Predicting Nocturnal Hypoglycemia from Continuous Glucose Monitoring Data with Extended Prediction Horizon

Long Vu, PhD¹, Sarah Kefayati, PhD², Tsuyoshi Idé, PhD¹, Venkata Pavuluri, MS¹, Gretchen Jackson, MD, PhD², Lisa Latts, MD, MSPH², Yuxiang Zhong, PhD³, Pratik Agrawal, MS³, Yuan-chi Chang, PhD¹
¹IBM Research AI, Yorktown Heights, NY, USA, ²IBM Watson Health, Cambridge, MA, USA, ³Medtronic, Northridge, CA, USA

Abstract

Nocturnal hypoglycemia is a serious complication of insulin-treated diabetes, which commonly goes undetected. Continuous glucose monitoring (CGM) devices have enabled prediction of impending nocturnal hypoglycemia, however, prior efforts have been limited to a short prediction horizon (~30 minutes). To this end, a nocturnal hypoglycemia prediction model with a 6-hour horizon (midnight-6 am) was developed using a random forest machine-learning model based on data from 10,000 users with more than 1 million nights of CGM data. The model demonstrated an overall nighttime hypoglycemia prediction performance of ROC AUC = 0.84, with AUC = 0.90 for early night (midnight-3 am) and AUC = 0.75 for late night (prediction at midnight, looking at 3-6 am window). While instabilities and the absence of late-night blood glucose patterns introduce predictability challenges, this 6-hour horizon model demonstrates good performance in predicting nocturnal hypoglycemia. Additional study and specific patient-specific features will provide refinements that further ensure safe overnight management of glycemia.

Introduction

Hypoglycemia – typically defined as a blood glucose (BG) below <70 mg/dL – is one of the acute complications of diabetes in patients treated with insulin (1). Over half of hypoglycemic episodes occur during nocturnal sleep with a higher prevalence for insulin-dependent diabetes patients that include patients with type 1 diabetes (i.e., insulin dependent) more so than type 2 diabetes (2, 3). Although the definition varies, nocturnal hypoglycemia is usually defined by a predetermined time window for which the international consensus recommendation is from midnight to 6 am, a period that usually includes the duration of nighttime sleep and the longest interprandial interval (1, 4).

Nocturnal hypoglycemia is often asymptomatic due to a sleep-induced effect that shifts the counterregulatory activation to hypoglycemia to a lower threshold(4, 5). Unless severe and followed by a seizure, coma, or other noticeable impairment at the awakening time, it can go unrecognized and be prolonged in patients with type 1 diabetes, and in worst cases, can lead to ‘dead-in-bed’ syndrome (5). Detection of nocturnal hypoglycemia is critically important; if recurrent and undetected, even mild asymptomatic episodes can lead to further impairment of defenses against future episodes due to hypoglycemia-associated autonomic failure and defective counterregulatory responses to subsequent events (4, 6). In addition to the increased risk for future episodes, the daytime effects after an episode of nocturnal hypoglycemia, such as fatigue, impaired mood and higher calorie intake and weight gain, considerably lower quality of life (5).

Several risk factors have been associated with nocturnal hypoglycemia including demographic and non-demographic factors: age, race, BMI, comorbidities, stress, and diabetic nephropathy; physical exercise – with greater risk associated with evening exercise; tight glycemic control; excessive insulin concentration during the day; previous episodes of nocturnal hypoglycemia; and low bedtime glucose levels (2, 7, 8).

With the recent availability of continuous glucose monitoring (CGM) devices, better diabetes management is possible that includes detection of asymptomatic episodes of nocturnal hypoglycemia that go unrecognized with self-monitoring blood glucose (SMBG) devices, which are not commonly used during sleep hours. In fact, the 2017 American Diabetes Association (ADA) consensus statement on measures beyond glycated hemoglobin emphasizes the role of CGM in diabetes management, especially for nighttime and asymptomatic hypoglycemic events, which SMBG devices typically fail to detect, and provides standardized recommendations for analyzing CGM data (1, 9).

The richness of CGM data has enabled extraction of many useful insights, such as retrospective analyses revealing trending patterns, that can increase the daily decision-making power of diabetes patients. Furthermore, CGM data
combined with artificial intelligence (AI) techniques can provide prediction of impending adverse events, such as hyperglycemia or hypoglycemia, within a time horizon leading to the event. Contreras and Vehl provide a review of 10,1849 articles from 2010-2018 applying AI techniques focused on predicting BG levels. The AI techniques were categorized into learning models that included neural networks and decision trees, information discovery models, such as clustering techniques, and models to extract reasoning from information, such as Bayesian methods. The review also reports a smaller group of studies on detection of adverse effects of diabetes, such as hypoglycemia or hyperglycemia. The majority of the reported studies in this category, however, address real-time detection of adverse effects rather than predicting future events within a prediction horizon.

