Improved Low-Glucose Predictive Alerts Based on Sustained Hypoglycemia: Model Development and Validation Study

Darpit Dave¹, MS; Madhav Erraguntla¹, PhD; Mark Lawley¹, PhD; Daniel DeSalvo², MD; Balakrishna Haridas³, PhD; Siripoom McKay², MD; Chester Koh⁴,⁵, MD

¹Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, United States
²Department of Pediatrics, Baylor College of Medicine / Texas Children’s Hospital, Houston, TX, United States
³Department of Biomedical Engineering, Texas A&M University, College Station, TX, United States
⁴Division of Pediatric Urology, Texas Children’s Hospital, Houston, TX, United States
⁵Scott Department of Urology, Baylor College of Medicine, Houston, TX, United States

Corresponding Author:
Madhav Erraguntla, PhD
Department of Industrial and Systems Engineering
Texas A&M University
4021 Emerging Technologies Building
College Station, TX, 77843
United States
Phone: 1 979 458 2336
Email: merraguntla@tamu.edu

Abstract

Background: Predictive alerts for impending hypoglycemic events enable persons with type 1 diabetes to take preventive actions and avoid serious consequences.

Objective: This study aimed to develop a prediction model for hypoglycemic events with a low false alert rate, high sensitivity and specificity, and good generalizability to new patients and time periods.

Methods: Performance improvement by focusing on sustained hypoglycemic events, defined as glucose values less than 70 mg/dL for at least 15 minutes, was explored. Two different modeling approaches were considered: (1) a classification-based method to directly predict sustained hypoglycemic events, and (2) a regression-based prediction of glucose at multiple time points in the prediction horizon and subsequent inference of sustained hypoglycemia. To address the generalizability and robustness of the model, two different validation mechanisms were considered: (1) patient-based validation (model performance was evaluated on new patients), and (2) time-based validation (model performance was evaluated on new time periods).

Results: This study utilized data from 110 patients over 30-90 days comprising 1.6 million continuous glucose monitoring values under normal living conditions. The model accurately predicted sustained events with >97% sensitivity and specificity for both 30- and 60-minute prediction horizons. The false alert rate was kept to <25%. The results were consistent across patient- and time-based validation strategies.

Conclusions: Providing alerts focused on sustained events instead of all hypoglycemic events reduces the false alert rate and improves sensitivity and specificity. It also results in models that have better generalizability to new patients and time periods.

(JMIR Diabetes 2021;6(2):e26909) doi:10.2196/26909

KEYWORDS
machine learning; quantile regression forests; sustained hypoglycemia; continuous glucose monitoring; false alert rate; model generalizability; diabetes; glucose monitoring; hypoglycemia

Introduction

Glucose measurements are critical for effective diabetes management. Real-time continuous glucose monitoring (CGM) devices allow for frequent, automated glucose readings from interstitial fluid in the subcutaneous tissue space. CGM has been shown to improve glycemic control and reduce glycemic excursions—decreasing both hypoglycemia and hyperglycemia [1]. An important feature of CGM devices is their ability to provide real-time auditory alerts for trending glucose excursions...
above or below customized threshold levels. The CGM data can also be used to develop models to predict future hypoglycemia [1,2].

Since the first attempt at predicting future glucose values based on CGM data in 1999 [3], numerous studies have been performed to provide alerts for hypoglycemia based on predictive models. Some of the initial works relied on classical time-series based forecasting methods such as autoregressive integrated moving average (ARIMA)—ARIMAX [4-9] and state-space models [10-19]. Later, different machine learning methods were explored to predict future glucose values [20-25]. More recently, deep learning techniques have been proposed toward solving the problem [26-33]. Some works have also explored incorporating contextual data such as meal information, insulin levels, heart rate, and physical activity as features in the predictive model [34-40]. However, currently integrated information on the data sources is not readily available for real-time predictions.

