of the various categories and then assigns it a label. Marling et al. [32], Mosquera-Lopez et al. [48], Seo et al. [43], Güemes et al. [60], Oviedo et al. [68], Vehi et al. [59], Chen et al. [30], Bertachi et al. [41], and Rodriguez et al. [79] have all used SVM.

Regression techniques in ML predict the result of a continuous output variable. In the case of LR, however, the output is often a discrete label. The various types of regression used by studies in this review are LR, GE, and MLR. Studies proposed by Chen et al. [30], Dave et al. [52], and Seo et al. [43] use LR. LR fits a logistic function to data and outputs the probability of one or more classes. KNN is another ML approach that was used in several studies, such as Tyler et al. [61], Zhang et al. [36], Aiello et al. [67], Seo et al. [43], and Hidalgo et al. [77]. KNN looks for the closest examples in the feature space and then assigns them a label. Jensen et al. [39] and Siegel et al. [27] used LDA for prediction purposes. LDA is often used for the purpose of dimensionality reduction in classification problems.

It is important to consider that ML models can be evaluated with a range of different performance metrics. It is, therefore, impossible to present a quantitative performance comparison of the reviewed literature since the performance metrics differ for different works. Sensitivity and specificity have been the researchers most favorite performance metric with 47% of the studies using it, followed by root mean square error (RMSE), in 21% of the research works reviewed. Accuracy was used as a performance metric in a total of 13% of the manuscripts, similar to the area under the ROC curve (AUC). The mean absolute percentage error and blood glucose risk index were each used to evaluate 3.3% of the total manuscripts.

3.4. Prediction Horizon: How Far Are the Current Systems Forecast in the Future?

In ML analysis of time series data, PH or forecasting horizon is the amount of time the user has before the occurrence of a predicted event. In biomedical applications, ideally, the PH should be large enough to give the patient apt time to take preventive measures and prevent an adverse event from happening. In the case of hypoglycemia, if the ML algorithm predicts the occurrence of a possible hypoglycemic event 30 min from the time of prediction, the user only has 30 min to take necessary actions in order to prevent the predicted hypoglycemic event from happening. In the mentioned case, whether the PH of 30 min is enough time or not is a debate that is dependent on various factors such as the severity of the hypoglycemic event, effectiveness of medications, and the amount of CHO consumption. The PH defined in a particular approach has two important effects on the achieved predictions: the time a patient has to respond and the error associated with estimations increase together. Therefore, it is extremely important to find a balance between the error we are willing to take and the requirements of our approach. There are a total of 14 different PHs reported in the literature reviewed with the PH of 30 min being the most common, followed by PH values of 6 h, 60 min, and 15 min. The PH values

are categorized based on short-term, medium-term, and long-term predictions. The PH categories based on their frequency of usage are provided in Figure 4. The details about all the PHs reported in the literature are given in Table 3.

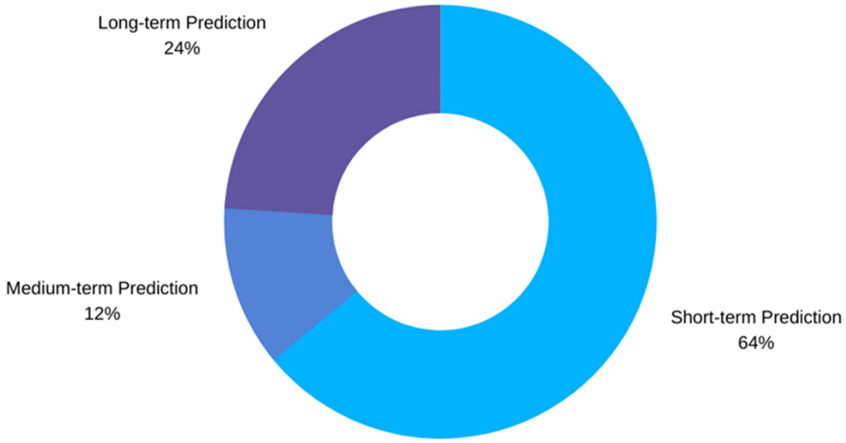


Figure 4. Different types of PH categories used in the studies.

Table 3. Prediction horizons used by different works.

Prediction Type	Manuscript	Prediction Horizon (PH)
Shor-Term Prediction	[52,54,57,79]	15 min
	[40,43,46,50–52,58,63,64,71,74–78,80,81]	30 min
	[39]	35 min
	[52,62]	45 min
	[49,51,52,59,74,75,78,80]	60 min
Medium-Term Prediction	[80]	90 min
	[60]	120 min
	[56]	210 min
	[49,53,68]	240 min
Long-Term Prediction	[41,42,47–49,69]	360 min
	[55]	420 min
	[44]	540 min
	[61]	1 Week
	[66,67,72]	Meal Duration

4. Discussion

After an organized analysis of high quality research work in PubMed and Google Scholar, we pinned down manuscripts with an aim of providing a thorough overview of the work done in the field of ML for predicting hypoglycemia. The review demonstrates that the use of ML models for the prediction of hypoglycemia has increased considerably over the last five years. It was observed that not all the manuscripts reviewed focused on predicting hypoglycemia. Some of the works only focused on the detection of the glycemic event. Technically, detection of an event could be referred to as description. It is important to understand that hypoglycemia prediction is BG level prediction in essence. Hypoglycemia is but a condition labelled on the predicted BG value graph. That is precisely

the reason why some of the works that we have reviewed are BG level prediction models and do not talk about hypoglycemia prediction explicitly. These works, however, do provide a framework for the prediction of hypoglycemia.

A correlation between the type of data used for training the ML model and the nature of output (description/prediction) suggests that ML time series prediction is only possible with time series data such as BG, insulin, and CHO, while detection could be performed using other types of data such as breath samples, EEG, PPG, etc. Data acquisition in biomedical applications suffers from multiple constraints such as hardware limitations, restricted clinical environments, failure of patients to comply with study protocols, and obstacles in the way of large biomedical data collection. These barriers compel ML designers to work with the available imperfect data and look for solutions. The issue of imperfect data may be tackled through different strategies. The missing values problem is often addressed by using some kind of interpolation or imputation method. Prediction of missing values based on other values is also a technique that has been used to address this problem. Different types of regression or classification models can be used to predict missing values. Deep learning-based imputation is often preferred because of its accuracy. In various studies a certain range of missing values is selected to perform interpolation. Any gap in the data that exceeds that limit of missing values is then termed missing data and no interpolation is done. Frameworks based on conditional probability such as the theory of belief function, evidence theory or linear belief functions can be used to address the problem of incomplete data or missing data.

In this review, the assessment of data used in training exhibits a slanted picture with BG data dominating most of the reviewed studies. It has been observed that studies that use data other than BG are almost all targeted at detecting hypoglycemia. There are claims by some works of predicting hypoglycemia while using data other than BG, but a thorough inspection revealed that the targeted PH was either too small to be classified as real prediction or it does not exist at all. It is also worth mentioning here that many works reported an issue with the acquired data in terms of size or completeness. The need for data that are both large in size and good in quality is ultimate. It is known that ML models map complex nonlinear relationships in physiological data to perform prediction or description. To perform this nonlinear mapping, the required data have to be complete and relevant. In particular, BG value detection and prediction feed on data obtained from various types of sensors, i.e., CGM and HR sensors, etc. Two of the most common issues with CGM sensors is the sensor delay and sensor malfunctions. Sensor delay in CGM is the inherited 10-min discrepancies, while sensor malfunctions are those periods in which no BG value is recorded. The quality of these sensors is one area that needs to be improved in future.

An in-depth analysis of ML models presented a broader picture of the preferred techniques for the purpose of hypoglycemia prediction. It is understood that the quality and quantity of data affect the choice of the ML model. Since the data in this case suffer from various issues, the choice of ML model should be made such that it makes up for the deficiencies in the data. Models such as SVM are preferred because of their ability to handle a relatively small amount of data with greater efficiency. ML models are also chosen based on the level of complexity. Simpler models such as RF and KNN are preferred because they give good results most of the time and are easier to implement. Moreover, the reason that the majority of the works use RF and SVM is that these algorithms provide a higher level of versatility in terms of the type of problem they are used for. On the other hand, ANNs are data hungry and in the case of hypoglycemia prediction it is observed that efforts are made to train the network on large datasets. DL is an area that has not been used extensively for hypoglycemia prediction and can be explored more in future. Gradient-boosted tree algorithms such as XGB are nowadays preferred by designers for time series analysis if the time series problem is a supervised learning problem. The use of multiple gradient-boosted algorithms can be observed in the review. The quest for improved results is analogous to

training different ML models. Designers, therefore, have trained various other models in the work reviewed.

It is understood from the review that, for a hypoglycemia DSS, the most important trait to have is to warn the patient about a hypoglycemic episode well before it happens. Early detection helps the patient cope with the hypoglycemic event in a better way. PH is a term closely associated with time series data forecasting. How far does an ML-based prediction model see into the future is a thing to consider before passing any verdict about the quality of the model. The length of a PH correlates with the amount of data an ML model is trained on. Ranging from a PH length of 15 min to 1 week, designers have tried to predict hypoglycemia in different future time frames. The race for a longer PH is always on but the desirable choice of PH mostly depends on the nature and type of application.

5. Conclusions

ML for the prediction of hypoglycemia has been trending topic among biomedical data engineers. Our review demonstrates the potential impact such predictive models could have in the field of diabetes healthcare. A highly efficient hypoglycemia predictor may prove life-changing for T1Ds. The timely prediction of a hypoglycemic episode can immensely improve the life quality of T1D patients and on top of that, save their lives. There has been a stark increase in the amount of research work done in the area of hypoglycemia prediction using ML. This is evident from the increasing number of studies published in this domain during the past five years as depicted by this review. Though ML models appear to be the right choice for figuring out the nonlinear relationships between different types of physiological data and the occurrence of hypoglycemia, there is still room for improvement. This review gives an insight into the challenges faced by the designers while dealing with imperfect data for hypoglycemia prediction and detection. The results obtained from this review provide an overview of the go-to ML models for researchers while predicting/detecting hypoglycemia. Discussion of the PH portrays a picture regarding how far the current systems predict hypoglycemia in the future. It is concluded that ML for hypoglycemia prediction holds considerable potential. Research in this domain must continue and more directions should be explored. Researchers are advised to further explore this domain by training different ML models on various types of sensor data. In the context of hypoglycemia prediction, it is paramount to come up with new strategies to train ML models with more data. ML engineers could use a two-phase training approach by first training the ML models with a huge amount of data (populational models, similar patients, virtual patients, generated data, etc.) and then training the ML models with more specific data such as cohort data, real time data, etc. Creativity in feature engineering and techniques for the acquisition of healthy datasets are areas that need to be worked on for the realization of accurate ML-based hypoglycemia predictors. The incorporation of such ML models in DSS should be ensured and made available for the benefit of patients.

Author Contributions: Conceptualization, O.M., I.C., and J.V.; methodology, O.M., I.C., and J.V.; investigation, O.M., I.C., and J.V.; writing—Original draft preparation, O.M.; writing—Review and editing, O.M., I.C., and J.V.; supervision, I.C. and J.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially funded by the Spanish Government (PID2019-107722RB-C22) and the Government of Catalonia under 2017SGR1551 and 2020 FI_B 00965.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alsahli, M.; Gerich, J.E. Hypoglycemia. *Endocrinol. Metab. Clin. North Am.* **2013**, *42*, 657–676. [CrossRef] [PubMed]
2. Kittah, N.E.; Vella, A. Management of endocrine disease: Pathogenesis and management of hypoglycemia. *Eur. J. Endocrinol.* **2017**, *177*, R37–R47. [CrossRef] [PubMed]
3. Yale, J.-F.; Paty, B.; Senior, P.A. Hypoglycemia. *Can. J. Diabetes* **2018**, *42*, S104–S108. [CrossRef] [PubMed]
4. International Diabetes Federation. *IDF Diabetes Atlas*, 9th ed.; International Diabetes Federation: Brussels, Belgium, 2019.
5. Cryer, P.E. Symptoms of hypoglycemia, thresholds for their occurrence, and hypoglycemia unawareness. *Endocrinol. Metab. Clin. North Am.* **1999**, *28*, 495–500. [CrossRef]
6. Szadkowska, A.; Czyżewska, K.; Pietrzak, I.; Mianowska, B.; Jarosz-Chobot, P.; Myśliwiec, M. Hypoglycaemia unawareness in patients with type 1 diabetes. *Pediatr. Endocrinol. Diabetes Metab.* **2018**, *2018*, 126–134. [CrossRef]
7. Anderbro, T.; Gonder-Frederick, L.; Bolinder, J.; Lins, P.-E.; Wredling, R.; Moberg, E.; Lisspers, J.; Johansson, U.-B. Fear of hypoglycemia: Relationship to hypoglycemic risk and psychological factors. *Acta Diabetol.* **2015**, *52*, 581–589. [CrossRef]
8. Gumprecht, J.; Nabrdalik, K. Hypoglycemia in patients with insulin-treated diabetes. *Pol. Arch. Med. Wewn.* **2016**, *126*, 870–878. [CrossRef]
9. Cryer, P.E.; Arbeláez, A.M. Hypoglycemia in diabetes. In *Textbook of Diabetes*; Wiley Online Library: Hoboken, NJ, USA, 2017; pp. 513–533.
10. Graveling, A.J.; Frier, B.M. The risks of nocturnal hypoglycaemia in insulin-treated diabetes. *Diabetes Res. Clin. Pract.* **2017**, *133*, 30–39. [CrossRef]
11. Galati, S.-J.; Rayfield, E.J. Approach to the patient with postprandial hypoglycemia. *Endocr. Pract.* **2014**, *20*, 331–340. [CrossRef]
12. Resalat, N.; El Youssef, J.; Reddy, R.; Castle, J.; Jacobs, P.G. Adaptive tuning of basal and bolus insulin to reduce postprandial hypoglycemia in a hybrid artificial pancreas. *J. Process Control* **2019**, *80*, 247–254. [CrossRef]
13. Miller, K.M.; Foster, N.C.; Beck, R.W.; Bergenstal, R.M.; DuBose, S.N.; DiMeglio, L.A.; Maahs, D.M.; Tamborlane, W.V. Current state of type 1 diabetes treatment in the U.S.: Updated data from the T1D Exchange clinic registry. *Diabetes Care* **2015**, *38*, 971–978. [CrossRef] [PubMed]
14. Shende, P.; Sahu, P.; Gaud, R. A technology roadmap of smart biosensors from conventional glucose monitoring systems. *Ther. Deliv.* **2017**, *8*, 411–423. [CrossRef] [PubMed]
15. Chen, C.; Zhao, X.-L.; Li, Z.-H.; Zhu, Z.-G.; Qian, S.-H.; Flewitt, A.J. Current and Emerging Technology for Continuous Glucose Monitoring. *Sensors* **2017**, *17*, 182. [CrossRef] [PubMed]
16. Cappon, G.; Vettoretti, M.; Sparacino, G.; Facchinetti, A. Continuous Glucose Monitoring Sensors for Diabetes Management: A Review of Technologies and Applications. *Diabetes Metab. J.* **2019**, *43*, 383–397. [CrossRef] [PubMed]
17. Lucidi, P.; Porcellati, F.; Bolli, G.B.; Fanelli, C.G. Prevention and Management of Severe Hypoglycemia and Hypoglycemia Unawareness: Incorporating Sensor Technology. *Curr. Diabetes Rep.* **2018**, *18*, 83. [CrossRef] [PubMed]
18. Weisman, A.; Bai, J.-W.; Cardinez, M.; Kramer, C.K.; Perkins, B.A. Effect of artificial pancreas systems on glycaemic control in patients with type 1 diabetes: A systematic review and meta-analysis of outpatient randomised controlled trials. *Lancet Diabetes Endocrinol.* **2017**, *5*, 501–512. [CrossRef]
19. Wiens, J.; Shenoy, E.S. Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. *Clin. Infect. Dis.* **2018**, *66*, 149–153. [CrossRef]
20. Roth, J.A.; Battegay, M.; Juchler, F.; Vogt, J.E.; Widmer, A.F. Introduction to Machine Learning in Digital Healthcare Epidemiology. *Infect. Control Hosp. Epidemiol.* **2018**, *39*, 1457–1462. [CrossRef]
21. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930. [CrossRef]
22. Contreras, I.; Vehi, J. Artificial intelligence for diabetes management and decision support: Literature review. *J. Med. Internet Res.* **2018**, *20*, e10775. [CrossRef]
23. Ngiam, K.Y.; Khor, I.W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* **2019**, *20*, e262–e273. [CrossRef]
24. Marks, V.; Teale, J.D. Drug-induced hypoglycemia. *Endocrinol. Metab. Clin.* **1999**, *28*, 555–577. [CrossRef]
25. Freeland, B. Hypoglycemia in Diabetes Mellitus. *Home Healthc. Now* **2017**, *35*, 414–419. [CrossRef] [PubMed]
26. Zhou, L.; Siddiqui, T.; Seliger, S.L.; Blumenthal, J.B.; Kang, Y.; Doerfler, R.; Fink, J.C. Text preprocessing for improving hypoglycemia detection from clinical notes—A case study of patients with diabetes. *Int. J. Med. Inform.* **2019**, *129*, 374–380. [CrossRef] [PubMed]
27. Siegel, A.P.; Daneshkhah, A.; Hardin, D.S.; Shrestha, S.; Varahramyan, K.; Agarwal, M. Analyzing breath samples of hypoglycemic events in type 1 diabetes patients: Towards developing an alternative to diabetes alert dogs. *J. Breath Res.* **2017**, *11*, 026007. [CrossRef]
28. Tronstad, C.; Elvebakk, O.; Staal, O.M.; Kalvøy, H.; Høgetveit, J.O.; Jenssen, T.G.; Birkeland, K.I.; Martinsen, Ø.G. Non-invasive prediction of blood glucose trends during hypoglycemia. *Anal. Chim. Acta* **2019**, *1052*, 37–48. [CrossRef]
29. Jin, Y.; Li, F.; Yu, H. HYPE: A High Performing NLP System for Automatically Detecting Hypoglycemia Events from Electronic Health Record Notes. *arXiv* **2018**, arXiv:181111945.
30. Chen, J.; Lalor, J.; Liu, W.; Druhl, E.; Granillo, E.; Vimalananda, V.G.; Yu, H. Detecting Hypoglycemia Incidents Reported in Patients' Secure Messages: Using Cost-Sensitive Learning and Oversampling to Reduce Data Imbalance. *J. Med. Internet Res.* **2019**, *21*, e11990. [CrossRef]

