Penalty Weighted Glucose Prediction Models Could Lead to Better Clinically Usage

Cichosz, Simon Lebech; Kronborg, Thomas; Jensen, Morten Hasselstrøm; Hejlesen, Ole

Published in:
Computers in Biology and Medicine

DOI (link to publication from Publisher):
10.1016/j.compbiomed.2021.104865

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Computers in Biology and Medicine 138 (2021) 104865

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Penalty weighted glucose prediction models could lead to better clinically usage

Simon Lebech Cichosz [a,*], Thomas Kronborg [a], Morten Hasselstrøm Jensen [a,b], Ole Hejlesen [a]

[a] Department of Health Science and Technology, Aalborg University, Denmark
[b] Steno Diabetes Center North Denmark, Aalborg University Hospital, Aalborg, Denmark

ARTICLE INFO

Keywords:
Neural network
Ensemble learning
Prediction
CGM
Type 1 diabetes
Glucose
Continuous glucose monitoring

ABSTRACT

Background and objective: Numerous attempts to predict glucose value from continuous glucose monitors (CGM) have been published. However, there is a lack of proper analysis and modeling of penalty for errors in different glycemic ranges. The aim of this study was to investigate the potential for developing glucose prediction models with focus on the clinical aspects.

Methods: We developed and compared six different models to test which approach were best suited for predicting glucose levels at different lead times between 10 and 60 min. The models were: last observation carried forward, linear extrapolation, ensemble methods using LSBoost and bagging, neural networks, one without error-weights and one with error-weights. The modeling and test were based on 225 type 1 diabetes patients with 315,000 h of CGM data.

Results: Results show that it is possible to predict glucose levels based on CGM with a reasonable accuracy and precision with a 30-min prediction lead time. A comparison of different methods shows that there are improvements on performance gained from using more advanced machine learning algorithms (MARD 10.26–10.79 @ 30-min lead time) compared to a simple modeling (MARD 10.75–12.97 @ 30-min lead time). Moreover, the proposed use of error weights could lead to better clinical performance of these models, which is an important factor for real usage. E.g., the percentages in the C-zone of the consensus error grid without error-weights (0.57-0.68%) vs including error-weights (0.28%).

Conclusions: The results point toward that using error weighting in the training of the models could lead to better clinical performance.

1. Introduction

It is estimated that more than 20 million people in the United States have diabetes and approximately 5–10% of people with diabetes have insulin dependent diabetes (type 1 diabetes)[1]. Diabetes and complications related to diabetes are a major burden for the individual patient and a heavy economic burden for the healthcare sector[1]. Adequate and sustained control of blood sugar levels is a key element in the prevention or delay of onset of diabetes-related complications[2].

Continuous glucose monitors (CGMs) have proven effective in the reduction of glucose variability, hypo- and hyperglycemic, and glycosylated hemoglobin (HbA1c) levels among people with diabetes[3–5]. With the increasing use of CGM, research have accelerated toward the development and usage of an artificial pancreas. Many proposed control algorithms for an artificial pancreas are based on predicting glucose levels ahead of time. Precise forecasts on glucose level behavior could substantially improve diabetes regulation, making it more robust and efficient. Moreover, a precise and timely forecast could lead to reduction in adverse metabolic events such as hypo and hyperglycemia[6–8]. Precise predictions is important because these could enable timely regulations of insulin doses or glucagon treatment either automatically (artificial pancreas) or manually (injections), potentially improving blood glucose time-in-range (TIR; 70–180 mg/dL; 3.9–10.0 mmol/L) for people with diabetes. Stable blood glucose within TIR has been proposed as a strong prognostic marker for disease progression and avoiding long-term complications such as retinopathy, neuropathy and nephropathy[9].

We have in a recently published study shown in a heterogeneous cohort how a neural network model could reduce the time-lag induced by the CGM[10]. However, to fully integrate a predictive glucose

* Corresponding author. Postal address: Fredrik Bajers Vej 7D2 - DK-9220 Aalborg, Denmark.
E-mail address: simcich@hst.aau.dk (S.L. Cichosz).

https://doi.org/10.1016/j.compbiomed.2021.104865
Received 2 August 2021; Received in revised form 27 August 2021; Accepted 10 September 2021
Available online 15 September 2021

algorithm with an artificial pancreas system, a longer prediction horizon is often needed to make timely adjustments due to the slower effect of insulin on blood glucose. Furthermore, many of the already published studies on glucose prediction have been modeled on in silico data, cohorts including few patients, and during a short observational time [11–14]. Glucose dynamics is complex, effected by intra- and inter-variability due to e.g. disease progression, medicine, infections, physical activity, smoking, ingested food, and physiological stress [15–17]. There is a need for future research to both examine a longer prediction horizon, clinical accuracy, the ability under free living conditions, and a thorough validation of the results in real patients to capture a realistic performance with the complexity of the intra- and inter-variability[14]. Additionally, from a clinical perspective, the potential for improving treatment is not only based on a general accuracy of glucose predictions. Different glucose levels, such as hypoglycemia, euglycemia, and hyperglycemia, would require a customized approach. Due to these aspects of the treatment, there is also a need to investigate prediction models accuracy in relation to the glucose level and potential tailoring the models to fit each scenario.

Therefore, the aim of this study was to investigate the potential for developing a generalizable glucose prediction model for usage under different regions of the glucose physiological span such as hypoglycemia, euglycemia, and hyperglycemia, utilizing CGM data from a large heterogeneous cohort. The findings presented in this study are also focusing on assessing the accuracy of several prediction horizons (lead time).

