AI and Machine Learning in Diabetes Management: Opportunity, Status, and Challenges

Marwa Qaraqe, Madhav Erraguntla, and Darpit Dave

Abstract Diabetes is a costly and burdensome metabolic disorder that occurs due to the elevated blood glucose levels. Poorly managed diabetes can lead to serious and life-threatening health complications. A person's glycated hemoglobin (HbA1C or A1C) measures the average blood glucose for the past 2–3 months by measuring how much glucose is bound to the hemoglobin cells in the blood. The HbA1C is used both to diagnose diabetes and assess the effectiveness of a person's management plan. Developing a model that can accurately predict a person's future HbA1C 2–3 months in advance holds immense potential for preventative and tailored medical care. With the new era of artificial intelligence (AI) it becomes increasing evident that some of unanswered health issues can be unlocked by leveraging on advanced AI and machine learning algorithms. In addition, sudden plummeted or elevated blood glucose levels also pose serious and life-threating consequences to diabetic people. The development of a detection and prediction model capable of detecting or predicting instances of hyperglycemia or hypoglycemia using new CGM technology is critical. This chapter discusses the consequences of poorly managed diabetes and how a more personalized treatment plan for diabetes may lie in the detection of hyper/hypoglycemic events and the prediction of a person's HbA1C using their current blood glucose values.

1 What is Diabetes?

Diabetes is a disorder that affects millions of people around the world. Diabetes impairs the body's ability to process blood glucose (blood sugar). The body breaks down the carbohydrates eaten into blood glucose which is then used to generate energy. Insulin is a hormone that the body needs to get glucose from the bloodstream into the cells of the body. Persons with diabetes are unable to produce insulin or do

M. Qaraqe (✉) · M. Erraguntla · D. Dave
Division of Information and Computing Technology, Hamad Bin Khalifa University, Doha, Qatar
e-mail: mqaraqe@hbku.edu.qa

not use it efficiently. Without careful management, diabetes can lead to a buildup of sugars in the blood, increasing the risk of serious complications, including stroke, heart disease, vision impairment, and infection.

1.1 Forms of Diabetes

For Type 1 diabetes, the exact cause is still unclear to doctors, but genetics and environmental aspects seem to play an important role. In this form of diabetes, the body produces little to no insulin, thereby requiring patients to use insulin therapy and other treatments to manage their condition.

Type 2 diabetes has a stronger link to family history and lineage than type 1, but it also depends on environmental and lifestyle factors. Type 2 patients still produce insulin, but the receptors at the cell are unable to capture the glucose from the blood stream. Insulin allows the glucose from a person's food to access the cells in their body to supply energy.

In the case of type 2 diabetes, insulin resistance takes place gradually. Leading a healthy and active lifestyle and eating well-balanced meals can help in delaying or offsetting the development of type 2 diabetes.

Gestational diabetes occurs in pregnant women due to their body becoming less sensitive to the insulin produced. This form of diabetes does not occur in all pregnant woman and usually is resolved after delivery. However, women who develop gestation diabetes are at increased risk in developing type 2 diabetes later.

1.2 Diagnosis

There are several ways that doctors diagnose diabetes. The first and most infamous is called the glycated hemoglobin (HbA1C or A1C) test. This test measures the average blood glucose for the past 2–3 months by measuring how much glucose is bound to the hemoglobin in the blood. This testing method does not require fasting or drinking a sugary solution. An A1C of greater than or equal to 6.5% indicates diabetes. Table highlights A1C levels and their corresponding diagnosis.

The fasting plasma glucose (FPG) test checks the fasting blood glucose levels. Fasting blood glucose levels of greater than or equal to 126 mg/dl indicates a high probability of diabetes. Table 1 shows the normal, prediabetes range, and diabetes FPG range. Another test that is commonly used by physicians in the diagnosis of diabetes, particularly for gestational diabetes, is the oral glucose tolerance test (OGTT). This test is a two-hour long test that requires patients to drink a special sugary drink when fasting. Blood glucose levels are checked prior to drinking the solution, one hour after, and two hour after consumption. This test enables doctors

Table 1 Hemoglobin A1C, FPG, and OGGT levels and corresponding diagnosis

Diagnosis	HbA1C (%)	FPG (mg/dl)	OGTT (mg/dl)
Normal	Less than 5.7	Less than 100	Less than 14
Prediabetes	5.7–6.4	100–125	140–199
Diabetes	6.5 or higher	126 or higher	200 or higher

to assess how the body processes glucose. A two hour blood glucose of great than or equal to 200 mg/dl indicates diabetes. The diagnostic range of the OGTT test is listed in Table 1.

