Internet of Things 24 (2023) 100945

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.sciencedirect.com/journal/internet-of-things

# Forecasting glycaemia for type 1 diabetes mellitus patients by means of IoMT devices

Ignacio Rodríguez-Rodríguez [a,*], María Campo-Valera [b], José-Víctor Rodríguez [b]

[a] Departamento de Ingeniería de Comunicaciones, Universidad de Málaga, Avda. Cervantes, 2 29071 Málaga, Spain
[b] Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Cartagena, Antiguo Cuartel de Antigones, Plaza del Hospital, 1, Cartagena 30202 Murcia, Spain

ARTICLE INFO

*Keywords:*
Diabetes
IoT
Wearable trackers
Continuous glucose monitoring
Machine Learning

ABSTRACT

The chronic metabolic condition, Type 1 diabetes mellitus (DM1), is marked by consistent hyperglycemia due to the body's inability to produce sufficient insulin. This necessitates the patient's daily monitoring of blood glucose fluctuations to discern a trend and predict future glycemia, subsequently dictating the amount of external insulin needed to regulate glycemia effectively. However, this technique often grapples with a degree of inaccuracy, presenting potential hazards. Nonetheless, contemporary advancements in information and communication technologies (ICT) coupled with novel biological signal sensors offer a refreshing perspective for DM1 management by enabling comprehensive, continual patient health evaluation. Herein, burgeoning technological disruptions such as Big Data, the internet of medical things (IoMT), cloud computing, and machine learning algorithms (ML) could serve pivotal roles in the effective control of DM1. This paper delves into the exploration of the latest IoMT-based methodologies for the unbroken surveillance of DM1 management, facilitating a profound characterization of diabetic patients. The fusion of wearable technologies with machine learning strategies has the potential to yield robust models for short-term blood glucose prediction. The ambition of this study is to develop precise, individual-centric prediction models harnessing an array of pertinent factors. The study applied modeling techniques to a comprehensive dataset comprising glycaemia-associated biological attributes, sourced from an expansive passive monitoring campaign involving 40 DM1 patients. Leveraging the Random Forest method, the resulting models can predict glucose levels over a 30-min time span with an average error as minimal as 18.60 mg/dL for six-hour data and 26.21 mg/dL for a 45-minute prediction horizon, offering also a good performance in the prediction delay.

## 1. Introduction

Type 1 diabetes mellitus (DM1) is marked by high blood glucose levels, due to the body's lack of ability to produce or utilize insulin, caused by the autoimmune destruction of the insulin-producing cells in the pancreas. Individuals with DM1 require an insulin pump or exogenous insulin injections, and constant monitoring of their glucose levels throughout the hours. Intelligent data analysis (IDA) can assist in glycemia level modeling, thereby enhancing the ability of individuals with diabetes to manage their condition [1,2].

Advancements in the development of an artificial pancreas (AP) occurred 50 years after the initial efforts. It is believed that APs,

* Corresponding author.
*E-mail address:* ignacio.rodriguez@ic.uma.es (I. Rodríguez-Rodríguez).

https://doi.org/10.1016/j.iot.2023.100945

Available online 22 September 2023

still under development, will comprise a continuous glucose-monitoring (CGM) device and an injection system of insulin guided by a mathematical model. The CGM device would monitor glucose levels in real-time, while the mathematical model would help to maintain optimal glycemic balance [3–5].

Alternative systems to AP and CGM devices are also being explored, with researchers looking into 24-hour patient monitoring using variables like heart rate, temperature, sleep quality, and physical activity. These systems could gather crucial health data to assist in the development of effective glycemia treatments. Such systems, however, would require robust information and communication technology (ICT) to manage the data collected effectively [6].

The internet of medical things (IoMT) suggests an appropriate framework for these systems, allowing data collection from various sensors connected to a person with diabetes. Machine Learning algorithms in this IoMT environment could enhance the prediction of hyper- or hypoglycemia situations. In a study, continuous data collection from 40 diabetic patients in real-life scenarios showed the potential to predict glycemic evolution, crucial for managing DM1 [7].

Overall, the internet of things (IoT) presents significant promise for developing sophisticated and reliable model systems for DM1 monitoring. Cloud computing facilitates the use of data-intensive intelligent data methods, providing insights into blood sugar discrepancies by analyzing extensive data collected from IoT connections.

The manuscript begins with an 'Introduction' (Section 1) that outlines the subject matter. This leads to 'Section 2: Internet of medical things (IoMT)', which discusses the role of IoMT in healthcare. 'Section 3: Continuous glucose monitoring (CGM) and commercial smart devices for DM1 monitoring' then delves into the monitoring technologies for DM1. 'Section 4: Variables related to DM1 in the context of IoMT' elucidates the key variables associated with DM1 within IoMT.

'Section 5: Connectivity and the communications environment' describes the networking requirements for IoMT solutions. 'Section 6: Intelligent Data Analysis for DM1 Management and Modelling Methods' presents data analysis techniques for managing DM1.

Main contributions are addressed in 'Section 7 ', followed by a concrete monitoring campaign in 'Section 8'. 'Section 9' explicates the methods, with '9.1′ detailing feature selection, and '9.2′ explaining the glucose level prediction methods.

Findings and interpretations are presented in 'Section 10: Results and discussion'. Finally, 'Section 11: Conclusions' rounds off the study by summarizing the main insights and proposing future research directions.

## 2. Internet of medical things (IoMT)

The internet of medical things (IoMT), consisting of interconnected medical equipment, allows remote collection, evaluation, and transfer of medical data using automated sensors and machine learning (ML) for improved healthcare monitoring [8]. Such technology reduces health costs and unnecessary hospital visits, with the potential to alert caregivers to anomalies in a patient's health [9]. This approach employs a network of sensing devices and intelligent platforms to record patient data, allowing medical staff to monitor patients and respond appropriately based on the severity of their condition.

Different studies have looked into various aspects of IoMT. Moosavi [10] examined end-to-end solutions of security for healthcare IoT, recommending updates to cryptographic key generation algorithms due to the limited memory, computing power, and bandwidth of healthcare IoT sensors. An energy-optimized algorithm for a smartphone-based heart disease monitoring system was formulated by Talpur et al. [11], where the severity of the data determined smartphone power management. Ali et al. [12] developed a system using IoT to identify signs of depression and provide suitable assistance, suggesting a microservices model for system functions.

In the realm of disease diagnosis and monitoring, Romero et al. [13], regarding Parkinson's disease, proposed using IoT to diagnose and monitor, while Bajaj et al. [14] investigated heart attack detection systems, and Hemalatha et al. [15] presented a novel method for identifying and quantifying frequent coughs. Matar et al. [16] introduced a technique for posture monitoring through body pressure distribution on mattresses, leveraging supervised learning SVM.

Other studies explored the interaction between IoT and emerging technologies. Magsi et al. [17] discussed the integration of 5 G with IoT, focusing on how 5 G could enhance IoT healthcare diagnosis and treatment. Fan et al. [18] considered RFID systems based on cloud technologies, proposing a lightweight authentication method to secure medical data from potential leaks.

Leveraging IoT-WSN to improve cancer treatment, Onasanya et al. [19] put forward an IoT-centric system for therapy and detection of cancer., bolstered by business analytics and cloud services. Through the generation of patient data streams, these cloud services and business analytics contribute to evidence-based decision-making.

## 3. Continuous glucose monitoring (CGM) and commercial smart devices for DM1 monitoring

The commencement of diabetes monitoring necessitates the selection of a continuous glucose monitoring (CGM) system. An artificial pancreas (AP), an innovative approach to revolutionizing diabetes management, relies heavily on CGM, making it an indispensable component [20]. CGM technology enables the comprehensive tracking of glucose fluctuations in terms of their magnitude, trend, frequency, and duration, thereby contributing significantly to effective diabetes management. The power of the MCG lies in its ability to generate glucose data points with a high rate, some models up to one data point per minute, equating to a staggering 1440 values over a 24-hour period. The frequency of data acquisition significantly exceeds that of traditional fingerstick or capillary blood glucose monitoring, which typically yields three to ten readings per day. Despite its advantages, it is crucial to acknowledge that CGM may have potential limitations, such as random noise and transient sensitivity failure, as pointed out by Kovatchev et al. [21].

To gage the accuracy of CGM devices, measures like the mean absolute relative difference (MARD) and Clarke error grid analysis are employed [22]. The MARD, derived from the average discrepancy between CGM readings and actual glucose values, is an effective indicator of device accuracy. Lower MARD values correspond to reduced average error, thus increasing the reliability of the CGM device. The most accurate CGM systems generally exhibit a MARD of 6–8 %.

With the advancements in smart device technology, there has been a significant shift in monitoring practices. The advent of smartphones, in particular, has been transformative, boasting unprecedented versatility in the realm of health monitoring. These devices have the ability to host software responsible for various functionalities, encompassing elements of solution optimization, process control, system dynamics, and glucose prediction. Their extensive connectivity, ranging from 4 G, Bluetooth, and WiFi to NFC and Ant+, facilitates CGM data collection and interaction with insulin pumps, thereby diversifying user options. Smartphones also allow data transfer to cloud storage, enable emergency calls, and facilitate software updates. Furthermore, they can measure physical activity through built-in accelerometers, gyroscopes, and pedometers, and some even estimate heart rate and ambient temperature, albeit with considerable error.

The potential of GPS technology in patient monitoring has also been explored, with studies conducted by Place et al. [23] and others investigating the feasibility of tracking patient movement from various locations and transmitting emergency calls using GPS coordinates [24]. It's imperative to underscore, nonetheless, that smartphones fall under the classification of non-medical devices and carry a high risk (Class III), require a reliable controller application to avoid interference with other apps, prevent battery discharge, and manage connection issues, as Rigla suggests [25].

Advancements in electronics have led to the miniaturization and improvement of biometric devices, allowing continuous monitoring of vital signs such as heart rate and exercise levels that influence blood glucose balance. Modern wearable technologies, like smart bands, have simplified the collection of day-long data, providing valuable insights into patient health [26]. Although these devices are limited by factors such as size, battery life, and professional application, they offer the promise of accurate data collection, enhanced by their Bluetooth connectivity [27].

Innovative technologies like accelerometers, electrocardiograms, thermistors, etc., have been trialed in isolation but have yet to be integrated into DM1 management systems. They have been used alongside CGM to explore the relationship between glycemia and other physiological parameters and predict blood glucose levels. The future of DM1 management thus holds immense potential, resting on the integration and synergistic functioning of these diverse devices.

## 4. Variables related to DM1 in the context of IoMT

The management of Type 1 diabetes mellitus (DM1) hinges on monitoring various biosignals that impact blood glucose levels, or glycemia. While current DM1 management systems often focus primarily on glycemia, insulin, and meal estimations, there are multiple variables that can influence glucose levels and thus warrant inclusion [28].

- The key variable in DM1 management is glycemia, representing the patient's blood glucose level. Contemporary continuous glucose monitoring (CGM) systems provide precise and frequent sampling, thus supplying valuable data for a continuous control algorithm [29]. This focus on glycemia is reinforced by the fact that some studies use only previous glycemia data for their analyses, indicating its importance [30].
- Insulin is another crucial variable that influences glycemia. The amount of insulin administered by the patient dictates the hypoglycemic response and understanding this variable aids in DM1 management.
- Exercise is a significant variable as it can affect glucose and insulin levels. Exercise increases the permeability of cellular barriers, leading to efficient glucose entry into cells and faster insulin usage [31]. This process can potentially result in hypoglycemia and insulin reduction. Regular exercise has a balancing effect on blood sugar and can decrease insulin needs [32].
- Food intake also impacts blood glucose levels. The absorption rates of fats, proteins, and carbohydrates vary, affecting how quickly glucose levels rise [33]. Furthermore, the digestion and absorption of food can increase blood sugar, necessitating accurate meal and carbohydrate counting for proper DM1 management. Factors such as the type of macronutrient consumed also play a role, as some can slow glucose absorption or reduce insulin sensitivity [34,35].
- Stress and sleep quality are variables that can significantly impact glycemia. Stress hormones can cause hyperglycemia [36], and poor sleep can alter glucose metabolism and insulin resistance, leading to increased blood glucose levels [37].
- Physiological variables such as heart rate and body temperature can serve as indicators of glycemic fluctuations. Elevated heart rate can indicate exercise or stress [38], both of which affect glycemia.
- Similarly, hypoglycemia can lower body temperature [39], while conditions like hyperthermia and hypothermia may cause hyperglycemia [40,41].
- Hypoglycemia can also trigger sweating [42], and blood pressure changes may signify poor glycemic management [43]. Monitoring these biosignals, along with keeping track of daily schedules [44], helps facilitate system prediction, thereby improving DM1 management.