Very few researchers have developed models to predict nocturnal hypoglycemia using CGM data (11-13). One model, for instance, was based on a Kalman filter method with earlier version of the work including voting algorithm approach (11). Another group of researchers used a support vector regression model originally developed for diurnal hypoglycemia detections with added features of physical exercise and sleep time to predict nocturnal hypoglycemic events (12). These studies were limited by small cohorts, dependency on additional sensors (e.g., motion, sleep), and unrealistically optimal patients’ data entry in controlled environments. In addition, a major limitation of these methods is the short prediction horizon, typically chosen at 30 minutes prior to impending hypoglycemic event. With the short prediction horizon, such models are usually used in the context of CGM-enabled insulin pump devices, which have an automatic corrective mechanism to control and suspend the basal insulin delivery. However, the long prediction horizon, such as the work proposed herein, is applicable to broader CGM users including the non-pump users.

Another important factor that has not been examined in the prediction of nocturnal hypoglycemia is the marked distinction between the early (normally defined by midnight-3 am) and late portion of nighttime sleep (3-6 am). The early stage of sleep is considered in the slow wave sleep (SWS) while late sleep is dominated by rapid eye movement (REM) portion of the sleep cycle. These phases of the sleep cycle dictate different glucose needs from the brain, which are markedly reduced in early sleep compared to late night sleep or awake time (14). It is reported that 60% to 70% of nocturnal hypoglycemic episodes occur during late sleep, between 3 am and 7 am (14). Compared to early nighttime hypoglycemia, the counterregulatory hormone response is significantly weaker during late night hypoglycemia (5). Moreover, due to distinct physiological differences, some of the treatment recommendations, such as bedtime snack, were found to be effective for only early nighttime hypoglycemia, but not for late night events (5).

The most ideal approach for minimizing nocturnal hypoglycemia is, therefore, to enable patients to manage their impending hypoglycemia well before going to sleep. This means extending the prediction horizon for the duration of sleep. This manuscript describes an effort to develop a prediction model from CGM data with an extended prediction horizon. The early performance results along with the challenges of early and late-night hypoglycemia prediction are reported.

**Methods**

Based on the ADA guideline of international consensus on using CGM device (1), nocturnal hypoglycemia was defined by the glucose level falling below 70 mg/dL and lasting at least 15 minutes. An overall nighttime hypoglycemic event was defined by the episode happening from midnight-6 am; early-night hypoglycemia was defined from midnight-3 am, and the late-night events from 3-6 am.

Over 1 million nights/days of CGM data were voluntarily uploaded by 9,800 individuals with type 1 diabetes using the MiniMed™ 530G system (Medtronic MiniMed Inc., Northridge, CA) Data from a 6 pm to 6 am window were analyzed. MiniMed™ 530G system users had the option of turning on the system’s threshold suspend feature and setting it to a value between 60-90 mg/dL in consultation with their physician. Upon suspension of insulin delivery, an alarm alerts the user of the pump suspension and, if not disabled by the user, the suspension can last up to 2 hours before insulin delivery resumes. All of the users in the present study had the threshold suspend feature activated with different predefined low threshold, which for the majority was set to 60-70 mg/dL.

The prediction was based on a random forest machine learning model composed of 100 trees and built using python scikit-learn library. Each decision tree spits out a prediction score and the final class is defined by the class with the highest ensemble-average probability score. The random forest classifier was selected for its robustness, accuracy, scalability, and ease of maintenance. For the testing purposes, prior to model finalization, a smaller dataset comprised of approximately 1000 users each with an average of 300 nights of uploaded data (total of 300K nights) was also considered. The positive label for the nights in the dataset was assigned if at least one nocturnal hypoglycemic event was detected per night. The total night samples were first sorted from the furthest night to the most recent night in increasing order by the timestamp of the users streaming the data. The first 80% of the dataset was then considered
for training (10-fold cross validation) and the latest 20% of the data was reserved for the holdout test. The model performance was assessed based on the night data in the test set using the area under the curve (AUC) of the receiver operating characteristics (ROC) curve.