2 Importance of Diabetes Management

Diabetes is a disorder that requires constant management. In addition to medication, self-management of diabetes is very important to prevent acute complications and minimize the risk of long-term complications. Management of diabetes includes efficiently inducing self-care behaviors among the patients, such as scheduling meals, counting carbohydrates intake, monitoring daily blood glucose trends, exercising, and tracking aim-oriented life behaviors on a daily basis. Nonadherence to any of the aforementioned activities may lead to lead to long-term complications such as heart disease, stroke, blindness, amputation, kidney disease, dental disease, and increased susceptibility to infections (Diabetes Prevention Program Research Group, op.cit.; CDC, National Diabetes Statistic Report 2017). As a consequence, diabetes management becomes a cumbersome and complex task, and should account for diverse factors, such as medications, personal behaviors, and life-related activities. These factors must be jointly be optimized in order to improve the quality of life a person with diabetes.

2.1 Retinopathy and Blindness

Uncontrolled blood glucose levels, over time, can cause damage to small blood vessels within the retina of the eye. This damage can cause vision loss by two common ways: (1) a disease known as proliferative retinopathy, and (2) macular oedema. Proliferative retinopathy occurs when weak and abnormal blood vessels develop on the surface of the retina and leak fluids onto the center of the eye. Macular oedema occurs when fluid leaks from the blood vessels into the center of the macula causing it to swell. If left untreated, people with diabetic retinopathy can potentially lose vision in the eye affected. Figure 1 shows the results of a study by CDC, National Diabetes Statistic Report (2017) that concludes that diabetes is the leading cause of
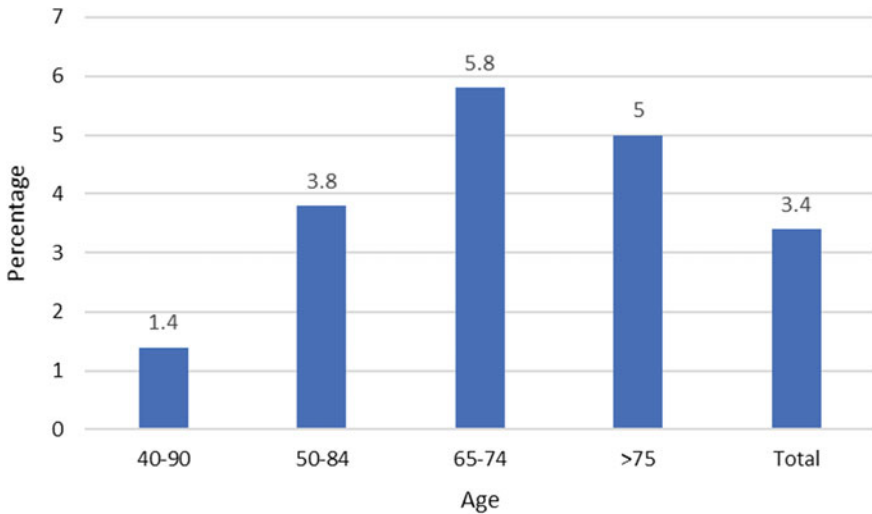
Fig. 1 Prevalence of diabetic retinopathy among adults 40 years or older (CDC, National Diabetes Statistic Report 2017)

new cases of blindness among young adults (aged 20–74) where 12,000–24,000 new cases of blindness each year is reported due to diabetic retinopathy.