2. Methods

To investigate the potential for predicting future glucose values with focus on clinical usability we developed, deployed, and compared six different (simple and advanced) models with a prediction lead time from 10 to 60 min based on a window of retrospective samples of CGM from each participant. This study adhere to the guideline for Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)[18].

2.1. Data material

We included 225 type diabetes 1 patients form the REPLACE-BG trial. The results were initially published by Aleppo et al. [19]. Participants used a CGM (Dexcom G4) and an insulin pump and was followed for up to 6 months. The inclusion criteria for the REPLACE-BG trial were the usage of insulin pump, hemoglobin A1c (HbA1c) $< 8.5\%$ (8.5% NGSP = 69 mmol/mol IFCC = 197 mg/dL eAG = 11.0 mmol/l eAG), no history of severe hypoglycemia and unawareness. The cohort characteristics presented as mean/(SD)/$\overline{x}(\sigma)$ were for age 44 years (14), BMI 27.7 kg/m^2 (4.1), diabetes duration 23 years (12). In this study, all participants with measurements of CGM were included in the analytic cohort.

2.2. Outcome

The assessed outcome were accuracy and precision of the predicted glucose levels compared the measured reference value measured by the CGM. The predictions had lead times of $\Delta = 10,20,30,40,50$, and 60 min between the point in time where the prediction calculated and the point in time where the predicted measurement was assessed.

2.3. Statistical assessment

Statistical assessment was used to compare the predicted glucose levels with the measured CGM levels. To quantify the prediction performance, several statistical metrics were assessed. We included general performance metrics such as root mean square error (RMSE) and mean absolute relative difference (MARD). The calculation of RMSE and MARD are presented in equations (1)–(3).

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

$$ARD = 100\frac{\left|Y_i - \widehat{Y}_i\right|}{Y_i} \quad \text{Equation 2}$$

$$MARD = \frac{1}{N}\sum_{i=1}^{N} ARD_i \quad \text{Equation 3}$$

RMSE and MARD are often reported in the literature concerning glucose prediction, these metrics assess a general performance and fit of the models. However, these metrics do not assess the performance with the focus of clinical relevance. The predicted glucose accuracy and precision tolerance is highly influenced by the level of blood glucose. For instance, a much higher error is generally tolerated in the range of hyperglycemia because, form a clinical perspective, this will less likely lead to an incorrect treatment of the subject. On the other hand, in the range of low euglycemia and hypoglycemia a much smaller error could lead to an incorrect or delayed treatment with serious consequences for the subject. Therefore, assessing the clinical aspect is important. In this study, we used the Consensus Error Grid Analysis[20,21]. This technique is used to evaluate the precision of new glucose measurements compared to a reference standard. It helps assess the clinical effect of the differences between the CGM measurements and predicted glucose levels. The analysis labels the predicted values into five zones A-E, depending on the clinical risk of predicting a glucose value in the particular zone. Described in short, predictions in zone A are considered clinically accurate; predictions in zone B are benign errors; predictions in zone C are characterized as the potential for over correction; predictions in zone D describes the potential for delayed treatment, and predictions in zone E presents clinical errors. The Consensus Error Grid with zones for people with type 1 diabetes is presented in Fig. 1.

2.4. Prediction approach

In this work, we developed and compared six different models to test which approach were best suited for predicting glucose levels at different lead times. The models consisted of a simple “dumb” approach,

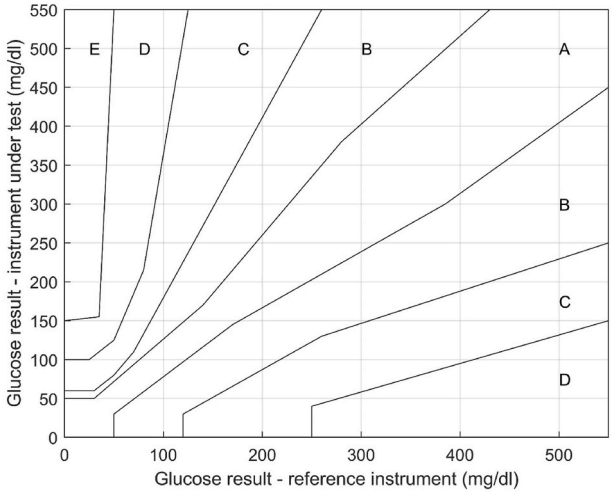

Fig. 1. Consensus grid for patients with type 1 diabetes. Zone A are considered clinically accurate; predictions in zone B are benign errors; predictions in zone C is characterized as the potential for over correction; predictions in zone D describes the potential for delayed treatment, and predictions in zone E presents clinical errors.

last observation carried forward (LOCF) [22]; a linear extrapolation (Extrap); two ensemble methods using Least-squares boosting (LSBoost) [23] and bagging (Bag) [24]; two neural networks, one without error weights (Net) and one with error weights (Net ew).

Simple models We included two straightforward and simple methods for glucose prediction. The first method was a last observation carried forward (LOCF) approach where the last obtained CGM level were carried xΔ minutes forward as an alternative to the predicted value. The second method was a linear extrapolation where retrospective CGM data (last 15 min) was used to calculate a predicted glucose level xΔ minutes into the future. The two simple approaches were included as a prediction reference and to assess the potential gain from using more advanced methods.