The feature vector was composed of three categories of features: SG features, count-based features, and temporal location features. SG features were extracted as statistical properties of the previous sensor glucose time series. Count-based features were simply a count of the number of hypoglycemic events in the past windows. Temporal location features captured the time of day and day of week, which are important because hypoglycemic events have a tendency to recur on a daily or weekly basis. For the same model, three types of performance for nocturnal hypoglycemia prediction were assessed with similar feature vectors and prediction timepoint at midnight, which included 1) overall nighttime with 6-hour prediction horizon (midnight-6 am); 2) early-night with 3-hour prediction horizon (3-6 am); and 3) late-night with 6-hour prediction horizon (midnight-6 am) only looking to predict if the hypoglycemic event happened 3-6 am.

Results

Study Population

The 9,800 users had a mean±SD age of 45.34±16.38 years (51% female and 49% male) and had been using the MiniMed™ 530G device on average for about 100 days and had been on insulin at least 1 year and maximum of 71 years. The CGM device enabled users to log their carbohydrate consumption as well as administer insulin boluses, although these auxiliary inputs were at the users’ discretion and not consistent throughout their uploads. The prevalence of nocturnal hypoglycemia among the users was 18.7% for the overall nighttime period (12-6 am), where 12.7% of the users exhibited early-night and 11.8% exhibited late-night hypoglycemia.

Overall Nighttime Prediction

The overall nighttime prediction covered the prediction horizon of 6 hours, from the midnight to 6 am, for each user. This prediction duration was for any hypoglycemic event that occurred regardless of its occurrence during early or late night. The overall performance of the model based on the total number of nights (collected from different users) in the test dataset was 0.84 as depicted in Figure 1 by the full AUC metric. The scatter plots show the AUC-based model performance for each user in the test dataset as a function of hypoglycemia prevalence in the totaled nights (about 300 nights) for each user. When the number of users was increased 10 fold, the performance of the model remained unchanged (Figure 1b). The performance of the model deviated less from the mean as the prevalence of hypoglycemia increased indicating more stable hypoglycemia detection for users with higher recurrence.

![Figure 1](image)

*Figure 1: Model performance for the overall nighttime hypoglycemia prediction with 6-hour prediction horizon for approximately 1000 users (a) and 9800 users (b). Scatter points represent AUC values for each user in the test dataset. Full AUC represents the value calculated over all the data points included in the test set.*

Early versus Late-Night Prediction

The counterregulatory hormone response to low BG during the early and late portion of the night makes differs for these two stages of sleep with respect to predictability of hypoglycemic occurrence. For the early-night hypoglycemia
prediction with prediction horizon of 3 hours from midnight, the model yields 0.90 AUC-based performance as calculated by the full AUC metric (Figure 2a). A separate prediction for the late night for the same user population, however, reveals a lower performance of 0.75 by the model (Figure 2b). For both predictions and for the users with greater hypoglycemia prevalence (referred to as skew in the plots), the user-based performance of the model approached the mean value.

![Image](image.png)

**Figure 2**: Model performance for early-night model with 3-hour prediction horizon (a) and late-night model with 6-hour prediction horizon (b) for 9800 users. Scatter points represent AUC values for each user in the test dataset. The full AUC represents the value calculated over all the data points included in the test-set.

**Late-Night Patterns**

To understand the more specific patterns associated with the late-night hypoglycemia, the smaller dataset of about 1000 users (approximately 300 thousand nights of data) was assessed more closely. The variable-wise probability density functions of two samples were compared for the same components of the feature vector as described in the Methods section: the sample size of 11% (of all the nights) with history of late-night hypoglycemia (positive labels) and the sample size of 77% with no late-night history of hypoglycemia (negative labels).

The comparison revealed some unique characteristics for user with late-night hypoglycemia which include: lower SG values at midnight, average downward SG trend before midnight, higher recent hypoglycemic event count, and, in general, a non-zero count of late-night hypoglycemia over the past couple of months.

**Trends in False Negatives Predictions of Late-night Hypoglycemia**

Comparing the early-night and late-night predictions, the rate of false positives was similar, but the false negatives were significantly higher in the late-night predictions. For the late-night false negative predictions, the users could be categorized into four distinct classes as described in Table 1. The illustration of an example for each class is shown in Figure 3 for class 1 (Figure 3a), class 2 (Figure 3b), class 3 (Figure 3c), and class 4 (Figure 3d).

**Table 1**: User classes for the false-negative predictions of late-night hypoglycemia. A user may belong to more than one class. The prevalence values are estimated from approximately 1000 users.