2.2 Kidney Disease

Long term high blood glucose levels also have damaging effects on the kidneys. In particular, uncontrolled blood glucose increases the risk of developing diabetic nephropathy. This disease begins long before any symptoms appear and slowly damages parts of the kidney that is responsible for filtering the blood. Left untreated, this disease can cause total kidney failure, requiring patients to undergo dialysis treatment. In the United States, diabetes is the leading cause for kidney failure accounting for 43% of new cases each year (NIDDK 2004).

2.3 Heart Disease and High Blood Pressure

Diabetes and heart disease are intricately connected. People with diabetes may have several underlying conditions, such as high blood pressure, high cholesterol, and obesity, which increases their risk for heart disease. Managing their blood glucose levels greatly decreases the risk of the development of heart disease. The prevalence of high blood pressure in diabetic people is approximately 73%. In addition, adults

with diabetes have four times increased risk for heart disease related death than adults without diabetes, and these statistics are predicted to increase in the upcoming years. Due to the link between poor management of diabetes and heart disease, it is imperative to take courses of actions to properly monitor and manage glucose levels (Diabetes Prevention Program Research Group, op.cit.).

2.4 Other Diabetes Associated Diseases

Along with the aforementioned diabetes related diseases, approximately 60–70% of diabetes patients suffer from mild to severe forms of nervous system damage (Diabetes Prevention Program Research Group, op.cit.). Long term high blood glucose levels often cause impaired sensations or pain the feet and hands, sometime causing amputation of lower-extremity limbs. In particular, in the US alone, 60% of non-trauma related lower-limb amputations are among persons with diabetes. Approximately 82,000 lower-limb amputations were performed among persons with diabetes just between the years 2000–2001 (CDC, National Diabetes Statistic Report 2017).

In general, people with uncontrolled diabetes have higher risks to develop other diseases. They are also more susceptible to have longer recovery times or worse symptoms from other illnesses such as the flu or pneumonia. Thus, it is evident that proper management of the disease is imperative to live a healthy and normal life.

2.5 Effective Management Technologies

Continuous Glucose Monitoring (CGM) is a method to track glucose observations at regular intervals (typically every few minutes) throughout the day (Hess 2019). CGM devices have a sensor that is inserted under the skin that measures glucose values. Typically a CGM device is composed of two main components:

(a) Sensor: The sensor is a small wire that is inserted under the skin which measures the interstitial glucose levels from the subcutaneous tissue space.
(b) Transmitter: The transmitter captures the readings from the sensor. This information is then transmitted wirelessly to an attached insulin pump device or a separate device like a reader or a phone via near field communication (NFC) or Bluetooth.

The development of these CGM devices revolutionized diabetes self-management. Traditional methods of using a manual fingerstick to measure blood glucose only provides a "snapshot" of the glucose level at a point in time, whereas, CGM devices allow better visibility of the glucose trends as the readings are continuously measured. As a consequence, it benefits the patient in gaining insight about their glucose trends throughout the day and helps them optimize their food intake and plan physical
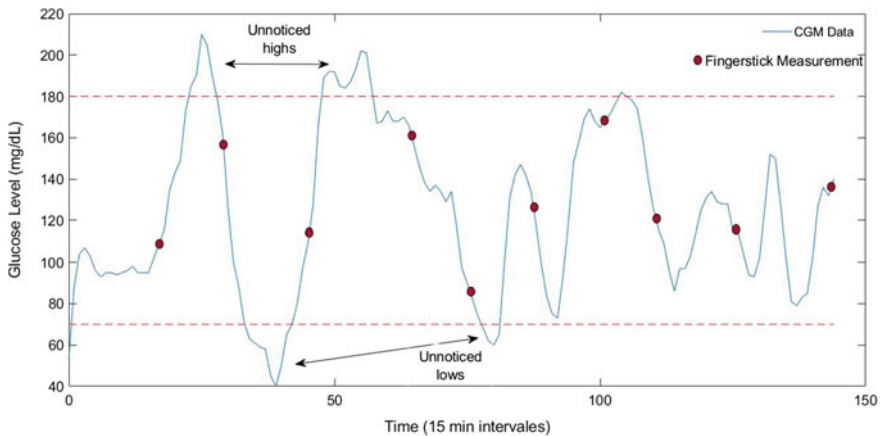
Fig. 2 CGM plot illustrating benefits of CGM over periodic fingerstick glucose measurements

activity. With full access to a patient's glucose trends, clinicians can prescribe a better treatment plan/therapy for a patient in Fig. 2 shows the benefits of analyzing CGM data versus traditional methods. In the figure, the fingerstick measures not only do not give a good representation of the patient's glucose trend, they fail to capture critical instances when the patient's blood glucose was above and below safe levels.