The customization of treatment solutions can be enhanced by considering variables like age, sex, height, weight, and BMI. Additionally, other patient characteristics, settings, and concurrent conditions can impact glucose levels. Particular attention must be given to situations like pregnancy in women [45] and the presence of mental illnesses [46]. Other concurrent disorders, specifically those linked to diabetes like cardiovascular and renal impairment [47], diabetic foot, retinopathy [48], need to be considered and incorporated into a coordinated strategy.

Mental health plays a pivotal role in diabetes management, providing a critical link between mind and body wellness. Chronic stress, which can be exacerbated by high environmental noise levels [49], has been demonstrated to trigger a cascade of physiological responses that may disrupt glucose metabolism, leading to poor glycemic control. It is, therefore, essential to consider not only physical factors but also psychological ones, such as stress and the overall mental wellbeing that can be now measured using wearables [50], when devising a comprehensive approach for diabetes management. Incorporating mental health considerations could significantly enhance the efficacy of diabetes care, thereby improving patient outcomes and quality of life.

In summary, while glycemia remains the central variable in DM1 management, a multitude of other variables exert influence on it. A comprehensive approach, which includes continuous monitoring and individual customization, can provide more accurate information and allow for a more effective, tailored diabetes treatment. This approach necessitates a focus on non-invasively assessed variables, supporting the advancement of the current state of DM1 management.

## 5. Connectivity and the communications environment

The emergence of continuous biometric sensors necessitates compatibility between clinical devices, management platforms, and healthcare digital records for effective data exchange. This is facilitated by Health Level 7 (HL7) [77], a standard ensuring real-time clinical data transfer. Technologies like mobile access to health information (MAHI) [78] and Wireless Sensor Networks provide initial connectivity between diabetic patients and healthcare professionals, although they lack real-time communication capabilities.

Data capture, crucial for Internet of Things (IoT) healthcare systems, employs Wireless Body Area Networks (WBANs) and environmental sensors [81]. WBANs are regulated by the IEEE 802.15.6 standard, ensuring reliable, low-power, short-range communication. Various WBAN and wireless personal area network (WPAN) protocols, including ZigBee, Bluetooth Low Energy (BLE), and WiFi, are utilized for numerous applications, transmitting sensor data for collection and analysis.

## 6. Intelligent data analysis for DM1 management and modelling methods

The amalgamation of wearable technology and machine learning algorithms is at the helm of innovation, playing a pivotal role in enhancing patient safety and self-management in Type 1 diabetes mellitus (DM1) through glycemia prediction. Various studies have highlighted the potential of different methodologies in leveraging this technology to predict glycemia accurately, including the new wearables [51].

Bent et al. [52] ignited the exploration of non-invasive methods by utilizing wearable devices to estimate hemoglobin A1c (HbA1c) and glucose variability. They demonstrated significant correlations with continuous glucose monitor (CGM) measurements by using machine learning models developed based on skin temperature, electrodermal activity, heart rate, and accelerometry [52]. Zhu et al. [53] extended this initiative by integrating data from CGM, meal, and bolus insulin entries with a clinically validated wearable sensor wristband into a smartphone-based platform named ARISES. The platform, armed with deep learning algorithms, exhibited an average RMSE of 35.28 +/- 5.77 mg/dL in predicting glucose levels and hypo- and hyperglycemia [53].

D'Antoni et al. [54] implemented a minimalist neural network approach that only required past glucose values for predicting future glycemic levels. Despite its simplicity, it outperformed other glucose level prediction models, particularly in terms of RMSE, setting a benchmark for models that prioritize computational efficiency without compromising accuracy [54].

Mosquera-Lopez et al. [55] and Mayo, Chepulis, and Paul [56] honed in on support vector regression (SVR) for prediction, training their model on continuous glucose measurements and insulin data. Impressively, Mosquera-Lopez et al.'s model achieved 94.1 % sensitivity in predicting nocturnal hypoglycemia [55], while Mayo et al.'s model proved optimal for forecasting blood glucose levels within the normal and hyperglycemic ranges [56].

Zhang et al. [57] emphasized the power of regression models in predicting long-term horizons, displaying robust performance and lower computational costs. The utility of these models offered a counterpoint to models focused on shorter prediction horizons such as the GluNet model by Li et al. [58]. The latter, a deep neural network model, showed remarkable accuracy in short-term (30–60 min) predictions, aiding immediate decision-making processes pertaining to insulin dosing [58].

On the clinical front, Katayama et al. [59] reported that sensor-augmented pump therapy with predictive low glucose management (PLGM) significantly reduced hypoglycemic events in Japanese patients with DM1 [59]. At the same time, Alfian et al. [60] proposed systems that utilized processing real-time data and utilizing sensors based on BLE for diabetic patients, demonstrating the evolving integration of technologies for real-time monitoring and early prediction of diabetes and blood glucose levels [60].

In recent research advancements concerning the prediction of glycemic levels in DM1 patients, D'Antoni et al. [61] presented a layered meta-learning approach predicated on multi-expert systems exclusively utilizing continuous glucose monitoring data. This sophisticated approach, composed of three specialized deep neural networks predicting hypoglycemia, hyperglycemia, or euglycemia, subsequently passes its output to a meta-learner for a final classification. Remarkably, this model showcased superior performance, anticipating hypoglycemic and hyperglycemic events with significant time gains and minimizing false positives. The flexibility of this method is evidenced as it can effectively adapt to new patient cohorts by merely training the meta-learner with limited data, aiming to optimize insulin therapy adjustments and dietary intakes, ultimately enhancing patient self-management and reducing complications. Concurrently, Quan et al. [62] offered an AI-based system specializing in low-periodic blood glucose environments, utilizing the Long Short-Term Memory (LSTM) network for time series data. Demonstrating a capability to predict hypoglycemia occurrences with approximately 80 % accuracy after a 30-minute window, the study further highlighted the enhancement in prediction accuracy by alternately performing learning and inference. Collectively, these contributions underscore the potential of advanced machine learning techniques in refining glucose level predictions, substantiating their crucial role in bettering diabetic patient care and management.

Integrating IoT paradigms has also seen significant focus in managing and evaluating biosensor data. Intelligent analysis is applied to commonplace devices such as smartphones in numerous proposed models [63]. The transformation of vast amounts of data observed in an IoT context into knowledge can be achieved through machine learning strategies. In particular, an Artificial Neural Network (ANN), which replicates the nervous system's functioning, can predict complex interactions in glucose metabolism components [64].

Bayesian regularized neural networks (BRNNs) [65] were used in diabetes and shown to be more robust than BP nets. BRNNs can eliminate cross-validation and make overtraining impossible due to Bayesian criteria for halting training, thereby dealing with the overfitting problem [66]. Another study used an MLP with BP for online glucose prediction from CGM data, achieving an RMSE of 27 mg/dl with past 20-minute glucose measurements [67].

SVMs have also found significant use in the field. Khanam and Foo [m] leveraged SVM in their diabetes prediction model using the PID dataset, contributing to early detection efforts. Moreover, SVM-based methods were also employed to predict hypoglycemia, with EEG data and a Fuzzy SVM model were deployed, successfully predicting low glucose levels with a 75 % rate [68].

In sum, SVM, RF, and BRNN emerge as the most promising methods for glycemia prediction. SVM and SVR, in particular, have shown their versatility across a range of studies [f][m][n] [68], providing consistent results in predicting glucose levels. RF's strength lies in its ability to handle large datasets with high dimensionality and irrelevant features, and BRNN's Bayesian regularization approach mitigates overfitting issues while preventing overtraining [65]. These attributes make SVM, RF, and BRNN instrumental in the progression of more personalized and efficacious treatment strategies for DM1 patients.

The importance of predicting blood glucose levels at multiple forecast horizons (predictive horizon, PH) cannot be overstated in managing diabetes. Short-term predictions, for example, those at 15 min, are essential to prevent rapid-onset hypoglycemia, which could result from insulin over-administration. On the other hand, medium-term forecasts, such as those at 45 min, provide necessary alerts to avert gradual glucose fluctuations including hyperglycemic and non-pronounced hypoglycemic episodes. Longer-term predictions guide the management of postprandial glucose spikes and aid the administration of slower-acting insulin. Furthermore, the ability to predict at various PH enables the anticipation of the effect of lifestyle factors on future glucose levels, assisting in the development of personalized diabetes management plans.

## 7. Contributions

The overarching landscape of healthcare has been dramatically revolutionized by the seamless integration of advanced information technology paradigms, especially in the sphere of chronic diseases like DM1. This research delves deeply into this novel intersection, crafting a blueprint for DM1 management that amalgamates the vast potentials of the IoMT with state-of-the-art machine learning mechanisms. Each stride made in this research illuminates a pathway towards more efficient, personalized, and predictive healthcare, making it indispensable for both academia and clinical applications. Below, we outline the seminal contributions offered by this study in detail:

- Novelty in DM1 Management: One of the standout accomplishments of this research lies in its avant-garde approach to DM1 management. By leveraging the IoMT in tandem with intricate machine learning algorithms, we've cultivated a mechanism that not only monitors but predicts short-term glycemia dynamics with significant accuracy, a feat that stands in stark contrast to traditional methodologies.
- In-depth Characterization of DM1 Variables within IoMT: The research goes above and beyond by dissecting the multitude of variables that play a role in DM1, especially within the IoMT environment. Traditional models might account for glucose concentrations and insulin dosages, but our approach broadens this horizon to encapsulate variables like patient's physical activities and dietary habits - elements previously underemphasized but now understood as crucial, thanks in part to the insights from our work.
- Cutting-Edge Data Collection Techniques: Grounding our research in empirical evidence, we've employed the Abbott Freestyle Libre CGM sensor. This wearable represents the zenith of IoT devices designed for health applications, capturing continuous glucose monitoring data. The scale of our data collection, encompassing 40 DM1 patients monitored over two weeks, provides a comprehensive dataset, enabling nuanced analyses that are both broad in scope and granular in detail. This extensive dataset was enriched with data collected from the Fitbit Charge 5 smart band.
- Advanced Predictive Modeling: The cornerstone of our research is the development of a sophisticated machine learning model tailored to the complexities of DM1. With the depth of data from our CGM devices, this model predicts blood glucose trajectories up to half an hour in advance, a crucial window for DM1 patients. The precision showcased by our results reaffirms the model's efficacy, and it sets a new benchmark in predictive healthcare.
- Future directions: While our study offers considerable advancements in the DM1 management arena, we also recognize the journey ahead. Our research postulates several intriguing future directions, informed by the potential to include a wider range of biological and lifestyle determinants, iterative model enhancement using real-time patient feedback, and the tantalizing prospect of applying similar methodologies to a myriad of other chronic diseases. These insights, especially when considered in light of the discussion on prediction delay from earlier in our conversation, offer a rich tapestry of opportunities for subsequent research.

The elucidation of our empirical implementation involving 40 DM1 patients over a 14-day monitoring campaign and the resulting insights stand testament to the viability and the transformative potential of the approach advocated in this research.