Predictive hypoglycemia alerts have the potential to be extremely helpful in reducing hypoglycemia risk; however, false alerts have been a major hindrance to the acceptance of predictive hypoglycemia alerts among users [41-43]. In our earlier work, we developed a machine learning–based hypoglycemia predictive model with a sensitivity and specificity of >95%, comparable with the best predictive models in the literature [44]. Typically, the number of hypoglycemic events is very small compared with that of nonhypoglycemic events. For example, only 2.13% (35,075/1,644,875) of readings in the data set used in this study were in the hypoglycemic range (ie, <70 mg/dL). The high-class imbalance resulted in a false alert rate (FAR) of around 85% (40,502/47,683), even with an impressive specificity of 95%. Improvement in specificity in such highly imbalanced class cases will reduce the FAR and therefore improve user experience and trust in the alerts, facilitating persuasive adoption of alerts.

Previous studies have found that hypoglycemia prediction model performance is reduced when applied to new patients and different time periods [45]. Improvement in model generalizability to new patients and time periods will facilitate ease of deployment and retention of performance postdeployment.

Thus, despite the many advances made in terms of hypoglycemia prediction models, the shortcoming of a high FAR makes the alerts ill-suited for real-world application [42,46,47]. These results indicate a need for the development of approaches to reduce the FAR, maintain high sensitivity, and improve the generalizability of the prediction model.

**Methods**

**Data Description**

The CGM data sets were obtained from 110 pediatric patients with type 1 diabetes over 30 to 90 days. The data comprised over 1.6 million CGM values under normal living conditions. Dexcom G6 CGM devices were used to collect the CGM readings. The cohort-level profile of patients in this study can be found in Table 1. Summary statistics of a patient hypoglycemia profile and a patient pump profile are presented in Multimedia Appendices 1 and 2, respectively. Of note, data were obtained from a mix of patients using multiple daily injections, sensor-augmented pump therapy without automated basal rate modulation, and sensor-augmented pump therapy with a predictive low-glucose suspend feature (ie, t:slim X2 insulin pump with Basal-IQ technology; Tandem Diabetes Care, Inc [48]). The t:slim X2 insulin pump uses a simple linear regression algorithm to predict glucose levels 30 minutes ahead and suspend basal insulin in the pump if glucose values are predicted to drop below 80 mg/dL in the next 30 minutes or if a CGM value falls below 70 mg/dL. Insulin delivery can remain suspended for a minimum of 5 minutes to a maximum of 2 hours and will then resume as soon as glucose values begin to rise.

**Table 1.** Demographic and diabetes profile of patients enrolled in the study.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>12.67 (4.84)</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Glycated hemoglobin A1c (%)</td>
<td>7.70 (1.63)</td>
<td>5.00</td>
<td>12.50</td>
</tr>
<tr>
<td>Duration of diabetes (years)</td>
<td>4.93 (4.09)</td>
<td>0.25</td>
<td>19.18</td>
</tr>
<tr>
<td><strong>Continuous glucose monitoring metrics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of hypoglycemic values per day per patient</td>
<td>6.20 (5.98)</td>
<td>0.10</td>
<td>23.73</td>
</tr>
<tr>
<td>Percentage of hypoglycemic values below 70 mg/dL</td>
<td>2.13 (2.10)</td>
<td>0.50</td>
<td>12.20</td>
</tr>
</tbody>
</table>

**Hypoglycemic Events**

A glucose threshold of 70 mg/dL is used to identify the hypoglycemic range [49,50]. Of the 1.6 million CGM readings in the study data, approximately 35,000 values representing 6010 events were in the hypoglycemic range. The period of time between a first CGM reading below the threshold and the point where the CGM value rises to ≥70 mg/dL is considered a single “hypoglycemic event.”

Table 2 shows the frequency distribution of the duration of the hypoglycemic events. The hypoglycemic events in this analysis were further classified as transient (lasting less than 15 minutes; 77.24% [4642/6010] of the hypoglycemic events) or sustained (lasting 15 minutes or longer; 22.76% [1368/6010] of the events) based on the recommendations in previous studies [51,52]. Nonhypoglycemic events are CGM observations of ≥70 mg/dL. Detailed information regarding the distribution of hypoglycemic
events and a breakdown of sustained events by day and night are presented in Multimedia Appendices 3-5.

Table 2. Frequency distribution of the duration of hypoglycemic events (n=6010).

