Sleep and Activity Prediction for Type 2 Diabetes Management using Continuous Glucose Monitoring

Anonymous Author(s)
Affiliation
Address
email

Abstract

Continuous glucose monitors (CGMs) generate frequent glucose measurements, and numerous studies suggest that these devices may improve diabetes management. These devices support behavior change and self-management by giving people with diabetes real-time visibility into how behavioral and lifestyle factors, i.e., meals, physical activity, sleep, stress, and medication adherence, drive their glucose levels. While earlier studies have shown that individual's actions can influence their CGM data, it has not been clear whether CGM data can provide information about these actions. This is the first study to show on a large cohort that CGM can provide information about sleep and physical activities. We first train a neural network model to determine the sequence of daily activities from CGM signals, and then extend the model to use additional data, such as individual demographics and medical claims history. Using data from 6981 participants in a Type 2 diabetes (T2D) management program, we show that a model combining an individual's CGM, demographics, and claims data is highly predictive of sleep (AUROC 0.947), and moderately predictive of a range of physical activities (AUROCs of 0.722-0.817). These results show that CGM may have wider utility as a tool for behavior change than previously known.

1 Introduction

As of 2021, an estimated 536.6 million people worldwide had diabetes, and the number is expected to increase [16]. Among people with diabetes, as many as 90-95% have T2D [5], which appears to be caused by complex interactions between genetic, environmental, and lifestyle factors. Daily actions, such as physical exercise, healthy eating, and sufficient sleep, are important factors in both preventing or delaying the onset of T2D as well as its management.

In this paper, we demonstrate the strong connection between daily activities and continuous glucose monitor (CGM) output by showing that the sequence of daily activities (sleep, walking, and elevated heart rate) can be predicted directly from the CGM signal. This may expand the utility of the device as a tool for behavior change. In addition, this can be used to analyze CGM and activity jointly to better understand their relationship and effects on health outcomes. We then develop the model further to include demographic information as well as medical claims data, and we show that this additional information further improves the prediction results.

There have been several studies using CGM data in predictive models. The most common use of this data is to predict if a person with type 1 diabetes will develop hypoglycemia [6, 4, 14, 2, 9] or hyperglycemia [9]. In addition, CGM data has been used to determine additional insulin dose requirement at mealtimes [10]. A few studies have also used CGM data to determine people's daily actions. For example, CGM data has been used to determine a person's adherence to their daily insulin injection [17], and other studies have detected eating activities [12][3]. To the best of our

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

knowledge, there has been only one study using CGM data to predict physical activity [1]. However, that study was limited to 11 people with type 1 diabetes, and its generalizability is therefore limited.

2 Methods

2.1 Dataset and Preprocessing

Our dataset consists of de-identified CGM and fitness tracker data collected from participants in a commercial diabetes care program. Participants in this program had type 2 diabetes at the time of program enrollment. As some participants do not wear the devices consistently, we ensure data quality by only using data for calendar days where the participant has at least 22 hours of overlapping CGM and fitness tracker data. In addition, we require them to have at least 1 hour of sleep during that day to ensure the fitness tracker is collecting data accurately. We use data for all participants who have at least one day of data. The resulting dataset consists of 6981 participants with an average of 65.1 days of data (median 28.0, standard deviation (SD) 85.6). The mean age of participants is 54.3 years (SD 9.1), distribution of female/male is 51.3%/48.7%, and the mean body mass index (BMI) is 34.8 (SD 8.6).

Glucose level is measured using a CGM device which provides interstitial glucose measurements every 5 minutes, and physical activity and sleep are tracked using a fitness watch which reports average heart rate and step count once a minute, as well as the intervals when the participant is predicted to be asleep. Due to the difference in time sampling frequencies, the fitness tracker data is aggregated for each 5-minute time block to match the CGM. Heart rates are categorized using the fitness tracker's personalized heart rate zones into resting, fat burn, and cardio zones. Walking activity is binarized using a threshold of 400 steps per five minutes, which is slightly below a typical moderate-intensity walking pace [15]. To reduce noise, we filter out individual positive values from the labels, thereby requiring each activity to last for at least 10 minutes at a time.

In addition, we account for participants' medical history by means of their medical claims data. This data includes the information necessary for the reimbursement of medical services including International Classification of Diseases (ICD10) diagnosis codes [11], and service dates. In addition, we group similar ICD codes together by only considering the first 3 letters/numbers of each code. On average, participants had 188.2 (SD 178.2) ICD codes in their claims data, with 40.7 (SD 23.8) unique codes.

