

Explainable machine learning for real-time hypoglycaemia and hyperglycaemia prediction and personalised control recommendations

Christopher Duckworth Ph.D¹; Matthew J Guy Ph.D^{2,4}, Anitha Kumaran MRCPCH, PhD, Aisling Ann O’Kane Ph.D^{4,5}, Amid Ayobi Ph.D⁴, Adriane Chapman Ph.D⁶, Paul Marshall DPhil^{4,5}, Michael Boniface CEng MIET¹

Author Affiliations:

¹IT Innovation Centre, Electronics and Computer Science, University of Southampton (Southampton SO17 1BJ); ²Department of Medical Physics, University Hospital Southampton NHS Foundation Trust (Southampton SO16 6YD); ³ Child Health. Department of Endocrinology, University Hospital Southampton NHS Foundation Trust (Southampton SO16 5YA); ⁴Human-Computer Interaction for Health, University of Bristol (75 Woodland Rd, Bristol BS8 1UB); ⁵Honorary Lecturer at UCL Interaction Centre, University College London; ⁶Electronics and Computer Science, University of Southampton (Southampton SO17 1BJ)

CD: C.J.Duckworth@soton.ac.uk

MG: Matthew.Guy@uhs.nhs.uk

AK: Anitha.Kumaran@uhs.nhs.uk

AO’K: a.okane@bristol.ac.uk

AA: amid.ayobi@bristol.ac.uk

AC: Adriane.Chapman@soton.ac.uk

PM: p.marshall@bristol.ac.uk

MB: M.J.Boniface@soton.ac.uk

Abbreviations:

(T1D) Type-1 Diabetes, (CGM) Continuous Glucose Monitor, (AUROC) Area Under the Receiver Operating Curve, (SHAP) SHapley Additive exPlanations

Keywords: continuous glucose monitoring, explainable and trustworthy AI, feature extraction, hypoglycaemia prediction, hyperglycaemia prediction, machine learning

Corresponding Author: Christopher Duckworth C.J.Duckworth@soton.ac.uk

Conflict-of-Interest Disclosure: None

Acknowledgements: We acknowledge funding from UKRI Trustworthy Autonomous Systems Hub (Grant code: RITM0372366).

4 Figures, 2 Tables, 2999 words Main text (excluding title page, abstract, figure legends, tables (& headings), and references), 229 words Abstract

Abstract

Background: The occurrences of acute complications arising from hypoglycaemia and hyperglycaemia peak as young adults with type 1 diabetes (T1D) take control of their own care. Continuous glucose monitoring (CGM) devices provide real-time blood glucose readings enabling users to manage their control pro-actively. Machine learning algorithms can use CGM data to make ahead-of-time risk predictions and provide insight into an individual's longer-term control.

Methods: We introduce explainable machine learning to make predictions of hypoglycaemia (<70mg/dL) and hyperglycaemia (>270mg/dL) 60 minutes ahead-of-time. We train our models using CGM data from 153 people living with T1D in the CITY survey totalling over 28000 days of usage, which we summarise into (short-term, medium-term, and long-term) blood glucose features along with demographic information. We use machine learning explanations (SHAP) to identify which features have been most important in predicting risk per user.

Results: Machine learning models (XGBoost) show excellent performance at predicting hypoglycaemia (AUROC: 0.998) and hyperglycaemia (AUROC: 0.989) in comparison to a baseline heuristic and logistic regression model.

Conclusions: Maximising model performance for blood glucose risk prediction and management is crucial to reduce the burden of alarm-fatigue on CGM users. Machine learning enables more precise and timely predictions in comparison to baseline models. SHAP helps identify what about a CGM user's blood glucose control has led to predictions of risk which can be used to reduce their long-term risk of complications.

Introduction

People with type-1 diabetes (T1D) face a daily balance to keep their blood glucose levels within safe levels (i.e. 'in-range'). Severe complications are prevalent and arise from glycaemic variability, low blood sugars (hypoglycaemia) and high blood sugars (hyperglycaemia)[1]. For hypoglycaemic incidents alone, the requirement for emergency assistance may be as high as 7.1% per year [2] and could account for 6-10% of deaths for those with T1D [3, 4]. Long-term impacts of hypoglycaemia include impacts on cognition and potential links with dementia[5]. In addition, frequent hyperglycaemia can lead to short-term risk such as diabetic ketoacidosis and long-term complications such as retinopathy, neuropathy, nephropathy, and cardiovascular disease[6-8]. Effective glucose management for adolescents and young adults living with T1D is challenging[9, 10], due to the multiple transitions taking place in their lives, including puberty, relationships, the move to more independent living and diabetes self-care, and also the transfer from paediatric to adult clinical care teams. Parental fear of severe complications is prevalent throughout these transitional years[11-13].