CGM devices have allowed patients to achieve good glycemic control and reduce glycemic excursion (fluctuations in blood sugar), thereby decreasing both hypoglycemia (low glucose) and hyperglycemia (high glucose) instances (Rodbard 2017). Modern-day CGM devices come with an inbuilt functionality that provides notifications if the glucose readings are reaching or are likely to reach below specified thresholds in the imminent future. This helps patients take preventative measures to avoid serious outcomes. In addition, CGM devices present opportunities for in-depth analysis to be performed on the data that is being captured. With the advancement of ML and AI methodologies, valuable insights on factors influencing glucose levels can be extracted and provide critical functionalities for improving patient care. The next section highlights how CGM data has been exploited using AI methods.

3 The Integration of AI and Machine Learning for Diabetes Care

The AI methodologies used in the area of health management in general and diabetes management in particular can be divided into two broad categories, namely, expert systems and machine learning. An expert system (ES) represents one of most common types of AI which assists care givers in their routine work by capturing expert knowledge, facts and reasoning techniques. The aim is to mimic clinician's expertise to support decision making. In the area of diabetes, the most common

ES used are rule-based reasoning (RBR), case-based reasoning (CBR) and fuzzy systems. RBR is based on transferring the knowledge of an expert to a computer in the form of conditions and rules, while CBR uses previous experience to find solutions to new problems similar to previously seen examples. However, fuzzy systems generally translate expert knowledge and account for ambiguity and degrees in class assignment. For instance, typically a blood glucose range <70 mg/dl is considered low and >180 mg/dl is considered high. However, this definition does not accommodate finer distinctions within low and high classes. A high value of 185 mg/dl is clinically different than 285 mg/dl and both cannot be simply classified in the same 'High' class. In fuzzy modeling, 185 mg/dl is high but can be acceptable, while 285 mg/dl can never be acceptable.

Machine learning (ML) is the ability of a machine to learn over time without being explicitly programmed. In the medical field, ML algorithms are extensively used to extract valuable knowledge from large databases, such as medical records. Methods of ML that are extensively applied in the field of diabetes management include, but limited to, decision trees (DT), support vector machines (SVM), artificial neural networks (ANN), genetic algorithms (GA), and deep learning. There has been significant work in the literature that leverage on ML methods for the prediction and management of diabetes. The work in Yu et al. (2010) implemented SVM to test its ability to classify individuals with diabetes mellitus. The authors of Lopez et al. (2018) used the random forest (RF) algorithm to select corresponding attributes single-nucleotide polymorphism (SNPs) responsible for diabetes mellitus. A modified LR model for detecting the most relevant predictor of T2DM was investigated in Devi et al. (2016). The work in Mhaskar et al. (2017) proposed a deep neural network-based approach for blood glucose monitoring. They used a semi-supervised method with three networks of the different clusters and a final layer to predict the output. Their model achieved accuracies of accuracies of 88.72% (hypoglycemia), 80.32% (euglycemia) and 64.88% (hyperglycemia). Because early detection or prediction of diabetes is important in its prevention or proper management, research has recently focused on using the power of AI for predicting diabetes (Barakat et al. 2010; Zhang et al. 2017; Xu et al. 2017; Malik et al. 2016; Thulasi et al. 2017; Alghamdi et al. 2017; Heikes et al. 2008; Stern et al. 2002; Abdul-Ghani et al. 2007, 2011; Tripathy et al. 2000) using a variety of clinical data, ranging from images to blood plasma levels.