2.2 Model

Our model is based on the U-Net architecture [13], which was originally designed for medical image segmentation. As our goal is to transform a sequence of glucose measurements to a sequence of daily activities, we modify the original model to use 1-dimensional convolutions instead of 2-dimensional convolutions. The input matrix $\mathbf{X}$ is a sequence of 288 glucose measurements (24 hours * 12 measurements/hour) along with the rate of change computed by the CGM device. These two values are combined as separate channels, i.e., $\mathbf{X} \in \mathbb{R}^{2\times 288}$. Prediction targets include four activities: sleep, walk, elevated heart rate (Fat burn zone), and high heart rate (Cardio zone). These values are measured at the same time intervals as the glucose values and treated as separate channels, therefore the target matrix $\mathbf{Y} \in \mathbb{R}^{4\times 288}$. The prediction is treated as a multi-label classification task, i.e., multiple target values can be positive simultaneously (e.g., walking and high heart rate).

Next, to incorporate claims data into the model, we convert the ICD codes that occur prior to the current time window into vectors $\mathbf{x_{icd}} \in \mathbb{R}^{300}$. This conversion is done using pretrained ICD2Vec embeddings [7]. As our hypothesis is that some ICD codes might be more relevant for the prediction task than others, we want the model to be able to learn which ICD codes to use. In addition, we want the model to be able to distinguish between acute and chronic diseases, as acute diseases might be relevant only for a few days or weeks while chronic ones could have effects across many years. Including ICD codes into the model is implemented using a multi-head attention layer [18], as shown in Figure 1. Time is first converted into larger time blocks to avoid overly sparse time values using conversion $f(t) = \lfloor log(t) \rfloor$, where t is the number of days between the current time window and the claim. This time block is then converted into a vector $\mathbf{t} \in \mathbb{R}^{50}$ using step encoding shown in [18] and concatenated to the ICD vector $\mathbf{x_{icd}}$. The resulting vector is used as both the key $\mathbf{k}$ and value $\mathbf{v}$ for the attention layer while a constant vector is used as the query $\mathbf{q}$.

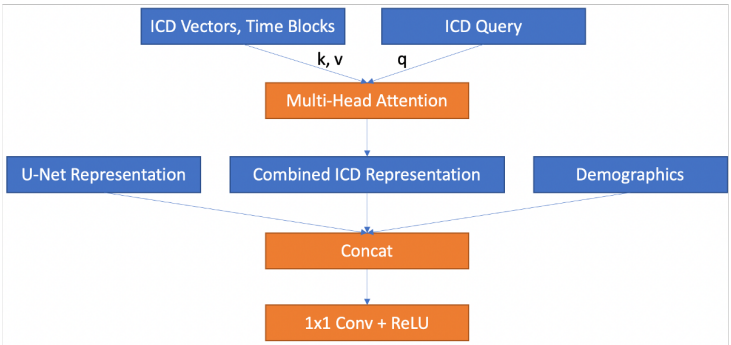


Figure 1: Combining ICD vectors using Multi-Head Attention and adding the combined vector along with demographics to U-Net's internal representation.

We evaluate including the combined ICD representations in three locations within the U-Net model: in the input layer, between encoder and decoder, as well as before the last convolutional layer. These vectors are concatenated with the U-Net's internal representation and passed through a convolution with size 1 and a rectified linear unit (ReLU) layer. The resulting vector is then returned to the same location of the U-Net model where the original vector was. Demographics (age, gender, height, weight, BMI) are standardized and combined with the internal representations the same way.

2.3 Training

Individual participants are randomly divided 60%-20%-20% to training, validation, and test sets respectively. Models are trained using AdamW [8] algorithm for 35 epochs, which is enough to reach convergence. In each training epoch, a random day of data is chosen for each participant such that each participant appears exactly once in each epoch, but different epochs might use different time windows for participants. In validation and test phases, one day of data is randomly chosen for each participant but this same day is used in every epoch and experiment. This ensures that the results are consistent across different experiments and each participant has an equal impact on the results.

3 Experiments and Results

We evaluated our model with and without using demographics and ICD codes. We also compared the predictive performance when changing the location where demographics and ICD codes were incorporated into the model. Results of these ablation experiments are shown in Table 1. As the results show, including ICD codes and demographics improved heart rate zone and walk predictions. Overall, including demographics and ICD codes later in the model provided better results than including them early or between the encoder and decoder. Sleep predictions were not affected by demographics or ICD codes.