31. Ruan, Y.; Bellot, A.; Moysova, Z.; Tan, G.D.; Lumb, A.; Davies, J.; Van Der Schaar, M.; Rea, R. Predicting the Risk of Inpatient Hypoglycemia With Machine Learning Using Electronic Health Records. *Diabetes Care* **2020**. [CrossRef]
32. Marling, C.; Xia, L.; Bunescu, R.; Schwartz, F. Machine learning experiments with noninvasive sensors for hypoglycemia detection. In Proceedings of the IJCAI Workshop on Knowledge Discovery in Healthcare Data, New York, NY, USA, 10 July 2016; Morgan Kaufmann Publishers Inc.: New York, NY, USA, 2016; pp. 1–6.
33. Vahedi, M.R.; MacBride, K.B.; Wunsik, W.; Kim, Y.; Fong, C.; Padilla, A.J.; Pourhomayoun, M.; Zhong, A.; Kulkarni, S.; Arunachalam, S.; et al. Predicting Glucose Levels in Patients with Type1 Diabetes Based on Physiological and Activity Data. In Proceedings of the 8th ACM MobiHoc 2018 Workshop on Pervasive Wireless Healthcare Workshop, Los Angeles, LA, USA, 25 June 2018. [CrossRef]
34. Ranvier, J.E.; Dubosson, F.; Calbimonte, J.P.; Aberer, K. Detection of hypoglycemic events through wearable sensors. In Proceedings of the International Workshop on Semantic Web Technologies for Mobile and PErvasive Environments 2016 (No. CONF), Heraklion, Germany, 29 May–2 June 2016; CEUR-WS: Heraklion, Germany, 2016.
35. San, P.P.; Ling, S.H.; Nguyen, H.T. Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3503–3506. [CrossRef]
36. Jin, Y.; Li, F.; Vimalananda, V.G.; Yu, H. Automatic Detection of Hypoglycemic Events from the Electronic Health Record Notes of Diabetes Patients: Empirical Study. *JMIR Med. Inform.* **2019**, *7*, e14340. [CrossRef]
37. Juhl, C.B.; Duun-Henriksen, J.; Sørensen, J.A.; Sejling, A.S.; Elsborg Madsen, R. Prevention of Severe Hypoglycemia by Continuous EEG Monitoring. In *Prediction Methods for Blood Glucose Concentration*; Kirchsteiger, H., Jørgensen, J.B., Renard, E., del Re, L., Eds.; Springer International Publishing: Cham, Germany, 2016; pp. 79–92. [CrossRef]
38. Zhang, Y.; Zhang, Y.; Siddiqui, S.A.; Kos, A. Non-invasive blood-glucose estimation using smartphone PPG signals and subspace kNN classifier. *Elektrotehniski Vestnik* **2019**, *86*, 68–74.
39. Jensen, M.H.; Dethlefsen, C.; Vestergaard, P.; Hejlesen, O. Prediction of nocturnal hypoglycemia from continuous glucose monitoring data in people with type 1 diabetes: A proof-of-concept study. *J. Diabetes Sci. Technol.* **2020**, *14*, 250–256. [CrossRef] [PubMed]
40. Reddy, R.; Resalat, N.; Wilson, L.M.; Castle, J.R.; El Youssef, J.; Jacobs, P.G. Prediction of hypoglycemia during aerobic exercise in adults with type 1 diabetes. *J. Diabetes Sci. Technol.* **2019**, *13*, 919–927. [CrossRef]
41. Bertachi, A.; Viñals, C.; Biagi, L.; Contreras, I.; Vehi, J.; Conget, I.; Giménez, M. Prediction of Nocturnal Hypoglycemia in Adults with Type 1 Diabetes under Multiple Daily Injections Using Continuous Glucose Monitoring and Physical Activity Monitor. *Sensors* **2020**, *20*, 1705. [CrossRef] [PubMed]
42. Ling, S.H.; San, P.P.; Nguyen, H.T. Non-invasive hypoglycemia monitoring system using extreme learning machine for Type 1 diabetes. *ISA Trans.* **2016**, *64*, 440–446. [CrossRef] [PubMed]
43. Seo, W.; Lee, Y.-B.; Lee, S.; Jin, S.-M.; Park, S.-M. A machine-learning approach to predict postprandial hypoglycemia. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 210. [CrossRef] [PubMed]
44. Kriukova, G.; Shvai, N.; Pereverzyev, S.V. Application of regularized ranking and collaborative filtering in predictive alarm algorithm for nocturnal hypoglycemia prevention. In Proceedings of the 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Bucharest, Romania, 21–23 September 2017; pp. 634–638.
45. Calhoun, P.; Levine, R.A.; Fan, J. Repeated measures random forests (RMRF): Identifying factors associated with nocturnal hypoglycemia. *Biometrics* **2020**. [CrossRef] [PubMed]
46. Seo, W.; Lee, J.; Lee, S.; Park, S.-M. An ensemble approach for accurately predicting hypoglycemia. In Proceedings of the 2019 Frontiers Medical Devices, Washington, DC, USA, 19–21 March 2019.
47. Vu, L.; Kefayati, S.; Idé, T.; Pavuluri, V.; Jackson, G.; Latts, L.; Zhong, Y.; Agrawal, P.; Chang, Y. Predicting Nocturnal Hypoglycemia from Continuous Glucose Monitoring Data with Extended Prediction Horizon. In *AMIA Annual Symposium Proceedings*; American Medical Informatics Association: Bethesda, MD, USA, 2019; p. 874.
48. Mosquera-Lopez, C.; Dodier, R.; Tyler, N.S.; Wilson, L.M.; El Youssef, J.; Castle, J.R.; Jacobs, P.G. Predicting and preventing nocturnal hypoglycemia in type 1 diabetes using big data analytics and decision theoretic analysis. *Diabetes Technol. Ther.* **2020**. [CrossRef]
49. Tkachenko, P.; Kriukova, G.; Aleksandrova, M.; Chertov, O.; Renard, E.; Pereverzyev, S.V. Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: Proof of concept for clinical application. *Comput. Methods Programs Biomed.* **2016**, *134*, 179–186. [CrossRef]
50. Quan, T.M.; Doike, T.; Bui, D.C.; Arata, S.; Kobayashi, A.; Islam, M.Z.; Niitsu, K. AI-based edge-intelligent hypoglycemia prediction system using alternate learning and inference method for blood glucose level data with low-periodicity. In Proceedings of the 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS), Taiwan, China, 18–20 March 2019; pp. 201–206.
51. Bertachi, A.; Biagi, L.; Contreras, I.; Luo, N.; Vehi, J. Prediction of Blood Glucose Levels and Nocturnal Hypoglycemia Using Physiological Models and Artificial Neural Networks. In Proceedings of the KHD@ IJCAI 2018, Stockholm, Schweden, 13 July 2018; pp. 85–90.

52. Dave, D.; DeSalvo, D.J.; Haridas, B.; McKay, S.; Shenoy, A.; Koh, C.J.; Lawley, M.; Erraguntla, M. Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction. *J. Diabetes Sci. Technol.* **2020**. [CrossRef]
53. Oviedo, S.; Contreras, I.; Bertachi, A.; Quirós, C.; Giménez, M.; Conget, I.; Vehi, J. Minimizing postprandial hypoglycemia in Type 1 diabetes patients using multiple insulin injections and capillary blood glucose self-monitoring with machine learning techniques. *Comput. Methods Programs Biomed.* **2019**, *178*, 175–180. [CrossRef] [PubMed]
54. Jung, M.; Lee, Y.-B.; Jin, S.-M.; Park, S.-M. Prediction of Daytime Hypoglycemic Events Using Continuous Glucose Monitoring Data and Classification Technique. *arXiv* **2017**, arXiv:170408769.
55. Noaro, G.; Cappon, G.; Vettoretti, M.; Sparacino, G.; Del Favero, S.; Facchinetti, A. Machine-learning based model to improve insulin bolus calculation in type 1 diabetes therapy. *IEEE Trans. Biomed. Eng.* **2020**, *68*, 247–255. [CrossRef] [PubMed]
56. Shifrin, M.; Siegelmann, H. Near-optimal insulin treatment for diabetes patients: A machine learning approach. *Artif. Intell. Med.* **2020**, *107*, 101917. [CrossRef] [PubMed]
57. Rodriguez-Rodriguez, I.; Chatzigiannakis, I.; Rodriguez, J.-V.; Maranghi, M.; Gentili, M.; Zamora-Izquierdo, M.-Á. Utility of Big Data in Predicting Short-Term Blood Glucose Levels in Type 1 Diabetes Mellitus Through Machine Learning Techniques. *Sensors* **2019**, *19*, 4482. [CrossRef]
58. Marcus, Y.; Eldor, R.; Yaron, M.; Shaklai, S.; Ish-Shalom, M.; Shefer, G.; Stern, N.; Golan, N.; Dvir, A.Z.; Pele, O.; et al. Improving blood glucose level predictability using machine learning. *Diabetes Metab. Res. Rev.* **2020**, *36*, e3348. [CrossRef]
59. Vehi, J.; Contreras, I.; Oviedo, S.; Biagi, L.; Bertachi, A. Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning. *Health Inform. J.* **2020**, *26*, 703–718. [CrossRef]
60. Güemes, A.; Cappon, G.; Hernandez, B.; Reddy, M.; Oliver, N.; Georgiou, P.; Herrero, P. Predicting Quality of Overnight Glycaemic Control in Type 1 Diabetes using Binary Classifiers. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 1439–1446. [CrossRef]
61. Tyler, N.S.; Mosquera-Lopez, C.M.; Wilson, L.M.; Dodier, R.H.; Branigan, D.L.; Gabo, V.B.; Guillot, F.H.; Hilts, W.W.; El Youssef, J.; Castle, J.R.; et al. An artificial intelligence decision support system for the management of type 1 diabetes. *Nat. Metab.* **2020**, *2*, 612–619. [CrossRef]
62. Dong, Y.; Wen, R.; Zhang, K.; Zhang, L. A Novel RNN-Based Blood Glucose Prediction Approach Using Population and Individual Characteristics. In Proceedings of the 2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB), Hangzhou, China, 21–23 March 2019; pp. 145–149.
63. Mhaskar, H.N.; Pereverzyev, S.V.; van der Walt, M.D. A deep learning approach to diabetic blood glucose prediction. *Front. Appl. Math. Stat.* **2017**, *3*, 14. [CrossRef]
64. Zhu, T.; Li, K.; Chen, J.; Herrero, P.; Georgiou, P. Dilated Recurrent Neural Networks for Glucose Forecasting in Type 1 Diabetes. *J. Healthc. Inform. Res.* **2020**, *4*, 1–17. [CrossRef]
65. Zhu, T.; Li, K.; Kuang, L.; Herrero, P.; Georgiou, P. An Insulin Bolus Advisor for Type 1 Diabetes Using Deep Reinforcement Learning. *Sensors* **2020**, *20*, 5058. [CrossRef] [PubMed]
66. Cappon, G.; Facchinetti, A.; Sparacino, G.; Georgiou, P.; Herrero, P. Classification of postprandial glycemic status with application to insulin dosing in type 1 diabetes—An in silico proof-of-concept. *Sensors* **2019**, *19*, 3168. [CrossRef] [PubMed]
67. Aiello, E.M.; Toffanin, C.; Messori, M.; Cobelli, C.; Magni, L. Postprandial glucose regulation via KNN meal classification in type 1 diabetes. *IEEE Control Syst. Lett.* **2018**, *3*, 230–235. [CrossRef]
68. Oviedo, S.; Contreras, I.; Quirós, C.; Giménez, M.; Conget, I.; Vehi, J. Risk-based postprandial hypoglycemia forecasting using supervised learning. *Int. J. Med. Inform.* **2019**, *126*, 1–8. [CrossRef]
69. Sampath, S.; Tkachenko, P.; Renard, E.; Pereverzev, S.V. Glycemic control indices and their aggregation in the prediction of nocturnal hypoglycemia from intermittent blood glucose measurements. *J. Diabetes Sci. Technol.* **2016**, *10*, 1245–1250. [CrossRef]
70. Daskalaki, E.; Diem, P.; Mougiakakou, S.G. Model-free machine learning in biomedicine: Feasibility study in type 1 diabetes. *PLoS ONE* **2016**, *11*, e0158722. [CrossRef]
71. Mayo, M.; Chepulis, L.; Paul, R.G. Glycemic-aware metrics and oversampling techniques for predicting blood glucose levels using machine learning. *PLoS ONE* **2019**, *14*, e0225613. [CrossRef]
72. Noaro, G.; Cappon, G.; Sparacino, G.; Del Favero, S.; Facchinetti, A. Nonlinear Machine Learning Models for Insulin Bolus Estimation in Type 1 Diabetes Therapy. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 5502–5505.
73. Zhu, T.; Li, K.; Herrero, P.; Georgiou, P. Basal Glucose Control in Type 1 Diabetes using Deep Reinforcement Learning: An In Silico Validation. *arXiv* **2020**, arXiv:200509059. [CrossRef]
74. Li, K.; Liu, C.; Zhu, T.; Herrero, P.; Georgiou, P. GluNet: A deep learning framework for accurate glucose forecasting. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 414–423. [CrossRef]
75. Amar, Y.; Shilo, S.; Oron, T.; Amar, E.; Phillip, M.; Segal, E. Clinically accurate prediction of glucose levels in patients with type 1 diabetes. *Diabetes Technol. Ther.* **2020**. [CrossRef] [PubMed]
76. Pérez-Gandia, C.; Garcia-Sáez, G.; Subias, D.; Rodriguez-Herrero, A.; Gómez, E.J.; Rigla, M.; Hernando, M.E. 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*, 243–250. [CrossRef] [PubMed]
77. Hidalgo, J.I.; Colmenar, J.M.; Kronberger, G.; Winkler, S.M.; Garnica, O.; Lanchares, J. Data based prediction of blood glucose concentrations using evolutionary methods. *J. Med. Syst.* **2017**, *41*, 142. [CrossRef] [PubMed]

78. Li, K.; Daniels, J.; Liu, C.; Herrero, P.; Georgiou, P. Convolutional recurrent neural networks for glucose prediction. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 603–613. [CrossRef]
79. Rodriguez-Rodriguez, I.; Rodriguez, J.-V.; Chatzigiannakis, I.; Zamora Izquierdo, M.Á. On the Possibility of Predicting Glycaemia 'On the Fly'with Constrained IoT Devices in Type 1 Diabetes Mellitus Patients. *Sensors* **2019**, *19*, 4538. [CrossRef]
80. Fiorini, S.; Martini, C.; Malpassi, D.; Cordera, R.; Maggi, D.; Verri, A.; Barla, A. Data-driven strategies for robust forecast of continuous glucose monitoring time-series. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Korea, 11–15 July 2017; pp. 1680–1683.
81. Mosquera-Lopez, C.; Dodier, R.; Tyler, N.; Resalat, N.; Jacobs, P. Leveraging a big dataset to develop a recurrent neural network to predict adverse glycemic events in type 1 diabetes. *IEEE J. Biomed. Health Inform.* **2019**. [CrossRef]
82. Emmert-Streib, F.; Dehmer, M. Evaluation of regression models: Model assessment, model selection and generalization error. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 521–551. [CrossRef]
83. Kumar, B.; Vyas, O.P.; Vyas, R. A comprehensive review on the variants of support vector machines. *Mod. Phys. Lett. B* **2019**, *33*, 1950303. [CrossRef]

2.2 Conditional synthesis of blood glucose profiles for T1D patients using deep generative models

This publication is focused on the generation of realistic BG profiles of T1D patients using a CGAN. The generated BG values are conditioned on the plasma insulin values of the T1D patient. The generated BG profiles are then statistically compared to the real profiles.

Title: Conditional synthesis of blood glucose profiles for T1D patients using deep generative models

Authors: **Omer Mujahid**, Ivan Contreras, Aleix Beneyto, Ignacio Conget, Marga Gimenez and Josep Vehí

Journal: MDPI Mathematics

Volume: 10, Number: 3741, Published: October 2022

DOI: https://doi.org/10.3390/math10203741

Quality index: JCR 2021 Mathematics, Impact factor: 2.592, Q1 (21/333)

mathematics

MDPI

Article

Conditional Synthesis of Blood Glucose Profiles for T1D Patients Using Deep Generative Models

Omer Mujahid [1], Ivan Contreras [1,†], Aleix Beneyto [1], Ignacio Conget [2,3,4], Marga Giménez [2,3,4] and Josep Vehi [1,2,*]

[1] Modeling, Identification and Control Laboratory, Institut d'Informatica i Applicacions, Universitat de Girona, 17003 Girona, Spain
[2] Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), 17003 Girona, Spain
[3] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), 08023 Barcelona, Spain
[4] Diabetes Unit, Endocrinology and Nutrition Department, Hospital Clínic, 08036 Barcelona, Spain
* Correspondence: josep.vehi@udg.edu; Tel.: +34-620131826
† Professor Serra Húnter.

Abstract: Mathematical modeling of the glucose–insulin system forms the core of simulators in the field of glucose metabolism. The complexity of human biological systems makes it a challenging task for the physiological models to encompass the entirety of such systems. Even though modern diabetes simulators perform a respectable task of simulating the glucose–insulin action, they are unable to estimate various phenomena affecting the glycemic profile of an individual such as glycemic disturbances and patient behavior. This research work presents a potential solution to this problem by proposing a method for the generation of blood glucose values conditioned on plasma insulin approximation of type 1 diabetes patients using a pixel-to-pixel generative adversarial network. Two type-1 diabetes cohorts comprising 29 and 6 patients, respectively, are used to train the generative model. This study shows that the generated blood glucose values are statistically similar to the real blood glucose values, mimicking the time-in-range results for each of the standard blood glucose ranges in type 1 diabetes management and obtaining similar means and variability outcomes. Furthermore, the causal relationship between the plasma insulin values and the generated blood glucose conforms to the same relationship observed in real patients. These results herald the aptness of deep generative models for the generation of virtual patients with diabetes.

Keywords: deep generative models; conditional data synthesis; type 1 diabetes simulators; blood glucose data generation

MSC: 68T07

Citation: Mujahid, O.; Contreras, I.; Beneyto, A.; Conget, I.; Giménez, M.; Vehi, J. Conditional Synthesis of Blood Glucose Profiles for T1D Patients Using Deep Generative Models. *Mathematics* **2022**, *10*, 3741. https://doi.org/math10203741

Academic Editors: Costin Badica, Nick Bassiliades, Kalliopi Kravari and Theodoros Kosmanis

Received: 15 September 2022
Accepted: 8 October 2022
Published: 12 October 2022

1. Introduction

Biomedical simulation is at the heart of modern medical research. Any biomedical phenomena that can be simulated has the liberty to be subjected to the hit and trial approach of the scientific method, which is impossible to apply on living beings [1]. Similarly, type 1 diabetes (T1D) simulators have been the reason for many groundbreaking discoveries in diabetes research [2,3]. However, the conventional biomedical simulators are based on physiological mathematical models that do not cover the entirety of human physiology. Hence, the scenarios simulated by these simulators are far from perfect [4]. Keeping in view the challenges in the path of a model that should describe the complete physiology of the human body, the realization of such a digital twin seems very unlikely in the near future [5]. Even though the advancement in computing technology and power, the availability of large data storage devices, low-cost smart devices, and the ease of data acquisition provides us

with a glimmer of hope, there is still a long way to go in the path of achieving a perfect physiological description of the human body [6].

Current diabetes simulators suffer from the same fate. They try to mimic the mechanism of a human pancreas using mathematical modeling and control algorithms [7]. Most of the current diabetes simulators employ a mathematical model of the glucose–insulin system. A typical glucose–insulin model contains sub compartments such as the glucose model, the insulin delivery model, the gastro-intestinal tract, muscle and adipose tissue, and the liver model. However, it is important to note that in a real human being the BG profile is affected by many more biological systems and processes than just these. Sometimes, the real biological systems are so complex that it is impossible to approximate them mathematically. Additionally, there are certain hidden biological phenomena occurring that affect the BG profile of a human being and hence cannot be approximated at all. A generalization over all those systems that are not modeled or cannot be modelled in a T1D simulator leads to an intrinsic error that is evident in the simulated scenario [8,9].

The majority of the existing diabetes simulators also lack support for scenarios such as exercise (most types of it), illness, menstrual cycles, sleep disorders, depression, etc. Having said that, there have been efforts from scientists/engineers to incorporate disturbances such as physical activities and patient behaviors in the existing simulators [10,11]. It is understood that patient behavior affects the glycemic profile of a diabetic patient immensely. Lifestyle choices and routine habits may cause the glycemic profile of a diabetic to go from good to worse and vice versa [12]. Eating habits, alcohol consumption, missed meals, and non-compliance with the doctor's instructions may prove decisive in achieving control of diabetes for a patient [13]. An ideal diabetes simulator should provide support for the simulation of such scenarios in order to capture the complete essence of the effects of diabetes on a person.

Regardless of the modeling issues, the conventional T1D simulators are still an effective way to simulate diabetes scenarios for in silico experiments. However, there is great room for improvement in these simulators and, consequently, the possibility of getting closer to a real patient with T1D. An ideal biomedical simulation task demands the simulator produce phenomena that fulfills certain performance criteria just like the real human body would. In addition, realistic simulators must have mechanisms that include a certain level of randomness to mimic the biological stochasticity of real-life phenomena. Though it is understood that the simulated scenarios of a T1D simulator must be reusable, understandable, and reproducible, a T1D simulator in no way should be completely deterministic. All this points to the need of such a T1D simulator that encompasses the true and whole physiology of the human glucose–insulin model, is capable of simulating customized scenarios, and produces unique outcomes. Figure 1 shows two ways of approximating the biological systems/processes of the human body in order to be incorporated into a simulator. It can be seen from the figure that the framing of a biological system could be done either by physiological modeling or by learning the distribution of data these biological systems produce. It could be understood intuitively that the compound effect of all the disturbances on the glucose–insulin system of a person is reflected in the data acquired from the system. When investigated for its probability distribution, this data contains the collective features of all the processes and disturbances that influence the glycemic profile of a person. The idea is to learn this underlying probability distribution of the data and then generate novel samples of data using the learned distribution.

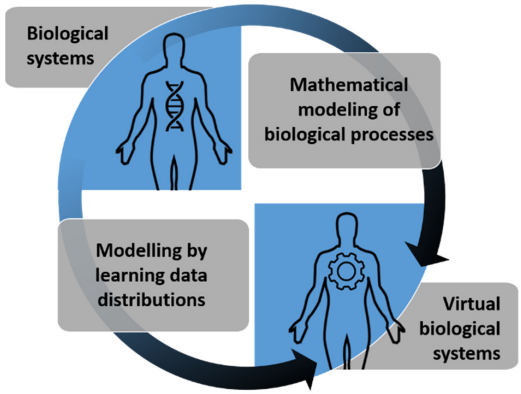


Figure 1. Two approaches to synthetic human biological systems.