<table>
<thead>
<tr>
<th>Hypoglycemic events</th>
<th>Duration of consecutive CGM values falling below 70 mg/dL (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>572 (9.51)</td>
</tr>
<tr>
<td>10</td>
<td>796 (13.24)</td>
</tr>
<tr>
<td>15</td>
<td>885 (14.73)</td>
</tr>
<tr>
<td>20</td>
<td>842 (14.01)</td>
</tr>
<tr>
<td>25</td>
<td>676 (11.25)</td>
</tr>
<tr>
<td>30</td>
<td>562 (9.35)</td>
</tr>
<tr>
<td>35</td>
<td>391 (6.51)</td>
</tr>
<tr>
<td>40</td>
<td>283 (4.71)</td>
</tr>
<tr>
<td>45</td>
<td>209 (3.48)</td>
</tr>
<tr>
<td>&gt;45</td>
<td>794 (13.21)</td>
</tr>
</tbody>
</table>

Evaluation Metrics
We define the following metrics for evaluating model performance: sensitivity, specificity, and FAR.

Sensitivity measures the proportion of true positives that are correctly identified. It is also known as the true-positive rate.

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

where TP are the true positives and FN are the false negatives.

Specificity measures the proportion of true negatives that are correctly identified. It is also known as the true-negative rate.

$$\text{Specificity} = \frac{TN}{TN + FP}$$

where TN are the true negatives and FP are the false positives.

FAR was defined based on the definition provided by Mosquera-Lopez et al [53] and measures the proportion of alerts that are not truly indicative of predicted hypoglycemic events.

$$\text{FAR} = \frac{FP}{TP + FP}$$

Machine Learning Methodologies

Random Forest
Random forest (RF) is a nonparametric approach that builds on an ensemble prediction of a “forest” of regression trees grown via bootstrap sampling. Model predictions are obtained from the mean of the predictions of the individual trees. RF performs well when dealing with nonlinear relationships among variables and makes no assumptions about data distributions. Owing to these characteristics, utilizing RF-based machine learning modeling resulted in good performance in our previous work [44] compared with other machine learning methods for hypoglycemia prediction. In this study, an RF-based model was used to classify events as sustained hypoglycemic events (positive class) or transient and nonhypoglycemic events (negative class).

Quantile Regression Forest
For the multistep prediction approach, future CGM values were predicted using quantile regression forests (QRFs). The concept of quantile regression was introduced by Koenker and Hallock [54] and is advantageous when quantile functions are of interest. Quantile functions provide information about the spread of the response variable beyond the conditional mean by estimating the full conditional distribution. This is particularly useful for predicting values other than the mean (eg, median or 90th quantile). QRFs are a generalization of RFs and provides an accurate way of estimating the conditional quantiles [55]. Since it is more important to accurately predict CGM values near the hypoglycemic range rather than within the euglycemic or hyperglycemic range, QRFs were used to predict future CGM values using the regression approach.

QRFs were used as a multistep forecasting method to predict the glucose values for every 5-minute interval in the prediction horizon (PH). This resulted in 6 predictions for the 30-minute PH and 12 predictions for the 60-minute PH. Based on these predictions, a sustained hypoglycemic event was detected if 3 or more consecutive predicted CGM values were <70 mg/dL.

Validation Mechanism
An appropriate validation mechanism is critical to assess the performance of a machine learning model [56,57]. This validation strategy helps to ascertain the generalizability of the model and ensures that the model performs well in real-world scenarios. This can be performed by sampling a subset of the data for model development and sampling a different sample of data for model validation [58]. Two validation strategies—patient-based and time-based—were used to evaluate model performance.

Patient-Based Approach
In this approach, the prediction model was developed on a subset of patients and validated on a different set of patients. Of the 110 patients, 70 patients (approximately 65% of the data) were randomly selected for training and the remaining 40 patients were used for performance evaluation. The final model performance reported is the mean of 5 replications of this procedure of 65%/35% split of training and validation data.

Time-Based Approach
In this approach, for each of the 110 patients, the first 70% of data was used for model development and sampling a different sample of data for model validation [58]. Two validation strategies—patient-based and time-based—were used to evaluate model performance.