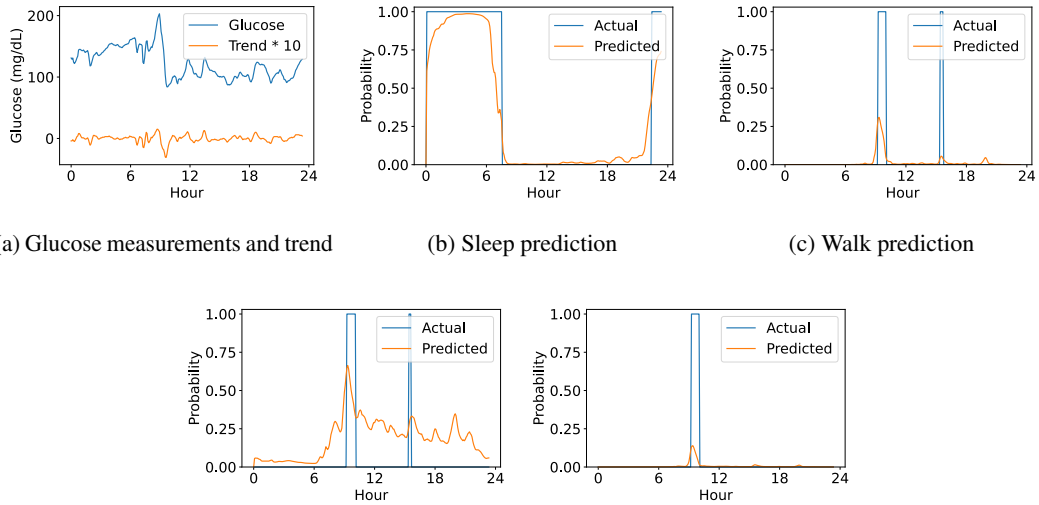
Figure 2 shows an example of predictions made for an individual participant. The model was typically very confident in the sleep predictions except in the beginning and end of the sleep period. The lower confidence in both ends can be partially explained by the noise in the ground truth data, as the sleep times were estimated by the fitness tracker based on movement and heart rate patterns. Walking probabilities were skewed toward zero because of the rarity of the event, but despite the miscalibration the model did give higher probabilities when the participant was truly walking. However, the probability sometimes increased slightly at other times also, possibly when the participant was doing other physical activities which had a similar impact on the glucose measurements as walking. Similarly, heart rate predictions often had higher values when the heart rate was elevated, but the predictions were slightly noisy.

4 Conclusion and Future Work

In this paper, we have shown for the first time using a large cohort of people with T2D that CGM data is highly predictive of sleep and moderately predictive of physical activities. To do so, we developed

Table 1: Comparison of AUROC scores using demographics and ICD codes at different parts of the model. The first row shows results without demographics or ICD codes, the following rows show results when either ICD codes or demographics are added at different locations (early, middle, late), and the last row shows results when both demographics and ICD codes are added into the model.

Demographics	ICD Codes	Sleep AUROC (SD)	HR Fat Burn AUROC (SD)	HR Cardio AUROC (SD)	Walk AUROC (SD)
-	-	0.946 (0.002)	0.708 (0.007)	0.727 (0.020)	0.776 (0.015)
-	Early	0.936 (0.002)	0.687 (0.005)	0.707 (0.017)	0.770 (0.020)
-	Middle	0.945 (0.003)	0.704 (0.007)	0.715 (0.020)	0.790 (0.018)
-	Late	0.945 (0.002)	0.719 (0.008)	0.744 (0.019)	0.804 (0.016)
Early	-	0.942 (0.003)	0.669 (0.006)	0.740 (0.018)	0.809 (0.012)
Middle	-	0.945 (0.003)	0.716 (0.007)	0.758 (0.018)	0.797 (0.013)
Late	-	0.946 (0.002)	0.716 (0.007)	0.752 (0.025)	**0.820 (0.015)**
Late	Late	**0.947 (0.002)**	**0.722 (0.007)**	**0.768 (0.019)**	0.817 (0.016)



(a) Glucose measurements and trend (b) Sleep prediction (c) Walk prediction

(d) Fat burn zone prediction (e) Cardio zone prediction

Figure 2: One participant's CGM signal and predictions for one day. X-axis shows hours starting from midnight.

a neural network architecture for predicting daily activities using CGM, demographics, and claims data. We first modified the U-Net architecture so that it can be used with 1-dimensional CGM signals, and then incorporated claims data using an attention layer which allowed the model to learn which claims were important for the task.