Data-driven generic function approximators such as deep neural networks (DNNs) have the ability to learn from data the underlying behaviors without deriving from first principle [14]. DNNs have been extensively used for the prediction, detection, and classification of adverse glycemic events in T1D management [15,16]. With sufficient good quality data, DNNs have the ability to perform better than conventional physiological models [17,18]. Models that can synthesize the observed data are known as generative models. Generative models learn the joint probability distribution of data, unlike discriminative models that learn the conditional probability distribution of data instead. When DNNs are used for this purpose, they are termed as deep generative models (DGMs). DGMs have the ability to learn from a complex distribution and then generate new samples from the learned distribution. Such deep learning models could be well suited for biomedical data simulation tasks because, firstly, they can learn almost all the characteristics of a cohort of patients given that there is sufficient data, and secondly, they can produce unique samples from the learned distribution that are statistically similar to original data [19]. One of the most successful techniques capable of generating synthetic data based on a training in real scenarios is the generative adversarial network (GAN), which in recent years have demonstrated its value in biomedical data generation [20,21]. A particular type of GAN is known as conditional GAN (CGAN). CGAN differs in its functionality from a common GAN in such a way that an input label can control the generated data. CGANs have been used extensively in image translation, i.e., converting one type of image in to another type of image. It is equivalent to saying that one type of image is used as an input label to produce the targeted output. The image-to-image translation task is often referred to as pixel-to-pixel translation and the CGAN used for it is known as a pixel-to-pixel (Pix2Pix) GAN [22]. This paper proposes a strategy for the generation of realistic BG profiles of T1D patients conditioned on input insulin values using a Pix2Pix GAN. Furthermore, this work also discusses the possibility of a DGM that may be trained for the generation of T1D scenarios and can emulate a real T1D patient. The reason GAN is preferred over other types of DGMs in the proposed work is the fact that it is explicitly set up to optimize the generative task. Other DGMs, such as variational autoencoders, flow-based models, and diffusion models, are all set up to model the latent variable and are avoided here because of the problems related to latent variable approximation.

The major contributions of this work are as follows:

- A GAN has been used for the conditional generation of realistic BG values for the first time.
- The rationale for the development of a GAN-based diabetes patient simulator is given.

The rest of the paper is structured as follows: Section 2 explains deep generative models in detail along with its various types. Section 3 elaborates the methodology of the proposed work, with Section 4 dwelling in on the implementation and results of the

proposed research work. Discussion about the proposed methodology, implementation, and results is presented in Section 5. Section 6 concludes the paper with various suggestions for future work.

2. Deep Generative Models

Deep generative models have been used extensively to generate synthetic data for numerous sorts of applications over the past few years [23,24]. Though the most famous of these applications are the ones that involve facial image generation [25] and are under scrutiny for various reasons [26,27], the generative ability of these models could be harnessed positively in a variety of other ways. The DeepMind's rain predictor is one such example out of the many practical applications of deep generative models [28]. The aim of a DGM is to approximate a complex and high dimensional probability distribution using neural networks [29]. It does so by training a generative neural network in such a way that the obtained generator model learns to transform a tractable probability distribution Z in to a complex probability distribution X. Since DGM is an ill-posed problem, it is almost impossible to uniquely identify the probability distribution of a finite set of samples and hence the generation accuracy of a DGM is highly dependent on the hyper-parameters involved in the design of the model. There are several types of DGMs being used by researchers. None could be said to have any advantage or disadvantage over the other but that each one performs better than others under certain defined circumstances. The most common types of DGMs are the generative adversarial networks (GANs), the variational auto-encoders, the Bayesian networks, and the normalizing flows. Since our system uses a variant of GAN, our main focus of explanation is this methodology.

2.1. Generative Adversarial Network

GANs are a type of DGM that use two competing models in a zero-sum game, trying to optimize the loss function [30]. These models are a discriminator model (D) and a generator model (G). The function of D is to tell real data from the data generated by G, while G keeps on improving its generated output in order to fool D. Both D and G are DNNs trained to outperform each other. D is trained on real data and hence can classify real data from fake data. G, on the other hand, is trained using the feedback from D. The training of D is an act of minimizing a negative log likelihood loss whereas G uses two types of losses, the adversarial loss from D and the L1 loss from the similarity between the generated and real data. Once G learns the probability distribution of the real data set well enough, it can create infinite samples that are statistically similar to the real data and fool the discriminator.

2.2. Conditional Generative Adversarial Networks

CGAN is a variant of GAN that takes labeled data to train the discriminator instead of solo data. Where in a simple GAN the generation of data is not in our control and G generates new samples from the latent space Z that are unknown to us, in a CGAN the generation of new samples can be controlled with the help of data labels. Data samples along with their labels are fed to D for training purpose. Data labels are also used to condition Z in order to generate outputs that are similar to the data associated with the label. CGAN has been used effectively in text-to-image GAN architectures [31]. These models generate images according to the textual description provided by the user. The labels are provided as an embedded layer to the neural network during training. While generating new samples, the label information is added to Z so that the output is dependent on the label.

The operation of a CGAN may serve the purpose of a diabetes virtual patient simulator aptly. First, it can draw random samples from a random latent space and convert them into realistic glucose values. Secondly, it can also condition the generation of these glucose values on any of the other data such as insulin, carbohydrates, physical activity, etc. The generation of BG values that are conditioned on input insulin values mean that the

Mathematics **2022**, *10*, 3741

generated values are insulin dependent and will follow the underlying distribution of the insulin profile provided to the model as input.

2.3. Image Translation Using Conditional Generative Adversarial Networks

A variant of CGAN, known as the Pix2Pix architecture, is implemented in this research work. Part of the reason why we opted for a Pix2Pix architecture was to capture not only the data distribution but also its internal representations in a structured way from the portrayal of the BG and insulin time-series in image form. The Pix2Pix GAN performs the task of translating one type of image into another type of image. The generator in a Pix2Pix GAN learns to map an input image to an output image. To carry out this task, the input image is taken as a label to the target image. Pairs of input image and target image are used to train the discriminator. Additionally, the generator latent space is conditioned on the input label image as well in order to generate the images of our choice. Once trained, G is used to generate new images that are similar to real images. Each time the trained G is fed with a new input/label image, it will produce an output/target from the latent space. Hence, it is synonymous to say that unique target images will be generated using unique inputs. Fundamentally, the mechanism of generation in the proposed methodology is a translation task from one type of data to another type of data. In our model, having the generation of BG values respond to the input insulin labels gives us the liberty to generate BG scenarios of our choice. Moreover, along with the quantity of input insulin, the generated BG values are also affected by how the insulin samples interact with each other, hence preserving the temporal characteristics of the data.

For synthetic BG value generation, the insulin data are used as input, whereas the BG values are used as output data. The numerical data is first converted into images to prepare it for training. The conversion of data from numerical form to a graphical form is carried out using the HSV color map. After the formation of the image data set, training of the model is performed. This can be observed in Figure 2. Once, the generator learns the underlying probability distribution of the real data set, it can produce unique BG images when provided with unique insulin images.

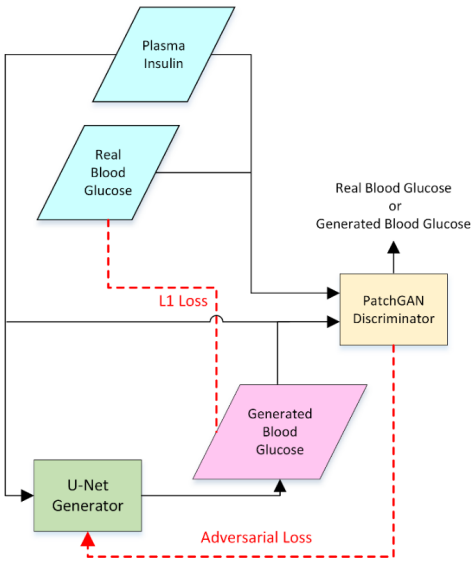


Figure 2. Pix2Pix generative adversarial network architecture in training configuration for translation of plasma insulin images into blood glucose images.

3. Methodology

The proposed system uses a Pix2Pix GAN together with a series of various other techniques to achieve the task of virtual patient generation. The proposed system consists of four main parts that are also its main operations. In the order of operation, these parts are:

- Numerical data to graphical data conversion;
- Pix2Pix GAN Training;
- Data generation using a trained Pix2Pix generator;
- Graphical data to numerical data conversion.

3.1. Experimental Data Sets

The proposed model is put to test in order to generate data that are statistically similar to two distinct data sets from separate studies. The first data set contains 29 patients on insulin pump therapy provided by the Hospital Clínic de Barcelona [32]. All the patients included in this data set wore Medtronic's MiniMed 640G insulin pumps. This data set contains a cumulative sum of 1169 days of insulin and BG data. By insulin data, here, we mean the plasma insulin approximation obtained using the Hovorka plasma insulin model of the total insulin injected to the patient [33]. The second data set used to train the Pix2Pix GAN is the Ohio T1DM data set [34]. We used data from the first 6 patients of this data set that consists of 264 days of BG and plasma insulin data. These Ohio T1DM data set patients were on insulin pump therapy as well and wore Medtronic 530G or 630G insulin pumps.

3.2. Model Architecture

A Pix2Pix GAN is implemented using the Keras deep-learning framework in Python. As discussed already, a Pix2Pix GAN is a type of CGAN that conditions the generation of target images on some sort of label images. In our case, the generated/target images were that of BG and the input labels were insulin. It helps understanding the phenomena better if the insulin images are imagined as inputs while the glucose images as output. The implementation of Pix2Pix GAN involves expert designing of the generator and discriminator model and the strict optimization of the system. Deciding the specifications of G and D is the most important step in moving towards the implementation of a GAN. We have used the same architecture as presented by [22] with some modifications to the model layers and hyper-parameters. The D model is a DNN that performs the binary classification task of classifying images as real or fake. The images are provided to D as input and output pair as shown in Figure 2. The D model receives an insulin-BG image pair and classify them as real or fake. However, the D model used in the proposed model is a special type of model known as the patchGAN discriminator. Figure 3 shows the formation of the D model. The patchGAN D works on the concept of the effective receptive field of the model. It essentially means that the discriminator neural network only classifies a portion of the input image and do not take the entire image for the classification purpose. In the end, the outcomes of all such decisions are convoluted to obtain the one combined output. The benefit of such an approach is that different sizes of images could be classified with the help of such a model. Moreover, in our application, the patchGAN discriminator learns about the relationship between a particular segment of insulin image and a particular segment of BG image. This in effect means that our model does not only learn the general relationship between the 6 h chunks but also approximates the relationships between the samples falling inside these 6 h chunks.

Since for our model the input is a 72×1 dimension image, the kernel (filter) size is also kept as one-dimensional (1D). Here, it is important to consider that the convolutional layer used is still a two-dimensional (2D) layer and not 1D. The concept of padding has been employed in order to keep the spatial dimensions of the layer's outputs same as the inputs. Leaky rectified linear unit (Leaky ReLU) is used as the activation function at each layer. A total of six 2D convolutional layers are used in D. The training pair of plasma insulin image and BG image are concatenated and labeled as 1 for the real image pair before feeding

Mathematics **2022**, *10*, 3741

them to the Pix2Pix model for training. For the generated images, the plasma insulin and generated BG images are concatenated and labeled as 0. Concatenation is performed in order to have both types of images in the batch when the model weights are updated and the mapping between the images is learned.

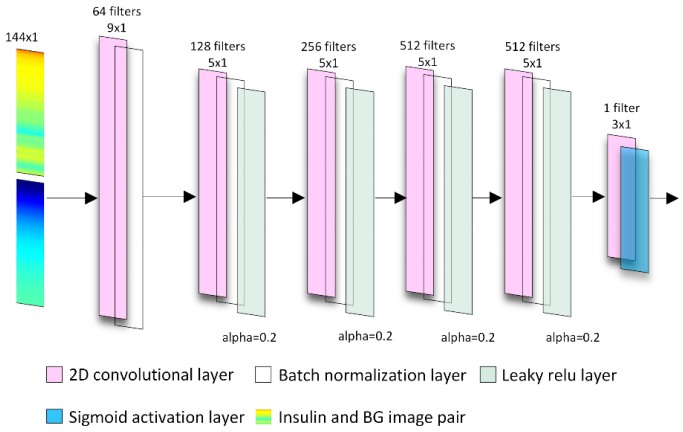


Figure 3. The patchGAN discriminator model that performs classification between real and generated BG images.

On the other hand, the *G* model, as shown in Figure 4, is comparatively more complex. This model is based on an encoder–decoder architecture which consists of two blocks of CNNs that are standardized for the purpose of replication: the encoder block and the decoder block. Both of these blocks are connected with the help of a bottleneck convolutional layer. Moreover, a slight modification is made to the conventional design of the encoder-decoder model by forming connections between layers of same size in both of these blocks [35]. This modification is usually referred to as the U-net configuration and is the reason why the generator in our designed Pix2Pix model is referred to as a U-Net generator. This design has been proven to facilitate the learning of minute details in an image [35] and can support the algorithm in learning transitions between colors for time-dependent characteristics of the insulin and BG time-series.

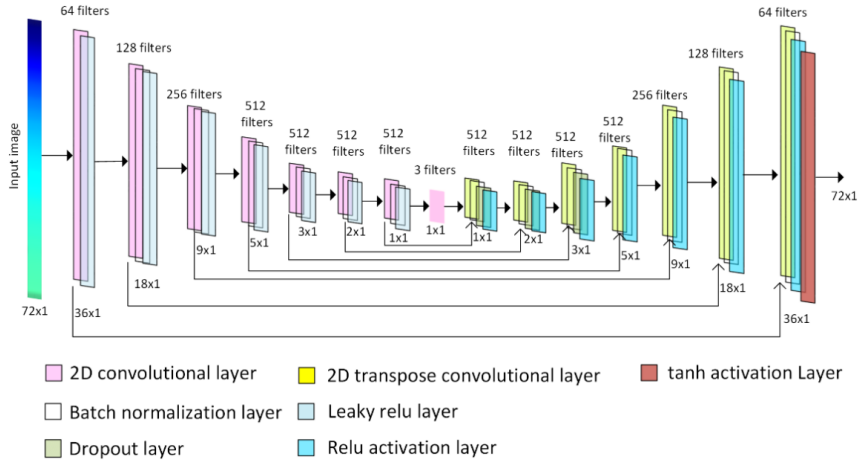


Figure 4. The U-Net generator model used to generate novel plausible BG images.

Mathematics **2022**, *10*, 3741

3.3. Translation between Numerical and Image Information

Since the GAN used in this work is a convolutional neural network-based image to image translator, the data used to train this model are image data. For this particular purpose, numerical data from the actual diabetes patient cohort are first converted into images. Conversion to image data is carried out through the HSV color map as shown in Figure 5. Six hours of insulin/BG data are converted into one HSV image using MATLAB. A duration of six hours is selected in order to completely approximate the effect of insulin and meals in the output BG profile. Each pixel in the HSV image represents the magnitude of the insulin/BG value in the actual 1D array. Six hours of data mean that there are 72 pixels in the obtained image since the sample time in the actual time-series is 5 min. The proposed system uses plasma insulin and glucose data for training and hence, only these two sets of data are converted into images.

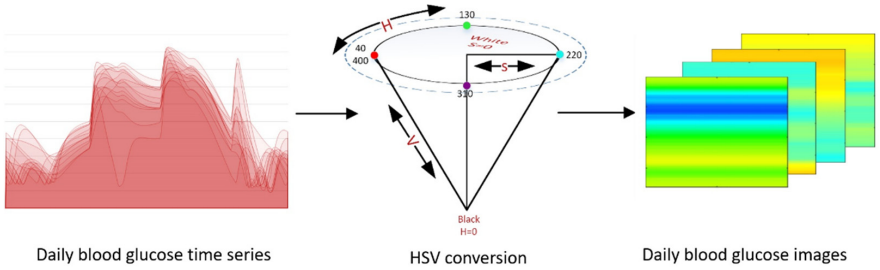


Figure 5. Numerical values to image translation using HSV color map: *H* stands for hue, *S* stands for saturation, *V* stands for value.

The HSV color map is chosen for image generation because of its simpler parametric composition. The HSV color scale is inspired by how the human eye perceives colors [36]. We see the same colors as those depicted by an RGB scale but do not comprehend them as a mixture of red, blue, and green. Instead, the human eye comprehends a color in terms of its saturation, tone, and brightness/darkness levels. The HSV color scale represents these parameters as the hue, the saturation, and the value (of brightness). The hue of the color is the color itself. The saturation of the color defines how saturated the quantity of color is, whereas the value of a color describes the brightness/darkness of a color.

It is important to note that even though its three channels, i.e., H, S, and V, represent the definition of a color on an HSV scale, it is still possible to depict any color correctly by using just the H parameter while keeping the S and V parameters constant (greater than 0). We have leveraged this property of the HSV scale by using just the H parameter for color portrayal while keeping the S and V parameters at a constant value of 1. This has made the translation from time-series data to color images a 1D-to-1D data translation in essence. Figure 5 shows an HSV cone. The H parameter represents the circumference of the circular plane in the cone. Each of its 360 degrees represents one color. The S parameter is represented by the radius of this plane, whereas the V parameter lies on the plane that connects the flat circular plane to the apex. To increase the saturation of a particular hue/color, we move away from the axis of the cone on the S parameter. On the other hand, increasing the V parameter takes us towards the lighter colors. Since, each color in the HSV color map can be represented by a single numerical value, i.e., the H parameter, it is easier to map an insulin/BG value to a pixel in the obtained image. The H parameter encodes each color as one of the 360 degrees of a circular plan. Since each color in the pallet is represented by one degree, there are a total of 360 colors on the H scale. The numerical time-series is transformed into its graphical representation by choosing one of these 360 colors for one magnitude value. Since for both insulin and BG time-series the possible values do not exceed the range of 360, transformation into images is a straightforward affair.

Furthermore, a translation mechanism is employed in order to translate the images back into numerical values. Since the output of the trained model is in image form, this translation mechanism is used to obtain the final BG time-series. As shown in Figure 6, while translating the generated images into BG values, the images are first given as an input to a μ-law expander in order to increase the dynamic range of colors. There are several reasons for the low dynamic range of colors in the generated images. One reason is the bottleneck layer in the U-Net generator model where significant image details maybe lost. Decrease in the dynamic range might be considered somewhat similar to a logarithmic compression process where the amplitudes are brought more or less to the same level and the dynamic range is compressed. The expander reverses this process and increases the difference between the smallest and largest values of the signal. After passing through a μ-law expander, the images are translated into BG values by using a pixel-to-BG scale. According to the pixel-to-BG scale, one pixel value is translated to one BG value. Since the HSV color map has 360 degrees representing 360 colors and the BG scale of 40–400 mg/dL too has 360 BG values, each color value can be mapped to each BG value. After the pixels are translated in to BG values, the resultant time-series is passed through a low-pass filter for further smoothing. A BG time-series is obtained at the output of the low-pass filter. Figure 6 demonstrates the conversion of generated BG images into numerical values using the translation mechanism discussed above.

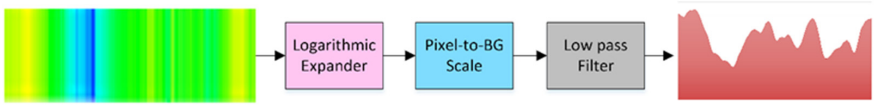


Figure 6. Image to BG translation mechanism depicted with example: 18 h of generated image data is translated into blood glucose time-series.

3.4. Model Training and Data Generation

The proposed Pix2Pix generator is trained using an entire batch of paired images, i.e., the insulin and BG pairs. Each input insulin image corresponds to one target BG image. The total number of steps required to complete the training depends on the number of images in the data set multiplied by the number of epochs. Since there are 29 patients in the first data set, the Pix2Pix GAN was trained 29 times by excluding data from one patient during each training instance. This way, a total of 29 models were obtained. Similarly, for the Ohio data set, the model was trained 6 times for 6 patients in order to obtain 6 separate models. The training task of the proposed GAN architecture involves optimizing for two types of loss functions, the discriminator loss and the generator loss. The model weights for both the discriminator and generator model are updated after training on each individual image. The model is trained for 50 epochs because an empirical testing phase has pointed out this value as a common framework where the models achieved the convergence. The trained models are then used to generate data for each patient individually. The training process could be understood by looking at Figure 2. Once the generator is able to generate fake images that are real enough to fool the discriminator completely, the generator weights stop updating because the discriminator adversarial loss becomes zero. For an optimally trained GAN model the discriminator must demonstrate theoretically equal loss values for both real data and generated data. This means that the discriminator is unable to distinguish between real and fake images. In other words, the generator is producing images that are plausible enough to be considered real. This is when a generator model is considered to be fully trained and its weights should not be updated further. A fully trained generator model is then used to generate new fake images that are considerably similar to the real images. Figure 7 shows three sample images that have been generated by our generator along with the corresponding real BG data. Each generated image represents 6 h of BG data.