3.1 Estimated HbA1C Versus Predictive HbA1C

The HbA1C is considered the "gold-standard" when it comes to diagnosing and managing diabetes. HbA1C is based on a laboratory test from a blood sample to measure the accumulated blood glucose over a 2–3-month span. As mentioned in Sect. 2, consistently elevated blood glucose levels cause a variety of health issues. The future prediction of the HbA1C based on the CGM data holds a critical significance in maintaining long term health of diabetes patients. There has been significant work

done on conversion formulas that estimate HbA1C using past average blood glucose levels. In particular, research from a clinical study, Diabetes Control and Complications Trial (DCCT)), derived the mathematical formula, $eHbA1C = \frac{AG+77.3}{35.6}$ as an appropriate estimate for laboratory tested HbA1C measures, where AG denotes the average blood glucose level. Variations for the estimated HbA1C have been proposed throughout the literature, however, the aforementioned mathematical models only provide instantaneous estimated HbA1C levels of past blood glucose and provide no information on future values. In order to assess whether a patient's current medical treatment and lifestyle plan is appropriate, predictions of HbA1C based on current trends is needed.

3.1.1 Challenges of HbA1C Prediction

Long-term predictive HbA1C measures using short-term CGM data is a revolutionary idea but is yet to be achieved due to the complexity of the problem. Although AI methods have evolved dramatically in the past decade, HbA1C prediction using only CGM data is still challenging even for the most robust AI techniques as the data provided (7–14 days of CGM data) is transient and is subject to various external influences/ interventions. In particular, three main challenges are identified when it comes to HbA1C prediction based on CGM readings, namely, (1) data samples over a short time duration, (2) highly varying nature of the data, and (3) missing data. The CGM sensors in the market usually measure blood glucose every 5–15 min thus generating 96–288 measurements per day. Occasionally, patients might remove sensors due to certain events or sensors might become dislodged and stop collecting data. This presents large time spans of missing blood glucose measurements. Developing algorithms that can accurately estimate missing blood glucose values is a challenge and usually suffer from high error rates. Ignoring the missing data creates misleading blood glucose trends and negatively affects the prediction accuracy.

Blood glucose measurements are highly variable and depend on a number of factors such as erythropoiesis (iron and vitamin B12 deficiency, liver disease, etc.) and altered hemoglobin glycation (alcoholism, renal failure, aspirin, vitamin C and E, etc.). Consequently, devising an accurate HbA1C prediction algorithm is difficult if a patient's full health history and lifestyle choices are not incorporated into the algorithm, but generally integrating such information is difficult and not-realistic. In addition, two patients may have similar HbA1C measures but their CGM trends may be drastically different. This phenomenon creates a "many-to-one" scenario where varying CGM trends can potentially equate to similar HbA1C measures, increasing the difficulty of accurate predictions.

The prediction of HbA1C thus boils down to extracting optimized features from CGM data and integrating these features with state-of-the-art AI techniques for prediction. Due to the increasing popularity of CGM sensors and improvements in data analytics field, this gap in research will soon be filled.

3.2   CGM Based Hyper and Hypoglycemia Predictions

3.2.1   Prediction and Challenges

Predicting a hypoglycemic (low glucose) and hyperglycemic (high glucose) at least 30–60 min in advance provides enough time for a patient to take corrective measures. The challenge of predicting impending episodes with high true positive rates (sensitivity) and high true negative rates (specificity) remains a key issue for patients and clinicians. Addressing this challenge could be a landmark achievement in the treatment of diabetic patients as it would lead to saving of many lives. Many of the existing alarm functionalities are plagued with giving too many "false alerts". As a result, patients are inclined to turn-off notifications, which defeats the purpose of the alerts. A typical patient is out of normal range only 1–10% of the time. The uneven class membership makes development of AI and Machine learning capabilities for accurate predictions difficult (Hu et al. 2009). Contextual information such as physical activity, sleep, driving, and food intake all have an effect on blood glucose values (Allen and Gupta 2019; Rodbard 2016) but relevant data is unavailable in real-time. Although devices such as wearables and Smartphone Apps are available to capture most of these data, integrated data is not currently available to facilitate real-time glucose predictions.