2.5. Advanced models

We choose to compare four nonlinear regression algorithms (LSBoost, Bag, Net and Net ew) for the prediction of future CGM measurements. This choice was motivated by the ensemble methods (EM) and artificial neural network (ANN) modeling capabilities of nonlinear and non-stationary problems. These methods have several advantages, i.e. flexible input, ability to indirectly identify dynamic non-linear interactions between dependent and independent predictors, ability to identify all potential interactions between predictors and have demonstrated high performance in solving medical prediction challenges [25–27].

The EM and ANN required training/validation of the models to find the optimal set of hyperparameters or number of hidden layers/neurons in each hidden layer/neuron's weights. The design of this study's model training, validation, and test is presented in Fig. 2. The data material was at patient level randomly divided such that 70% was used for training/validating of the models and 30% of the data were reserved for testing only. The proportion reserved for training was again split into a proportion used for training and one used for validation of the model. This procedure ensured that the results were not prone to overfit and would be transferable to a similar cohort using the same type of CGM sensor [28].

The ANN implementation was performed in Matlab R2020b (The Mathworks Inc., Natick, Massachusetts) using the Levenberg-Marquardt algorithm.

To investigate if the use of error weights could be beneficial in optimizing the clinical performance of the models, we implemented one ANN without and one with error weights. The weights were constructed in a such way that prediction error below 100 mg/dL glucose would be penalized 5:1 and error in the hypoglycemic range below 70 mg/dL would be penalized 10:1. The reason for penalizing these ranges were based on the clinical relevance of accurate prediction of low glucose levels. As also seen from Fig. 1 (consensus grid), even small percentages of errors in zone C, D, or E could result in improper treatment and potentially dangerous situation for the patient. The error weights were implemented in the mean square error (MSE) performance function used to train the network such that each squared error contributes an individual amount, based on the individual error weight w_i^e, to the performance function:

$$F = MSE = \frac{1}{N}\sum_{i=1}^{N} w_i^e \left(Y_i - \widehat{Y}_i\right)^2$$

The reported ANNs has five layers: an input layer, three hidden layers where weight adaptations were conducted using a hyperbolic tangent sigmoid transfer function, and an output layer modeled by a linear activation function. The number of layers and size was optimized using the training dataset. The ANNs model was not recalibrated based on the test datasets. This architecture have previous shown to capture the complexity of the CGM signal with good results[29].

The two EMs were also implemented using Matlab R2020b (The Mathworks Inc., Natick, Massachusetts). The hyperparameters were optimized regarding performance and transferability using the training/validation split of the data. LSBoost and Bag were both optimized for finding the best number of trees, learning cycles, and leaf size. The following search space for hyperparameter optimization was; (number of trees 10–150; learning cycles 1–100; leaf size 10–100).

Input features for the four models were derived in a sliding window of 120 min retrospectively to a point in time where the prediction was computed. The features were provided to the models as matrix of absolute readings in the timely order of measurement:

$$input = \left[f_{i=-23}\,; f_{i=-22}\,; \ldots f_{i=0;}\right]$$

To concretize this: each time the CGM reading is performed the models will use this reading and 23 adjacent readings prior to the current reading as input to the model, in order to make a glucose level prediction xΔ minutes into the future. This imply that the models will need a warm-up of 120 min wear-time before the first prediction can be computed. Every CGM reading were qualified for prediction if there were adjacent readings as described above and subsequently reading 60

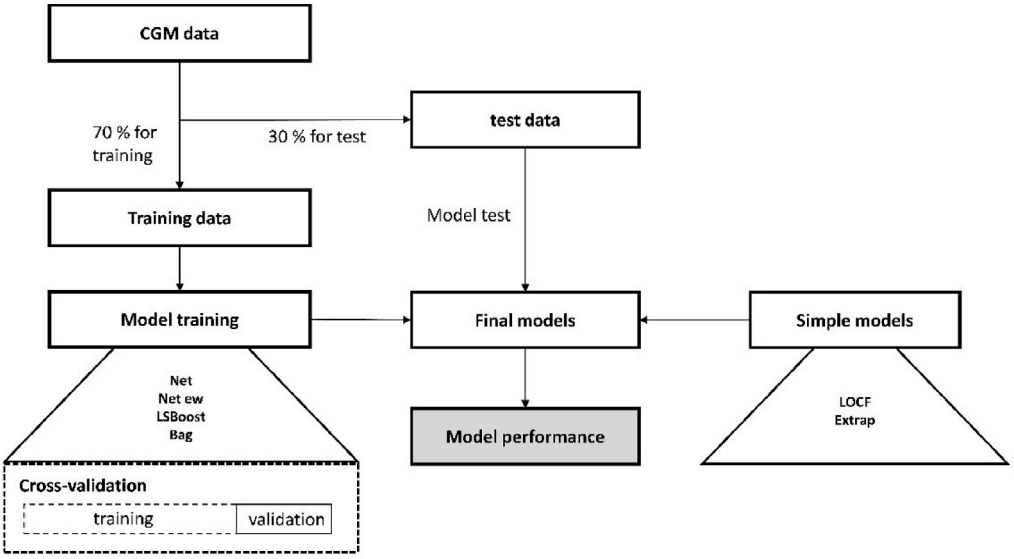

Fig. 2. Flow diagram for splitting data material, training, validating and testing the models.

min post the point used for prediction, in order to assess the accuracy of the prediction.