## 8. Monitoring campaign

In validating our thesis, we launched an initial monitoring campaign focused on patients. This operation unfolded by implementing an interconnected system equipped with a portable Abbott Freestyle Libre and a smartphone that was configured to transmit collected data to a central hub for subsequent examination. This setup incorporated a CGM sensor with a local memory that can retain up to eight hours of past data. The acquisition of the most recent measurements from the CGM apparatus is facilitated through the use of a secure, NFC-enabled short-range wireless linkage, supported by software installed on the patient's personal electronic device, such as a smartphone or a tablet. Some devices, acting as NFC-Bluetooth transformers, can establish a connection with the Libre, ensuring a consistent data transfer to the smartphone. The information is subsequently dispatched securely to a central database for additional processing and generation of predictive models. Further insight into the information and communication technology (ICT) system utilized for data collection can be found in [7].

Inserted beneath the skin, the CGM sensor exhibits the capability to evaluate blood glucose levels (reported in mg/dL) on a minute-to-minute basis [69]. These readings, which display a mean absolute relative difference (MARD) of 11.4 % [70], offer an estimated depiction of the authentic glucose concentrations in the blood, as per the manufacturer's disclosure. The sensor's sensitivity range is broad, extending from 40 mg/dL (at which point it indicates "low") to 500 mg/dL (beyond which it merely represents "high"). There exists a slight delay, approximately 5 to 10 min, between the CGM sensor data and the actual glucose levels in the patient's bloodstream [71]. It is worth noting that this time discrepancy can be effectively minimized, potentially down to six minutes, using mathematical techniques [72]. The calibration period for the CGM sensor is relatively brief, lasting only a few hours [73], and the sensor itself can function optimally for a period of fourteen days. The collection process culminated in a dataset consisting of 13,440 h of data.

This comprehensive dataset was complemented with information gathered via the Fitbit Charge 5 smart band. Each participant was outfitted with the device, which persistently recorded physical activity (step count), heart rate, and sleep duration. Despite the fact that these apparatuses do not constitute specialized medical tools, they supply valuable data inputs while exerting minimal energy consumption. An expanded overview of these devices is available on the manufacturer's website.

This advanced technology was trialed with a cohort of 40 DM1 diabetics in 2021, in collaboration with local hospitals. The study adhered to the guidelines of the Helsinki Declaration, with data privacy regulations strictly enforced to protect stored data. Clinical characteristics of the participating patients are presented in Table 1. All participants were properly briefed on the nature of the study, its implications, and data management, following which they provided their informed consent in accordance with national regulations.

Participants were required to have at least four years of familiarity with their disease process to ensure adequate engagement with diabetic equipment such as glucose meters, insulin pumps, or CGM for effective blood glucose management. All participants were proficient in the usage of the Abbott Freestyle Libre CGM sensor and had unrestricted access to their CGM data to enable informed decision-making regarding glycemic control.

The overall DM1 management of the patients was commendable. All reported maintaining a healthy lifestyle with regular exercise, at least three times per week. Additionally, they adhered to structured schedules, ensuring a largely consistent daily routine devoid of unexpected disruptions.

During the passive monitoring period, patients were encouraged to maintain their regular routines and consume a balanced diet. Throughout the monitoring phase, all patients were advised to abide strictly by their doctors' recommendations. All participants adopted a basal-bolus regimen, utilizing insulins with a flat action curve, such as Levemir or Tresiba, along with fast-acting insulins like Humalog Lispro. The former supplies a basal dose for over 24 h, while the latter serves to counteract any spikes in blood glucose levels resulting from meal consumption or other hyperglycemic triggers.

In this paper, we will critically evaluate and compare various methods of modeling glycemic oscillations. The following parameters, gathered during our experimental phase, will be considered: Blood sugar levels, administration of insulin, meals, physical activity, cardiac rhythm, and slumber duration. These parameters encapsulate historical and current values, ranging from past glycemia measurements, insulin dosages, meals consumed, steps taken, heart rate values, and sleep status ("asleep" or "awake"). These values were carefully collected from selected DM1 patients with sufficient experience in carbohydrate estimation.

**Table 1**
Information on the patients included in the trial.

| Population Feature | Value | | |
|---|---|---|---|
| Subjects (number) | 40 | | |
| Sex | 24 men – 16 women | | |
| Population feature | Median | Min | Max |
| Age (years) | 22.53 | 18 | 56 |
| Body mass index (BMI, kg/m2) | 21.30 | 18.25 | 23.71 |
| Duration of diabetes (years) | 12 | 4 | 29 |
| Insulin units per day (fast insulin + slow insulin, median) | 45 | 34 | 53 |
| HbA1C (%) | 6.7 | 6.1 | 7.8 |

## 9. Methods

### 9.1. Features selection

As you add more features to the model, the model becomes more complex and may overfit the data. Some features can be noise and potentially damage the model. If you remove these unimportant features, the model can be generalized better.

Variable selection is made using a voting-based approach based on the idea of using various techniques to select variables. When an algorithm selects a variable, a vote is cast for that variable. In the end, the total number of votes for each variable is calculated, and then the best variables are selected based on the votes (Fig. 1). This way, the best variables are selected with the least effort in variable selection [74].

Once this process is completed, we select the best variables selected by each algorithm (voting). We count the total number of votes and then perform a multicollinearity test for the selected variables.

Thus, in this voting strategy, different methods are used for feature selection. In this work, the following methods are used to select important features:

- Information value (IV) using weight of evidence (WOE): WOE indicates the predictive power of an independent variable concerning the dependent variable [75]. It helps to convert a continuous independent variable into a set of groups or bins based on the similarity of the distribution of the dependent variable, i.e., the number of events and non-events. WOE can handle outliers, deal with missing values, and there is no need for dummy variables [76]:

$$WOE = \ln\left(\frac{Event\%}{Non\ Event\%}\right) \tag{1}$$

$$IV = \Sigma[(Event\% - Non\ Event\%) * WOE] \tag{2}$$

If the IV statistic is higher than 0.3, then the predictor strongly correlates with the event/non-event likelihood ratio.

- Variable importance with Random Forest/Extra trees classifier: tree-based estimators can be used to compute impurity-based feature importance, which can be used to discard irrelevant features.
- The standard method for computing variable importance is the mean decrease in impurity (or Gini importance [77]) mechanism: At each split in each tree, the improvement in the splitting criterion is the measure of importance attributed to the splitting variable and is accumulated over all trees in the forest for each variable separately. Similarly, this measure is comparable to the R2 in regression on the training set.
- Recursive Feature Elimination: given an external estimator (a linear regression model) that assigns weights to features, the goal of recursive feature elimination (RFE) is to select features by recursively considering smaller and smaller feature sets [78]. First, the estimator is trained on the initial set of features, and the importance of each feature is determined by a particular attribute (e.g., feature importance). Then, the least important features are removed from the current set of features. This procedure is repeated recursively with the cleaned set until the desired number of features to be selected is finally reached.
- Chi-square best variables: In this method, a statistical test known as chi-square (usually written as $\chi 2$) is used to measure correlations between the different features of a dataset and then detect multicollinearity between them. The basic idea of the chi-square

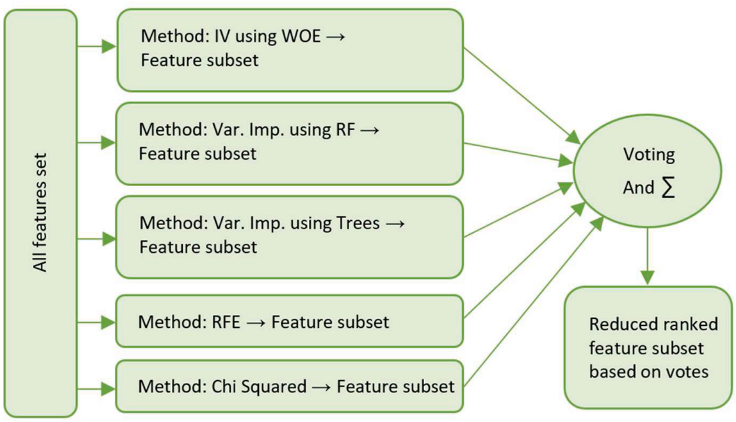


**Fig. 1.** Feature selection.

test is to compare the dependent variable with each of the independent variables and determine whether they are related or not [79].

The chi-square test for independence tests whether two variables are correlated or not. $H_0$ (null hypothesis) assumes that the two compared variables are independent, while $H_1$ (alternative hypothesis) expects dependence. We set $\alpha=0.05$, and a p-value of 0.05 or greater is considered critical. Anything less means that the deviations are significant, and the tested hypothesis must be rejected.

### 9.2. Methods for glucose level prediction

The partitioning of collected data into input windows for the personalized prediction model is illustrated in Fig. 2. Every five minutes, a singular reading was collected, sampling data from the CGM sensor. This sampled information was then utilized to construct a retrospective sliding window (PSW), containing historical data spanning from 3 to 36 h. The quantity of data employed by the model for prediction is governed by the PSW. Continuously, the model employed data from the sliding window to forecast glucose levels for 15, 30, and 45 min in advance at set prediction horizons (PH).

Time series cross-validation is a pivotal statistical technique in modeling, especially when data points have chronological or sequential order, such as patient medical records. Its application in our study was both nuanced and tailored to the unique challenges of DM1 data. Here, we elucidate the methodology with particular attention to our approach's granularity and specificity. Time series data fundamentally differ from cross-sectional data, mainly because of the inherent temporal structure. Thus, traditional cross-validation techniques that randomly partition the data might disrupt this order, leading to potential information leakage where future information might inadvertently be used to predict past events. This is especially problematic for our study, where DM1 patient data evolves over time, and each data point potentially influences subsequent ones. In our work, the dataset for each patient was divided into three distinct sets: training, validation, and test, with a split ratio of 70–20–10 % respectively. To clarify:

- Training Set (70 % of data): This portion of the data was used for model learning. The model internalized patterns, trends, and relationships in this data subset to make predictions.
- Validation Set (20 % of data): After training the model, we utilized the validation set to tune hyperparameters and optimize the model's performance, according to Table 2. This iterative process ensures that the model generalizes well to new, unseen data while preventing overfitting.
- Test Set (10 % of data): Once the model was refined using the validation set, it was evaluated on the test set, which it hadn't seen before. This assessment provides an unbiased performance metric, offering insights into the model's real-world efficacy.

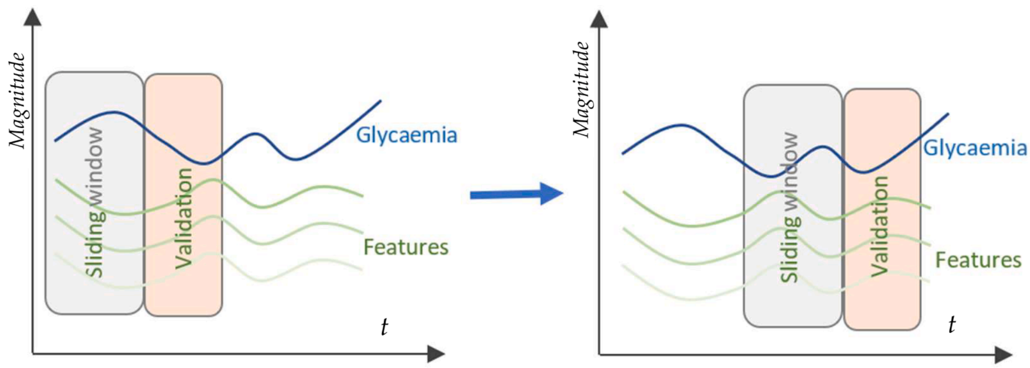


**Fig. 2.** Cross-validation and time series analysis using a slide window.

**Table 2**
Hyper-tunning parameters.

| Algorithm | Parameter | Range |
|---|---|---|
| Random forest (RF) | Max depth | 10 to 70 |
| | Min samples leaf | 1 to 4 |
| | Min samples split | 2 to 10 |
| | n estimators | 200 to 1200 |
| Support vector machines (SVM) | C | 0.1 to 1000 |
| | gamma | 1 to 0.0001 |
| | kernel | 'rbf' |
| Bayesian regularized | Hidden layers | 1 |
| Neural networks (BRNN) | Dense layer size | 50–350 |
| | Optimizer | 'adam' |
| | Learning rate | 0.1 to 0.001 |

The sliding window is a cornerstone of our TSCV strategy. At each iteration of the validation:

A fixed-size segment of the training data (the window) was used to train the model. The model was then validated on the next consecutive data segment, ensuring temporal integrity. The sliding window operates as described. Every time a new CGM value is received, the training dataset is restructured by deleting the oldest observation, shifting all values up by one place, and then inserting the newly arrived value as the newest value. Consequently, the size of the dataset and the order of the observations are always maintained. The window then "slides" forward, including the validation segment into the training data for the next iteration. For instance, in the first iteration, the model might be trained on the first 60 % of the training data and validated on the subsequent 10 %. In the next iteration, it's trained on the first 70 % (including the previous validation set) and validated on the next 10 %. This method ensures that every data point serves as both training and validation data at different iterations, bolstering the model's robustness.