3.2.2   Current Literature

Researchers have tried to solve the glucose prediction problem using two main approaches:

(a)   regression-based approach: Predicting the exact glucose value into the future
(b)   classification-based approach: Predicting a probabilistic estimate of the risk of low or high glucose levels at a future time point.

   Existing literature for CGM prediction has generally looked into a prediction horizon between 15 and 60 min, giving ample time for a patient to take corrective measures. The first professional CGM device was approved by the United States F.D.A. in 1999. Since then, many studies have been published in the diabetes literature about the prediction of glucose levels. The earlier methods relied more on classical statistical modelling such as Autoregressive Integrated Moving Average (ARIMA), linear regression, etc., but as machine learning became more popular and accessible, researchers adopted sophisticated machine learning algorithms like Random Forests, Support Vector Machines, Boosting, Neural Networks for addressing this prediction problem. The work in Gadaleta (2018) provides a summary of the different methodologies used in this application area.

   Despite the progress made, there are some fundamental issues that need to be tackled to facilitate practical, robust, and universal AI and ML based solution to this serious health problem:

Table 2 Sample RMSE results for various glycemic ranges

CGM range	Reported RMSE	Glycemic state
<55	20.12	Severe hypoglycemia
≥55 and <70	15.75	Hypoglycemia
≥70 and <80	1.08	Normal range
≥180 and <300	15.87	Hyperglycemia
≥300	18.8	Severe hyperglycemia
Overall results	6.65	

(a) Standard data: Results reported in the literature have been obtained through analysis on different datasets. Some studies have relied on simulated data (Cappon 2018; Dassau 2010; Li, et al. 2019; Mahmoudi 2014; Reddy, et al. 2019; Zecchin 2012), UVA/ Padova Type 1 Diabetes simulator being the more popular, for obtaining data for analysis. Some studies have also collected data through a controlled pilot, where data is collected from patients through camps ranging from a few hours to few days. Only a handful of the existing studies have based their results on data collected from subjects in real-world settings. Being human specific, a lot of influencing factors such as age, gender, glycemic profile, lifestyle, etc. determine the glucose variability within a patient's body. Because of the relatively small data volumes, performance of machine learning algorithms are highly dependent on the dataset used. In the absence of standard datasets, it becomes difficult to unbiasedly evaluate different approaches in the literature. A large, open and standardized data that researchers can use to objectively test and evaluate performance of algorithms will be helpful to address this need.

(b) Standard comparison metrics: The most widely used metrics for reporting regression results is the root-mean-squared error (RMSE) which is the square root of the mean of squares of differences between the predicted and actual CGM values. However, it will be critical to evaluate the RMSE of the results in different target ranges. For example, in Table 2, though the overall RMSE is very low, a deeper look into the results among various glycemic ranges will show that the particular method doesn't work too well for prediction in lower and higher ranges which are more important to diabetes control. Due to higher number of observations in the normal range, the overall RMSE appears to be misleadingly low. There is a need to evaluate RMSE in critical glucose ranges for effective diabetes management.

In the classification approaches, due to the presence of imbalanced classes, it is required to evaluate performance using both sensitivity and specificity or alternately both precision and recall. Some studies report only one of these metrics or use non-standard metrics such as "number of false alarms per week" (Dassau 2010). There is a need to use standard classifier evaluation criterion such as sensitivity and specificity to compare different AI and ML approaches.

(c) Hypoglycemic/ Hyperglycemic definition: Majority of the studies in the literature consider a CGM reading less than 70 mg/dL as a hypoglycemic event and

a reading above 200 mg/dL as hyperglycemic, but there are instances where a different criterion is used. Some studies (Cameron 2008; Georga 2013; Jensen 2013, 2014) need more than 2 consecutive readings below a threshold to define a hypoglycemic/hyperglycemic event, and a few other studies combine all the readings within a time window below the threshold value as a single hypoglycemic/hyperglycemic event. Such variations make it difficult to evaluate different approaches in the literature.