Features Extracted for the Prediction
A rich combination of demographic, dynamic, snowball, interaction, and contextual features were extracted from the data. An optimal set of features for hypoglycemia prediction was identified in our previous work [44] and these features were used for the model development in this study (Multimedia Appendix 6). Records associated with missing data were eliminated from the analysis; that is, if a feature was dependent on a missing CGM value, that record—as well as all dependent time-lagged records—were eliminated from the analysis.
**Results**

**Model Performance**

Table 3 summarizes the performance of the model based on patient-based and time-based validation strategies. The total number of false alerts when considering transient and nonhypoglycemic events as false alerts and when considering nonhypoglycemic events as false alerts are provided along with sensitivity and specificity metrics.

In the patient-based validation approach, for both 30-minute and 60-minute PHs, the QRF method provided a significant advantage over the RF method with high sensitivity, high specificity, and low FAR. The patient-based validation approach indicated that the sustained hypoglycemic model developed using QRFs is generic and can be applied to new patients without performance degradation.

In a time-based validation setting, the RF method performed well for both 30-minute and 60-minute predictions with high sensitivity, high specificity, and low FAR, but the QRF method still outperformed it. The time-based validation methodology indicated that both models retain performance when applied to new time periods and in postdeployment.

**Table 3.** Comparison of model performance based on sensitivity, specificity, and false alerts with patient-based and time-based validation for 30-minute and 60-minute prediction horizons (PHs).

<table>
<thead>
<tr>
<th>Metrics</th>
<th>30-minute PH</th>
<th>60-minute PH</th>
<th>Time-based validation</th>
<th>30-minute PH</th>
<th>60-minute PH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient-based validation</td>
<td>Time-based validation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Method 1: RF&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Method 2: QRF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Method 1: RF</td>
<td>Method 2: QRF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Method 1: RF</td>
</tr>
<tr>
<td>Sensitivity, % (SD)</td>
<td>39.11 (2.25)</td>
<td>99.09 (0.16)</td>
<td>49.27 (3.03)</td>
<td>97.61 (0.41)</td>
<td>96.17</td>
</tr>
<tr>
<td>Specificity, % (SD)</td>
<td>98.96 (0.09)</td>
<td>98.19 (0.10)</td>
<td>98.63 (0.12)</td>
<td>98.09 (0.11)</td>
<td>98.3</td>
</tr>
<tr>
<td>False alerts, n (SD)</td>
<td>6936 (356)</td>
<td>9339 (459)</td>
<td>7043 (317)</td>
<td>9672 (431)</td>
<td>6476</td>
</tr>
<tr>
<td>Considering transient and</td>
<td>3907 (200)</td>
<td>5368 (162)</td>
<td>4109 (156)</td>
<td>5677 (201)</td>
<td>3324</td>
</tr>
<tr>
<td>nonhypoglycemic events as false</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False alert rate, % (SD)</td>
<td>26.32 (2.56)</td>
<td>26.50 (2.41)</td>
<td>26.44 (2.37)</td>
<td>26.36 (2.57)</td>
<td>22.79</td>
</tr>
</tbody>
</table>

<sup>a</sup>RF: random forest.

<sup>b</sup>QRF: quantile regression forest.

**Comparison of Sustained Versus All Hypoglycemic Events Prediction Models**

Table 4 provides a comparison of model performance between sustained hypoglycemia and all-hypoglycemia prediction models. Even though the all-hypoglycemia model had high sensitivity, a specificity of 93% on a large number of nonhypoglycemic events resulted in an FAR of 85%. Focusing the alerts on sustained hypoglycemic events resulted in an increase in specificity to 98% and reduced the FAR to approximately 20% to 30%. Also, the performance of the all hypoglycemic events prediction model was adversely affected when evaluated on new patients and new time periods (drop in sensitivity of 5%). On the other hand, prediction models based on sustained hypoglycemic events retained their performance for new patients and new time periods, indicating better generalizability of the sustained hypoglycemic events model.
Table 4. Comparison of model performance based on sensitivity, specificity, and false alert rate with different characterizations of hypoglycemic events and different validation strategies (patient-based and time-based) for giving predictive alerts.