It's paramount to mention that our models are patient-specific. Given the heterogeneity in DM1 manifestations, a model built for one patient might not perform optimally for another. This individualized modeling approach ensures the highest degree of accuracy as it tailors to the unique data characteristics of each patient.

The prediction methodologies for glucose levels examined in this study include the following:

*Random Forest* (RF): Regarded for its versatility and minimal processing resource requirements, RF is one such approach. RF, a Bagging technique, is employed by various algorithms and is distinguished by its iterative sampling of data points, thereby facilitating the creation of diverse training subsets originating from the same training data [80]. Subsequent to the formation of each training subset, decision trees are created, culminating in an ensemble. The ultimate determination of an incoming data instance's class label hinges upon the collective vote cast by each tree.

*Support Vector Regression* (SVR): Recognized as dual learning algorithms, SVRs process instances solely by calculating their dot products. This process can be effectively executed through a kernel function, without the necessity to iterate across all pertinent features [81]. Armed with the kernel function, the Support Vector Machine (SVM) learner endeavors to pinpoint a hyperplane that distinctly separates positive from negative samples, defining a margin. The method stands out for its resistance to overfitting and its powerful generalization performance, owing to the maximum-margin criteria employed in the optimization process. Further enhancing its effectiveness, SVR takes advantage of an appropriate convex optimization formulation, thereby ensuring its ultimate convergence to a global optimum.

*Bayesian Regularized Neural Networks* (BRNN): these neural networks are more robust than standard backpropagation networks and can reduce or eliminate the need for tedious cross-validation. Bayesian regularization is a mathematical process that converts a nonlinear regression into a "well-posed" statistical problem in the manner of a ridge regression. The advantage of BRNNs is that the models are robust, and the validation process plays a minor role. It is difficult to overtrain them because evidence procedures provide an objective Bayesian criterion for terminating training. Overfitting is also difficult because BRNN computes and trains a set of effective network parameters or weights, effectively eliminating those that are not relevant.

### 9.3. The measure of predictive performance

Prediction Error: The root mean square error (RMSE) is the metric of choice for gaging the accuracy of predictions in our analysis. This metric is widely recognized and frequently applied to assess prediction performance in relevant scholarly literature. Eq. (3) describes RMSE for a time series where we predict blood glucose levels, if we have N prediction points and $y_i$ is the actual blood glucose value at time $t_i$ and $y_i'$ is the value predicted by the model.

$$RMSE = \sqrt[2]{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_i'\right)^2} \quad (3)$$

Prediction Delay: defined as minimum time-shift between the predicted and observed signals. We consider positive (upward) and negative (downward) trends. Given a predictive horizon PH, suppose at time $t$, we predict that the glucose level will reach a certain

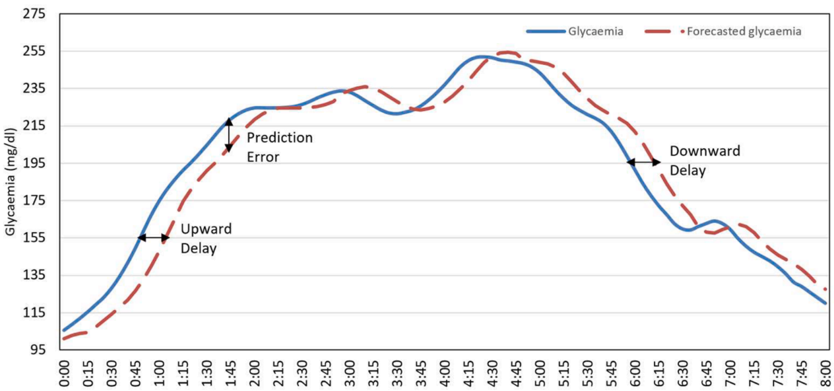


**Fig. 3.** Metrics to assess prediction performance.

value $y'$ at time $t$+PH. However, the actual time at which the glucose level reaches $y'$ is $t$+PH+$\Delta t$. Therefore, the Prediction delay $\Delta t$ can be defined as:

$$\Delta t = t_{actual} - (t + PH) \tag{4}$$

Where $t_{actual}$ is the actual time the glucose level reaches the predicted value $y'$ and $t$+PH is the time we initially predicted the glucose level would reach $y'$. Fig. 3 shows the different metrics to evaluate the performance of the forecasting algorithm.

The following actions were taken in order to develop suitable models:

- *Transformation*: The transformation process was a key step in model development. In this stage, all identified parameters were reconfigured to act as input data for our models. Factors such as frequency of changes in each parameter and its historical values could potentially serve as these representative features.
- *Normalization*: Once transformation was completed, normalization was undertaken. The aim of this stage was to ensure that all values fall within a standard range, namely the interval [0, 1]. This is an essential step in data preparation as it brings all data onto a common scale without distorting differences in ranges of values or losing information.
- *Evaluating Metrics*: For model evaluation, multiple metrics were employed.

  The root-mean-square error (RMSE) provided an estimate of the average error magnitude, offering a quantitative measure of model accuracy.
  The Prediction Delay metric, which is defined as the minimum time-shift between the predicted and observed signals, was also used to assess the model's predictive speed.

- *Validation Method*: The reliability and robustness of the models were further tested using a 10-fold cross-validation method, which was performed five times on the training dataset. This validation method is widely used as it provides a comprehensive assessment of model performance and generalizability.
- *Test*: The final step was the execution of tests, intended to verify the model's performance and its predictive capability under various scenarios. These tests were essential to assess the overall effectiveness of the developed models.

## 10. Results and discussion

Fig. 4 shows an example of the prediction process. By choosing a PSW, the biomedical data contained in that time window are collected, and with them and by means of a model, future glycemia is predicted at a predictive horizon (PH). The figure, constructed with real data, shows a PSW = 6 h and a PH of 45 min.

We investigate the accuracy with which future glucose levels may be predicted. The findings expressed in the root mean square error (RMSE) between planned and actual measured values are shown in Fig. 5. First, we note that the prediction error for all created patient models rises as the prediction horizon (PH) lengthens, which makes sense since the gathered data travel more and further away from the prediction horizon. 45 min could be adequate to evaluate the trend of glucose levels.

The importance of previous data volume in accurately forecasting future glucose levels will next be investigated. It is believed that

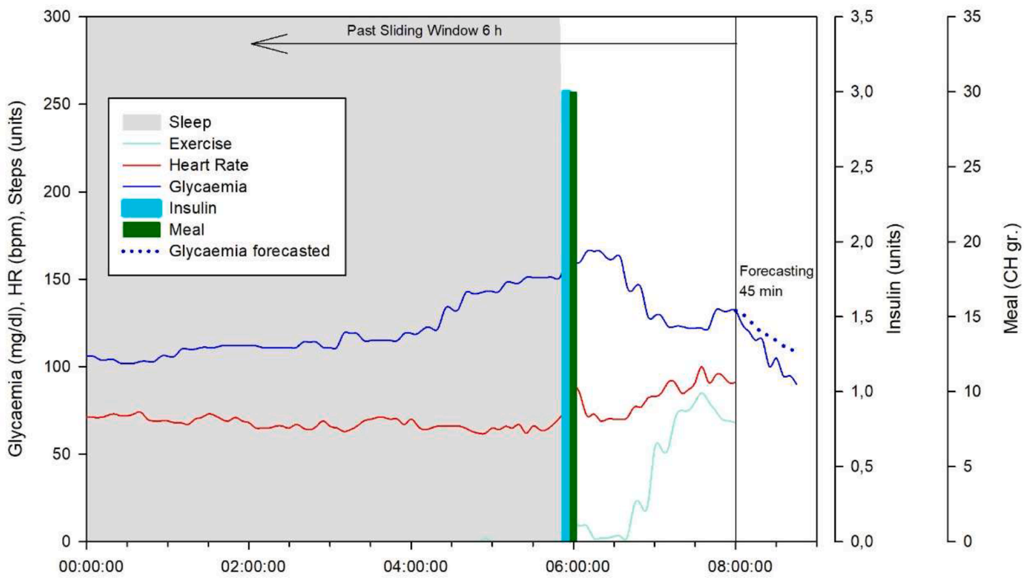


**Fig. 4.** Prediction process.

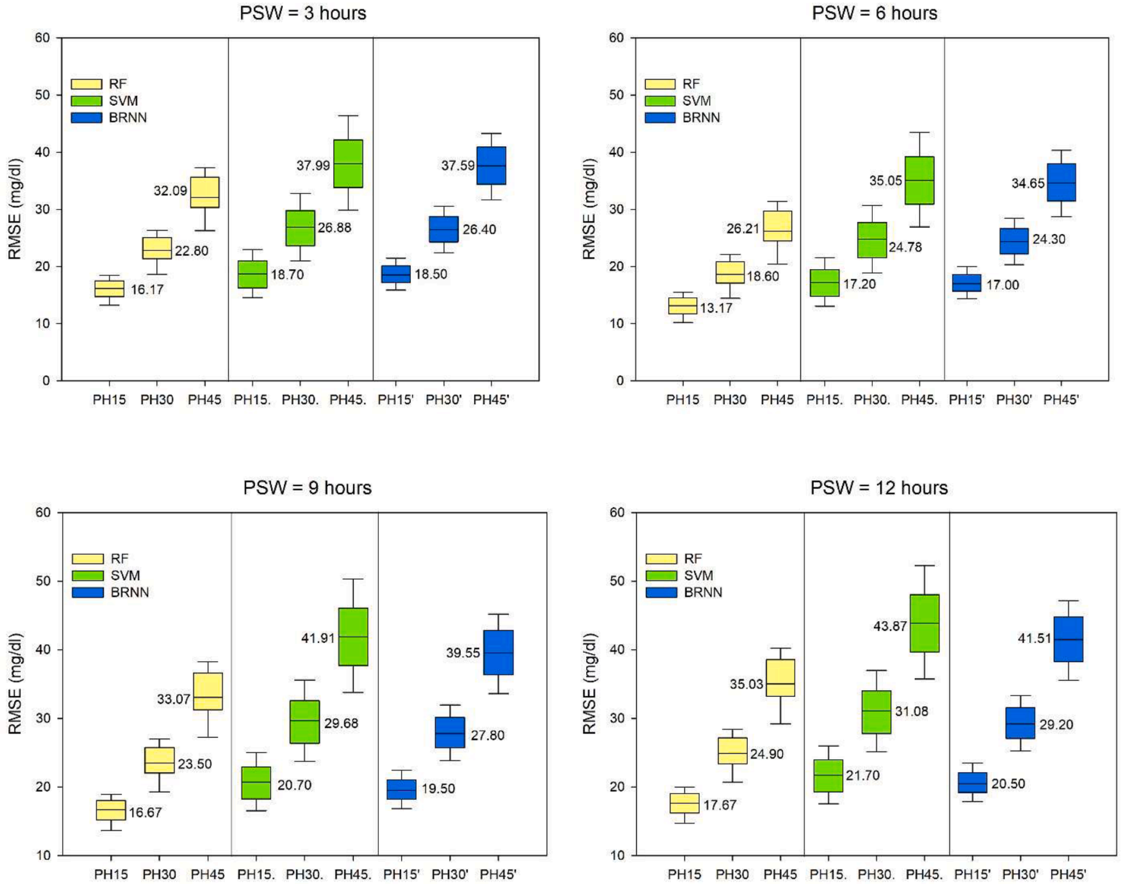


**Fig. 5.** RMSE expressed in mg/dL: RF, SVM, and BRNN using past sliding window lengths (PSWs) of 3, 6, 9, 12 h; predictive horizons (PHs) in minutes, 15, 30, and 45.

increasing the quantity of previous data would enhance the forecast's ultimate precision. The preceding data window size assists the model in capturing the temporal component of the time series. Using previous windows (PSW), we trained patient models using data from the prior 3, 6, 9, and 12 h. In this experiment, samples are taken every five minutes, and the prediction horizons are 15, 30, and 45 min.