Mathematics **2022**, *10*, 3741

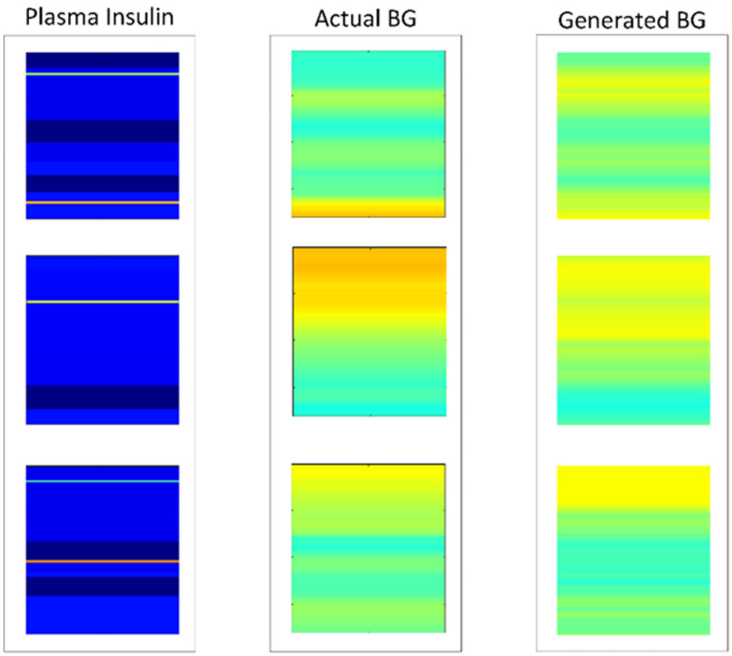


Figure 7. Generated images using a Pix2Pix GAN trained on blood glucose and plasma insulin data.

At least 4 weeks of data is generated for each patient in order to achieve the optimal amount of data required for reporting the glycemic performance parameters such as times in ranges and coefficient of variation (CV) etc. [37]. The generated BG data is then analyzed for various clinical parameters and times in standardized ranges of BG levels. These ranges are classified into three groups, namely the time below range (TBR), the time in range (TIR), and the time above range (TAR). Other metrics such as mean, CV, and maximum and minimum values of BG are also calculated for each patient. Means of all these parameters for the generated cohort are then compared against the means of the actual cohort. Moreover, the p-values from the Wilcoxon rank sum test are computed for each of these metrics of the generated data set against the metrics of the real data set. The 25% and 75% quartiles are computed for each metric as well and given as such.

4. Results

Tables 1 and 2 contain all the results for both data sets. Table 1 shows the times in standardized ranges of BG levels, mean, CV, minimum and maximum values, and their corresponding p-values for the 29 patients real and generated cohorts. It could be clearly understood from the table that p-values depict strong statistical similarity between the two cohorts. In other words, the generated cohort and the real cohort display statistical significance since all the p-values qualify the 0.05 threshold of hypothesis confirmation. Similarly, Table 2 shows the results for the Ohio T1D data set. It could be observed for this data set, too, that the p-values for all the parameters qualify the null hypothesis of the two-sample test except for the p-value of the maximum value.

The inter quartile ranges (IQ) for the Ohio data set hints at space for improvement. The reason for this is understood to be the lesser quantity of data. Just like it has been a trend with DNNs, an increase in the amount of training data improves the performance of the model.

Figures 8 and 9 present a graphical comparison of the generated data against the real data for both experimental data sets in terms of the standardized ranges of BG levels. The bar graphs show the mean values with the p-values that show statistical similarities are

Mathematics **2022**, *10*, 3741

written on top of the quantities in comparison. It can be seen that the generated data sets are significantly similar in terms of both the means of data and the statistical p-values. This graph also provides us an insight on the standard deviations of the two data sets.

Table 1. A comparison of generated cohort and real cohort: Medtronic's MiniMed 640G data set (29 patients).

Ranges	Real Cohort	Generated Cohort	p-Values
Below 54 (mg/dL)	1.26(0.16–1.03)	0.90 (0.34–1.15)	0.32
54 to 69 (mg/dL)	3.14 (1.41–4.03	3.41 (2.43–4.56)	0.15
70 to 140 (mg/dL)	38.72 (29.32–46.81)	42.40 (37.98–47.99)	0.35
70 to 180 (mg/dL)	62.68(55.64–69.45	63.39 (56.91–70.71)	0.99
180 to 250 (mg/dL)	24.70 (17.52–30.46)	22.46 (18.34–27.51)	0.39
Above 250 (mg/dL)	8.94 (3.43–11.75)	8.25 (4.31–11.21)	0.85
Mean CGM (mg/dL)	159.46 (148.06–173.86)	157.64 (147.76–169.16)	0.86
CV (Percentage)	35.96 (32.89–39.05)	38.56 (35.85–41.88)	0.12
Maximum CGM (mg/dL)	373.03 (358–400)	389.27 (387.93–400)	0.06
Minimum CGM (mg/dL)	43.86 (40–43)	46.21 (40.00–50.20)	0.25

Table 2. A comparison of generated cohort and real cohort: Ohio T1D data set (6 patients).

Ranges	Real Cohort	Generated Cohort	p-Values
Below 54 (mg/dL)	0.63 (0.19–0.81)	0.63 (0.06–0.84)	0.55
54 to 69 (mg/dL)	2.20(1.45–2.43)	2.18(0.08–3.64)	0.82
70 to 140 (mg/dL)	34.52(32.07–40.77)	45.24(26.33–63.36)	0.48
70 to 180 (mg/dL)	55.91(50.98–61.97)	55.65(34.92–74.06)	0.94
180 to 250 (mg/dL)	24.45(20.23–28.49)	28.34(15.94–38.83)	0.99
Above 250 (mg/dL)	7.66(4.75–10.25)	12.65(4.65–18.52)	0.94
Mean CGM (mg/dL)	146.53(133.25–156.02)	158.63(127.73–190.56)	0.99
CV (Percentage)	49.28(43.60–56.20)	40.44(34.47–45.89)	0.31
Maximum CGM (mg/dL)	395.67(397.75–400.00)	340.43(331.15–347.81)	0.004
Minimum CGM (mg/dL)	41.00 (40–40)	51.55 (40.08–62.53)	0.12

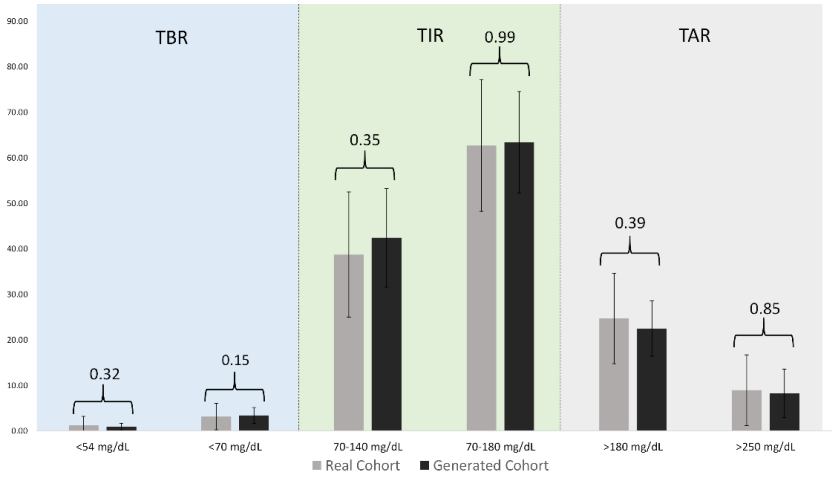


Figure 8. Times in standardized ranges of blood glucose for real and generated cohorts with corresponding p-values of the Midtronic's MiniMed 640G 29 patient data set.

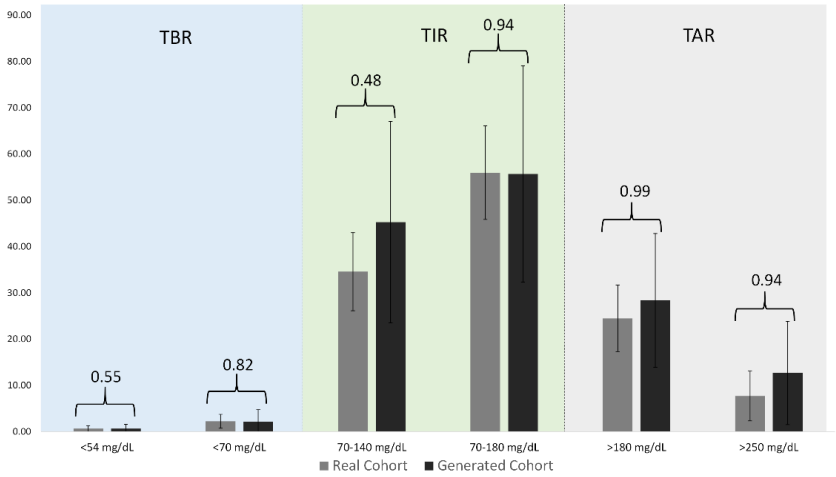


Figure 9. Times in standardized ranges of blood glucose for real and generated cohorts with corresponding *p*-values of the Ohio 6 patient data set.

5. Discussion

The proposed system demonstrated that, first, our Pix2Pix GAN learned from the distributions of BG profiles of diabetic patients labeled with plasma insulin estimation and, second, that it is able to generate new BG values conditioned on plasma insulin values. The generated BG values showed statistical similarity with the original data set. Since the model was trained and tested with two different data sets, the results for both data sets confirmed the effectiveness of our method. However, it is important to note that the approximation was better for the 29 patients' data set as compared to the 6 patients' data set, confirming that the amount of data used for training is a significant performance constraint. The need for more data, though paradoxical when it comes to data generation, is also evident from the theory of GANs [38].

Since, we used a variant of GAN that performs the task of image translation, the numerical data was transformed into graphical data in order to be used with the GAN. This strategy proved efficient for our cause because fundamentally, graphical data are a depiction of numerical data and the transformation was a number-to-number transformation in essence, i.e., the translation task of numerical values to images and then back to numerical values; however, this needs to be carried out with utmost care as slight neglect may lead to information loss and induction of errors. For instance, the expander parameters in the proposed model needs to be selected after rigorous testing or it could induce errors in the measurements. Keeping this in mind, the translation task could well be coupled with a correction factor of some sort to correct for the induced errors. Having said that, there can never be complete compensation for all sorts of losses and errors. Furthermore, the Pix2Pix GAN used in the proposed study learns from an original distribution as well as trains a loss function. This approach makes it possible for it to be used in any sort of image-to-image translation task without the need of designing a customized loss function. Additionally, this GAN is found to be efficient at pushing the output color distribution closer to the input color distribution [22]. Furthermore, the patchGAN discriminator used in our model encourages greater color diversity, which makes it a suitable choice of discriminator for the problem of translating a particular type of heat map into another type.

Since the aim of this work is to move towards a working environment that could simulate scenarios mimicking T1D patients, the correction factors used in the translation task should be evaluated thoroughly for all possible errors and documented accordingly. Furthermore, the translation algorithm should be tested for multiple cohorts and any errors should be quantified. In case of the proposed work, the same translation algorithm worked

equally good for both the data sets. In future, the generation of BG can be conditioned on various other types of data such as carbohydrates or physical activity data. In order to enable such DGMs to be able to approximate the whole of the human glucose–insulin mechanism, they must be trained on a large amount of versatile data. The proposed methodology may prove advantageous when a variety of data is used for glucose conditioning. In the current form, another column of pixel data may be sufficient for a new type of data to be added to the input image. Pix2Pix GAN's adaptability to various types and sizes of images makes it suitable for such kind of a problem where the formation of images may change because of the addition or removal of data.

From the clinical perspective, the proposed methodology shows promise in the creation of simulation environments that could be used for the development and testing of new treatment methodologies in T1D management. One of the strengths of conventional T1D simulators is the control of various phenomena affecting the glycemic profile, such as the amount and type of meal and insulin, among other factors. These simulators are also capable of reproducing results for a particular combination of the affecting factors. All of this points towards the approximation of causal relationships between these factors and the generation of BG time-series. One way to approximate causal relationships is through recurrence. Introducing snippets of recurring relationships to a CGAN enables it to generate data that depicts a causal relationship to the data it is conditioned on. Once an understanding of the causality is established, generative models could be utilized to generate data in future and hence provide predictions. The predictions, when tested experimentally, may give us a deeper understanding of the biological system at hand and help us devise treatments that are effective in real life.

6. Conclusions

Existing T1D simulators are based on physiological models and fail to provide an all-inclusive synthetic T1D patient platform for researchers to work with in the field. Data-driven models have the capability to learn those underlying characteristics of the human body from data that are hard to represent through physiological models. This research work proposed a Pix2Pix GAN model for the generation of BG values conditioned on plasma insulin approximation of T1D patients. The glucose data generated by the proposed system are realistic enough to be used as an alternative to real data. The BG data generated by our model could also be added to real data in order to improve the effectiveness of certain other data driven models, i.e., predictive models or classification models. In future, the generation of BG values could be conditioned on other types of data, such as carbohydrates or physical activity data, for more realistic results.

Author Contributions: Conceptualization, O.M., I.C. (Ivan Contreras) and J.V.; data curation, O.M. and A.B.; formal analysis, O.M., I.C. (Ivan Contreras) and A.B.; funding acquisition, I.C. (Ivan Contreras) and J.V.; investigation, O.M., I.C. (Ivan Contreras) and A.B.; methodology, O.M., I.C. (Ivan Contreras), A.B. and J.V.; project administration, I.C. (Ivan Contreras) and J.V.; resources, I.C. (Ignacio Conget), M.G. and J.V.; software, O.M.; supervision, I.C. (Ivan Contreras), I.C. (Ignacio Conget), M.G. and J.V.; validation, O.M., I.C. (Ivan Contreras), A.B., I.C. (Ignacio Conget) and M.G.; visualization, O.M. and I.C. (Ivan Contreras); writing—original draft, O.M., I.C. and A.B.; writing—review and editing, O.M., I.C. (Ivan Contreras), A.B., I.C. (Ignacio Conget), M.G. and J.V. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the Spanish Ministry of Universities, the European Union through Next GenerationEU (Margarita Salas), the Spanish Ministry of Science and Innovation through grant PID2019-107722RB-C22/AEI/10.13039/501100011033, PID2020-117171RA-I00 funded by MCIN/AEI/10.13039/501100011033 and the Government of Catalonia under 2017SGR1551 and 2020 FI_B 00965.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Hospital Clínic de Barcelona (protocol code HCB/2015/0683 and date of approval 22 September 2015).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from the Spanish Consortium on Artificial Pancreas and Diabetes Technology and are available on request from the corresponding author with the permission of the Spanish Consortium on Artificial Pancreas and Diabetes Technology.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Sá-Couto, C.; Patrão, L.; Maio-Matos, F.; Pêgo, J.M. Biomedical simulation: Evolution, concepts, challenges and future trends. *Acta. Med. Port.* **2016**, *29*, 860–868. [CrossRef] [PubMed]
2. Man, C.D.; Micheletto, F.; Lv, D.; Breton, M.; Kovatchev, B.; Cobelli, C. The UVA/PADOVA type 1 diabetes simulator: New features. *J. Diabetes Sci. Technol.* **2014**, *8*, 26–34. [CrossRef]
3. Wilinska, M.E.; Chassin, L.J.; Acerini, C.; Allen, J.M.; Dunger, D.; Hovorka, R. Simulation Environment to Evaluate Closed-Loop Insulin Delivery Systems in Type 1 Diabetes. *J. Diabetes Sci. Technol.* **2010**, *4*, 132–144. [CrossRef] [PubMed]
4. Alkhalifah, T.; Wang, H.; Ovcharenko, O. MLReal: Bridging the gap between training on synthetic data and real data applications in machine learning. *Eur. Assoc. Geosci. Eng.* **2021**, *2021*, 1–5. [CrossRef]
5. Jan, S.V.S.; Geris, L. Modelling towards a more holistic medicine: The Virtual Physiological Human (VPH). *Morphologie* **2019**, *103*, 127–130. [CrossRef]
6. Shengli, W. Is Human Digital Twin possible? *Comput. Methods Programs Biomed. Updat.* **2021**, *1*, 100014. [CrossRef]
7. Kovatchev, B.P.; Breton, M.; Man, C.D.; Cobelli, C. In SilicoPreclinical Trials: A Proof of Concept in Closed-Loop Control of Type 1 Diabetes. *J. Diabetes Sci. Technol.* **2009**, *3*, 44–55. [CrossRef]
8. Cobelli, C.; Man, C.D.; Sparacino, G.; Magni, L.; De Nicolao, G.; Kovatchev, B.P. Diabetes: Models, Signals, and Control. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 54–96. [CrossRef]
9. Hester, R.L.; Iliescu, R.; Summers, R.; Coleman, T.G. Systems biology and integrative physiological modelling. *J. Physiol.* **2011**, *589*, 1053–1060. [CrossRef]
10. Vettoretti, M.; Facchinetti, A.; Sparacino, G.; Cobelli, C. Type-1 Diabetes Patient Decision Simulator for In Silico Testing Safety and Effectiveness of Insulin Treatments. *IEEE Trans. Biomed. Eng.* **2017**, *65*, 1281–1290. [CrossRef]
11. Roversi, C.; Vettoretti, M.; Del Favero, S.; Facchinetti, A.; Sparacino, G. Modeling Carbohydrate Counting Error in Type 1 Diabetes Management. *Diabetes Technol. Ther.* **2020**, *22*, 749–759. [CrossRef] [PubMed]
12. American Diabetes Association. 15. Diabetes Care in the Hospital: Standards of Medical Care in Diabetes—2021. *Diabetes Care* **2021**, *44* (Suppl. S1), S211–S220. [CrossRef] [PubMed]
13. Karimy, M.; Koohestani, H.R.; Araban, M. The association between attitude, self-efficacy, and social support and adherence to diabetes self-care behavior. *Diabetol. Metab. Syndr.* **2018**, *10*, 86. [CrossRef] [PubMed]
14. Bertachi, A.; Biagi, L.; Contreras, I.; Luo, N.; Vehí, J. Prediction of Blood Glucose Levels And Nocturnal Hypoglycemia Using Physiological Models and Artificial Neural Networks. *KHD@ IJCAI* **2018**, 85–90.
15. Mujahid, O.; Contreras, I.; Vehi, J. Machine Learning Techniques for Hypoglycemia Prediction: Trends and Challenges. *Sensors* **2021**, *21*, 546. [CrossRef]
16. Contreras, I.; Vehi, J. Artificial Intelligence for Diabetes Management and Decision Support: Literature Review. *J. Med. Internet Res.* **2018**, *20*, e10775. [CrossRef]
17. Nalisnick, E.; Matsukawa, A.; Teh, Y.W.; Gorur, D.; Lakshminarayanan, B. Do deep generative models know what they don't know? *arXiv* **2018**, arXiv:1810.09136.
18. Nikzad, M.; Movagharnejad, K.; Talebnia, F. Comparative Study between Neural Network Model and Mathematical Models for Prediction of Glucose Concentration during Enzymatic Hydrolysis. *Int. J. Comput. Appl.* **2012**, *56*, 43–48. [CrossRef]
19. Noguer, J.; Contreras, I.; Mujahid, O.; Beneyto, A.; Vehi, J. Generation of Individualized Synthetic Data for Augmentation of the Type 1 Diabetes Data Sets Using Deep Learning Models. *Sensors* **2022**, *22*, 4944. [CrossRef]
20. Buczak, A.L.; Babin, S.; Moniz, L. Data-driven approach for creating synthetic electronic medical records. *BMC Med. Inform. Decis. Mak.* **2010**, *10*, 59. [CrossRef]
21. Frid-Adar, M.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. Synthetic data augmentation using GAN for improved liver lesion classification. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 289–293. [CrossRef]
22. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
23. Jordon, J.; Yoon, J.; van der Schaar, M. PATE-GAN: Generating synthetic data with differential privacy guarantees. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
24. Frid-Adar, M.; Diamant, I.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* **2018**, *321*, 321–331. [CrossRef]

25. Lee, C.-H.; Liu, Z.; Wu, L.; Luo, P. Maskgan: Towards diverse and interactive facial image manipulation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5549–5558.
26. Meskys, E.; Liaudanskas, A.; Kalpokiene, J.; Jurcys, P. Regulating deep fakes: Legal and ethical considerations. *J. Intellect. Prop. Law Pract.* **2020**, *15*, 24–31. [CrossRef]
27. Dornis, T.W. Artificial Creativity: Emergent Works and the Void in Current Copyright Doctrine. *Yale J. Law Technol.* **2020**, *22*, 1.
28. Ravuri, S.; Lenc, K.; Willson, M.; Kangin, D.; Lam, R.; Mirowski, P.; Fitzsimons, M.; Athanassiadou, M.; Kashem, S.; Madge, S.; et al. Skillful Precipitation Nowcasting using Deep Generative Models of Radar. *arXiv* **2021**, arXiv:2104.00954.
29. Ruthotto, L.; Haber, E. An introduction to deep generative modeling. *GAMM-Mitteilungen* **2021**, *44*, e202100008. Available online: https://arxiv.org/abs/2103.05180 (accessed on 20 August 2022). [CrossRef]
30. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative Adversarial Networks: An Overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [CrossRef]
31. Mirza, M.; Osindero, S. Conditional Generative Adversarial Nets. *arXiv* **2014**, arXiv:1411.1784.
32. Ahmad, S.; Ramkissoon, C.; Beneyto, A.; Conget, I.; Giménez, M.; Vehi, J. Generation of Virtual Patient Populations that Represent Real Type 1 Diabetes Cohorts. *Mathematics* **2021**, *9*, 1200. [CrossRef]
33. Hovorka, R.; Canonico, V.; Chassin, L.J.; Haueter, U.; Massi-Benedetti, M.; Federici, M.O.; Pieber, T.R.; Schaller, H.C.; Schaupp, L.; Vering, T.; et al. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol. Meas.* **2004**, *25*, 905–920. [CrossRef]
34. Marling, C.; Bunescu, R. The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020. *CEUR Workshop Proc.* **2020**, *2675*, 71–74.
35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
36. Vadivel, A.; Sural, S.; Majumdar, A.K. Human color perception in the HSV space and its application in histogram generation for image retrieval. *Proc. SPIE* **2005**, *5667*, 598–609. [CrossRef]
37. PHerrero, P.; Alalitei, M.A.; Reddy, M.; Georgiou, P.; Oliver, N. Robust Determination of the Optimal Continuous Glucose Monitoring Length of Intervention to Evaluate Long-Term Glycemic Control. *Diabetes Technol. Ther.* **2021**, *23*, 314–319. [CrossRef]
38. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661. [CrossRef]

2.3 Generative deep learning for the development of a type 1 diabetes simulator

In this publication, we have presented a framework for the development of a type 1 diabetes simulator with open-loop and closed-loop insulin therapies using generative deep learning.