3. Results

A total of 158 participants with type 1 diabetes were included in the training data; combined they had 13,267,540 min of eligible CGM wear time which were used for training the six prediction methods. Further, 68 participants with type 1 diabetes and eligible wear time of 5,686,085 min were used to internally validate the six prediction methods. Each of the six methods (LOCF, Extrap, LSBoost, Bag, Net, Net ew) were used to train a model for prediction lead times between 10 and 60 min. Hence, six models were trained for each method with lead time of $\Delta =$ 10,20,30,40,50, and 60 min. The performance was assessed for each of the trained models using the test data. The test results are presented in Table 1. Fig. 3 is illustrating a comparison of the models in relation to prediction time (10–60 min) for MARD %, RMSE mg/dL, the % of prediction in the C zone of the Parkers Error Grid, and the % of prediction in the D zone of the Parkers Error Grid.

The performance measured by all statistical assessments is negatively correlated with the lead time. This is of course as expected as longer lead time would induce a large uncertainty into the prediction of glucose levels. In general, the LOCF method is performing significantly worse than the other assessed methods. Looking at the performance measured by MARD and RMSE between the remaining methods, overall, less significant differences is observed between the methods.

In the Clark Error Grid Analysis, the ANN with error weights (Net ew) had the highest percentages of prediction in zone A + B for all lead times assessed. The absolute differences percentages are relatively small, but never the less clinical important. This is especially observed when comparing the performance between methods in zone C and D from Fig. 3. The performance of the Net ew is significantly better for both zones compared to the other methods. This is also shown in Fig. 4, where a slice of data (sliced for viewing purposes) is presented in the Clark Error Grid. The figure presents the difference in performance between the Net and the Net ew with a lead time of 30 min. The zone C is strongly populated using the Net model compared to the Net ew model. These predictions could from a clinical perspective lead to harmful or delayed treatment for hypoglycemia.

Table 1
Results from the models grouped by the different prediction horizon from 10 to 60 min.

	Models					
	LOCF	Extrap.	Net	Net ew	LSBoost	Bag
10 min prediction						
RMSE	12.04	11.57	8.76	9.01	9.01	8.52
MARD	5.84	5.15	4.05	4.06	4.29	3.94
Parker error grid						
- Zone A + B	99.96	99.95	99.97	99.98	99.97	99.97
- Zone A	97.04	97.63	98.97	98.79	98.92	99.06
- Zone B	2.92	2.32	1.01	1.19	1.05	0.92
- Zone C	0.04	0.04	0.03	0.02	0.03	0.03
- Zone D	0	0	0	0	0	0
- Zone E	0	0	0	0	0	0
20 min prediction						
RMSE	19.44	16.1	14.95	15.32	15.21	14.54
MARD	9.62	7.57	7.43	7.26	7.63	7.17
Parker error grid						
- Zone A + B	99.72	99.88	99.85	99.91	99.83	99.83
- Zone A	89.97	93.7	94.85	94.01	94.58	95.14
- Zone B	9.75	6.18	5.01	5.9	5.24	4.69
- Zone C	0.28	0.11	0.15	0.09	0.17	0.17
- Zone D	0.01	0	0	0	0	0
- Zone E	0	0	0	0	0	0
30 min prediction						
RMSE	25.98	22.15	20.65	21.63	21.01	20.31
MARD	12.97	10.75	10.58	10.26	10.79	10.3
Parker error grid						
- Zone A + B	99.17	99.61	99.42	99.72	99.31	99.33
- Zone A	82.71	87.52	88.72	87.23	88.19	89.03
- Zone B	16.46	12.09	10.7	12.48	11.12	10.3
- Zone C	0.8	0.37	0.57	0.28	0.68	0.66
- Zone D	0.03	0.01	0.01	0.01	0.01	0.01
- Zone E	0	0	0	0	0	0
40 min prediction						
RMSE	31.77	28.08	26.46	27.23	26.21	25.54
MARD	15.97	13.83	13.82	12.87	13.68	13.21
Parker error grid						
- Zone A + B	98.34	99.03	98.52	99.34	98.45	98.5
- Zone A	76.26	81.04	81.62	80.78	81.84	82.74
- Zone B	22.09	17.99	16.9	18.56	16.61	15.76
- Zone C	1.58	0.93	1.44	0.65	1.51	1.46
- Zone D	0.08	0.03	0.03	0.02	0.04	0.04
- Zone E	0	0	0	0	0	0
50 min prediction						
RMSE	36.92	33.58	30.58	33.28	30.84	30.21
MARD	18.69	16.7	16.18	15.62	16.32	15.88
Parker error grid						
- Zone A + B	97.37	98.17	97.58	98.66	97.46	97.54
- Zone A	70.73	75.1	76.37	74	76.08	76.95
- Zone B	26.63	23.06	21.21	24.66	21.39	20.59
- Zone C	2.48	1.74	2.34	1.29	2.45	2.37
- Zone D	0.16	0.08	0.09	0.05	0.09	0.09
- Zone E	0	0	0	0	0	0
60 min prediction						
RMSE	41.53	38.58	34.63	37.68	34.88	34.33
MARD	21.18	19.37	18.59	17.54	18.72	18.32
Parker error grid						
- Zone A + B	96.28	97.1	96.46	97.96	96.38	96.5
- Zone A	65.93	69.75	71.11	69.07	70.83	71.77
- Zone B	30.35	27.36	25.35	28.89	25.54	24.73
- Zone C	3.46	2.72	3.38	1.96	3.45	3.33
- Zone D	0.26	0.16	0.16	0.08	0.17	0.17
- Zone E	0	0	0	0	0	0

4. Discussion

In this study, we developed, tested, and compared performance from different approaches to glucose prediction based on CGM. The prediction methods were a mix of simple methods to more advanced machine learning algorithms. We also proposed the usage of error weights in the training of the models to compensate for the heterogeneous clinical decision space associated with glucose predictions. Moreover, we developed the models and tested them based on a large data material with over 315,000 h of CGM data.