4   Long Term Unmet Challenges

Technological advancements have been very beneficial to patients with diabetes—be it measuring glucose levels in real-time or the predictive capabilities incorporated in these devices, patients are being benefitted in improving their overall glycemic profile. However, there are some areas that need critical improvements. Firstly, though glucose observations are available in real-time, currently data related to insulin delivery from insulin pumps is not available in real-time. Especially for Type 1 diabetes patients, the CGM devices are often used in association with insulin pump (Pettus and Edelman 2017) which injects insulin at preset times or at user initiated times during a day (insulin bolus). Secondly, food intake and its macronutrient breakdown, especially carbohydrates have a substantial impact on glucose levels. Many Smartphone applications are available on different platforms for tracking a person's food intake and calculating the associated nutrition value for different food items. But integrated datasets covering CGM and food intake are currently not available. Physical activity is another important factor influencing blood glucose values. With the plethora of fitness devices available today, measuring physical activity with good accuracy isn't a hurdle anymore, but integrated CGM and physical activity data are also not available. There is a need to perform clinical studies to facilitate collection of CGM and associated contextual data (sleep, food intake, insulin intake, and physical activity) to facilitate next generation AI and ML solutions.

5   Future Work

Significant progress has been made in the CGM technology in regards to the device accuracy and predictive capabilities of AI/ML algorithms (Dave et al. 2019). These have been very beneficial to clinicians and patients. We believe that, the next round of innovations would come in addressing some of the challenges we discussed earlier. Especially integrating contextual information will help catapult existing predictive models for primetime usage and position them to better address diabetes management. Though most patients with type 1 diabetes use insulin pump in conjunction with the CGM device, the necessary settings to inject insulin are currently preset and

doesn't change dynamically based on real-time glucose readings. Integrated AI and ML based analysis of CGM, insulin pump, and contextual data will result in dynamic calibration of insulin to meet real-time needs of the patient, thus achieving the vision of artificial pancreas (Allen and Gupta 2019).

References

Abdul-Ghani MA, Williams K, DeFronzo RA, Stern M (2007) What is the best predictor of future type 2 diabetes? Diabetes Care 30(6):1544–1548

Abdul-Ghani MA, Abdul-Ghani T, P. Stern M, Karavic J, Tuomi T, Bo I, DeFronzo RA, Groop L (2011) Two-step approach for the prediction of future type 2 diabetes risk. Diabetes Care p DC 102201

Alghamdi M, Al-Mallah M, Keteyian S, Brawner C, Ehrman J, Sakr S (2017) Predicting diabetes mellitus using smote and ensemble machine learning approach: the henry ford exercise testing (fit) project. PloS one 12(7):e0179805

Allen N, Gupta A (2019) Current diabetes technology: striving for the artificial pancreas. Diagnostics 9(1):31

Barakat N, Bradley AP, Barakat MNH (2010) Intelligible support vector machines for diagnosis of diabetes mellitus. IEEE Trans Inf Technol Biomed 14(4):1114–1120

Cameron F et al (2008) Statistical hypoglycemia prediction. J Diabetes Sci Technol 2(4):612–621

Cappon G et al (2018) A neural-network-based approach to personalize insulin bolus calculation using continuous glucose monitoring. J Diabetes Sci Technol 12(2):265–272

CDC (2017) National Diabetes Statistic Report

Dassau E et al (2010) Real-time hypoglycemia prediction suite using continuous glucose monitoring: a safety net for the artificial pancreas. Diabetes Care 33(6):1249–1254

Dave D et al (2019) Feature selection based machine learning algorithm for real-time hypoglycemia prediction. T.A.M. University, Hospital TCS, Medicine BCO (eds) J Biomed Health Inf (JBHI) Submitted: 16 Sep 2019

Devi MN, Balamurugan AA, Kris MR (2016) Developing a modified logistic regression model for diabetes mellitus and identifying the important factors of type ii dm. Indian J Sci Technol 9(4)

Diabetes Prevention Program Research Group, op.cit.