Title: Generative deep learning for the development of a type 1 diabetes simulator
Authors: **Omer Mujahid**, Ivan Contreras, Aleix Beneyto, and Josep Vehí
Journal: Nature Communications Medicine
Status: Under Review, March 2023
Quality index: Not indexed

Generative deep learning for the development of a type 1 diabetes simulator

Omer Mujahid[1], Ivan Contreras[1], Aleix Beneyto[1] and Josep Vehi[1,2*]

[1]Modeling, Identification and Control Laboratory, Institut d'Informatica i Applicacionst, Universitat de Girona, Girona, 17003, Girona, Spain.
[2*]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Spain.

*Corresponding author(s). E-mail(s): josep.vehi@udg.edu;
Contributing authors: omer.mujahid@udg.edu;
ivan.contreras@udg.edu; aleix.beneyto@udg.edu;

Abstract

Type 1 diabetes (T1D) simulators employ mathematical physiological models of the glucose-insulin system. Used in the development and refinement of new treatments for diabetes, these simulators fail at encompassing the entire physiology of the glucose-insulin system because of the imprecise approximation of the physiological models. We present a simulation environment based on a conditional sequence-to-sequence deep generative model that performs the task of causally synthesizing virtual T1D patients. The generated patients display statistical similarity to the real patients when evaluated for time-in-range results for each of the standard blood glucose ranges in T1D management along with means and variability outcomes. When tested for causality, authentic causal links are identified between the insulin, carbohydrates, and blood glucose of the virtual patients. The trained generative model demonstrates behaviors that are closer to reality as compared to conventional T1D simulators when subjected to closed-loop insulin therapy using a state-of-the-art controller.

Keywords: virtual type 1 diabetes patients, type 1 diabetes simulators, causal blood glucose data synthesis, deep generative models

1 Introduction

The terms modeling and simulation (M&S) go hand in hand. Though defined differently by different researchers depending mostly on the field of study, the majority of these definitions are some variant of the interpretation presented by Kaizer et al. [1] which states that a model is a "representation of a system, entity, phenomenon, or process" whereas a simulation is "the imitation of a behavior of a system, entity, phenomenon or process through the exercise or use of a model". The purpose of M&S is the emulation and approximation of physical phenomena that cannot be directly observed, for the purpose of better understanding. It is the process of explaining how an object of interest behaves in an environment. Biomedical simulation tasks may employ some model(s) of a biological system to emulate the physics underlying biological organs in coordination with certain other mechanisms to form a complete simulation environment.

The majority of biomedical simulators use mathematical physiological or pharmacokinetic models to simulate biological phenomena [2]. Such simulators hold significance in the development and testing of new treatments and therapeutic strategies for different diseases because they offer a cheap alternative to patient and animal testing both in terms of time and money. Moreover, they can quickly help identify parts of the device design that may not be effective and prevent the development of adverse circumstances because they are easy to interpret [3]. Though efficient to some extent in the task of approximating phenomena resulting from biological organs, the physiological models do not capture the entirety of a biological process because of various elements that are simply not possible to model [4]. These elements may be certain unmeasured variables affecting the outcomes of a process or external influencing factors such as the patient lifestyle choices, routine habits, and environmental factors [5]. Consequently, an intrinsic error is induced in the approximation of physiological models in the form of a non-random grey noise that compromises the effectiveness of these models in an unavoidable fashion. Since, a model is a combination of various parts such as compartments, failure to approximate the real scenarios is often unexplained in terms of error-inducing components [6]. Similarly, in the case of an accurate approximation, the reason for success may not be measured deterministically either.

Existing type 1 diabetes (T1D) simulators employ physiological models of the glucose-insulin system for the purpose of generating glycemic scenarios emulating real-life T1D patients [7] [8] [9]. Integration of such physiological models with open-loop (OL) or closed-loop (CL) control methodologies enables the creation of simulation environments that imitate the glucose-insulin relationship of T1D patients. These simulators have made the development and testing of several treatment methodologies possible for T1D patients [10]. However, as discussed already, the physiological descriptions of the glucose-insulin system are far from perfect and the T1D simulators based on these models suffer from the induced inaccuracies of the mathematical descriptions. Having no provision regarding patients' life choices and habits or other life

disturbances such as menstruation, depression, medication, etc, these simulators often crumble in scenarios that require higher precision into the real-life factors affecting diabetics. This forms the basis of the rationale for the exploration of approximation techniques that will take a wholesome picture of the factors affecting the glycemic trends of a T1D patient into account. Such an approximation technique needs to take external factors like exercise, illness, menstrual cycles, sleep disorders, depression, and medication into consideration. Furthermore, the existing T1D simulators have little support for patient behavior such as eating habits, alcohol consumption, and lifestyle choices. These factors have a significant impact on the glycemic profile of a person and having these disturbances included in a simulator will ensure the simulation of highly accurate scenarios. It is understood that modeling each one of these disturbances individually is an impossible task; however, the effect of these disturbances can be approximated from the data they generate. The proposed methodology leverages the concept of modeling from data using generic function approximators.

Generic function approximators have been proven to learn complex nonlinear relationships from data. According to the universal approximation theorem, an artificial neural network (ANN) with a hidden layer is capable of learning almost any function given that it is sufficiently wide [11] [12]. This means that a neural network designed to learn the probability distribution of data may learn any complex distribution provided that the network is apt enough and the data is available in sufficient quantity. Moreover, given the data is available in sufficient quantity deep neural networks (DNNs) are capable of surpassing mathematical models in terms of better approximation of systems [13] [14]. Deep generative models (DGMs) are DNNs that are capable of learning the underlying probability distribution of data and then generating novel samples from the learned distribution. The effectiveness of DGMs in approximating distributions accurately has been demonstrated by several recent studies in various areas of research including biomedical applications [15] [16] [17] [18].

With reference to probability theory, deep learning models are often divided into the following categories: the generative models, the discriminative models, and the composite models [19]. Discriminative models learn the conditional probability distribution from data whereas generative models learn the combined probability distribution from data. Composite models on the other hand are a combination of discriminative models and generative models. DGMs are a type of composite model. Evidently, the data obtained from the human glucose-insulin system depicts a complex underlying probability distribution and DGMs are one of the most suitable methodologies for learning this type of distribution. Starting from the restricted Boltzmann machines the field of DGMs has evolved quite a bit and there are now several types of DGMs available that can perform the task of unsupervised learning of the hidden features in a particular data set [20] [21]. The performance quality of these models is context dependent such that none could be deemed superior to the other

and vice versa. Some of the most trendy DGMs include variational autoencoders, generative adversarial networks (GAN), normalizing flows, diffusion models, and autoregressive models. In this work, we have employed a GAN for the task of conditional BG generation. GAN is chosen over other types of DGMs because it has been specifically set up to optimize the generation tasks. Other DGMs, such as variational autoencoders, flow-based models, and diffusion models, are all set up to model the latent variable and are avoided here because of the problems related to latent variable approximation [22] [23].

This research work proposes a strategy of utilizing DGMs for the task of simulating T1D patients' profiles by learning glycemic trends in the form of glucose-insulin and glucose-carbs relationships from the data. The proposed methodology employs a conditional generative adversarial network (CGAN) for the generation of BG values conditioned on the plasma insulin approximation (PI) and the carbohydrate rate of appearance (RA) of T1D patients. We have demonstrated through a multi-therapeutic validation approach that the generated BG values exhibit a dependency on the input PI and RA values which is consistent with the glycemic relationships of real T1D patients. Moreover, the generated BG values depict the input insulin and carbohydrates taking effect in the glycemic profile over time. This is achieved by building on the idea of shifted input/output pairs to approximate the glucose-insulin model of a cohort of T1D patients. These properties show the appropriacy of our methodology for the purpose of simulating T1D patients and devising treatment regimens based on OL and CL therapies. The main contributions of this work are:

- The first-of-its-kind AI-based T1D simulator with open-loop and closed-loop insulin therapies
- Causal generation of BG values using DGMs
- Generation of realistic T1D patients conditioned on insulin and carbohydrates values

We have tried to include as many details as possible about the tools and data set used in the study, problem formulation, and results based on the benchmarks of robustness for AI in healthcare proposed by Diana Mincu and Subhrajit Roy [24]. The rest of the paper is focused on explaining the methodology to generate causal data using DGMs, the results containing the statistical similarity tests and causality analysis, and the discussion about the limitations and possibilities of this work. Section no. 2 unfolds the proposed methodology of setting up a CGAN for the generation of BG values that are realistic and causally dependent on inputs. Section no. 3 throws light on the results while section no. 4 discusses the results in light of the T1D simulation tasks. Section 5 concludes the paper with suggestions and possible future work directions.

2 Methodology

This section dwells on the approaches, techniques, and models required to devise the proposed methodology. The section starts by first explaining the proposed GAN model that constitutes the major portion of the proposed methodology and after that describes the training of this model for the task of learning causal relationships between BG, insulin, and carbohydrates data of T1D patients. The conditional generation of BG data for the synthesis of virtual T1D patients using the trained models is demonstrated at the end of the section.

2.1 Sequence-to-Sequence Generative Adversarial Network

This study proposes the use of CGAN, which is a type of GAN, for the generation of BG samples [25]. A CGAN conditions the generated samples on some other type of input. The conditions are provided as labels to the GAN. Such a setup could also be imagined as a translational model where one form of data (labels) is translated into another form of data (generated samples). When slightly altered, such a model can be transformed into a sequence-to-sequence (S2S) or a pixel-to-pixel (P2P) model [26]. For an S2S model, the CGAN is required to translate an input vector into an output vector. For a P2P model, the CGAN deals with image pixel data. The proposed methodology uses the S2S approach to condition the output array of BG values on the input arrays of insulin and carbohydrates. Figure 1 shows the S2S GAN architecture utilized in the proposed methodology for the generation of T1D patients.

As mentioned above, GAN is a hybrid deep architecture constituted of two DNNs working in a zero-sum game to achieve optimization. These DNNs are the discriminator (D) model and the generator (G) model. The D model in our proposed architecture is presented in Figure 2. Just as the name suggests, the D model discriminates between two sets of samples such as the real and generated samples. This model is trained on real samples from the data set and fake samples generated by the generator. For classification purposes, the real samples are labeled as '1' whereas the fake samples as '0'. As seen in Figure 2, the D model is constituted of five 1D convolutional layers of different sizes. 1D convolutional layers are chosen in order to learn the temporal characteristics of the time-dependent data. The input to the D model is a concatenated signal of one sample of PI, one sample of RA, and 18 samples of BG. This combination of signals serves as the input/output configuration required for the S2S GAN. The PI and RA are the input pair whereas BG is the output. The classification task of the D model is to distinguish between the real and synthetic configurations of these three signals. On the other hand, the G model tries to learn the underlying probability of the real data set based on the feedback it receives from the D model. Figure 3 displays the architecture of the G model in our S2S GAN. It takes a three-signal formation of PI, RA, and the latent space (Z) samples as input that are passed through a dense layer before

being reshaped and concatenated. Z is chosen to be a normal distribution. The G model is constituted of a total of four 1D transpose convolutional layers. For the G model, the PI and RA are obtained from the real data and condition the transformation of Z into BG values. Both D and G are trained using the Wasserstein loss scheme. The losses for both D and G are averaged over a mini-batch of data. The complete S2S model is trained using paired data with one PI and one RA value mapped to 18 BG values in each pair.

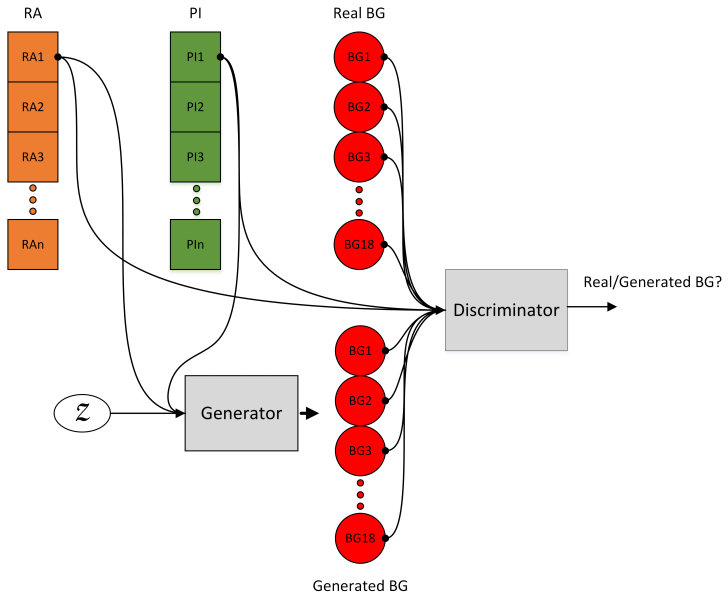


Fig. 1: S2S model used for the generation of blood glucose values conditioned on insulin and carbohydrates. (PI:Plasma Insulin Approximation, RA: Carbohydrates Rate of Appearance, BG: Blood Glucose, Z: Latent Space).

2.1.1 Experimental Data Set

We used data from 27 patients on insulin pump therapy of the Hospital Clínic de Barcelona T1D data set [27]. The data used in our implementation contained the BG profiles of each patient along with the insulin and carbohydrates information. Basal and bolus insulin values were first converted into the patient's PI approximation using Hovorka's insulin pharmacokinetic model [7]. On the other hand, carbohydrate values that were given in grams in the original data set were depicted as the RA using the mixed meal libraries from Ernesto et al [28]. RA of meals that fitted the meal description based on the time of the day the meal was taken and the number of grams of the meal was chosen to represent a particular meal. In the end, the time series obtained from carbohydrates was a combination of the carbohydrates RA of all the meals taken by the patient in a particular time frame.

The inputs and outputs were introduced as pairs to the GAN model. This essentially means that for one input pair there were a total of 18 values of output BG. The shift was achieved using the 'roll' function in python. The training data was normalized before being subjected to training which was achieved using the MinMaxScaler in scikit-learn. This shifted normalized pair data was then used to train the S2S GAN model.

2.2 Training With Recurring Data

Our proposed S2S model was trained on PI, RA, and BG values from a T1D patient cohort. To introduce causality between the inputs and the outputs, a shift of 90 minutes was introduced between the PI/RA pair and BG values. This was done in order to completely capture the effect of insulin and carbohydrates taking place in the BG profile of a T1D patient. Since the sampling rate of the signals was 5 minutes, to introduce a shift of 90 minutes, the shifted pairs were formed by mapping one PI/RA pair value to 18 future BG values. The recurring pairs of data acted as snapshots of the cause-and-effect relationship between them. The entire data set was first converted into groups of 1 PI/RA pair and 18 BG values. A visualization of the shifted pairs of insulin, carbohydrates, and BG can be seen in Figure 4. The batch size was kept at 1 because smaller batch sizes are often associated with training stability and lower generalization error. This enabled the model to learn the causal relationships among the data trickling through each recurrent sample of data. The S2S GAN model was then trained with the shifted samples for a total of 50 epochs because empirical testing suggested this as the optimal value for all the models to achieve convergence. Since the batch size was chosen to be 1, the time taken in training for one epoch directly depended on the total number of samples included in the training data. A leave-one-out scheme was used to train the S2S GAN for the entire cohort. Since there were 27 patients in the cohort, the S2S model was trained 27 times excluding 1 patient each time. A total of 27 trained models were obtained at the end. Each trained model was then used to generate a new T1D patient.

2.3 Synthesis of Virtual T1D Patients

The trained models were then used to generate novel BG samples when provided with unseen PI and RA values from T1D patients. During generation, 1 sample of input PI and RA each produced 18 samples of BG. For a series of insulin and carbohydrate values, the outputs were shifted by one sample for every input pair and then averaged to compute a recurrent output that depicted a causal relationship between the input and the output. This phenomenon can be understood by looking at Figure 5. With the help of this generation technique, a virtual patient was generated for every real patient, utilizing the real patient's PI and RA information. A total of 27 virtual T1D patients were generated.

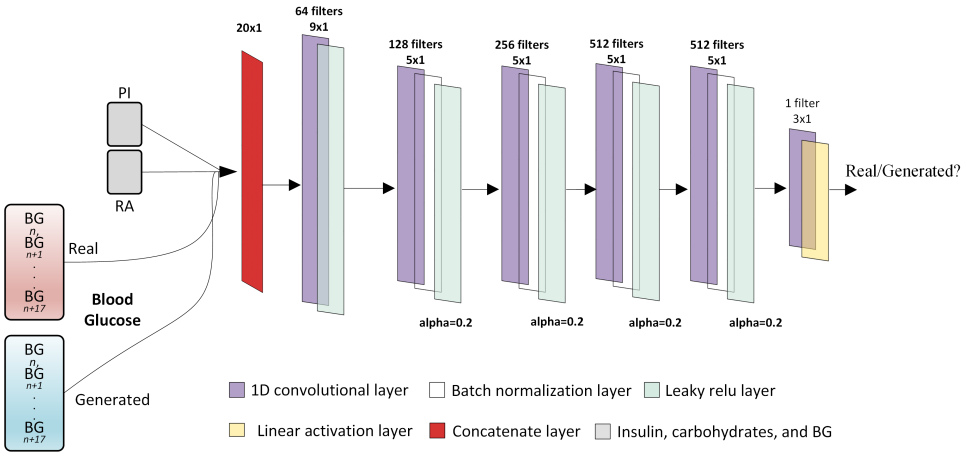


Fig. 2: Graphical depiction of the discriminator model. (PI: Plasma Insulin Approximation, RA: Carbohydrates Rate of Appearance, BG: Blood Glucose)

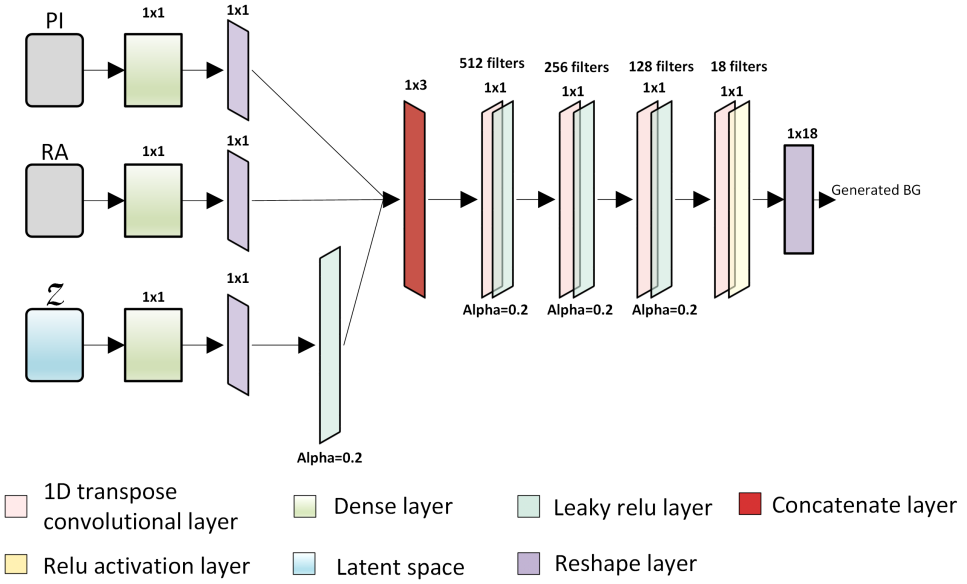


Fig. 3: Graphical depiction of the generator model. (PI: Plasma Insulin Approximation, RA: Carbohydrates Rate of Appearance, Z: Latent Space).