<table>
<thead>
<tr>
<th>False-Negative Classes</th>
<th>Description</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class #1</td>
<td>SG measurement above 180 mg/dL at midnight, sliding below 70 mg/dL between 3 am and 6 am</td>
<td>45%</td>
</tr>
<tr>
<td>Class #2</td>
<td>SG trajectory largely outside of 1-σ band from recent nights</td>
<td>56%</td>
</tr>
<tr>
<td>Class #3</td>
<td>SG trajectory exhibited U or W shaped bounce in the middle of night</td>
<td>33%</td>
</tr>
<tr>
<td>Class #4</td>
<td>SG trajectory not fitting above categories</td>
<td>15%</td>
</tr>
</tbody>
</table>
Discussion

This manuscript describes one of the first attempts to use CGM data to predict nocturnal hypoglycemia with an extended time horizon on a large cohort (~10,000 individuals) with more than 1 million nights of sample data from a wide demographic population (Insert some of the demographic attributes). The main goal of this study was to assess the feasibility of predicting nocturnal hypoglycemia solely based on CGM sensor glucose readings. The presented model can predict hypoglycemia events throughout the nocturnal sleep interval with the prediction horizon of 6 hours at the midnight timepoint with an AUC of 0.84. One previously reported predictive model using a prediction horizon greater than one hour showed that performance decreased with horizon extension, but was optimized when set on 30 minutes (15). To our knowledge, this is the first study using CGM data with a 6-hour extended prediction horizon and high performance, which might enable patients to better address potential hypoglycemia prior to going to sleep.

The main limitation of models with short prediction horizons that are typically around 30 minutes is that they are normally coupled with an automatic insulin suspension (shut-off). This automatic control was designed to reduce hypoglycemic exposure and to help individuals prone to hypoglycemia and/or with impaired hypoglycemia awareness (14). Although, overall, predictive low glucose suspension has been shown to effectively reduce exposure and frequency to nocturnal hypoglycemia (16-18), it can result in some elevation of glucose levels during the night and in the morning, or increase the time spent in moderate hyperglycemia (17-19).

It is also important to note that the use of threshold suspend is not particularly designed to eliminate the hypoglycemic events since it normally becomes active at the start point of the episode; however, depending on the threshold setting, this feature can be helpful in shortening a hypoglycemic event. As a result, the use of threshold suspend in our subjects did not eliminate the nocturnal hypoglycemic episodes. Similar observation has been reported by the Juvenile Diabetes
Research Foundation study (6) where frequent and prolonged hypoglycemic episode were noticed despite the activated nighttime CGM profile to adjust overnight basal delivery rate. Thus, the ultimate objective of this research is to benefit large spectrum of CGM users by leveraging the large data set collected from insulin pump users to evaluate the potential of nocturnal hypoglycemia prediction.

**Challenges of Late-Night Hypoglycemia Prediction**

Due to differences in counterregulatory responses to low-glucose levels during the early portion of the sleep cycle versus the late portion (dominated by REM sleep), as well as influences from circadian factors, it is expected that hypoglycemia prediction follows a different pattern in each portion of sleep. In fact, this study showed that, when using the same feature parameters, prediction of early-night hypoglycemia results in significantly a higher performance (AUC = 0.90) compared to late-night prediction (AUC = 0.75). While the false-positive rate was similar for both predictions, late-night prediction suffered from significantly higher false negatives and thus, a greater chance of missing an impending late-night hypoglycemic episode.

This observation indicates that a different or additional set of features that can more specifically relate to late-night sleep characteristics may be required for improved prediction. More specific features, in turn, implies having thorough understanding of the relationship between the late-night portion of sleep and hormonal response to BG regulations. Unfortunately, for insulin-dependent patients, this relationship is not fully understood as the bolus insulin and body’s natural insulin secretion do not follow similar glucose regulation (2). In fact, when it comes to insulin and body counterregulatory responses, contradictory findings have been reported. For example, the Somogyi effect or rebound hyperglycemia, is the phenomenon that causes the BG to rise to the hyperglycemic level in early morning in counterreaction response to a severe drop in insulin level to hypoglycemic levels. This Somogyi effect could perhaps be recognized in some of subjects in this study classified as #1 and # 3 in Table 1. However, some other researchers have challenged the Somogyi effect and in fact, reasoned a lower rather than higher BG level in the morning due to sleep-induced impaired counterregulatory responses (5). Adding to the complexity of the matter is the fear of encountering a hyperglycemia episode in patients that face an elevated SG reading at bedtime; this result could motivate them to administer bolus insulin or skip the valuable bed time snack. For instance, in our cohort, 30% of the nights SG readings started above 180 at midnight; subsequently 6% of those nights had late-night hypoglycemia. This adds to the challenges of predicting impending hypoglycemia solely based on SG readings if at the prediction timepoint an SG reading is at a hyperglycemic range.