<table>
<thead>
<tr>
<th>Model</th>
<th>30-minute prediction horizon</th>
<th>60-minute prediction horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
</tr>
<tr>
<td>All hypoglycemic events prediction (5-fold validation)</td>
<td>93.61</td>
<td>93.50</td>
</tr>
<tr>
<td>All hypoglycemic events prediction (new time periods)</td>
<td>87.10</td>
<td>92.66</td>
</tr>
<tr>
<td>All hypoglycemic events prediction (new patients)</td>
<td>87.60</td>
<td>92.47</td>
</tr>
<tr>
<td>Sustained hypoglycemic events prediction (QRF—a—new patients)</td>
<td>99.08</td>
<td>97.79</td>
</tr>
<tr>
<td>Sustained hypoglycemic events prediction (QRF—a—new time periods)</td>
<td>98.54</td>
<td>98.57</td>
</tr>
</tbody>
</table>

QRF: quantile regression forest.

A graphical comparison between the classifiers at different threshold values using receiver operating characteristic (ROC) curves can be found in Multimedia Appendix 7. The ROC plots show that the QRF model outperformed the RF models over the entire range of sensitivity and specificity levels. Table 5 shows the QRF model's performance metrics at different threshold levels. The table also presents the average time required to predict a hypoglycemic event at different threshold levels.

Table 5. Performance of the quantile regression forest model at different thresholds and the average time to predict a hypoglycemic event.

<table>
<thead>
<tr>
<th>Metric</th>
<th>30-minute prediction horizon</th>
<th>60-minute prediction horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Threshold 1</td>
<td>Threshold 2</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>98.54</td>
<td>99.27</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>98.57</td>
<td>97.56</td>
</tr>
<tr>
<td>False alerts (n)</td>
<td>6932</td>
<td>11,960</td>
</tr>
<tr>
<td>False alerts with transient events as positives (n)</td>
<td>3736</td>
<td>8149</td>
</tr>
<tr>
<td>False alert rate (%)</td>
<td>22.36</td>
<td>35.36</td>
</tr>
<tr>
<td>Average time to predict an event (minutes)</td>
<td>18.78</td>
<td>22.95</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

We present a robust prediction model for providing high-quality alerts for sustained hypoglycemic risk in patients with type 1 diabetes. The final model (QRF model) was demonstrated to be robust to different validation approaches that best represent real-world application scenarios (new patients and new time periods). The primary research contributions of this work are (1) the development of a prediction model that focused on sustained hypoglycemic events and resulted in high sensitivity, high specificity, and a low FAR; and (2) improved generalizability of the model to new patients and new time periods. The model makes use of only CGM data in the past 4 hours and contextual information about the current hour of the day and day of the week to make predictions. A methodology contribution is the use of glucose predictions at multiple time points to facilitate inference of sustained hypoglycemia. The model was built using data collected from 110 patients over a range of 30 to 90 days under normal living conditions, ensuring validity of the results. The QRF model proposed in this work had sensitivity and specificity >97% for both 30- and 60-minute PHs. The FAR was also kept low at 22% and 29% for 30-minute and 60-minute PHs, respectively, which will lead to improved user trust in and adoption of CGM-based alerts.

Comparison with the Literature

A comparative analysis of different hypoglycemia prediction methodologies can be found in the literature [11,20,59,60]. A straightforward comparison between different hypoglycemia prediction studies is complicated due to differences in CGM sensors used, sampling intervals, and data collection (synthetically generated, controlled study, free-living conditions). In addition, different studies have used different definitions of hypoglycemia, which makes their findings difficult to compare [21,61-63]. When using a regression approach, the majority of the works present an overall root-mean-square error (RMSE) value, but accuracy pertaining to the observations in the hypoglycemic range might be more relevant. Thus, providing an overall RMSE for the entire data set could misrepresent the model's performance. On the classification side, sensitivity and specificity provide accurate information on the TPs and FPs, respectively. However, due to high-class imbalance, even a
moderately high specificity can lead to a high FAR. It becomes important to consider the FAR, in addition to sensitivity and specificity, in such class-imbalanced applications.