2.3.1 Latent Space Exploration

Since we know from the generation operation of GAN, "which draws a sample from a random number space and then transforms it into the value of choice", latent space is a parameter of great importance in our model. The choice of latent space was a normal distribution in our implementation. As evident from the literature, the quality of generated data of GAN depends on the

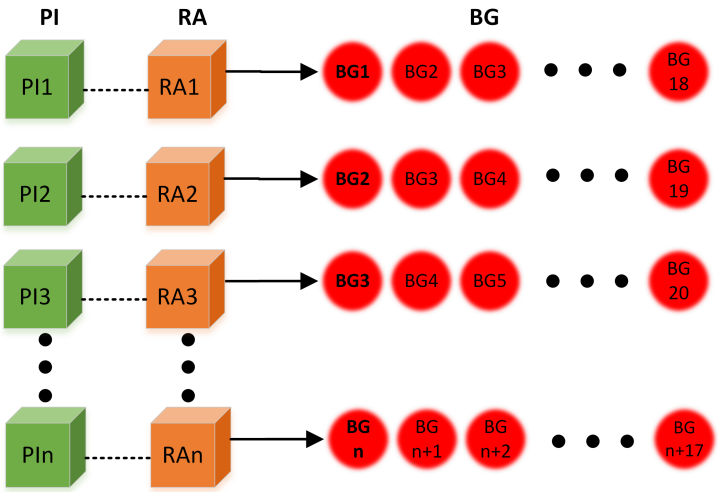


Fig. 4: The shifted plasma insulin approximation/carbohydrates rate of appearance and blood glucose pairs used to train the S2S GAN: each input sample pair of insulin and carbohydrates corresponds to 18 output samples of blood glucose. (PI:Plasma Insulin Approximation, RA: Carbohydrates Rate of Appearance, BG: Blood Glucose).

dimensions of the latent space [29] [30]. Even though there are no standards for latent space dimensions, a size of 100 or 512 is preferred in image generation tasks. For our application, it was observed that latent dimensions significantly smaller than 100 produced plausible results. Moreover, the exploration of latent space was performed during the inference phase when novel BG profiles were generated for each patient. The selection of the model along with the latent space configuration is an important parameter in our methodology since we strive for the generation of BG samples that are as close to the real samples as possible. The latent space exploration was done using vector arithmetic. Along with generating realistic samples, varying effects of latent space exploration were observed. It was observed during this exploration that the variability of the generated BG profile could be controlled by changing the magnitude of the random samples acquired from the latent space. Larger magnitudes tended to produce outputs with a higher coefficient of variation (CV) and vice-versa. It is important to mention here that all the virtual patients were generated using the same latent space configuration.

2.4 The Simulation Environment

The simulation environment shown in Figure 6 was set up for the evaluation of OL and CL insulin delivery strategies in patients with T1D. This also served as the validation phase for the trained DGM. The evaluation was done by observing the glycemic outcomes of the generated virtual patients. These glycemic outcomes include time-in-range results for each of the standard blood glucose

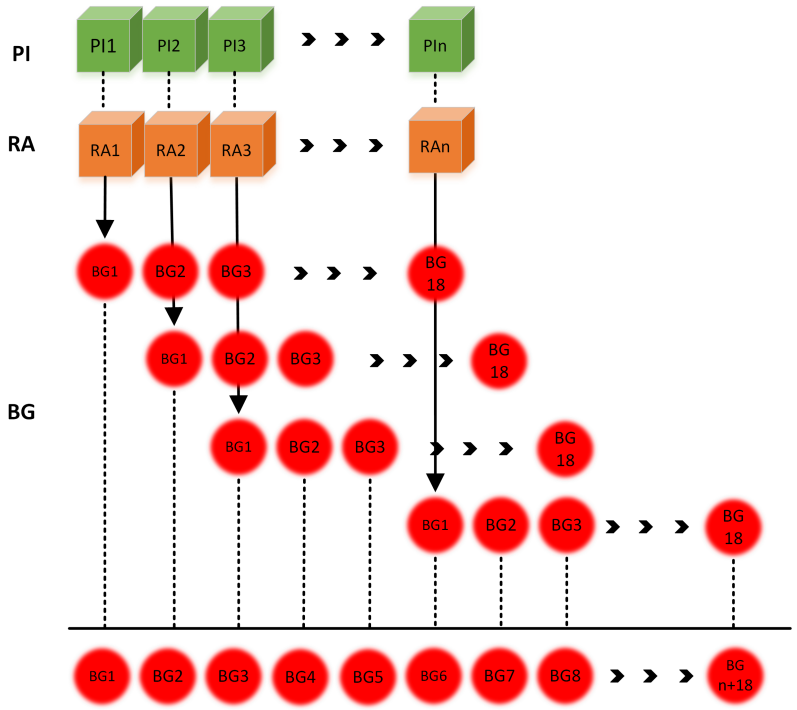


Fig. 5: The phenomena of recurrence in the generated blood glucose samples: each output blood glucose stream is shifted 1 sample for every input sample pair. (PI:Plasma Insulin Approximation, RA: Carbohydrates Rate of Appearance, BG: Blood Glucose).

ranges (<54 mg/dL, 54-69 mg/dL, 70-140 mg/dL, 70-180 mg/dL, 180-250 mg/dL, >250 mg/dL) in T1D management along with means and variability outcomes [31]. In the proposed work, the main reason to opt for two insulin delivery therapies was to check the practicality of the proposed methodology and to further validate the approximation of the trained DGM in terms of causality and exactness. In the future, such therapies could be employed in the proposed simulation environment for the sole purpose of validating the therapy. The T1D patients in the cohort used in this study administered insulin to their bodies using insulin pumps under an OL therapy. The insulin values from these patients, when provided as input to the proposed generative model, produced BG values with outcomes similar to the real patients.

Afterward, a CL insulin delivery strategy was adopted. During this strategy, the generated BG data was subjected to a state-of-the-art controller to emulate the closed-loop behavior of a human pancreas. The controller's control action focused on increasing the time-in-range of the generated BG by adjusting the insulin delivery to the generative model. The CL insulin delivery part of the controller proposed by Beneyto et.al [32] was utilized for this purpose. The feedback control action of the controller is composed of two loops, i.e. an insulin feedback loop which is also referred to as the inner loop comprising of

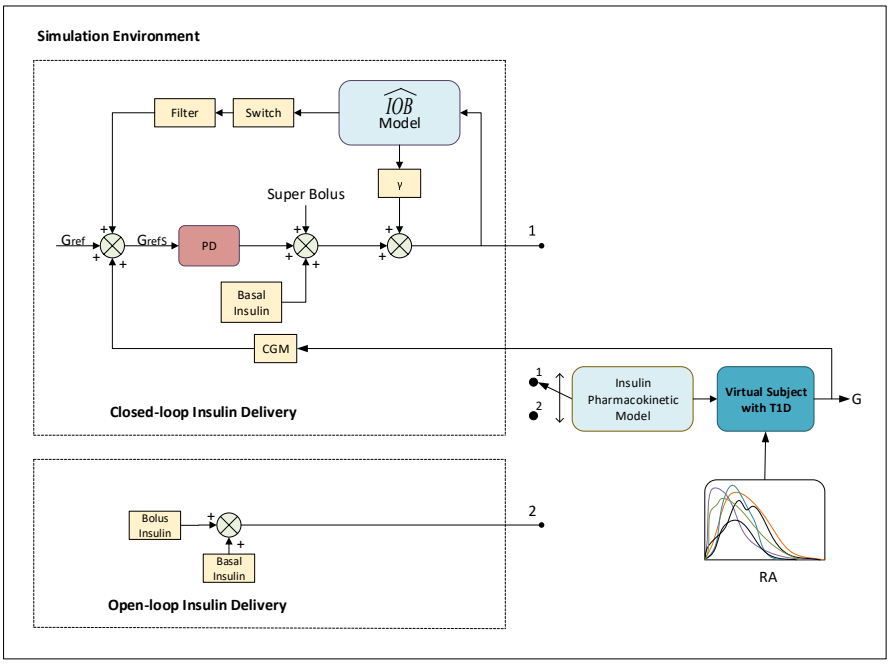


Fig. 6: Simulation environment to evaluate the closed-loop and open-loop insulin delivery systems in virtual T1D subjects. (IOB: Insulin on Board, Gref: Reference Blood Glucose, GrefS: Adjusted Blood Glucose, CGM: Continuous Glucose Monitor, γ: Insulin Feedback Gain Parameter, RA: Carbohydrates Rate of Appearance, PI: Plasma Insulin Approximation).

a proportional-derivative (PD) controller and an outer safety loop with insulin on board constraints and a sliding mode reference conditioning. Three insulin signals constitute the inner control loop of the controller: the basal insulin profile of the patient, the super bolus, and the PD control action. Basal profiles from the actual patients were used in this loop. The outer safety loop is defined to compute the conditions under which the reference glucose $Gref$ needs to be changed. This is done in order to cease insulin infusion to keep the insulin on board (IOB) bounded i.e., IOB $\in [0, \overline{IOB}]$, where $\overline{IOB}$ is the maximum allowed IOB. Correction factor and carbohydrate ratio parameters from the actual patient cohort were used during patient synthesis under CL therapy.

Since the generated BG data is conditioned on both insulin and carbohydrate values, the trained generative model has to have an input for insulin and carbohydrates each in order to generate new BG samples. In the CL therapy case, the carbohydrate data is taken from the actual cohort of patients such as the quantity of meals in carbohydrates units taken by a patient in real life being translated in the RA using the meal library referenced in the methodology section. This is done in order to generate comparable profiles under both OL and CL therapies using the same meal information. Moreover, the insulin input of the generative model is fed by the CL controller. Because the input

insulin and the generated BG depict a causal relationship, altering the amount of insulin alters the output BG values.

To illustrate the dependence of generated BG profiles on the input PI and RA, Figure 7 shows a total of 2 days of BG data for four different patients. It is clear from these plots that the generated BG values take on trajectories guided by the input insulin and carbohydrate values. Similar relationships are observed in the BG data of real-life patients. These results make our proposed model unique in terms of the lifelike behaviors of the generated BG values.

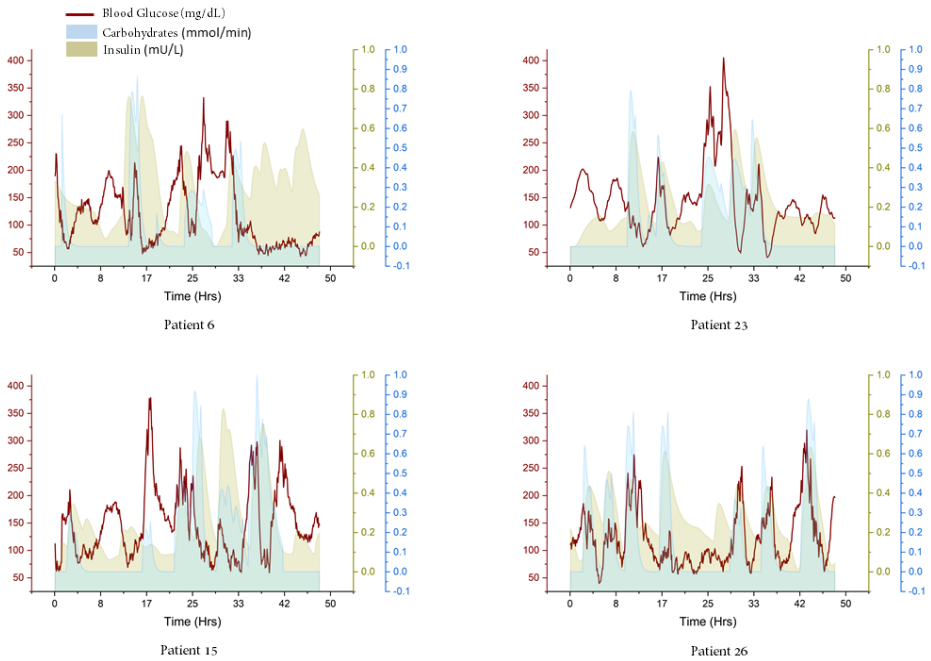


Fig. 7: Two days each of generated blood glucose data of Patient 6, 15, 23, and 26 conditioned on plasma insulin approximation and carbohydrates rate of appearance under open-loop therapy.

3 Results

The results section has been divided into two subsections. The first subsection is dedicated to the results of statistical tests performed to check the statistical similarity of the generated data to the real data. The second subsection contains results about the causal relationships between the input and output data.

3.1 Statistical Similarity

A total of 4 weeks of BG data was generated for each patient under both OL and CL therapies for the purpose of attaining the optimal amount of data required to report glycemic outcomes [33]. After computing the glycemic outcomes for each patient in both real and generated cohorts, the medians of these values along with the inter-quartile ranges (IQR) were computed for both cohorts. It was observed that the medians of both data sets were comparable for each of the glycemic outcomes. The Wilcoxon signed-rank test was used to evaluate statistical similarity by assuming that both the real and generated glycemic outcomes came from the same population of patients. Based on the rejection hypothesis, all the glycemic outcomes qualified the test by demonstrating P-values of 0.05 or more. Table no. 1 presents the results for the generated and real patients along with their P-scores under OL therapy. The obtained statistical scores demonstrated that the glycemic outcomes such as BG time in standardized ranges, mean, CV, maximum value (max), minimum value (min), and standard deviation (STD) all showed significant statistical similarity to those of the real T1D patients. These results confirm the accurate approximation of the T1D cohort by the generative model since the same insulin and carbohydrate inputs yielded almost similar BG outcomes.

Glycemic metrics	Real patients (open-loop therapy)	Generated patients (open-loop therapy)	P-Values
% time CGM < 54	0.43 (0.13 - 0.90)	0.40 (0.0 - 2.89)	0.25
% time CGM 54-69	2.39 (1.24 - 3.65)	2.45 (0.43 - 5.67)	0.50
% time CGM 70-140	40.57 (30.07 - 45.68)	42.80 (30.32 - 54.94)	0.98
% time CGM 70-180	64.02 (55.08 - 68.87)	67.28 (52.85 - 75.70)	0.74
% time CGM 180-250	24.28 (18.79 - 30.08)	19.89 (11.59 - 29.57)	0.63
% time CGM > 250	6.41 (3.45 - 11.43)	7.19 (1.93 - 13.29)	0.39
Mean CGM	157.06 (148.66 - 172.32)	154.20 (127.92 - 178.52)	0.85
STD	56.57 (50.20 - 64.49)	61.49 (49.07 - 70.48)	0.34
% CV	35.53 (32.94 - 39.03)	38.58 (31.81 - 42.60)	0.22
Max	400 (359.50 - 400.00)	400 (398.87 - 400.00)	0.12
Min	40.00 (40.00 - 44.50)	41.75 (40.0 - 55.02)	0.12

Table 1: Statistical comparison of the glycemic metrics of real patients against the generated patients: standard BG ranges in T1D management, average BG (Mean), standard deviation (STD), coefficient of variation (CV), maximum (Max), minimum (Min).

During the CL therapy, the PD controller provided insulin as input to the generative model while the carbohydrate data used was from real patients. The glycemic outcomes from the CL therapy in terms of the medians and IQR of all the metrics for each generated patient are presented in table 2. As per the discussion above, the control action of the controller was designed to increase the percentage time in the 70-180 mg/dL range of generated BG profiles. The glycemic outcomes showed that the proposed model behaved in a

Glycemic metrics	Generated patients (closed-loop therapy)
% time CGM < 54	0.79 (0.07 - 1.88)
% time CGM 54-69	1.85 (0.57 - 4.23)
% time CGM 70-140	35.46 (21.47 - 49.84)
% time CGM 70-180	74.67 (57.03 - 80.70)
% time CGM 180-250	13.99 (8.30 - 32.37)
% time CGM > 250	5.31 (2.45 - 9.29)
Mean CGM	155.13 (143.21 - 173.41)
CV	35.60 (31.92 - 41.01)
STD	58.11 (47.18 - 62.62)
Max	400.00 (400.00 - 400.00)
Min	40.00 (40.00 - 47.20)

Table 2: Glycemic outcomes of the generated patients under closed-loop insulin therapy: standard BG ranges in T1D management, average BG (Mean), standard deviation (STD), coefficient of variation (CV), maximum (Max), minimum (Min).

way similar way to real T1D patients by demonstrating TIR results that were closer to real scenarios. Furthermore, it has been observed in T1D patients that tight glycemic control is often the reason for the occurrence of iatrogenic hypoglycemia [34]. A similar phenomenon was observed in our generated patients under CL therapy with an increase in the percentage time in the level 2 hypoglycemia.

3.2 Causality Analysis

In the context of a simulation environment, the model is expected to respond to the effects of certain inputs in a causal manner. The causal relationship between the inputs and outputs of a model is the basis for validating real-life phenomena using a simulation environment. There are several ways to test for causality in systems. Two of the more famous causality tests are the Granger causality test and convergent cross mapping (CCM). The Granger causality test has been used extensively by researchers for causality analysis in a variety of applications. The intuition behind this test is that if variable X causes variable Y in a system, there should be a model of such a system that improves the prediction of Y after the inclusion of X. This means that X should be separable from the rest of the system's variables. This is, however, a limitation of the Granger test since in many systems that have interacting variables, the information of a variable may not separable from other variables. This poses a problem when causality is checked using the Ganger causality test in highly complex dynamical systems such as the human glucose-insulin system. CCM, on the other hand, is a causality testing methodology that identifies causalities in a system whose variables are inseparable. In addition, CCM can quantify weak to moderate causalities that other causality tests may miss. In the past, several studies have utilized CCM for causal analysis in nonlinear systems [35] [36]. A recent study by Hoda et al. demonstrated the use of CCM for the causality analysis in T1D [37].

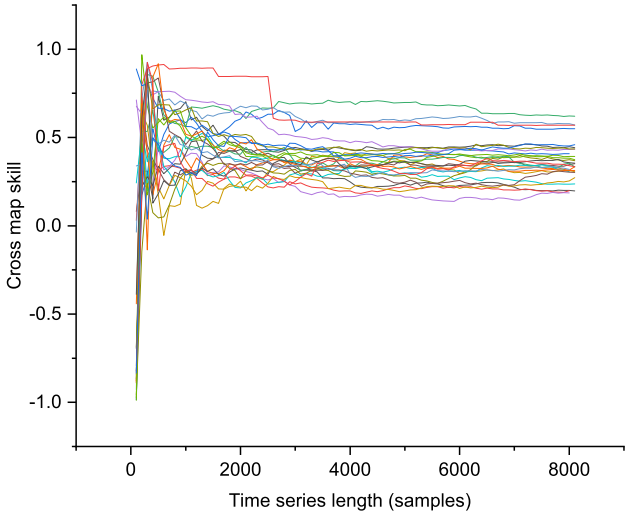


Fig. 8: The cross map skills for the effect of insulin on the generated blood glucose profiles of all 27 patients as a function of time series length.