In our future work, we will evaluate using additional data sources (such as medications) as features to further improve our results. In addition, we will investigate using alternative approaches to determine physical activity, such as combining heart rate and walking into a more informative label, as neither of these labels is a perfect measure of physical activity on their own. For example, walking is not the only type of physical activity for many people and it might have a similar CGM pattern as other physical activities. On the other hand, estimated heart rate zones can be imprecise for some of the participants. We expect these changes to improve our physical activity predictions even further.

Our results open up new opportunities for the utilization of CGM devices for diabetes management and allow us to gain a better understanding of CGM signal during activities.

References

[1] Mohammad Reza Askari et al. "Meal and Physical Activity Detection from Free-Living Data for Discovering Disturbance Patterns of Glucose Levels in People with Diabetes". In: *BioMedInformatics* 2.2 (2022), pp. 297–317. ISSN: 2673-7426. DOI: 10.3390/biomedinformatics2020019.

[2] Vladimir B. Berikov et al. "Machine Learning Models for Nocturnal Hypoglycemia Prediction in Hospitalized Patients with Type 1 Diabetes". In: *Journal of Personalized Medicine* 12.8 (2022). ISSN: 2075-4426. DOI: 10.3390/jpm12081262.

[3] Lauriane Bertrand, Nathan Cleyet-Marrel, and Zilu Liang. "The Role of Continuous Glucose Monitoring in Automatic Detection of Eating Activities". In: *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*. 2021, pp. 313–314. DOI: 10.1109/LifeTech52111.2021.9391849.

[4] Darpit Dave et al. "Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction". en. In: *Journal of Diabetes Science and Technology* 15.4 (July 2021), pp. 842–855. ISSN: 1932-2968, 1932-2968. DOI: 10.1177/1932296820922622.

[5] Centers for Disease Control and Prevention. *Type 2 diabetes*. Dec. 2021. URL: https://www.cdc.gov/diabetes/basics/type2.html (visited on 09/20/2022).

[6] Yazan Jian et al. "A Machine Learning Approach to Predicting Diabetes Complications". en. In: *Healthcare* 9.12 (Dec. 2021), p. 1712. ISSN: 2227-9032. DOI: 10.3390/healthcare9121712.

[7] Yeong Chan Lee et al. *ICD2Vec: Mathematical representation of diseases*. July 2021. DOI: 10.21203/rs.3.rs-692012/v1.

[8] Ilya Loshchilov and Frank Hutter. "Decoupled Weight Decay Regularization". In: *ICLR*. 2019.

[9] Yonit Marcus et al. "Improving blood glucose level predictability using machine learning". In: *Diabetes/Metabolism Research and Reviews* 36.8 (2020), e3348. DOI: https://doi.org/10.1002/dmrr.3348.

[10] Giulia Noaro et al. "Machine-Learning Based Model to Improve Insulin Bolus Calculation in Type 1 Diabetes Therapy". In: *IEEE Transactions on Biomedical Engineering* 68.1 (2021), pp. 247–255. DOI: 10.1109/TBME.2020.3004031.

[11] World Health Organization. *International Classification of Diseases (ICD)*. URL: https://www.who.int/standards/classifications/classification-of-diseases (visited on 09/20/2022).

[12] Victor Palacios, Diane Myung-Kyung Woodbridge, and Jean L. Fry. "Machine Learning-based Meal Detection Using Continuous Glucose Monitoring on Healthy Participants: An Objective Measure of Participant Compliance to Protocol". In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2021, pp. 7032–7035. DOI: 10.1109/EMBC46164.2021.9630408.

[13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[14] Wonju Seo et al. "A machine-learning approach to predict postprandial hypoglycemia". en. In: *BMC Medical Informatics and Decision Making* 19.1 (Dec. 2019), p. 210. ISSN: 1472-6947. DOI: 10.1186/s12911-019-0943-4.

[15] J. Slaght et al. "Walking Cadence to Exercise at Moderate Intensity for Adults: A Systematic Review". en. In: *Journal of Sports Medicine* 2017 (2017), pp. 1–12. ISSN: 2356-7651, 2314-6176. DOI: 10.1155/2017/4641203.

[16] Hong Sun et al. "IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045". In: *Diabetes Research and Clinical Practice* 183 (2022), p. 109119. ISSN: 0168-8227. DOI: https://doi.org/10.1016/j.diabres.2021.109119.

[17] Daniel N. Thyde et al. "Machine Learning-Based Adherence Detection of Type 2 Diabetes Patients on Once-Daily Basal Insulin Injections". In: *Journal of Diabetes Science and Technology* 15.1 (2021). PMID: 32297804, pp. 98–108. DOI: 10.1177/1932296820912411.

[18] Ashish Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).