Numerous attempts to predict glucose value from CGM have been published since the introduction of CGM[30–38]. However, as reported in a recent review on the subject by Woldaregay et al.[14], there is a lack of proper analysis and modeling of penalty for errors in different glycemic regions. This is an important concern as modeling with the focus on general performance metrics such as RMSE is not sufficient in a clinical context. There is a clear need to consider the clinical aspects of glucose treatment based on glucose prediction. For instance, the risk profile for a prediction in the hyperglycemic region is much wider than the risk profile for predictions in the lower region of euglycemia and hypoglycemia. Some studies have reported the clinical accuracy using grid analysis. Zarkogianni et al.[39] reported the results from a Clark Error Grid analysis where 2.29% of the prediction were in zone D with a 30 min lead-time and 5.90% with a lead-time of 60 min. Others[40,41] have reported zero predictions in the C-E zone of the CEG, but these studies are based on few samples and cannot be considered representative for the common patient with type 1 diabetes. From our results, the use of a simple prediction method Extrap did perform with comparable results to more advanced methods in regard to overall performance measured in RMSE and MARD. However, we did see clinically relevant improvement of using the advanced method with error weights. Especially, the improvement in the zone C-E of the consensus grid. Avoiding these predictions are of highly importance for the patients as this could lead to delayed treatment and severe hypoglycemia. In patients with hypoglycemia unawareness and during the night this is

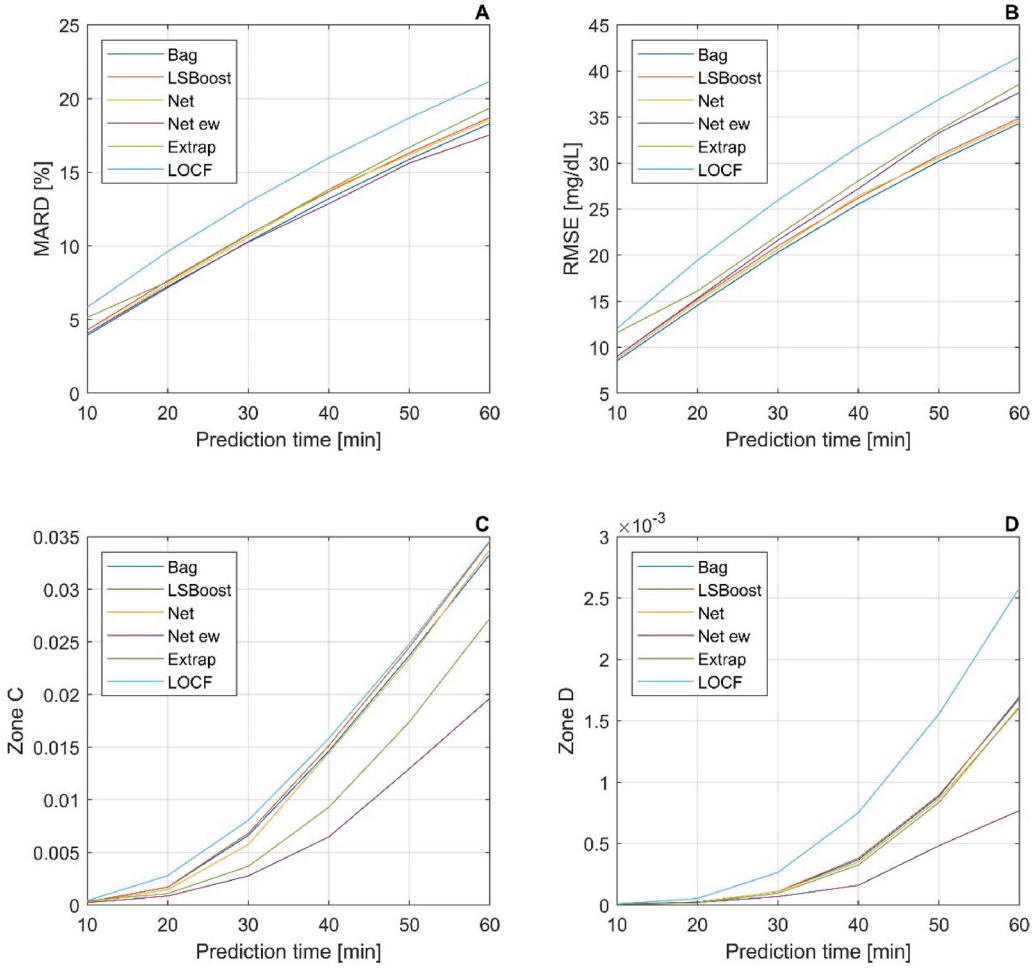

Fig. 3. Comparison of the models in relation to prediction time (10–60 min); (A) MARD %; (B) RMSE mg/dL; (C) the % of prediction in the C zone of the Parkers error grid; (D) the % of prediction in the D zone of the Parkers error grid.