Taking into account the different PSW sizes and the quantity of historical data necessary for accurate prediction, the findings indicate that all available person-centric models present a minimum RMSE in a 6-hour window length. When it is first raised the amount of historic data from 3 to 6 h, prediction accuracy increased. However, increasing the window size has a detrimental impact on

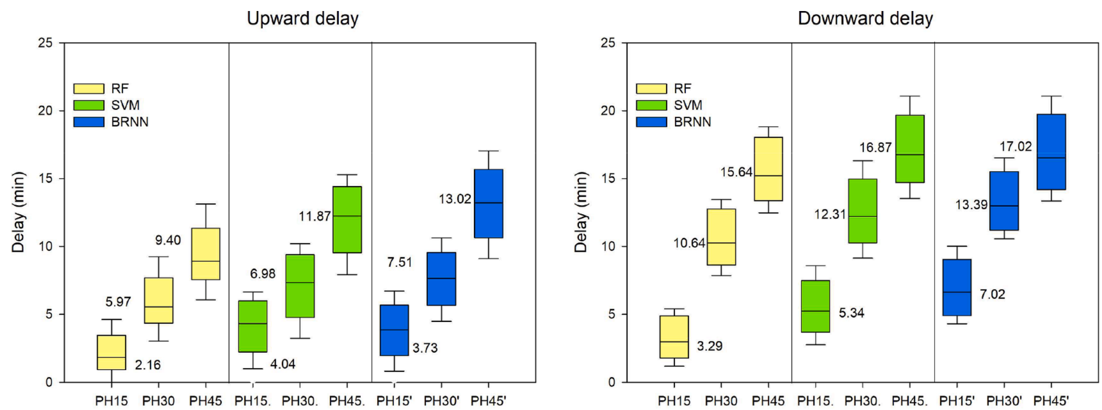


**Fig. 6.** Prediction delay expressed in minutes: random forest RF, SVM, and BRNN using past sliding window fixed in 6 h; predictive horizons (PHs) in minutes, 15, 30, and 45.

performance generally. The RMSE was greater for a 9-hour window than for a 3-hour timeframe. Any increase in the quantity of previous data reduces the acquired precision.

It is demonstrated by the evidence observed in experiments that, unrelatedly of the method used to create the models and for every forecast horizon, there is a boundary (6 h) that shifts backwards when considering past data to improve forecast accuracy; thus, there is an optimal value after which former data become irrelevant and reduce forecast accuracy. Prior research has linked the notion of circadian cycles [82] with the daily morning, afternoon, and evening time windows to establish this essential hierarchy.

Regarding the performance of the three distinct methods utilized to build person-centric models, the findings point that the RF technique is the most precise, giving high-quality forecasts for all PSWs and PHs investigated. Moreover, RF has a less significant standard deviation, indicating hence it is less dependent on the unique characteristics of each patient. We must believe that every subject's performance is influenced by their habits and characteristics.

Utilizing a constant desilting window of six hours, delays in prediction (both upward and downward) were assessed across various predictive horizons, as delineated in Fig. 6. The analysis reveals that the RF model exhibits superior performance in terms of both upward and downward times, aligning with the RMSE accuracy results previously examined. In the context of hyperglycemia prediction, the prediction error—or delay—is tolerable even at the longest predictive horizon when employing the BRNN.

It is essential to acknowledge that the implications of blood glucose levels exceeding the normal range are not as immediately perilous as their decrease; thus, there is a broader window for intervention. Conversely, the hypoglycemic interval is narrower and adjacent to euglycemia, necessitating swift action due to its instant impact on the patient's wellbeing. This factor accentuates the criticality of the downward delay over the upward delay and consequently influences the selection of the predictive method.

In this scenario, the outcomes are acceptable for predictive horizons extending up to 30 min, exclusively for the RF model. Nevertheless, the 45-minute predictive horizon results employing the BRNN model verge on potentially hazardous values.

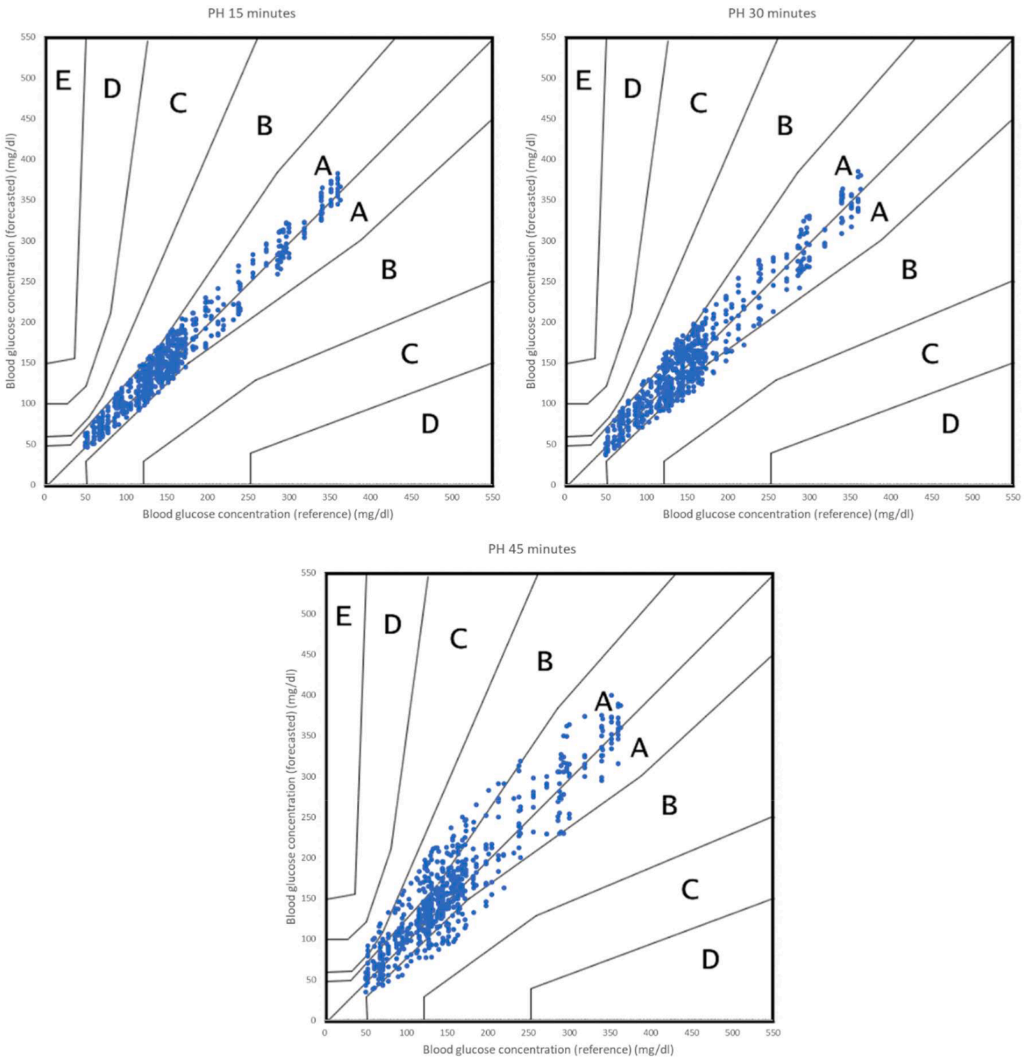


**Fig. 7.** Parkes grid forecasting performance. RF (algorithm), 6 h (PSW).

Remarkably, both models demonstrate swifter responses to increasing slopes compared to decreasing ones. The physiological characteristics typically cause faster escalation in trends than their reductions, and the shifts from ascent to descent are generally more pronounced than the inverse. Consequently, this causes both models to excessively extend beyond the peak in areas transitioning from an increase to a decrease, subsequently influencing the following downward phase. As such, there's a noticeable lag in the models as they cross the threshold during the ensuing decrease, a result of the initial overextension.

In order to evaluate more accurately, from the DM1 point of view, the performance of the RF algorithm, we are going to present the prediction results according to the so-called Parkes Grid [83]. The Parkes error grid was published in 2000 based on a survey of 100 physician attendees at the June 1994 American Diabetes Meeting. This metric has been intended for use in assessing clinical accuracy of blood glucose (BG) meters for patient self-measurement.

The grid is divided into 5 zones ranging from −20 % to +20 % in relation to the reference values, taking into account the direction and magnitude of the error.

Here's a brief overview of what each zone represents:

- *Zone A*: This zone encompasses values within 20 % of the reference value. The device's readings in this region are considered clinically accurate, and the treatment based on these readings would be appropriate.
- *Zone B*: This region represents readings that deviate by more than 20 % from the reference value but would still lead to benign or no treatment errors.
- *Zone C*: In this zone, readings could lead to an overcorrection of a hypo- or hyperglycemic event, which could potentially pose a moderate risk to the patient.
- *Zone D*: This zone represents values that would result in a failure to detect and treat hypoglycemia or hyperglycemia. The values in this region pose significant risk.
- *Zone E*: This zone is the most dangerous one and represents values that would confuse the treatment of hypoglycemia for hyperglycemia, and vice versa, posing severe risk.

Overall, the goal of glucose monitoring technologies is to have most of their readings fall within zones A and B, and minimize the values in zones C, D, and E.

By developing this new grid, the authors intended to provide an alternative to the original Clarke error grid, which was originally developed as a teaching tool rather than a BG meter clinical accuracy assessment tool and had been criticized for the placement of its risk boundaries. In our case, it will be used as a measure of the validity of the prediction, comparing it with the actual measurement obtained by the CGM for the predicted time.

Fig. 7 shows the results obtained for the different predictive horizons. We can see that for 15 and 30 min practically all the predictions are in zone A (clinically accurate measurements, no effect on clinical action), and in the case of 45 min only a part of the predictions are in zone B (altered clinical action, little or no effect on clinical outcome). The incursion into zone C is minimal for this PH (altered clinical action, likely to affect clinical outcome) and there are no results in zone D and E (altered clinical action, could have significant/dangerous clinical risk).

Before the prediction phase, a variable selection was made, as described in Section 9.1. Once we have decided that the best performance is achieved with RF for a PSW of 6 h, we will discuss the results of the variable selection ranking for that algorithm and that historical data. Fig. 8 depicts the ranking obtained. In any case, with other selections, the ranking obtained is broadly similar.

Initiating with glycaemia, our principal variable, we initially delve into it as an isolated feature. Subsequently, it functions as a predictive model's anticipated characteristic and an input parameter, incorporating historical data. In terms of significance, glycaemia holds the topmost position. Although there are some autoregressive models which only take into account glycaemia and the preceding half-hour [84], these choices lack adequate explanation. It's imperative to acknowledge that this feature somewhat embodies several other conditions potentially impacting glycaemia; hence the previous six hours can be interpreted as a condensed version of significant past events.