We have used both the Granger causality test and CCM to demonstrate that the BG profiles generated using our proposed model are causally dependent on input insulin and carbohydrate values. The results of both these tests are provided in Table 3 and 4. For the Granger causality test, it could be upheld from the p-values in Table 3 that there exists an intervention for both insulin and carbohydrates in the generated BG values. On the other hand, as the name suggests, CCM checks for two parameters while establishing causality between two quantities i.e. cross-mapping and convergence. Cross-mapping is measured using the correlation strength whereas convergence is checked by observing the cross map skill against the increasing amount of data. Even though a stronger correlation suggests stronger causation, a relationship is only deemed causal if the correlation converges as the amount of data increases. The causal strengths for individual patients are given in Table 4, whereas the convergence is shown for all the patients in Figure 8 and 9. It could be observed from these figures that for both insulin and carbohydrates the cross map skill converges as the amount of data increases. This confirms the existence of a causal relationship between the generated BG and insulin and carbohydrates. Figure 10 shows the average causal strength of insulin and carbohydrates on the generated BG values obtained using CCM and the average P-values obtained using the Granger causality test.

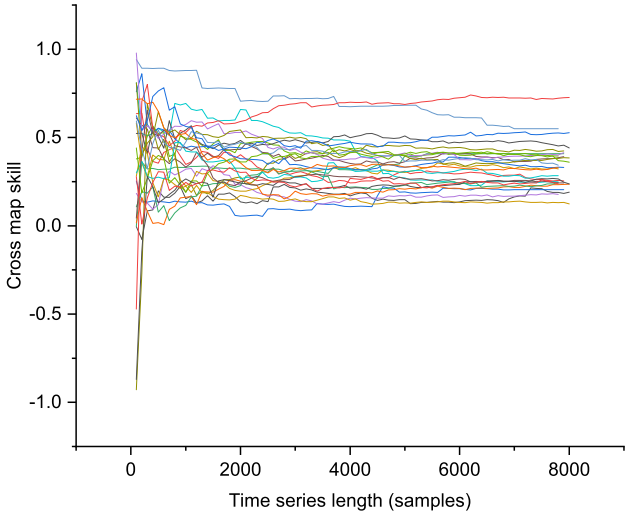


Fig. 9: The cross map skills for the effect of carbohydrates on the generated blood glucose profiles of all 27 patients as a function of time series length.

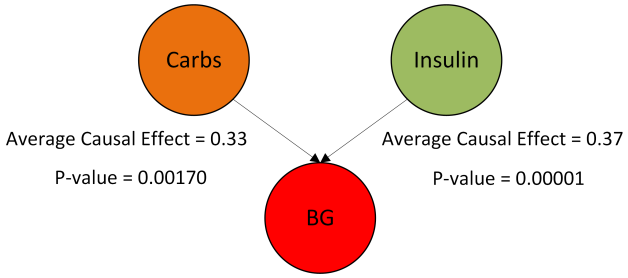


Fig. 10: The average causal effect of carbohydrates and insulin on the generated blood glucose profiles of all 27 patients.

4 Discussion

The aptness of the proposed methodology for the purpose of generating glycemic trends similar to real-life T1D patients has been demonstrated by the results. Approximating the human glucose-insulin system is understood to be the hardest part of a diabetes simulator. Formerly, the techniques used for these approximations were heavily based on mathematical physiological models. Apart from the advantages they offer, these models suffer from shortcomings that may lead to inaccurate approximation. In DGMs, we have a solution for this problem. The ability of DGMs to learn from the distribution of data obtained from a system gives it a significant advantage over the physiological models in a way that it can encompass almost all the characteristics

Patient	Carbohydrates → BG	Insulin → BG
P1	0.00960	0.00001
P2	0.04060	0.00000
P3	0.04140	0.00000
P4	0.00170	0.00000
P5	0.00010	0.00002
P6	0.08220	0.00000
P7	0.03210	0.00000
P8	0.00000	0.00005
P9	0.01260	0.00000
P10	0.00050	0.00001
P11	0.00000	0.00000
P12	0.00060	0.00004
P13	0.00090	0.00000
P14	0.00170	0.00000
P15	0.00000	0.00000
P16	0.00010	0.00000
P17	0.01670	0.00000
P18	0.01690	0.00000
P19	0.00000	0.00000
P20	0.03680	0.00000
P21	0.00070	0.00000
P22	0.00140	0.00005
P23	0.00840	0.00000
P24	0.00000	0.00000
P25	0.00000	0.00000
P26	0.03470	0.00000
P27	0.00280	0.00000
Average	0.00170	0.00001

Table 3: P-values obtained using Granger causality test considering the alternative hypothesis is true.

of the system including the minor behaviors. This makes a wholesome approximation of the system possible. Moreover, since the proposed methodology not only generates T1D patients that are similar to a cohort of real patients but can also demonstrate the behaviors of insulin, carbohydrates, and BG relationships found in real patients, the generated patients can be used in various simulation scenarios. In addition, the causality between insulin, carbohydrates, and glucose generation is shown by generating BG values in the future, which fulfills the benchmark of prediction in simulation environments.

Even though the generated BG profiles are tested for similarity in glycemic outcomes using statistical tests, the generative model is put under further scrutiny using a CL controller to imitate a CL insulin therapy setup. This was done in order to further validate the causal generation of BG values by observing whether the generated BG values demonstrated the same response behavior to the CL therapy as that observed in real patients. Since it was not possible to validate the CL glycemic outcomes for the real patients' cohort used in this study because the patients were under OL insulin therapy, the only alternative way we had was to observe the glycemic outcomes for general trends and

Patient	Carbohydrates → BG	Insulin → BG
P1	0.19	0.20
P2	0.25	0.20
P3	0.20	0.55
P4	0.41	0.62
P5	0.18	0.19
P6	0.32	0.31
P7	0.25	0.24
P8	0.39	0.35
P9	0.38	0.37
P10	0.24	0.33
P11	0.55	0.57
P12	0.40	0.39
P13	0.44	0.44
P14	0.72	0.57
P15	0.33	0.41
P16	0.24	0.33
P17	0.41	0.43
P18	0.12	0.27
P19	0.28	0.31
P20	0.24	0.30
P21	0.42	0.44
P22	0.33	0.31
P23	0.34	0.32
P24	0.36	0.37
P25	0.24	0.35
P26	0.24	0.34
P27	0.53	0.46
Average	0.33	0.37

Table 4: Causality strength values for each patient obtained using CCM.

response behaviors seen under CL therapy patients. It was observed under the CL therapy that the proposed simulator exhibits similar glycemic outcomes as observed in real-life patients under CL therapy [38]. This is in contrast to what is observed in conventional physiological T1D simulators where the glycemic outcomes of patients under CL are unrealistically optimistic. According to the literature, tight glycemic control has often been associated with the occurrence of hypoglycemia and this behavior was noted in the generated patients under CL insulin therapy.

The proposed methodology shows promise in setting up a T1D simulation environment for the generation of novel cohorts of patients, validation of treatment methodologies, and the formation of new therapies. Considering that the BG generation is caused by the input insulin and carbohydrate values, the generation of BG data with glycemic responses of our choice may be made possible with the help of this simulator. Moreover, by exploring and quantifying the latent space configurations the generation of cohorts with desirable characteristics may be ensured. The proposed system may also provide us the leverage to augment data for specific T1D patients in order to obtain more data in lesser time or reduced cost. This is important because the acquisition

of diabetes-related data suffers from the issue of intrapatient variability since the condition of a patient's disease keeps varying with time. Hence, classification/prediction techniques based on data often fail to work efficiently when trained on T1D data collected over a long period of time.

By introducing improved CL control strategies like carbohydrate recommendation, the glycemic outcomes of the generated BG profiles may even further be improved and used in devising various other treatments. Furthermore, the approximation of the DGM may also be improved by conditioning the generation of BG data on more variables such as physical activity, stress, etc. A proficient CL control system along with an accurate T1D patient generation model may make the in-silico trials of different treatment methodologies as close to real scenarios as possible. Moreover, BG profiles generated by a model with a high level of approximation accuracy will be as challenging to the techniques/models based on data to learn from as the BG profiles from real patients. This will ensure the development of more robust models possible. It will also allow designers and practitioners to be more creative having the confidence that the simulated scenarios will stay closer to real scenarios and won't diverge towards the unattainable. Furthermore, the proposed methodology enables us to replicate any cohort. This allows the creation of individualized treatments possible. The intrinsic random nature of the generated data with the proposed model allows the generation of patients with different BG profiles and similar glycemic characteristics. Also, altering the latent space allows the generation of patient cohorts that may have similar glycemic characteristics but different variability outcomes. In the future, the authors of this work are confident to produce the cohorts of choice at will in order to challenge the control system techniques so that they are more robust for real-life scenarios.

As evident from the theory of DNNs, the larger the quantity of training data the better the approximation. Since they are based on DNNs, the same is observed for DGMs. However, it is important to realize that there is no optimal amount of data when it comes to training DNNs to achieve realistic outcomes. In the proposed methodology, realistic BG generation was made possible using 1120 days of data from 27 T1D patients. Nonetheless, it was learned empirically that increasing the amount of training data resulted in a better approximation of the data set. This has also been proven by our prior works on the generation of BG profiles using DGMs [39] [15]. As the glycemic behaviors of T1D patients vary greatly over the course of their lives, in a particular time phase the generation of data replicating a patient's physiology will address the issue of scarce data. In long term, the approximation of a patient's glucose-insulin system, however, will suffer from glycemic variability. This could pose a problem for insilico longitudinal trials. Current evidence shows that DGMs are capable of learning the underlying distribution of any sort of data and are robust to mistaken confidence errors. This suggests that the proposed method could be utilized equally well for any sort of cohort. As patients from a single cohort often exhibit similar characteristics, it is understood that training the

proposed DGM on data from a single T1D cohort led to a good approximation of the cohort. We believe that heterogeneous training data such as data from various different cohorts or radically different patients may affect the approximation adversely. This essentially means that for a DGM-based T1D simulator to replicate the outcomes of a particular cohort of patients, it is best to train it on data from the specific cohort. However, for a more diverse simulator, the training should be performed on data from a diverse set of cohorts. This would, however, mean that the approximation of a particular cohort in the set of cohorts might not be as good.

5 Conclusion

To conclude, this research study presents the very first AI-based T1D simulation environment based on the distribution approximation capability of deep generative models. The proposed simulator employs a sequence-to-sequence generative adversarial network for the generation of synthetic Type 1 diabetes patients and shows realistic results under both open-loop and closed-loop therapies. The generated data display causal relationships between the input insulin and carbohydrate values and the output blood glucose values. Moreover, the data is generated for 90 minutes in the future for each input which introduces predictability in the simulation environment. In the future, the generation of blood glucose data may be conditioned on physical activity and stress, etc. for improved approximations. Moreover, improved control therapies could be integrated with the proposed generative model for the testing and development of new therapeutic strategies in the field of artificial pancreas.

References

[1] Kaizer, J.S., Heller, A.K., Oberkampf, W.L.: Scientific computer simulation review. Reliability Engineering System Safety **138**, 210–218 (2015). https://doi.org/10.1016/j.ress.2015.01.020

[2] Kadota, R., Sugita, K., Uchida, K., Yamada, H., Yamashita, M., Kimura, H.: A mathematical model of type 1 diabetes involving leptin effects on glucose metabolism. Journal of Theoretical Biology **456**, 213–223 (2018). https://doi.org/10.1016/j.jtbi.2018.08.008

[3] Farmer Jr, T., Edgar, T., Peppas, N.: Pharmacokinetic modeling of the glucoregulatory system. Journal of drug delivery science and technology **18**(6), 387 (2008)

[4] Nath, A., Biradar, S., Balan, A., Dey, R., Padhi, R.: Physiological models and control for type 1 diabetes mellitus: A brief review. IFAC-PapersOnLine **51**(1), 289–294 (2018). https://doi.org/10.1016/j.ifacol.2018.05.077. 5th IFAC Conference on Advances in Control and Optimization of Dynamical Systems ACODS 2018

[5] Mansell, E.J., Docherty, P.D., Chase, J.G.: Shedding light on grey noise in diabetes modelling. Biomedical Signal Processing and Control **31**, 16–30 (2017). https://doi.org/10.1016/j.bspc.2016.06.007

[6] Mari, A., Tura, A., Grespan, E., Bizzotto, R.: Mathematical modeling for the physiological and clinical investigation of glucose homeostasis and diabetes. Frontiers in Physiology **11** (2020). https://doi.org/10.3389/fphys.2020.575789

[7] Hovorka, R., Canonico, V., Chassin, L.J., Haueter, U., Massi-Benedetti, M., Federici, M.O., Pieber, T.R., Schaller, H.C., Schaupp, L., Vering, T., *et al.*: Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. Physiological measurement **25**(4), 905 (2004)

[8] Man, C.D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., Cobelli, C.: The uva/padova type 1 diabetes simulator: New features. Journal of Diabetes Science and Technology **8**(1), 26–34 (2014) https://arxiv.org/abs/https://doi.org/10.1177/1932296813514502. https://doi.org/10.1177/1932296813514502. PMID: 24876534

[9] BERGMAN, R.N., URQUHART, J.: The pilot gland approach to the study of insulin secretory dynamics. In: Proceedings of the 1970 Laurentian Hormone Conference, pp. 583–605 (1971). Elsevier

[10] Franco, R., Ferreira de Loza, A., Ríos, H., Cassany, L., Gucik-Derigny, D., Cieslak, J., Henry, D., Olçomendy, L.: Output-feedback sliding-mode controller for blood glucose regulation in critically ill patients affected by type 1 diabetes. IEEE Transactions on Control Systems Technology **29**(6), 2704–2711 (2021). https://doi.org/10.1109/TCST.2020.3046420

[11] Nielsen, M.: A visual proof that neural nets can compute any function. URL: http://neuralnetworksanddeeplearning. com/chap4. html (2016)

[12] Zhou, D.-X.: Universality of deep convolutional neural networks. Applied and Computational Harmonic Analysis **48**(2), 787–794 (2020). https://doi.org/10.1016/j.acha.2019.06.004

[13] Nikzad, M., Movagharnejad, K., Talebnia, F.: Comparative study between neural network model and mathematical models for prediction of glucose concentration during enzymatic hydrolysis. International Journal of Computer Applications **56**(1) (2012)

[14] Nalisnick, E., Matsukawa, A., Teh, Y.W., Gorur, D., Lakshminarayanan, B.: Do deep generative models know what they don't know? arXiv preprint arXiv:1810.09136 (2018)

[15] Noguer, J., Contreras, I., Mujahid, O., Beneyto, A., Vehi, J.: Generation of individualized synthetic data for augmentation of the type 1 diabetes data sets using deep learning models. Sensors **22**(13) (2022). https://doi.org/10.3390/s22134944

[16] Thambawita, V., Isaksen, J.L., Hicks, S.A., Ghouse, J., Ahlberg, G., Linneberg, A., Grarup, N., Ellervik, C., Olesen, M.S., Hansen, T., *et al.*: Deepfake electrocardiograms using generative adversarial networks are the beginning of the end for privacy issues in medicine. Scientific reports **11**(1), 1–8 (2021)

[17] Marouf, M., Machart, P., Bansal, V., Kilian, C., Magruder, D.S., Krebs, C.F., Bonn, S.: Realistic in silico generation and augmentation of single-cell rna-seq data using generative adversarial networks. Nature communications **11**(1), 1–12 (2020)

[18] Festag, S., Denzler, J., Spreckelsen, C.: Generative adversarial networks for biomedical time series forecasting and imputation. Journal of Biomedical Informatics **129**, 104058 (2022). https://doi.org/10.1016/j.jbi.2022.104058

[19] Xu, J., Li, H., Zhou, S.: An overview of deep generative models. IETE Technical Review **32**(2), 131–139 (2015) https://arxiv.org/abs/https://doi.org/10.1080/02564602.2014.987328. https://doi.org/10.1080/02564602.2014.987328

[20] Wan, C., Jones, D.T.: Protein function prediction is improved by creating synthetic feature samples with generative adversarial networks. Nature Machine Intelligence **2**(9), 540–550 (2020)

[21] Choudhury, S., Moret, M., Salvy, P., Weilandt, D., Hatzimanikatis, V., Miskovic, L.: Reconstructing kinetic models for dynamical studies of metabolism using generative adversarial networks. bioRxiv (2022)

[22] Dieng, A.B., Kim, Y., Rush, A.M., Blei, D.M.: Avoiding latent variable collapse with generative skip models. In: Chaudhuri, K., Sugiyama, M. (eds.) Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 89, pp. 2397–2405. PMLR, ??? (2019). https://proceedings.mlr.press/v89/dieng19a.html

[23] Ruthotto, L., Haber, E.: An introduction to deep generative modeling. GAMM-Mitteilungen **44**(2), 202100008 (2021)

[24] Mincu, D., Roy, S.: Developing robust benchmarks for driving forward ai innovation in healthcare. Nature Machine Intelligence, 1–6 (2022)

[25] Mirza, M., Osindero, S.: Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014)

[26] Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134 (2017)

[27] Ahmad, S., Ramkissoon, C.M., Beneyto, A., Conget, I., Giménez, M., Vehi, J.: Generation of virtual patient populations that represent real type 1 diabetes cohorts. Mathematics **9**(11) (2021)

[28] Estremera, E., Cabrera, A., Beneyto, A., Vehi, J.: A simulator with realistic and challenging scenarios for virtual t1d patients undergoing csii and mdi therapy. Journal of Biomedical Informatics **132**, 104141 (2022). https://doi.org/10.1016/j.jbi.2022.104141

[29] Marin, I., Gotovac, S., Russo, M., Božić-Štulić, D.: The effect of latent space dimension on the quality of synthesized human face images. Journal of Communications Software and Systems **17**(2), 124–133 (2021)

[30] Into the latent space. Nature Machine Intelligence **2**(3), 151–151 (2020). https://doi.org/10.1038/s42256-020-0164-7

[31] Battelino, T., Alexander, C.M., Amiel, S.A., Arreaza-Rubin, G., Beck, R.W., Bergenstal, R.M., Buckingham, B.A., Carroll, J., Ceriello, A., Chow, E., Choudhary, P., Close, K., Danne, T., Dutta, S., Gabbay, R., Garg, S., Heverly, J., Hirsch, I.B., Kader, T., Kenney, J., Kovatchev, B., Laffel, L., Maahs, D., Mathieu, C., Mauricio, D., Nimri, R., Nishimura, R., Scharf, M., Del Prato, S., Renard, E., Rosenstock, J., Saboo, B., Ueki, K., Umpierrez, G.E., Weinzimer, S.A., Phillip, M.: Continuous glucose monitoring and metrics for clinical trials: an international consensus statement. The Lancet Diabetes Endocrinology (2022). https://doi.org/10.1016/S2213-8587(22)00319-9

[32] Beneyto, A., Bertachi, A., Bondia, J., Vehi, J.: A new blood glucose control scheme for unannounced exercise in type 1 diabetic subjects. IEEE Transactions on Control Systems Technology **28**(2), 593–600 (2020). https://doi.org/10.1109/TCST.2018.2878205

[33] Herrero, P., Alalitei, A., Reddy, M., Georgiou, P., Oliver, N.: Robust determination of the optimal continuous glucose monitoring length of intervention to evaluate long-term glycemic control. Diabetes Technology & Therapeutics **23**(4), 314–319 (2021) https://arxiv.org/abs/https://doi.org/10.1089/dia.2020.0387. https://doi.org/10.1089/dia.2020.0387. PMID: 33064025

[34] Cryer, P.E.: Glycemic Goals in Diabetes: Trade-off Between Glycemic Control and Iatrogenic Hypoglycemia. Diabetes **63**(7), 2188–2195 (2014) https://arxiv.org/abs/https://diabetesjournals.org/diabetes/article-pdf/63/7/2188/578013/2188.pdf. https://doi.org/10.2337/db14-0059

[35] Ma, H., Aihara, K., Chen, L.: Detecting causality from nonlinear dynamics with short-term time series. Scientific reports **4**(1), 1–10 (2014)

[36] Verma, A.K., Garg, A., Xu, D., Bruner, M., Fazel-Rezai, R., Blaber, A.P., Tavakolian, K.: Skeletal muscle pump drives control of cardiovascular and postural systems. Scientific reports **7**(1), 1–8 (2017)

[37] Nemat, H., Khadem, H., Elliott, J., Benaissa, M.: Causality analysis in type 1 diabetes mellitus with application to blood glucose level prediction. Computers in Biology and Medicine **153**, 106535 (2023). https://doi.org/10.1016/j.compbiomed.2022.106535

[38] Breton, M.D., Kovatchev, B.P.: One year real-world use of the control-iq advanced hybrid closed-loop technology. Diabetes Technology & Therapeutics **23**(9), 601–608 (2021) https://arxiv.org/abs/https://doi.org/10.1089/dia.2021.0097. https://doi.org/10.1089/dia.2021.0097. PMID: 33784196

[39] Mujahid, O., Contreras, I., Beneyto, A., Conget, I., Giménez, M., Vehi, J.: Conditional synthesis of blood glucose profiles for t1d patients using deep generative models. Mathematics **10**(20) (2022). https://doi.org/10.3390/math10203741

Acknowledgements

This work was partially supported by the Spanish Ministry of Universities, the European Union through Next GenerationEU (Margarita Salas), the Spanish Ministry of Science and Innovation through grant PID2019-107722RB-C22/AEI/10.13039/501100011033, PID2020-117171RA-I00 funded by MCIN/AEI/10.13039/501100011033 and the Government of Catalonia under 2017SGR1551 and 2020 FI_B 00965.