Although several risk factors have been associated with nocturnal hypoglycemia, the predictive power of those risk factors still remains to be fully investigated and trusted (5). Some researchers have found weak link between patients demographics, such as age, sex, years of diabetes, and hypoglycemia (20). While the Somogyi effect seems to be correlated with a higher sum of bolus insulin, Woodward et al. (21) found neither an association of total insulin bolus to nocturnal hypoglycemia, nor any link between type of insulin, and frequency of insulin injection. The only predictive factor reported was bedtime SG level.

Besides patients’ demographic and CGM-based SG trajectory measurements, other studies on nocturnal hypoglycemia have looked to include more auxiliary input data provided by the patients to strengthen the prediction power. These additional data inputs include nutritional intake logs and insulin regimen (type, dose, time). Although valuable inputs, and possible to collect under controlled study conditions, the lack of consistency of log entries introduces a challenge for building a predictive model in free-living situations.

In addition to nutritional intake and insulin regimen, physical exercise and emotional state (such as stress and anxiety) have shown to affect glucose hemostasis. For instance, it has been reported that daily stress can increase the instability of glucose levels and risk of hypoglycemia (22). Physical exercise for insulin-dependent diabetes patients has been among the factors most strongly associated with severe hypoglycemia. Exercise enhances the insulin sensitivity in the muscles which can result in lowering of BG levels up to 8-10 hours after exercise. Thus, if performed during the evening, strenuous activity can pose a great risk for late-night hypoglycemia (23). In fact, Campbell et al. (7) reported that the risk of hypoglycemia following an evening exercise can be alleviated by consuming the snack only for the early stage of sleep but not in the late portion of the night. Due to the strong association of exercise and emotional state of the patients with insulin and glucose metabolism, some studies have used additional sensors to track these stimuli and include them in the feature space of the predictive models (12, 13). Eren-Oruklu et al. (13) showed improved performance of their hypoglycemia predictive model with inclusion of emotion and physical activity compared to their previously developed model without these factors. Although important and beneficial, the inclusion
of these factors and the use of additional devices besides CGM can also introduce a challenge in predicting late-night hypoglycemia in free-living conditions.

Besides exercise and stress, it is important to recognize the impact of the sleep pattern on the body’s physiological hormonal response to BG regulation. According to the report by the Centers for Disease Control and Prevention, about 1 in 3 adults do not get enough sleep and 15% of Americans work the night shift and many more work at night in addition to their day work thus, having irregular circadian rhythm. Irregular sleep pattern, short duration of sleep, and poor-quality sleep have been associated with prevalence of diabetes and impaired insulin response, and directly impact the nocturnal counterregulatory response to hypoglycemia. Although a predefined window of sleep from midnight to 6 am was considered in this study according to the recommendation by International Consensus on Use of CGM, it is important to recognize that not all the subjects sleep during this window and for the full duration.

Future work will focus mostly on the nocturnal hypoglycemia and on developing methods to include more patient input, most importantly emotion, exercise, and sleep pattern. This will also include methods to detect meals consistently since food logs could be inconsistent. Moreover, since some of the false-negative cases revealed some of the late-night glucose profile characteristics, such as Somogyi effect, we anticipate prior detection of these events could also help with the prediction modeling. Some of the long-term effort will also focus on developing personalized models to help patients better manage their diabetes according to their living conditions, such as their unique sleep pattern and BG profile.

Conclusion
This manuscript describes a model using CGM data to predict nocturnal hypoglycemia with acceptable performance with an extended prediction horizon. Early night prediction was more accurate than late night prediction. We attribute challenges in late-night hypoglycemia prediction mainly to lack of clinical understanding of some of the residual effects during the day manifesting at this portion of the sleep. We anticipate inclusion of additional BG-altering stimuli, such as emotion, exercise, and sleep patterns, that may improve late-night hypoglycemia predictability. However, inclusion will require additional devices that may not be convenient for patients in free-living conditions. Future work will focus on more data-driven methods to predict patient-specific inputs in order to enrich the model feature vector and improve the predictability.

References