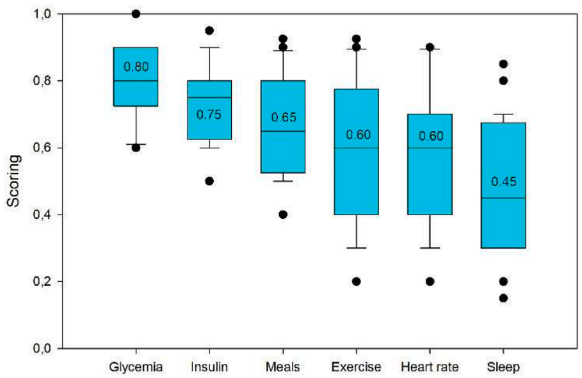


**Fig. 8.** Feature Selection ranking.

Next in the order of importance is insulin, closely followed by meal intake. It's notable that insulin exhibits a smaller standard deviation, given that its behavior is generally more consistent across individuals. Rapid-acting insulin, or boluses, mainly exhibit activity for approximately two and a half hours, peaking at 90 min [85]. However, evidence suggests a prolonged residual effect. The term "Insulin on Board" frequently surfaces in scientific literature, referring to the amount of insulin currently active in the body. This comprises basal insulin and remaining rapid insulin, both of which can have a low-intensity (yet detectable) effect for several hours. An exceptional influence has been observed for insulin active since the past eight hours [86], though other research indicates an acting range of five to eight hours [87].

Our third significant variable is meal intake, which appears to vary greatly between individuals. This aligns logically with the idea that individual dietary intake and metabolism diversely impact food absorption. For instance, meals rich in fat tend to influence glycaemia at a later stage [33,34], while fiber consumption can slow down the absorption of carbohydrates [88].

Exercise's effect on glycemic progression is evaluated as it enhances glucose demand and insulin sensitivity. It exhibits a high correlation with heart rate, suggesting that exercise data could be disregarded as it's likely included within the heart rate time series. However, this assumption warrants careful consideration and depends largely on the specific dataset, patients, and their habits. This aligns with the sports impact literature [31], demonstrating high variability across various sports and intensity levels. Nonetheless, the influence of exercise can persist up to 48 h later [89]. According to the data, individuals can be broadly categorized into those engaging in high-to-moderate intensity sports and a smaller group participating in less intense physical activities.

Heart rate stands as the sixth variable. If a heart rate increase isn't attributed to physical activity, it could be stress-induced [38]. Additionally, pulse variability may signal a hypoglycaemic episode [90], but heart rate responses tend to be extremely diverse among individuals.

Lastly, sleep duration is incorporated as a feature. Given the connection between sleep deprivation and elevated hyperglycemic stress hormones, we scrutinize the impact of sleep duration. Although its influence is minimal and varies widely from person to person, sleep deprivation has been linked to insulin resistance [91], leading to hyperglycemia in diabetics.

Blood glucose levels, or glycemia, is the principal variable in managing Type 1 diabetes mellitus (DM1), serving as a reflection of other influential variables and an essential predictor of its trajectory. Glycemia's pivotal role derives from its direct relevance to the disease's core pathology - the body's inability to regulate blood glucose levels.

New continuous glucose monitoring (CGM) systems offer robust data for glycemia, presenting the means for a continuous control algorithm that informs clinical decisions in real-time. Other factors, such as insulin input, exercise, meals, stress, sleep quality, and heart rate, can potentially impact glycemia, thereby underlying the need for their measurement or estimation. However, the commonality among these variables is their indirect effect on glycemia, reinforcing its centrality in DM1 management.

Importantly, the prediction of future glucose levels can be facilitated by previous glycemia data, particularly in auto-regressive model approaches. This prediction value is heightened when considering the temporal nuances of DM1, such as insulin activity periods, exercise impacts, meal absorption rates, and stress-induced hyperglycemia. The ripple effects of these variables on glycemia underscore its importance as a real-time gage of patient status and a foresight tool for disease progression.

Therefore, glycemia sits at the convergence point of these variables, reflecting their cumulative influence and providing an insightful, accurate basis for predicting its future course, thereby reinforcing its fundamental significance in DM1 management.

It can be concluded that insulin and food are, to a certain extent, variables of primary importance, with exercise, heart rate and sleep being secondary variables whose level of influence will be greater or lesser depending on the patient's lifestyle. Glycemia, which is the very variable we wish to predict, in a way encapsulates the information contained in all the others and is therefore the most relevant. In any case, as shown in Fig. 8, the order of importance depends on the patient and his or her specific characteristics, which reinforces the idea that these models must be specific to each individual.

The addition of more variables opens a debate as to whether to adopt a univariate/multivariate strategy. Within the myriad of approaches available for glycaemia prediction, there's a perennial debate between the efficacy of univariant methods, which rely solely on past glycaemia, and multivariant ones, which incorporate a wider array of physiological parameters.

Deep learning models have recently underscored their capacity in forecasting glucose levels, achieving an accuracy marked by an RMSE of 9.38±0.71 mg/dL over a 30-minute horizon [92]. In certain experimental frameworks, to mitigate computational overload, univariant options are adopted, effectively simplifying the prediction algorithm. The seminal work by Palumbo et al. [93], involving differential equations, though remarkable, diverges from the machine learning focus of this paper.

While there are undeniable advantages to univariant models in terms of computational efficiency, it's often at the expense of accuracy and richness of prediction. Pérez Gandía et al. [67], for instance, managed to predict glucose levels for 45 min using only glycaemia data within an Artificial Neural Network. Despite its reasonable error level, this methodology demands vast amounts of data, thereby extending training times and computational efforts. Such a narrow focus also amplifies the risk of overfitting.

In contrast, when we look at the work of Plis et al. [94], they utilized Autoregressive Integrated Moving Average (ARIMA) models, observing diabetic patients over a period and achieving RMSEs of 22.9 and 42.2 mg/dl for 30 and 60-minute prediction horizons respectively. When juxtaposed against Support Vector Machines (SVM) for the same dataset, RMSEs reduced to 19.6 and 35.7 mg/dl. This enhancement in SVM's performance can be attributed to the incorporation of other physiological features, hinting at the potential of multivariant models. This is further illustrated by Hamdi et al. [81], whose algorithm, though intricate, yielded superior results when blending past glycaemia values with SVM and Differential Evolution algorithms.

AR and ARIMA models have historically been the preferred choice for predictions based solely on past glycaemia. Woldaregay et al.'s [95] study corroborates this, where an autoregressive (AR) model achieved an accuracy of 85.3 % and 66.2 % for 30 and 60-minute horizons, respectively.

Beyond the above methodologies, the Random Forest (RF) algorithm deserves a notable mention. Unlike traditional decision trees which are prone to overfitting, RF, an ensemble learning method, offers a remedy by constructing multiple decision trees. The effectiveness of RF, particularly when contrasted with SVM, is evident in [96], where both demonstrated exceptional accuracy in predicting hypoglycaemia.

BGL prediction models are multifaceted, encompassing physiological, hybrid, and data-driven models. Especially, machine learning and classical time-series methodologies have been gaining traction [97,84]. For instance, Mirshekarian et al. [98] leveraged both simulated and real T1DM datasets, introducing variables like skin conductance and heart rate, and found improvements in prediction accuracy. In a similar exploration, Martinsson et al. [99] leaned on the Ohio T1DM dataset and RNN models, yielding an RMSE of 18.867 ± 1.794 mg/dl and 31.403 ± 2.078 mg/dl for 30 and 60-minute horizons, respectively.

A comprehensive study by Xie and Wang [100], which encompassed a multitude of variables, positioned the ARX and Ridge regression models at the forefront in terms of accuracy. Jeon et al.'s [101] research further accentuates the merit of multivariant methods, revealing that an ensemble model derived from 19 physiological and monitoring variables outshined individual models. On another front, Zhu et al. [92] championed the use of dilated RNNs, surpassing the performance of traditional models. Meanwhile, Guemes et al. [102] effectively harnessed multiple parameters to alert T1DM patients of nocturnal adverse glycaemia.

In summary, while univariant models, with their singular focus on past glycaemia, offer computational simplicity, multivariant models embrace a holistic approach, more accurate. By assimilating diverse physiological parameters, they deliver richer and more accurate predictions, emphasizing their indispensable role in modern BGL forecasting.

## 11. Conclusions

As a result of technological advancement, our world is always experiencing tremendous changes. All of these technologies are groundbreaking and life-improving in a variety of disciplines. New technology, such as CGMs and insulin pumps, have aided the quest for comprehensive, thorough, and definitive diabetes management. Hence, science has progressed, and new systems are now accessible.

In DM1 patients, glucose dynamics canbe unstable and they are affected by, among other things, insulin responses, nutrition, and lifestyle.

In this sense, the amount of information that can be generated, can be appropriately collected in an IoMT environment. The current connection systems lead to the existence of suitable proposals for data management. Only recently has it been possible to integrate, on the one hand, the application of biosensors with the successful management of an enormous volume of data, as well as the generation of knowledge from this data by means of machine learning applications. In the field of DM1, the application has proven to be straightforward and suitable.

To produce an accurate forecast, it is vital to determine how much information about a person's prior physiological state is required. This article also examines the necessary amount to accurately forecast future blood glucose levels in diabetics.

Forty DM1 patients were passively observed for a maximum of fourteen days, resulting in the development of many glycemia-related characteristics. Notably, the measured dataset comprises not only the CGM readings of a large number of people over an extended period of time and in real-world scenarios, but also insulin, meals, and cleverly derived factors such as exercise, heart rate, and sleep duration. These characteristics improve glycemia prediction. Using the Random Forest technique, the developed models can forecast blood sugar levels in a 30-minute horizon presenting an average RMSE of 18.60 mg/dL for six-hour data and 26.21 mg/dL for a 45-minute prediction horizon. Other models can also accurately forecast the future.

As a limitation, the present study has taken into account the measurements of 40 diabetics, which may be a limitation given the characteristics of the volunteers. Although our sample reflects a wide range of glycemic control, it is possible that patients who experience very poor disease care may suffer from poor accuracy performance.

Our research reveals that the best accurate predictions are generated from historical data encompassing six hours. Adding more data does not improve accuracy. In fact, the performance of all created forecasting models declines as the amount of past measures grows across all horizons of forecast under analysis. Lastly, our analysis demonstrates that the most accurate and effective glucose prediction models are those constructed using the RF approach.

The proposed method presents promising potential for future developments in the area of diabetes management and personalized medicine. The primary area of interest for advancing this research is the refinement of the predictive models, focusing on exploiting the power of machine learning techniques. With the advent of more sophisticated algorithms and increasing computational capabilities, it would be worthwhile to explore more advanced techniques such as deep learning methods, which could capture intricate patterns and relationships in the data more effectively.

For future study, it is crucial to expand our monitoring program beyond 14 days in order to assess the models' validity over time. In addition, contemporary prediction algorithms enable reliable forecasts of blood glucose levels for a particular prediction horizon. In addition, the IoMT smart environment is expanding every day, so future proposals can certainly be made with other contributions in this area.

Another future direction could involve expanding the range of biological and lifestyle variables considered by the model. While this study focused on a select number of features, there are myriad other factors which could potentially influence blood glucose levels. Incorporating additional parameters like stress levels, detailed dietary information, or even genetic data, could contribute to creating a more holistic and accurate predictive model.

Also, the integration of real-time patient feedback and personalized model tuning could be investigated. The incorporation of a patient feedback loop could provide the opportunity to fine-tune the model based on individual experiences and responses, increasing its accuracy and reliability for the individual user.

Finally, the application of this approach to other chronic diseases should be explored. The fundamental principles of continuous monitoring, data analysis, and predictive modeling could be adapted and applied to a range of other conditions, thus extending the reach and impact of this research.

These future developments could offer significant benefits to patients, healthcare providers, and the wider medical community, by enabling more effective, personalized disease management strategies and improving patient outcomes.