Author contributions

All authors edited and reviewed the final manuscript. O.M., I.C., and J.V. conceptualized and designed this research work. O.M. and A.B. conducted the analysis and contributed to manuscript preparation. I.C. and J.V. were involved with funding, interpretation, manuscript preparation, and overall supervision of the study.

Competing interests

The authors declare no competing interests.

CHAPTER 3

DISCUSSION

The works presented in this thesis revolve around modeling in T1D using data from patients with T1D. This task is achieved by leveraging the distribution learning capability of deep generative models (DGMs). The thesis comprises three main publications that are presented as a compendium and one coauthored publication investigating ML models' potential to learn glucose-insulin dynamics from data obtained from T1D patients. The advances in diabetes technology have made the acquisition of diabetes-related data in large quantities possible, which in turn has prompted a significant shift toward data-driven modeling. Data-driven models are becoming a thing of choice for engineers because of the several advantages they offer over conventional mathematical models. The research work conducted in this compendium of publications has cashed in on the benefits of data-driven models and shown its advantages over other methods.

The first publication in the compendium is a systematic review of the state-of-the-art data-driven models for the task of hypoglycemia prediction. Evidently, the high complexity of glycemic trends has led to the extensive use of data-driven models in recent years for the prediction of adverse glycemic events (Gadaleta et al., 2019) (Contreras and Vehi, 2018).

Hypoglycemia is an adversity experienced by patients with T1D, caused by a decrease in blood glucose (BG) below a certain level and characterized by dizziness, reduction in cognitive ability, shaking, sweating, and in extreme cases, death. The inability of patients to act while experiencing a hypoglycemic episode leads to the fear of hypoglycemia and significantly reduces the life quality of patients (Przezak et al., 2022). All these factors make the accurate prediction of hypoglycemia extremely important. The findings of this review put forward a comprehensive picture of the trends and challenges associated with data-driven modeling for hypoglycemia prediction. It was identified from the results of the review that the availability of suitable data is a major performance constraint in data-driven models. Apart from the insufficiency of desirable instances for training supervised learning models, clinical data suffers from issues like missing values, and time-dependent variations that induce bias in the models. To counter the problem of data availability, an exploration of the data synthesis techniques was carried out. Due to their ability to generate realistic data, DGMs were pinned down as potential tools for data synthesis. Further empirical evidence showed the aptness of DGMs for T1D data and T1D model approximation. The following sections provide an elaborate discussion of the use of DGMs for modeling purposes in T1D.

3.1 Deep Generative Models for Modeling in T1D

DGMs have demonstrated their effectiveness in efficiently learning complex probability distributions from data and generating novel instances of data based on the learned distributions. This makes them a suitable tool for approximating the probability distribution of a highly complex data distribution like the human BG profile. This hypothesis is tested in the publications presented in this compendium. A data augmentation strategy is devised using generative adversarial networks to enhance the performance of a hypoglycemia prediction model in (Noguer et al., 2022). The results of this publication show that by augmenting personalized BG profiles of T1D patients with data from GANs, the classification accuracy of the prediction model

improves. It is also observed during this study that the model performance stops improving after a specific amount of data is added to the data set, whereas, this amount of data varies for distinct data sets. Data set augmentation presents an apt solution for solving problems with incomplete and insufficient data. Moreover, assuming any rate of improvement at all, the generated data from GANs may completely replace real data to be used in data-driven models in the future. This provides a potential solution for issues like data privacy and security.

The next publication in this compendium is an extension of the earlier work and introduces conditional variables for the generation of BG values using a conditional GAN (CGAN) (Mujahid et al., 2022). This paper puts forward a methodology based on a type of CGAN known as the pixel-to-pixel (P2P) GAN that performs the task of image-to-image translation. The idea is to generate novel BG images by providing input insulin images. The generated images are then converted into numerical BG values using a sophisticated translation mechanism. Nonetheless, great emphasis should be put on the accuracy of the translation mechanism employed in a methodology like this, since, a flawed translation mechanism may lead to errors and biases that can compromise the accuracy of the generative model. A novel translation mechanism is proposed in this study that has been shown to work efficiently on two distinct data sets. The results of this work confirmed the suitability of DGMs for the generation of diabetes-related data that is conditioned on other variables. The results also validated that not only does the DGM learn the underlying probability distributions of the BG and insulin data but it is also capable of learning the relationship between the two unique probability distributions. The translation quality of the CGANs provides the basis for the hypothesis of learning the dynamic relationships between inputs and outputs of a glucose-insulin system. Moreover, it was also hypothesized from this work that increasing the number of conditional variables may improve the approximation of the system and lead to more realistic scenarios. This hypothesis was tested in the final work presented in this compendium.

Modeling from data comes with its own challenges. DNNs have been shown to have a

strong correlation between their performance and the amount of data available for training. The data-hungry nature of DNNs presents a potential paradox when it comes to the task of data synthesis using DGMs, as it is necessary to have a sufficient quantity of data in the first place to obtain an acceptable approximation of the system. However, the question of what constitutes sufficient data is ambiguous and lacks a definitive answer. Literature states that the quantity of data required to learn varies depending on the specific application and the nature of the data itself. Moreover, the quality of the data is a critical factor in determining the performance of DNN models. It was observed empirically during the course of this work that increasing the amount of data resulted in a better approximation of the data set. It was also learned during the course of this research that homogeneous data from one cohort might be easier for the DGM to learn from. As the variability in the data increases, the learning capability of the DGMs decreases. While heterogeneous data sets may represent an efficient means of increasing the amount of training data available, it is essential that sufficient instances of each cohort be present to improve the model's learning capability. For a comprehensive approximation of the glucose-insulin system, the proposed models need to be trained on a large amount of versatile data. If the aim of the task is to approximate a populational model, a balanced data set of the entire population with every individual sufficiently represented should be used for training; however, in the case of a personalized model approximation, data representative of a particular patient should be used. It is important to consider that since diabetes-related is acquired over a period of time, it is susceptible to variations in time, and using such a data set for training a DGM may affect the performance of the generative model adversely. The longitudinal variability of diabetes-related data, however, provides the rationale for the use of DGMs to generate data with desired characteristics.

As reported in the literature, there are several other challenges associated with DGMs (Chen, 2021). GANs, for one, suffer from mode collapse, which is the problem of generating the same instances each time. Moreover, the performance of GANs is also affected by the failure to

converge. One or both of these problems were encountered during the implementation of the models presented in this thesis. Different strategies were adopted to tackle these problems. Wasserstein loss was used in the implementation of these GANs in order to tackle the problem of instability and non-convergence. Furthermore, the conditional generation of BG data addressed the problem of mode collapse. In addition, different normalization techniques accompanied by hyperparameter exploration were performed in order to achieve the best possible results.

It is understood that the majority of the trendy GANs are designed to generate visual and aural data and hence, the evaluation measures they use are perceptive (Borji, 2022). This essentially means that anything that appears good is good enough. However, in our application, the quality of the data cannot be based solely on perceptual metrics. Hence, the validation of our proposed methodologies is carried out in two phases, namely numerical validation, which is based on statistical methods, and clinical validation, which is based on identifying the physiological relationships embedded in the generated data. In numerical validation, the idea is to check the statistical similarity of generated data against the real data by testing a statistical hypothesis using some kind of paired statistical test, whereas, in clinical validation, the plausibility of the input/output relationship is checked against the real data. In both validation phases, the results confirmed that the generated instances were realistic and plausible.

3.2 T1D Simulations Using Deep Generative Models

The final work in this compendium provides a framework for setting up a T1D simulation environment using DGMs. Building on the work conducted previously, the translation of one variable to another variable is replaced by the translation of two variables (carbohydrates and insulin) into one variable (BG). This task is achieved by using a conditional sequence-to-sequence (S2S) GAN that is trained on insulin, carbohydrates, and BG data from a cohort of 27 patients with T1D. The trained model is then integrated with a state-of-the-art closed-loop (CL) controller to establish a simulation environment capable of emulating both open-

loop (OL) and CL insulin therapies. The results demonstrate that the proposed simulation environment produces more realistic outcomes when compared to traditional mathematical models, highlighting the potential of DGM-based simulations for T1D research.

In the context of a simulation environment, the model is expected to respond to the effects of certain inputs in a causal manner. This research work took upon the task of inculcating causality in the S2S model for realistic BG data generation. The simulation environment presented in this thesis utilizes the idea of recurrence by introducing shifted input/output pairs to the generative model in order to approximate the glucose-insulin dynamics of a cohort of T1D patients. Since the output BG values are caused by input insulin and carbohydrate values, desired glycemic responses can be obtained by using this methodology. This was confirmed by comparing the OL and CL outcomes of the model with real-life clinical outcomes of patients under similar therapies. It was observed that the model responses under OL and CL therapies closely resembled the real-life responses of patients with T1D. This implies that the model can facilitate the generation of highly realistic scenarios, thus affording designers greater freedom to execute their modeling tasks with enhanced creativity. Furthermore, these models possess the capacity to simulate complex scenarios, which may aid in the development of robust CL methodologies for use in automated insulin delivery systems. Even though the CL methodology employed in this work was solely based on CL insulin delivery, given that the model takes both insulin and carbohydrate as inputs, other sophisticated CL strategies like carbohydrate recommendation could be assessed using this model.

The proposed methodology also gives us the ability to replicate any cohort and devise individualized treatments. Building on the idea of individualized treatments, the proposed DGM-based techniques could be used in forming complete digital twins of patients with T1D. These digital twins could then be incorporated into educational platforms and recommendation systems to inform the patients better and enable them to make smart decisions about their condition. By enabling the patients to observe how their glycemic profile reacts to various

conditioning variables and giving them the liberty to control these variables ensures trust and eliminates the fear of unknown adversities in the patients.

The stochastic nature of the generative models also enables us to generate distinct BG profiles with similar glycemic trends and outcomes, hence, allowing the synthesis of a cohort of patients with similar characteristics. These models may also facilitate the generation of patients with similar glycemic profiles but different variability outcomes. This trait was investigated during the latent space exploration of the GANs used in this research work. The ability to control the variability of the generated glycemic profiles may be significant in testing various control therapies. The preceding discussion suggests the aptness of DGM-based T1D simulators for synthesizing novel cohorts of our choice, validation of treatment methodologies, and the formation of new therapies. Furthermore, empirical evidence suggests that the approximation of DGM models in T1D may further be improved by improving the amount and quality of data used to train these models and incorporating other conditioning variables. On top of that, the advancements in model architecture and hardware resources only suggest promising things in the future.

CHAPTER 4

CONCLUSIONS

This thesis aims to investigate data-driven models in the context of T1D, with a particular focus on their utility in synthesizing realistic T1D data and scenarios. To this end, a simulation environment has been established leveraging the aforementioned models, and the effectiveness of such simulators is demonstrated.

4.1 Contributions

The primary contributions of this research endeavor are outlined as follows:

- **Literature review and analysis of data-driven models:** A literature review of the state-of-the-art of data-driven models for the prediction of hypoglycemia is conducted. This review provides an analysis of data-driven models with regard to the data used for training, the prediction horizon, and the type of model itself.
- **Conditional synthesis of BG values:** A P2P and an S2S GAN architecture is proposed for the synthesis of realistic BG samples using plasma insulin approximation and carbohydrates rate of appearance as the conditioning variables. A part of this work also yielded

a translation mechanism for the conversion of generated images into numerical BG time series.

- **Causal generation of BG values:** A novel training scheme is proposed in which the training data is provided as shifted pairs of input insulin/carbohydrates and output BG values. A shift of 90 minutes is introduced between the pairs so that the effect of insulin and carbohydrates taking place in the output BG values can be captured. This enabled the trained model to respond causally to the input values. The causality was then tested using multiple causality tests.

- **T1D simulation environment:** A simulation environment based on generative models for the generation of T1D patients is presented for the first time. The models are coupled with a state-of-the-art CL controller to develop a simulator with OL and CL insulin therapies. The results showed that simulated scenarios mimicked real T1D patients under both OL and CL therapies.

- **Generation of realistic T1D patients:** A cohort of realistic virtual T1D patients is synthesized using the T1D simulation environment proposed in this work. The results showed that generated T1D patients were statistically similar to the real BG values. Moreover, it was noted visually that the generated BG values took on trajectories defined by the input insulin and carbohydrate values.

4.2 Future Work

Researchers may take on one of the following directions as future work in the context of the research work presented in this thesis.

- Incorporating more conditional variables in the conditional DGMs to further improve the approximation of the T1D models. For instance, variables such as physical activity, stress,

and external factors like weather, etc. may pave the way for more realistic synthesized scenarios.

- The techniques proposed in this thesis could be utilized for creating personalized treatment regimens and with sufficient data and conditioning variables, the models could be trained to imitate a particular patient, hence, setting the scene for a digital twin.

- Textual information could also be used as a conditional variable for the generation of diabetes-related data conditioned on text prompts. This could enable the user to generate data with the desired characteristics written in the text. Such models could enable the translation of clinical study protocols in a cohort that fulfills the criteria defined in a given protocol.

- The proposed models could be incorporated into a web-based application for real-time simulation of patients' glycemic profiles. An interactive GUI will enable patients to monitor and predict the effects of their lifestyles and behaviors on their glycemic profiles. Such applications could be used to educate patients with T1D and be integrated into other educational approaches, such as serious games.

BIBLIOGRAPHY

Alam, U., Asghar, O., Azmi, S. and Malik, R. A. (2014), 'Chapter 15 - general aspects of diabetes mellitus', **126**, 211–222.
URL: *https://www.sciencedirect.com/science/article/pii/B9780444534804000151*

Beck, R. W., Bergenstal, R. M., Laffel, L. M. and Pickup, J. C. (2019), 'Advances in technology for management of type 1 diabetes', *The Lancet* **394**(10205), 1265–1273.
URL: *https://www.sciencedirect.com/science/article/pii/S0140673619311420*

Bergman, R. N. and Urquhart, J. (1971), The pilot gland approach to the study of insulin secretory dynamics, *in* 'Proceedings of the 1970 Laurentian Hormone Conference', Elsevier, pp. 583–605.

Borji, A. (2022), 'Pros and cons of gan evaluation measures: New developments', *Computer Vision and Image Understanding* **215**, 103329.
URL: *https://www.sciencedirect.com/science/article/pii/S1077314221001685*

Chen, H. (2021), Challenges and corresponding solutions of generative adversarial networks (gans): a survey study, *in* 'Journal of Physics: Conference Series', Vol. 1827, IOP Publishing, p. 012066.

Contreras, I. and Vehi, J. (2018), 'Artificial intelligence for diabetes management and decision support: literature review', *Journal of medical Internet research* **20**(5), e10775.

Coutinho-Almeida, J., Rodrigues, P. P. and Cruz-Correia, R. J. (2021), Gans for tabular healthcare data generation: A review on utility and privacy, *in* 'Discovery Science: 24th International Conference, DS 2021, Halifax, NS, Canada, October 11–13, 2021, Proceedings 24', Springer, pp. 282–291.

Digitale, A. E., Leggett, A. H. and Armitage, A. H. (2019), 'New research shows how to keep diabetics safer during sleep'.
URL: *https://scopeblog.stanford.edu/2014/05/08/new-research-keeps-diabetics-safer-during-sleep/*

Dovc, K. and Battelino, T. (2020), 'Closed-loop insulin delivery systems in children and adolescents with type 1 diabetes', *Expert Opinion on Drug Delivery* **17**(2), 157–166.

Engelhardt, B., Frőhlich, H. and Kschischo, M. (2016), 'Learning (from) the errors of a systems biology model', *Scientific reports* **6**(1), 1–9.

Fox, B. Q., Benjamin, P. F., Aqeel, A., Fitts, E., Flynn, S., Levine, B., Maslak, E., Milner, R. L., Ose, B., Poeschla, M., Ray, M., Serino, M., Shah, S. S. and Close, K. L. (2021), 'Continuous Glucose Monitoring Use in Clinical Trials for On-Market Diabetes Drugs', *Clinical Diabetes* **39**(2), 160–166.
URL: *https://doi.org/10.2337/cd20-0049*

Gadaleta, M., Facchinetti, A., Grisan, E. and Rossi, M. (2019), 'Prediction of adverse glycemic events from continuous glucose monitoring signal', *IEEE Journal of Biomedical and Health Informatics* **23**(2), 650–659.

Gonzalez-Abril, L., Angulo, C., Ortega, J. A. and Lopez-Guerra, J.-L. (2022), 'Statistical validation of synthetic data for lung cancer patients generated by using generative adversarial networks', *Electronics* **11**(20).
URL: *https://www.mdpi.com/2079-9292/11/20/3277*

Harline, K., Martínez-Gómez, J., Specht, C. D. and Roeder, A. H. K. (2021), 'A life cycle for modeling biology at different scales', *Frontiers in Plant Science* **12**.
URL: *https://www.frontiersin.org/articles/10.3389/fpls.2021.710590*

Hovorka, R., Canonico, V., Chassin, L. J., Haueter, U., Massi-Benedetti, M., Federici, M. O., Pieber, T. R., Schaller, H. C., Schaupp, L., Vering, T. et al. (2004), 'Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes', *Physiological measurement* **25**(4), 905.

Liu, J., Ren, Z.-H., Qiang, H., Wu, J., Shen, M., Zhang, L. and Lyu, J. (2020), 'Trends in the incidence of diabetes mellitus: results from the global burden of disease study 2017 and implications for diabetes mellitus prevention', *BMC public health* **20**, 1–12.

Lovic, D., Piperidou, A., Zografou, I., Grassos, H., Pittaras, A. and Manolis, A. (2020), 'The growing epidemic of diabetes mellitus', *Current vascular pharmacology* **18**(2), 104–109.

Man, C. D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B. and Cobelli, C. (2014), 'The uva/padova type 1 diabetes simulator: New features', *Journal of Diabetes Science and Technology* **8**(1), 26–34.
PMID: 24876534.
URL: *https://doi.org/10.1177/1932296813514502*

Mujahid, O., Contreras, I., Beneyto, A., Conget, I., Giménez, M. and Vehi, J. (2022), 'Conditional synthesis of blood glucose profiles for t1d patients using deep generative models', *Mathematics* **10**(20).
URL: *https://www.mdpi.com/2227-7390/10/20/3741*

Nalisnick, E., Matsukawa, A., Teh, Y. W., Gorur, D. and Lakshminarayanan, B. (2018), 'Do deep generative models know what they don't know?', *arXiv preprint arXiv:1810.09136* .

Nath, A., Biradar, S., Balan, A., Dey, R. and Padhi, R. (2018), 'Physiological models and control for type 1 diabetes mellitus: A brief review', *IFAC-PapersOnLine* **51**(1), 289–294. 5th IFAC Conference on Advances in Control and Optimization of Dynamical Systems ACODS 2018.
URL: *https://www.sciencedirect.com/science/article/pii/S2405896318302416*

Newton, D. (2019), 'Generative deep learning in architectural design', *Technology| Architecture+ Design* **3**(2), 176–189.

Nikzad, M., Movagharnejad, K. and Talebnia, F. (2012), 'Comparative study between neural network model and mathematical models for prediction of glucose concentration during enzymatic hydrolysis', *International Journal of Computer Applications* **56**(1).

Noguer, J., Contreras, I., Mujahid, O., Beneyto, A. and Vehi, J. (2022), 'Generation of individualized synthetic data for augmentation of the type 1 diabetes data sets using deep learning models', *Sensors* **22**(13).
URL: *https://www.mdpi.com/1424-8220/22/13/4944*

Przezak, A., Bielka, W. and Molęda, P. (2022), 'Fear of hypoglycemia—an underestimated problem', *Brain and Behavior* **12**(7), e2633.

van den Boom, L., Karges, B., Auzanneau, M., Rami-Merhar, B., Lilienthal, E., von Sengbusch, S., Datz, N., Schröder, C., Kapellen, T., Laimer, M., Schmid, S. M., Müller, H., Wolf, J. and Holl, R. W. (2019), 'Temporal Trends and Contemporary Use of Insulin Pump Therapy and Glucose Monitoring Among Children, Adolescents, and Adults With Type 1 Diabetes Between 1995 and 2017', *Diabetes Care* **42**(11), 2050–2056.