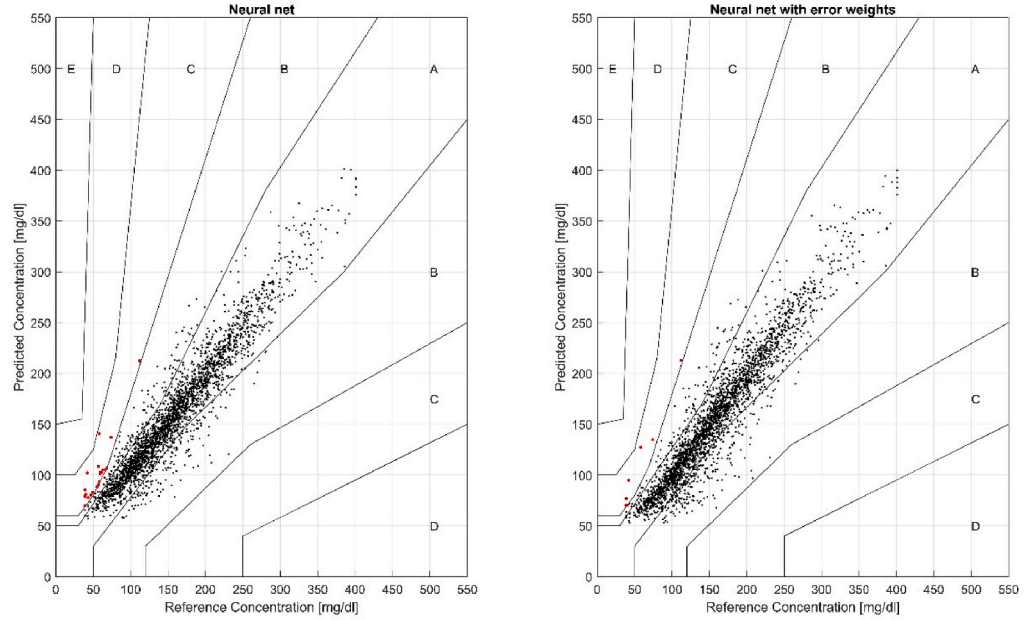

Fig. 4. Parkers error plot analysis for the neural network models with and without error weights for a 30 min prediction. The plot contains 1:400 CGM measurement from the test cohort (n = 2844). The red dots are predictions which are in the zone C.

even more problematic. Severe hypoglycemia can lead to short- and long-term complication, also hospitalization and death[42,43].

Firstly, our results show that is possible to predict glucose levels based on CGM with a reasonable accuracy and precision with a relative short prediction lead time. Our analysis was based on a large data material obtain on type 1 diabetes patients and under normal living conditions. This makes the results and performance more realistic than many of the previously reported algorithms, which are often based on in silico data or few patients[14]. Secondly, our comparison of different methods for modeling glucose predictions shows that there are some improvements on performance from using more advanced machine learning algorithms compared to a simple modeling approach. Thirdly, the proposed use of error weights could lead to better clinical performance of these models, which is an important factor for implementation in clinical practice.

Even though our study has a strong design with a large data material and clear separation between training and test, there are some limitations which are important to mention. Firstly, participants wore a CGM Dexcom G4 sensor – it is well known that CGM sensors have different accuracy and limitations in their usage. Therefore, we cannot fully generalize our results to the use with other sensors. Secondly, the cohort used in this study were included based on hemoglobin A1c (HbA1c) $<$ 8.5%, no history of severe hypoglycemia and unawareness. This means that the cohort studied were a group with low prevalence of complications related to diabetes, such as neuropathy. Glycemic control is correlated to HbA1c and the presence of neuropathy, which means that the results again cannot fully be transferred to a group of diabetes patients with higher HbA1c levels and higher prevalence of complications. In future work, validating the results in patients with different characteristics should be a priority. This could be done in longitudinal studies and using reference CGM databases (such as the SCGMS database[44]). Also, the potential of adjusting the generalized models to fit the individual patient could be interesting. However, it is a fine line between improved performance and complexity of using the models in clinical practice. Personalization of the models would typically require prior data from the patients before the prediction can be used in clinical decision making.

In conclusion, we proposed and tested six glucose prediction methods at six different lead time, spanning from 10 to 60 min. The models utilize input from a CGM system to facilitate glucose prediction. The results point toward that using error weighting in the training of the models could lead to better clinical performance. This is important if prediction algorithms should be implemented in a clinical decision system such as an artificial pancreas system.

Disclaimer

The source of the data is the T1D Exchange, but the analyses, content, and conclusions presented herein are solely the responsibility of the authors and have not been reviewed or approved by the T1D Exchange.

Author contributions

SLC had access to all of the data analyzed in this study. SLC take responsibility for the integrity and the accuracy of the study data analysis and results. SLC, TK, MHJ, and OH were involved in the study design, concept, analysis, and interpretation of data. SLC drafted the manuscript and performed the statistical analysis. TK, MHJ, and OH were involved in critical revision of the manuscript.

Funding received

None to disclose.

Declaration of competing interest

None to disclose.

Acknowledgements

N/A.