## Declaration of Competing 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.

## Data Availability

Data will be made available on request.

## Acknowledgments

Ignacio Rodríguez-Rodríguez would like to thank Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI), Junta de Andalucía, Spain.

María Campo-Valera is grateful for postdoctoral program Margarita Salas - Spanish Ministry of Universities (financed by European Union - NextGenerationEU).

The authors would like to acknowledge project PID2022-137461NB-C32 financed by MCIN /AEI /10.13039/501100011033 / FEDER(“Una manera de hacer Europa”), EU.

## References

[1] M.J. Haller, M.A. Atkinson, D. Schatz, Type 1 diabetes mellitus: etiology, presentation, and management, Pediatr. Clin. 52 (2005) 1553–1578.
[2] M.C. Riddell, I.W. Gallen, C.E. Smart, C.E. Taplin, P. Adolfsson, A.N. Lumb, A. Kowalski, R. Rabasa-Lhoret, R.J. McCrimmon, C. Hume, Exercise management in type 1 diabetes: a consensus statement, Lancet Diabetes Endocr. 5 (2017) 377–390.
[3] A.M. Albisser, B.S. Leibel, T.G. Ewart, Z. Davidovac, C.K. Botz, W. Zingg, An artificial endocrine pancreas, Diabetes 23 (1974) 389–396.
[4] T. Peyser, E. Dassau, M. Breton, J.S. Skyler, The artificial pancreas: current status and future prospects in the management of diabetes, Ann. N. Y. Acad. Sci. 1311 (2014) 102–123.
[5] C. Cobelli, E. Renard, B. Kovatchev, Artificial pancreas: past, present, future, Diabetes 60 (2011) 2672–2682.
[6] I. Rodríguez-Rodríguez, J.-V. Rodríguez, M.-Á. Zamora-Izquierdo, Variables to be monitored via biomedical sensors for complete type 1 diabetes mellitus management: an extension of the “on-board” concept, J. Diabetes Res. 2018 (2018).
[7] I. Rodríguez-Rodríguez, M.-Á. Zamora-Izquierdo, J.-V. Rodríguez, Towards an ICT-based platform for type 1 diabetes mellitus management, Appl. Sci. 8 (2018) 511.
[8] D. Karagiannis, K. Mitsis, K.S. Nikita, Development of a low-power iomt portable pillbox for medication adherence improvement and remote treatment adjustment, Sensors 22 (2022) 5818.
[9] C. Kaushal, M.K. Islam, A. Singla, M. Al Amin, An IoMT-based smart remote monitoring system for healthcare, IoT-Enabled Smart Healthcare Syst., Services. Applic. (2022) 177–198.
[10] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, A. Jamalipour, Wireless body area networks: a survey, IEEE Comm. surveys tutorials 16 (2014) 1658–1686.
[11] M.S.H. Talpur, M.Z.A. Bhuiyan, G. Wang, Energy-efficient healthcare monitoring with smartphones and IoT technologies, Int. J. High Perform. Comput. Networking 8 (2015) 186–194.
[12] S. Ali, M.G. Kibria, M.A. Jarwar, S. Kumar, I. Chong, Microservices model in WoO based IoT platform for depressive disorder assistance, in: 2017 International Conference on Information and Communication Technology Convergence (ICTC, 2017.
[13] L.E. Romero, P. Chatterjee, R.L. Armentano, An IoT approach for integration of computational intelligence and wearable sensors for Parkinson’s disease diagnosis and monitoring, Health Technol (Berl) 6 (2016) 167–172.
[14] A. Bajaj, M. Bhatnagar, A. Chauhan, Recent trends in internet of medical things: a review, Adv. Machine Learning Comput. Intell. (2021) 645–656.
[15] M.P. Hemalatha, M.R. Vidhyalakshmi, A study on chronic cough detection using IoT and machine learning, Inter. J. Res. Arts Sci. 5 (2019) 151–160.
[16] G. Matar, J.-M. Lina, J. Carrier, A. Riley, G. Kaddoum, Internet of Things in sleep monitoring: an application for posture recognition using supervised learning, in: 2016 IEEE 18th International conference on e-Health networking, applications and services (Healthcom), 2016.
[17] H. Magsi, A.H. Sodhro, F.A. Chachar, S.A.K. Abro, G.H. Sodhro, S. Pirbhulal, Evolution of 5 G in Internet of medical things, in: 2018 international conference on computing, mathematics and engineering technologies (iCoMET), 2018.
[18] K. Fan, Q. Luo, H. Li, Y. Yang, Cloud-based lightweight RFID mutual authentication protocol, in: 2017 IEEE Second International Conference on Data Science in Cyberspace (DSC, 2017.
[19] A. Onasanya, M. Elshakankiri, IoT implementation for cancer care and business analytics/cloud services in healthcare systems, in: Proceedings of the10th International Conference on Utility and Cloud Computing, 2017.
[20] G. Cappon, G. Acciaroli, M. Vettoretti, A. Facchinetti, G. Sparacino, Wearable continuous glucose monitoring sensors: a revolution in diabetes treatment, Electronics) 6 (2017) 65.
[21] B.P. Kovatchev, E. Renard, C. Cobelli, H.C. Zisser, P. Keith-Hynes, S.M. Anderson, S.A. Brown, D.R. Chernavvsky, M.D. Breton, Feasibility of outpatient fully integrated closed-loop control: first studies of wearable artificial pancreas, Diabetes Care. 36 (2013) 1851–1858.
[22] D.J. Cox, W.L. Clarke, L. Gonder-Frederick, S. Pohl, C. Hoover, A. Snyder, L. Zimbelman, W.R. Carter, S. Bobbitt, J. Pennebaker, Accuracy of perceiving blood glucose in IDDM, Diabetes Care. 8 (1985) 529–536.
[23] J. Place, A. Robert, N. Ben Brahim, P. Keith-Hynes, A. Farret, M.-J. Pelletier, B. Buckingham, M. Breton, B. Kovatchev, E. Renard, DiAs Web monitoring: a Real-Time Remote Monitoring System Designed For Artificial Pancreas Outpatient Trials, SAGE Publications Sage CA, Los Angeles, CA, 2013.
[24] E. Dassau, L. Jovanovic, F.J. Doyle III, H.C. Zisser, Enhanced 911/global position system wizard: a telemedicine application for the prevention of severe hypoglycemia—Monitor, alert, and locate, J. Diabetes Sci. Technol 3 (2009) 1501–1506.
[25] M. Rigla, Smart telemedicine support for continuous glucose monitoring: the embryo of a future global agent for diabetes care, J. Diabetes Sci. Techno.l 5 (2011) 63–67.

[26] S. Ding, M. Schumacher, Sensor monitoring of physical activity to improve glucose management in diabetic patients: a review, Sensors 16 (2016) 589.
[27] M.A. Case, H.A. Burwick, K.G. Volpp, M.S. Patel, Accuracy of smartphone applications and wearable devices for tracking physical activity data, JAMA 313 (2015) 625–626.
[28] C. Marling, L. Xia, R. Bunescu, F. Schwartz, 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, 2016.
[29] D. Rodbard, Continuous glucose monitoring: a review of successes, challenges, and opportunities, Diabetes Technol. Ther. 18 (2016) S2–S3.
[30] G.C. Estrada, L. del Re, E. Renard, Nonlinear gain in online prediction of blood glucose profile in type 1 diabetic patients, in: 49th IEEE Conference on Decision and Control (CDC), 2010.
[31] L.B. Borghouts, H.A. Keizer, Exercise and insulin sensitivity: a review, Int. J. Sports Med. 21 (2000) 1–12.
[32] C. Zecchin, A. Facchinetti, G. Sparacino, C. Dalla Man, C. Manohar, J.A. Levine, A. Basu, Y.C. Kudva, C. Cobelli, Physical activity measured by physical activity monitoring system correlates with glucose trends reconstructed from continuous glucose monitoring, Diabetes Technol. Ther. 15 (2013) 836–844.
[33] K.J. Bell, C.E. Smart, G.M. Steil, J.C. Brand-Miller, B. King, H.A. Wolpert, Impact of fat, protein, and glycemic index on postprandial glucose control in type 1 diabetes: implications for intensive diabetes management in the continuous glucose monitoring era, Diabetes Care. 38 (2015) 1008–1015.
[34] T. Hung, J.L. Sievenpiper, A. Marchie, C.W.C. Kendall, D.J.A. Jenkins, Fat versus carbohydrate in insulin resistance, obesity, diabetes and cardiovascular disease, Current Opinion Clinical Nutrition Metabolic Care 6 (2003) 165–176.
[35] J.C. Lovejoy, The influence of dietary fat on insulin resistance, Curr. Diab. Rep. 2 (2002) 435–440.
[36] M. Housiaux, H. Dorchy, O. Luminet, Influence of an emotional level of conscience on the glycemic equilibrium in type 1 diabetic children and adolescents, Diabetes Metab. 34 (2008) A33.
[37] K.L. Knutson, K. Spiegel, P. Penev, E. Van Cauter, The metabolic consequences of sleep deprivation, Sleep. Med. Rev. 11 (2007) 163–178.
[38] J. Taelman, S. Vandeput, A. Spaepen, S.V. Huffel, Influence of mental stress on heart rate and heart rate variability, in: 4th European conference of the international federation for medical and biological engineering, 2009.
[39] K. Johansen, S. Ellegaard, S. Wex, Detection of nocturnal hypoglycemia in insulin-treated diabetics by a skin temperature-skin conductance meter, Acta Med. Scand. 220 (1986) 213–217.
[40] J. Feldman, E. Gellhorn, The influence of fever on the vago-insulin and sympathetico-adrenal systems, Endocrinology 29 (1941) 141–143.
[41] T. Melhuish, Linking hypothermia and hyperglycemia, Nurs. Manage. 40 (2009) 42–45.
[42] V. McAulay, I.J. Deary, B.M. Frier, Symptoms of hypoglycaemia in people with diabetes, Diabet. Med. 18 (2001) 690–705.
[43] A. Lurbe, J. Redon, J.M. Pascual, J. Tacons, V. Alvarez, D.C. Batlle, Altered blood pressure during sleep in normotensive subjects with type I diabetes, Hypertension 21 (1993) 227–235.
[44] B. Guerci, L. Meyer, I. Delbachian, M. Kolopp, O. Ziegler, P. Drouin, Blood glucose control on Sunday in IDDM patients: intensified conventional insulin therapy versus continuous subcutaneous insulin infusion, Diabetes Res. Clin. Pract. 40 (1998) 175–180.
[45] P. Kekäläinen, M. Juuti, T. Walle, T. Laatikainen, Pregnancy planning in type 1 diabetic women improves glycemic control and pregnancy outcomes, J. Matern. Fetal. Neonatal. Med. 29 (2016) 2252–2258.
[46] J.R. Larsen, V.D. Siersma, A.S. Davidsen, F.B. Waldorff, S. Reventlow, N. de Fine Olivarius, The excess mortality of patients with diabetes and concurrent psychiatric illness is markedly reduced by structured personal diabetes care: a 19-year follow up of the randomized controlled study diabetes care in general practice (DCGP), Gen. Hosp. Psychiatry 38 (2016) 42–52.
[47] L. Tong, S. Adler, Glycemic control of type 2 diabetes mellitus across stages of renal impairment: information for primary care providers, Postgrad. Med. 130 (2018) 381–393.
[48] E.Y. Chew, There is Level 1 Evidence For Intensive Glycemic Control For Reducing the Progression of Diabetic Retinopathy in Persons With Type 2 Diabetes, 49, Springer, 2015, pp. 1–3.
[49] P. Wang, N. Jin, W.L. Woo, J.R. Woodward, D. Davies, Noise tolerant drift detection method for data stream mining, Inf. Sci. (Ny) 609 (2022) 1318–1333.
[50] D. Yang, B. Gao, W.L. Woo, H. Wen, Y. Zhao, Z. Gao, Wearable Structured Mental-Sensing-Graph Measurement, IEEE Trans. Instrum. Meas. (2022).
[51] B. Bogue-Jimenez, X. Huang, D. Powell, A. Doblas, Selection of noninvasive features in wrist-based wearable sensors to predict blood glucose concentrations using machine learning algorithms, Sensors 22 (2022) 3534.
[52] B. Bent, P.J. Cho, A. Wittmann, C. Thacker, S. Muppidi, M. Snyder, M.J. Crowley, M. Feinglos, J.P. Dunn, Non-invasive wearables for remote monitoring of HbA1c and glucose variability: proof of concept, BMJ Open Diabetes Res. Care 9 (2021), e002027.
[53] T. Zhu, C. Uduku, K. Li, P. Herrero, N. Oliver, P. Georgiou, Enhancing self-management in type 1 diabetes with wearables and deep learning, NPJ Digital Med. 5 (2022) 78.
[54] F. D'Antoni, M. Merone, V. Piemonte, G. Iannello, P. Soda, Auto-regressive time delayed jump neural network for blood glucose levels forecasting, Knowl Based Syst 203 (2020), 106134.
[55] C. Mosquera-Lopez, R. Dodier, N.S. Tyler, L.M. Wilson, J. El Youssef, J.R. Castle, P.G. Jacobs, Predicting and preventing nocturnal hypoglycemia in type 1 diabetes using big data analytics and decision theoretic analysis, Diabetes Technol. Ther. 22 (2020) 801–811.
[56] M. Mayo, L. Chepulis, R.G. Paul, Glycemic-aware metrics and oversampling techniques for predicting blood glucose levels using machine learning, PLoS One 14 (2019), e0225613.
[57] M. Zhang, K.B. Flores, H.T. Tran, Deep learning and regression approaches to forecasting blood glucose levels for type 1 diabetes, Biomed. Signal Process. Control 69 (2021), 102923.
[58] K. Li, C. Liu, T. Zhu, P. Herrero, P. Georgiou, GluNet: a deep learning framework for accurate glucose forecasting, IEEE J. Biomed Health Inform. 24 (2019) 414–423.
[59] A. Katayama, A. Tone, M. Watanabe, S. Teshigawara, S. Miyamoto, J. Eguchi, A. Nakatsuka, K. Shikata, J. Wada, The hypoglycemia-prevention effect of sensor-augmented pump therapy with predictive low glucose management in Japanese patients with type 1 diabetes mellitus: a short-term study, Diabetology Intern. 11 (2020) 97–104.
[60] G. Alfian, M. Syafrudin, M.F. Ijaz, M.A. Syaekhoni, N.L. Fitriyani, J. Rhee, A personalized healthcare monitoring system for diabetic patients by utilizing BLE-based sensors and real-time data processing, Sensors 18 (2018) 2183.
[61] F. D'Antoni, L. Petrosino, A. Marchetti, L. Bacco, S. Pieralice, L. Vollero, P. Pozzilli, V. Piemonte, M. Merone, Layered meta-learning algorithm for predicting adverse events in type 1 diabetes, IEEE Access 11 (2023) 9074–9094.
[62] T.M. Quan, T. Doike, D.C. Bui, S. Arata, A. Kobayashi, M.Z. Islam, K. Niitsu, AI-based edge-intelligent hypoglycemia prediction system using alternate learning and inference method for blood glucose level data with low-periodicity, in: 2019 IEEE international conference on artificial intelligence circuits and systems (AICAS), 2019.
[63] I. Rodríguez-Rodríguez, J.-V. Rodríguez, I. Chatzigiannakis, M.A. Zamora Izquierdo, On the possibility of predicting glycaemia 'on the fly'with constrained IoT devices in type 1 diabetes mellitus patients, Sensors 19 (2019) 4538.
[64] G. Cappon, M. Vettoretti, F. Marturano, A. Facchinetti, G. Sparacino, A neural-network-based approach to personalize insulin bolus calculation using continuous glucose monitoring, J. Diabetes Sci. Technol. 12 (2018) 265–272.
[65] F. Burden, D. Winkler, Bayesian regularization of neural networks, Artificial neural networks (2008) 23–42.
[66] H.T. Nguyen, N. Ghevondian, T.W. Jones, Detection of nocturnal hypoglycemic episodes (natural occurrence) in children with type 1 diabetes using an optimal Bayesian neural network algorithm, in: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008.
[67] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, C. Cobelli, E.J. Gómez, M. Rigla, A. de Leiva, M.E. Hernando, Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring, Diabetes Technol. Ther. 12 (2010) 81–88.
[68] N. Nuryani, S.S.H. Ling, H.T. Nguyen, Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection, Ann. Biomed. Eng. 40 (2012) 934–945.