Gadaleta M et al (2018) Prediction of adverse glycemic events from continuous glucose monitoring signal. IEEE J Biomed Health Inform 23(2):650–659

Georga EI et al (2013) A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions. Diabetes Technol Ther 15(8):634–643

Heikes KE, Eddy DM, Arondekar B, Schlessinger L (2008) Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and prediabetes. Diabetes Care 31(5):1040–1045

Hess A (2019) What is a Continuous Glucose Monitor (CGM)? Available from: https://www.ontrackdiabetes.com/type-1-diabetes/what-continuous-glucose-monitor-cgm. Accessed 15 Aug 2019

Hu S et al (2009) MSMOTE: improving classification performance when training data is imbalanced. In: 2009 second international workshop on computer science and engineering, IEEE

Jensen MH et al (2013) Real-time hypoglycemia detection from continuous glucose monitoring data of subjects with type 1 diabetes. Diabetes Technol Ther 15(7):538–543

Jensen MH et al (2014) Evaluation of an algorithm for retrospective hypoglycemia detection using professional continuous glucose monitoring data. J Diabetes Sci Technol 8(1):117–122

Lopez BL, Torrent-Fontbona F, Vinas R, Fernandez-Real JM (2018) Single nucleotide polymorphism relevance learning with random forests for type 2 diabetes risk prediction. Artif intell med 85:43–49

Li K et al (2019) GluNet: a deep learning framework for accurate glucose forecasting. IEEE J Biomed Health Inf 24(2):414–423

Mahmoudi Z et al (2014) Accuracy evaluation of a new real-time continuous glucose monitoring algorithm in hypoglycemia. Diabetes Technol Ther 16(10):667–678

Malik S, Khadgawat R, Anand S, Gupta S (2016) Non-invasive detection of fasting blood glucose level via electrochemical measurement of saliva. Springerplus 5(1):701

Mhaskar HN, Pereverzyev SV, van der Walt MD (2017) A deep learning approach to diabetic blood glucose prediction. Front Appl Math Stat 3:14

National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (2004) National diabetes statistics. NIH Publication No. 04–3892. Available at: http://diabetes.niddk.nih.gov/dm/pubs/statistics/index.htm

Nai-arun N, Moungmai R (2015) Comparison of classifiers for the risk of diabetes prediction. Procedia Computer Science 69:132–142

Pettus J, Edelman SV (2017) Recommendations for using real-time continuous glucose monitoring (rtCGM) data for insulin adjustments in type 1 diabetes. J Diabetes Sci Technol 11(1):138–147

Reddy R et al. (2019) Prediction of hypoglycemia during aerobic exercise in adults with type 1 diabetes. Journal of Diabetes Science and Technology 13(5):919–927

Rodbard D (2016) Continuous glucose monitoring: a review of successes, challenges, and opportunities. Diabetes Technol Ther 18(S2):S2-3–S2-13

Rodbard D (2017) Continuous glucose monitoring: a review of recent studies demonstrating improved glycemic outcomes. Diabetes Technol Ther 19(S3):S-25–S-37

Stern MP, Williams K, Haffner SM (2002) Identification of persons at high risk for type 2 diabetes mellitus: do we need the oral glucose tolerance test? Ann Intern Med 136(8):575–581

Thulasi K, Ninu E, Shiva KK (2017) Classification of diabetic patients records using naïve bayes classifier. In: 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT), IEEE, pp 1194–1198

Tripathy D, Carlsson M, Almgren P, Isomaa B, Taskinen M-R, Tuomi T, Groop LC (2000) Insulin secretion and insulin sensitivity in relation to glucose tolerance: lessons from the Botnia study. Diabetes 49(6):975–980

Xu W, Zhang J, Zhang Q, Wei X (2017) Risk prediction of type ii diabetes based on random forest model. In: 2017 Third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB), IEEE, pp 382–386

Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ (2010) Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. BMC Med Inform Decis Mak 10(1):16

Zecchin C et al (2012) Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration. IEEE Trans Biomed Eng 59(6):1550–1560

Zhang J, Xu J, Hu X, Chen Q, Tu L, Huang J, Cui J (2017) Diagnostic method of diabetes based on support vector machine and tongue images. BioMed Res Int 2017