References

[1] A.D. Deshpande, M. Harris-Hayes, M. Schootman, Epidemiology of diabetes and diabetes-related complications, Phys. Ther. 88 (11) (2008) 1254–1264, https://doi.org/10.2522/ptj.20080020.
[2] Centers for Disease Control and Prevention, Preventing diabetes-related complications | diabetes | CDC, Accessed March 9, 2021, https://www.cdc.gov/diabetes/data/statistics-report/preventing-complications.html.
[3] T.S. Bailey, H.C. Zisser, S.K. Garg, Reduction in hemoglobin A1c with real-time continuous glucose monitoring: results from a 12-week observational study, Diabetes Technol. Therapeut. 9 (3) (2007) 203–210, https://doi.org/10.1089/dia.2007.0205.
[4] Juvenile Diabetes Research Foundation Continuous Glucose Monitoring Study Group, Continuous glucose monitoring and intensive treatment of type 1 diabetes, N. Engl. J. Med. 359 (14) (2008) 1464–1476, https://doi.org/10.1056/NEJMoa0805017.
[5] R.W. Beck, The effect of continuous glucose monitoring in well-controlled type 1 diabetes, Diabetes Care 32 (8) (2009) 1378–1383, https://doi.org/10.2337/dc09-0108.
[6] S.L. Cichosz, J. Frystyk, O.K. Hejlesen, L. Tarnow, J. Fleischer, A novel algorithm for prediction and detection of hypoglycemia based on continuous glucose monitoring and heart rate variability in patients with type 1 diabetes, J Diabetes Sci Technol 8 (4) (2014) 731–737, https://doi.org/10.1177/1932296814528838.
[7] S.L. Cichosz, J. Frystyk, L. Tarnow, J. Fleischer, Combining information of autonomic odulation and CGM measurements nables prediction and improves etection of spontaneous hypoglycemic vents, J Diabetes Sci Technol 9 (1) (2015) 132–137, https://doi.org/10.1177/1932296814549830.
[8] D. Dave, M. Erraguntla, M. Lawley, et al., Improved low glucose predictive alerts based on sustained hypoglycemia (preprint), JMIR Diabetes 6 (2) (2021), https://doi.org/10.2196/26909.
[9] T. Battelino, T. Danne, R.M. Bergenstal, et al., Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range, Diabetes Care 42 (8) (2019) 1593–1603, https://doi.org/10.2337/dci19-0028.
[10] S.L. Cichosz, M.H. Jensen, O. Hejlesen, Short-term prediction of future continuous glucose monitoring readings in type 1 diabetes: development and validation of a neural network regression model, Int. J. Med. Inf. 151 (2021) 104472, https://doi.org/10.1016/j.ijmedinf.2021.104472.
[11] K. Zarkogianni, E. Litsa, K. Mitsis, et al., A review of emerging technologies for the management of diabetes mellitus, IEEE Trans. Biomed. Eng. 62 (12) (2015) 2735–2749, https://doi.org/10.1109/TBME.2015.2470521.
[12] 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 (6) (2017), e2833, https://doi.org/10.1002/cnm.2833.
[13] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, I. Chouvarda, Machine learning and data mining methods in diabetes research, Comput. Struct. Biotechnol. J. 15 (2017) 104–116, https://doi.org/10.1016/j.csbj.2016.12.005.
[14] A.Z. Woldaregay, E. Årsand, S. Walderhaug, et al., Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes, Artif. Intell. Med. 98 (2019) 109–134, https://doi.org/10.1016/j.artmed.2019.07.007.
[15] J. Fleischer, S.L. Cichosz, P. Hoeyem, et al., Glycemic variability is associated with reduced cardiac autonomic modulation in women with type 2 diabetes, Diabetes Care 38 (4) (2015) 682–688, https://doi.org/10.2337/dc14-0654.
[16] J. Fleischer, E. Laugesen, S.L. Cichosz, et al., Continuous glucose monitoring adds information beyond HbA1c in well-controlled diabetes patients with early cardiovascular autonomic neuropathy, J. Diabet. Complicat. 31 (9) (2017) 1389–1393, https://doi.org/10.1016/j.jdiacomp.2017.06.013.
[17] S.L. Cichosz, M.H. Jensen, O. Hejlesen, Associations between smoking, glucose metabolism and lipid levels: a cross-sectional study, J. Diabet. Complicat. 34 (10) (2020) 107649, https://doi.org/10.1016/j.jdiacomp.2020.107649.
[18] K.G.M. Moons, D.G. Altman, J.B. Reitsma, et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration, Ann. Intern. Med. 162 (1) (2015) W1–W73, https://doi.org/10.7326/M14-0698.
[19] G. Aleppo, K.J. Ruedy, T.D. Riddlesworth, et al., REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes, Diabetes Care 40 (4) (2017) 538–545, https://doi.org/10.2337/dc16-2482.
[20] J.L. Parkes, S.L. Slatin, S. Pardo, B.H. Ginsberg, A new consensus error grid to evaluate the clinical significance of inaccuracies in the measurement of blood glucose, Diabetes Care 23 (8) (2000) 1143–1148, https://doi.org/10.2337/diacare.23.8.1143.