In machine learning, a standard approach to validate prediction models is to split the data into a training set (to train the model) and a validation set (to evaluate model performance) [64]. This random partitioning of the data into training and validation subsets and repeating the process across multiple folds is called cross-validation. Studies across the literature have used different validation strategies such as random sampling [20,65-67], time-based splitting [4,5,67-69], patient-specific splitting [6,32,53,70,71], or a combination of these methods to estimate predictive model performance. Simple random sampling–based cross-validation [72,73] may not fully address the generalizability aspect of the model to new patients and new time periods. Some studies [6] using a patient-based validation strategy used a part of their test data for tuning model parameters, which affected the validity of the performance estimation. The model presented in this paper had high performance in both patient- and time-based validation methods.

Mosquera-Lopez et al [28] used a patient-specific validation approach in which patient data in the test set were exclusive from the training data. Performance was reported on metrics such as sensitivity, RMSE, and FARs. However, leveraging some of the preprocessing and postprediction error-correction steps to improve performance made it difficult to achieve similar results in a real-world setting. Also, the test performance was evaluated on a small sample size of 10 patients (in a 4-week study). This might affect the generalizability of the presented results.

Dave et al [44] recently showed good results with respect to sensitivity and specificity using a random sampling–based validation approach and a threshold of 70 mg/dL for hypoglycemia. However, it was observed that performance of this model was reduced when applied to new patients and new time periods (Table 3). In addition, even with sensitivity and specificity of >95%, the model resulted in an FAR of 80% due to a large number of nonhypoglycemic events relative to the number of hypoglycemic events. From a user experience perspective, this will lead to false alert fatigue. The model presented in this paper reduced the FAR to 22%.

Having an accurate and actionable hypoglycemia prediction model with low FARs is essential to the durability of CGM in diabetes management. Furthermore, a patient-facing hypoglycemia prediction algorithm may give patients the confidence to aim for in-range glucose values without fear of hypoglycemia, potentially leading to lower glycated hemoglobin A1c (HbA1c) values and increased time in range. Of note, 22.3% of patients analyzed were using sensor-augmented pump therapy with a predictive low-glucose suspend feature (ie, Basal-IQ technology). Patients using this system are still at risk for hypoglycemia because of insulin on board, exercise, overdosing on carbohydrates, and/or hyperglycemia, so a notification for predicted hypoglycemia using advanced machine learning models with good performance could still be clinically useful.

Limitations

A limitation of our approach is that transient hypoglycemic events were ignored in generating alerts. Ignoring the transient events helped the machine learning model better learn the more stable patterns of sustained events. Even though the alerts were focused on detecting sustained events, 61% of the transient events were still classified as FARs. This resulted in just 39% of the transient events (representing 13% of the total hypoglycemic events) not being detected. This trade-off was justified because transient events are not as serious as sustained hypoglycemic events. Transient events may occur because of random variations in glycemic levels (ie, noise) or temporal lag in the effect of an intervention taken by the patient (eg, consuming fast-acting carbohydrates). In either case, ignoring transient events will help in learning the stable patterns of sustained hypoglycemia. The improved FAR, sensitivity, specificity, and generalizability of the sustained hypoglycemia model presented in this paper justify this trade-off.

This study was based on patients with pediatric type 1 diabetes in the age range of 0 to 20 years using Dexcom G6 CGM devices. As such, the results are directly applicable to this population. The model may need to be recalibrated to other CGM devices such as the Guardian (Medtronic) or FreeStyle Libre (Abbott Laboratories Co.); however, the performance measures should be generalizable to other platforms provided the accuracy and frequency of incoming glucose readings remain the same. Similarly, while no specific activity profile of pediatric patients was explicitly used in the model development, the model may need to be calibrated to an adult cohort by retraining on adult CGM data [74-76]. Pediatric patients were selected as the focus of the study because of our collaboration in the United States Food and Drug Administration (FDA)–funded Southwest National Pediatric Device Innovation Consortium [77]. Additionally, there is a need for a paradigm shift in diabetes management in pediatrics to avoid risk of hypoglycemia to ameliorate parental and patient fear and move toward optimizing time in range and lowering HbA1c.