[69] M.J. Fokkert, P.R. Van Dijk, M.A. Edens, S. Abbes, D. De Jong, R.J. Slingerland, H.J.G. Bilo, Performance of the freestyle libre flash glucose monitoring system in patients with type 1 and 2 diabetes mellitus, BMJ Open Diabetes Res. Care 5 (2017), e000320.
[70] T. Bailey, B.W. Bode, M.P. Christiansen, L.J. Klaff, S. Alva, The performance and usability of a factory-calibrated flash glucose monitoring system, Diabetes Technol. Ther. 17 (2015) 787–794.
[71] K.J.C. Wientjes, A.J.M. Schoonen, Determination of time delay between blood and interstitial adipose tissue glucose concentration change by microdialysis in healthy volunteers, Int. J. Artif. Organs 24 (2001) 884–889.
[72] A. Basu, S. Dube, S. Veettil, M. Slama, Y.C. Kudva, T. Peyser, R.E. Carter, C. Cobelli, R. Basu, Time lag of glucose from intravascular to interstitial compartment in type 1 diabetes, J. Diabetes Sci. Technol 9 (2014) 63–68.
[73] U. Hoss, E.S. Budiman, Factory-calibrated continuous glucose sensors: the science behind the technology, Diabetes Technol. Ther. 19 (2017) S–44.
[74] L. Chamakura, G. Saha, An instance voting approach to feature selection, Inf. Sci. (Ny) 504 (2019) 449–469.
[75] D.L. Weed, Weight of evidence: a review of concept and methods, Risk Analysis: An Intern. J. 25 (2005) 1545–1557.
[76] J. Kolácek, M. Rezác, Assessment of scoring models using information value, in: 19th International Conference on Computational Statistics, Paris France, 2010.
[77] S. Nembrini, I.R. König, M.N. Wright, The revival of the Gini importance? Bioinformatics 34 (2018) 3711–3718.
[78] B.F. Darst, K.C. Malecki, C.D. Engelman, Using recursive feature elimination in random forest to account for correlated variables in high dimensional data, BMC Genet. 19 (2018) 1–6.
[79] A. Jović, K. Brkić, N. Bogunović, A review of feature selection methods with applications, in: 2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO), 2015.
[80] L. Breiman, Bagging predictors, Mach Learn 24 (1996) 123–140.
[81] T. Hamdi, J.B. Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, J.-M. Ginoux, Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm, Biocybern. Biomed. Engin. 38 (2018) 362–372.
[82] E. Van Cauter, K.S. Polonsky, A.J. Scheen, Roles of circadian rhythmicity and sleep in human glucose regulation, Endocr. Rev. 18 (1997) 716–738.
[83] A. Pfützner, D.C. Klonoff, S. Pardo, J.L. Parkes, Technical aspects of the Parkes error grid, J. Diabetes Sci. Technol 7 (2013) 1275–1281.
[84] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, C. Cobelli, Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series, IEEE Trans. Biomed. Eng. 54 (2007) 931–937.
[85] F. Holleman, J.B.L. Hoekstra, Insulin lispro, N. Engl. J. Med. 337 (1997) 176–183.
[86] S.D. Patek, L. Magni, E. Dassau, C. Karvetski, C. Toffanin, G. De Nicolao, S. Del Favero, M. Breton, C. Dalla Man, E. Renard, Modular closed-loop control of diabetes, IEEE Trans. Biomed. Eng. 59 (2012) 2986–2999.
[87] C. Ellingsen, E. Dassau, H. Zisser, B. Grosman, M.W. Percival, L. Jovanovič, F.J. Doyle III, Safety constraints in an artificial pancreatic β cell: an implementation of model predictive control with insulin on board, J. Diabetes Sci. Technol. 3 (2009) 536–544.
[88] T.G. Kiehm, J.W. Anderson, K. Ward, Beneficial effects of a high carbohydrate, high fiber diet on hyperglycemic diabetic men, Am. J. Clin. Nutr. 29 (1976) 895–899.
[89] J.T. Devlin, M. Hirshman, E.D. Horton, E.S. Horton, Enhanced peripheral and splanchnic insulin sensitivity in NIDDM men after single bout of exercise, Diabetes 36 (1987) 434–439.
[90] C. Alexakis, H.O. Nyongesa, R. Saatchi, N.D. Harris, C. Davies, C. Emery, R.H. Ireland, S.R. Heller, Feature extraction and classification of electrocardiogram (ECG) signals related to hypoglycaemia, Comput. Cardiol. (2003) 2003.
[91] E. Donga, M. van Dijk, J.G. van Dijk, N.R. Biermasz, G.-J. Lammers, K.W. van Kralingen, E.P.M. Corssmit, J.A. Romijn, A single night of partial sleep deprivation induces insulin resistance in multiple metabolic pathways in healthy subjects, J. Clin. Endocrinol. Metabol. 95 (2010) 2963–2968.
[92] T. Zhu, K. Li, J. Chen, P. Herrero, P. Georgiou, Dilated recurrent neural networks for glucose forecasting in type 1 diabetes, J. Healthcare Inform. Res. 4 (2020) 308–324.
[93] P. Palumbo, S. Ditlevsen, A. Bertuzzi, A. De Gaetano, Mathematical modeling of the glucose–insulin system: a review, Math. Biosci. 244 (2013) 69–81.
[94] K. Plis, R. Bunescu, C. Marling, J. Shubrook, F. Schwartz, A machine learning approach to predicting blood glucose levels for diabetes management, in: Workshops at the Twenty-Eighth AAAI conference on artificial intelligence, 2014.
[95] A.Z. Woldaregay, E. Årsand, T. Botsis, D. Albers, L. Mamykina, G. Hartvigsen, Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes, J. Med. Internet Res. 21 (2019) e11030.
[96] B. Sudharsan, M. Peeples, M. Shomali, Hypoglycemia prediction using machine learning models for patients with type 2 diabetes, J. Diabetes Sci. Technol. 9 (2014) 86–90.
[97] S. Oviedo, J. Vehí, R. Calm, J. Armengol, A review of personalized blood glucose prediction strategies for T1DM patients, Int. J. Numer Method Biomed. Eng. 33 (2017) e2833.
[98] S. Mirshekarian, R. Bunescu, C. Marling, F. Schwartz, Using LSTMs to learn physiological models of blood glucose behavior, in: 2017 39th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2017.
[99] J. Martinsson, A. Schliep, B. Eliasson, O. Mogren, Blood glucose prediction with variance estimation using recurrent neural networks, J. Healthcare Inform. Res. 4 (2020) 1–18.
[100] J. Xie, Q. Wang, Benchmarking machine learning algorithms on blood glucose prediction for type I diabetes in comparison with classical time-series models, IEEE Trans. Biomed. Eng. 67 (2020) 3101–3124.
[101] J. Jeon, P.J. Leimbigler, G. Baruah, M.H. Li, Y. Fossat, A.J. Whitehead, Predicting glycaemia in type 1 diabetes patients: experiments in feature engineering and data imputation, J. Healthcare Inform. Res. 4 (2020) 71–90.
[102] A. Güemes, G. Cappon, B. Hernandez, M. Reddy, N. Oliver, P. Georgiou, P. Herrero, Predicting quality of overnight glycaemic control in type 1 diabetes using binary classifiers, IEEE J. Biomed. Health Inform. 24 (2019) 1439–1446.