[21] W.L. Clarke, D. Cox, L.A. Gonder-Frederick, W. Carter, S.L. Pohl, Evaluating clinical accuracy of systems for self-monitoring of blood glucose, Diabetes Care 10 (5) (1987) 622–628, https://doi.org/10.2337/diacare.10.5.622.
[22] J.M. Lachin, Fallacies of last observation carried forward analyses, Clin. Trials 13 (2) (2016) 161–168, https://doi.org/10.1177/1740774515602688.
[23] Jerome H. Friedman, Greedy function approximation: a gradient boosting machine on JSTOR, Ann. Stat. 29 (5) (2001) 1189–1232. https://www.jstor.org/stable/2699986?casa_token=3Qj7zWiZIEEAAAAA%3AnFxtUHOmATEy6Oh7ZvZ7CmVRbK_PMp-jq2fq4WLMMPe6G5ykrKkIZCncTwSbalzIBn0kq6ZAoYS0yM5rWMMrCe4MkjZ2IxY1-eOm8D4EcVMFykhSGq0&seq=1#metadata_info_tab_contents. (Accessed 25 August 2021). Accessed.
[24] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140, https://doi.org/10.1007/bf00058655.
[25] S. Dreiseitl, L. Ohno-Machado, Logistic regression and artificial neural network classification models: a methodology review, J. Biomed. Inf. 35 (5–6) (2002) 352–359, https://doi.org/10.1016/S1532-0464(03)00034-0.
[26] J.V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, J. Clin. Epidemiol. 49 (11) (1996) 1225–1231, https://doi.org/10.1016/S0895-4356(96)00002-9.
[27] T.G. Dietterich, Ensemble methods in machine learning, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 1857 LNCS, Springer Verlag, 2000, pp. 1–15, https://doi.org/10.1007/3-540-45014-9_1.
[28] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection. http://robotics.stanford.edu/~ronnyk, 1995. (Accessed 9 March 2021). Accessed.
[29] S.L. Cichosz, M.H. Jensen, O. Hejlesen, Short-term prediction of future continuous glucose monitoring readings in type 1 diabetes: development and validation of a neural network regression model, Int. J. Med. Inf. 151 (2021) 104472, https://doi.org/10.1016/J.IJMEDINF.2021.104472.
[30] J. Reifman, S. Rajaraman, A. Gribok, W.K. Ward, Predictive monitoring for improved management of glucose levels, J Diabetes Sci Technol 1 (4) (2007) 478–486, https://doi.org/10.1177/193229680700100405.
[31] C. Zecchin, A. Facchinetti, G. Sparacino, C. Cobelli, Jump neural network for online short-time prediction of blood glucose from continuous monitoring sensors and meal information, Comput. Methods Progr. Biomed. 113 (1) (2014) 144–152, https://doi.org/10.1016/j.cmpb.2013.09.016.
[32] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, C. Cobelli, Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration, IEEE Trans. Biomed. Eng. 59 (6) (2012) 1550–1560, https://doi.org/10.1109/TBME.2012.2188893.
[33] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, et al., Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring, Diabetes Technol. Therapeut. 12 (1) (2010) 81–88, https://doi.org/10.1089/dia.2009.0076.
[34] E.I. Georga, V.C. Protopappas, D. Ardigò, D. Polyzos, D.I. Fotiadis, A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions, Diabetes Technol. Therapeut. 15 (8) (2013) 634–643, https://doi.org/10.1089/dia.2012.0285.
[35] 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 (5) (2007) 931–937, https://doi.org/10.1109/TBME.2006.889774.
[36] Chen J, Li K, Herrero P, Zhu T, Georgiou P. Dilated Recurrent Neural Network for Short-Time Prediction of Glucose Concentration.
[37] Meijner C, Persson S. reportBlood Glucose Prediction for Type 1 Diabetes Using Machine Learning Long Short-Term Memory Based Models for Blood Glucose Prediction Master's Thesis in Computer Science.
[38] W. Gu, Z. Zhou, Y. Zhou, M. He, H. Zou, L. Zhang, Poster abstract: predicting blood glucose dynamics with multi-time-series deep learning, 2017, https://doi.org/10.1145/3131672.3136965. Published online.
[39] Zarkogianni K, Mitsis K, Arredondo MT, Fico G, Fioravanti A, Nikita KS. Neuro-fuzzy based glucose prediction model for patients with Type 1 diabetes mellitus. In: 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014. IEEE Computer Society; 2014:252-255. doi:10.1109/BHI.2014.6864351.
[40] F. Allam, Z. Nossai, H. Gomma, I. Ibrahim, M. Abdelsalam, A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients, in: IFIP Advances in Information and Communication Technology, 363 AICT, Springer New York LLC, 2011, pp. 254–259, https://doi.org/10.1007/978-3-642-23957-1_29.
[41] S.M. Pappada, M.J. Borst, B.D. Cameron, et al., Development of a neural network model for predicting glucose levels in a surgical critical care setting, Patient Saf. Surg. 4 (1) (2010) 15, https://doi.org/10.1186/1754-9493-4-15.
[42] P.E. Cryer, The barrier of hypoglycemia in diabetes, Diabetes 57 (12) (2008) 3169–3176, https://doi.org/10.2337/db08-1084.
[43] C.D. Miller, L.S. Phillips, D.C. Ziemer, D.L. Gallina, C.B. Cook, I.M. El-Kebbi, Hypoglycemia in patients with type 2 diabetes mellitus, Arch. Intern. Med. 161 (13) (2001) 1653–1659, https://doi.org/10.1001/archinte.161.13.1653.
[44] Cichosz SL, Xylander AAP. A conditional generative adversarial network for synthesis of continuous glucose monitoring signals: https://doi.org/101177/19322968211014255. Published online May 30, 2021. doi:10.1177/19322968211014255.