Conclusions

Providing predictive alerts for hypoglycemia focused on sustained events instead of all hypoglycemic events reduces FARs and improves sensitivity and specificity. It also results in models that have better generalizability to new patients and time periods. This has important implications for sustaining CGM use and optimizing glycemic control with fewer hypoglycemic events, improved confidence, and potentially lower HbA1c. To that end, the predictive model presented in this paper will be implemented in a smartphone app in an upcoming clinical pilot study at Texas Children’s Hospital.
Acknowledgments

This study was supported by FDA P50 Pediatric Device Consortia Grant #5P50FD006428 (SWPDC) (Dr Koh—Contact Principal Investigator). The authors would like to thank the contributions of Achu Byju, Department of Biomedical Engineering, Texas A&M University, for designing the overall architecture for the implementation of the data collection system.

This study involves the use of secondary analysis of deidentified data that were not collected specifically for this project and is not human subject research (Texas A&M IRB number 2019-0710).

Authors' Contributions

DDave, ME, and ML conceived the idea for the study. DDave implemented the model and analyzed the data. DDave and ME wrote the manuscript. All authors provided input and helped in revising the manuscript.

Conflicts of Interest

DDeSalvo serves as an independent consultant for Dexcom.

Multimedia Appendix 1
Patient hypoglycemia profile.
[PDF File (Adobe PDF File), 57 KB - diabetes_v6i2e26909_app1.pdf ]

Multimedia Appendix 2
Patient pump profile.
[PDF File (Adobe PDF File), 68 KB - diabetes_v6i2e26909_app2.pdf ]

Multimedia Appendix 3
Breakdown of sustained events by daytime and nighttime.
[PDF File (Adobe PDF File), 858 KB - diabetes_v6i2e26909_app3.pdf ]

Multimedia Appendix 4
Breakdown of sustained and transient events.
[PDF File (Adobe PDF File), 46 KB - diabetes_v6i2e26909_app4.pdf ]

Multimedia Appendix 5
Breakdown of sustained events.
[PDF File (Adobe PDF File), 48 KB - diabetes_v6i2e26909_app5.pdf ]

Multimedia Appendix 6
Features extracted for prediction.
[PDF File (Adobe PDF File), 59 KB - diabetes_v6i2e26909_app6.pdf ]

Multimedia Appendix 7
Receiver operating characteristic plot showing a comparison between different classifiers for giving out predictive alerts for 30-minute (top) and 60-minute (bottom) prediction horizons.
[PDF File (Adobe PDF File), 782 KB - diabetes_v6i2e26909_app7.pdf ]

References


47. Wiklund M. Medical device alarm systems can serve as trusted advisors. In: Advancing Safety in Medical Technology, AAMI Medical Device Alarm Summit; Henndon, VA. 2011 Presented at: AAMI Medical Device Alarm Summit; October 4-5, 2011; Henndon, VA.


Abbreviations

ARIMA: autoregressive integrated moving average
CGM: continuous glucose monitoring
FAR: false alert rate
FDA: Food and Drug Administration
FN: false negatives
FP: false positives
HbA1c: glycated hemoglobin A1c
PH: prediction horizon
QRF: quantile regression forest
RF: random forest
RMSE: root-mean-square error
ROC: receiver operating characteristic
TN: true negatives
TP: true positives

Edited by D Griaucze, K Mizokami-Stout; submitted 03.01.21; peer-reviewed by P Prahalad, YK Lin; comments to author 25.02.21; revised version received 09.03.21; accepted 17.03.21; published 29.04.21.

Please cite as:
Dave D, Erraguntla M, Lawley M, DeSalvo D, Haridas B, McKay S, Koh C
Improved Low-Glucose Predictive Alerts Based on Sustained Hypoglycemia: Model Development and Validation Study
JMIR Diabetes 2021;6(2):e26909
URL: https://diabetes.jmir.org/2021/2/e26909
doi:10.2196/26909
PMID:33913816

©Darpit Dave, Madhav Erraguntla, Mark Lawley, Daniel DeSalvo, Balakrishna Haridas, Siripoom McKay, Chester Koh. Originally published in JMIR Diabetes (https://diabetes.jmir.org), 29.04.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Diabetes, is properly cited. The complete bibliographic information, a link to the original publication on https://diabetes.jmir.org/, as well as this copyright and license information must be included.