Continuous glucose monitoring (CGM) enables regular automated readings of estimated blood glucose levels, providing immediate insight into blood glucose control. CGM has been demonstrated to reduce the risk of both hypoglycaemia and hyperglycaemia, along with reducing daily glycaemic variability for users

with type-1 diabetes[14-16]. In addition to mitigating short-term risk of severe hypoglycaemia and hyperglycaemia, compliance of wearing CGM devices has been shown to improve glycosylated haemoglobin HbA1c levels, which, if sustained, reduce long-term complication risks[17, 18]. The magnitude of reduction in HbA1c from CGM usage is dependent on the user's original HbA1c value; i.e. those at highest risk of complications from poorer control are likely to benefit the most [16]. Specific to young adults, Laffel et al. [19] demonstrate a clear improvement in HbA1c for those utilising CGM.

Real-time CGM devices provide alerts for users when their blood glucose falls above or below a desired range. T1D management can be aided further by having *ahead-of-time* predictions so individuals can identify risk early and better plan self-care activities, such as insulin dosages. Simple threshold-based algorithms have been able to successfully predict hypoglycaemia 30 minutes in advance (e.g. Medtronic-640 'SmartGuard'[20]). More complex statistical models and machine learning algorithms enable more accurate prediction and are able to extend this prediction horizon[21-28]. Dave et al. [23] emphasize the importance of feature extraction when generating predictions of hypoglycaemia in CGM data. Generating features that are both predictive in models and insightful for understanding a user's blood glucose control is a difficult balance.

In this work, we make two novel contributions: algorithms tailored to young adults and explanations. First, we introduce machine learning models to predict hypoglycaemia (<70mg/dL) and hyperglycaemia (>270mg/dL)[29] with a trustworthy 60-minute prediction horizon for young adult users of CGM. While CGM risk prediction is a well explored topic, more must be done to understand what led to increased risk for an individual so they can be proactive. We introduce using *explainable* machine learning, to not only predict risk, but to automatically identify the most important factors in an individual's CGM data that led to increased risk. Explanations have no detrimental impact on model performance. We provide a framework in which machine learning can be used to:

- 1) Provide real-time predictions of hypoglycaemia and hyperglycaemia (Results - Model Evaluation) using intuitive features (Methods – Features) generated from CGM data (Methods – Data).
- 2) Automatically identify the most important features that have led to predictions of risk for each CGM user over a given time-period (Results – Model Explanation).
- 3) Provide personalised control recommendations for each CGM user to help with their T1D management (Results – User Interface).

Methods

Data

We make use of publicly available data from “A Randomized Clinical Trial to Assess the Efficacy and Safety of Continuous Glucose Monitoring in Young Adults 14-<25 with Type 1 Diabetes” (CITY)[19]. By design, the study recruited adolescents and young adults with T1D (duration > 12 months) exhibiting poorer glycaemic control (HbA1c 7.5-<11.0%), most likely to benefit from CGM usage [16]. Study participants were randomly assigned to either CGM (Dexcom

G5) or regular blood glucose meter (finger-prick) monitoring. The CGM users were compared to the control group using HbA1c levels after six months of usage. After six months, all study participants were provided with CGM devices and HbA1c tracked for a further six months.

We make use of CGM data from 153 people living with T1D in the CITY study, where users were provided CGM devices for 6-12 months; totalling over 28,000 days of usage data. In addition to CGM data, basic screening information and the most recently recorded HbA1c test result were used to generate predictions.

Features

To utilise CGM data for hypoglycaemic and hyperglycaemic prediction, we generate a total of 30 features which summarise a young adult's CGM data on different timescales. Blood glucose control is summarised on short-term (one hour), medium-term (one day) and long-term (one week) baselines prior to the current CGM reading. This is combined with six features that characterise basic patient information. A complete description of all generated features are given in Table 1. Features are generated at the point of each unique CGM reading. Features are only used in modelling if the CGM device has been used for $\geq 80\%$ for the prior week.

Feature	Description	Time-period(s)
Current reading	Most recent CGM blood glucose reading	N/A
Time of day	Hour (0-24) at which reading was reported	N/A
Day of week	Day on which reading was reported	N/A
Gender		N/A
Diagnosis Age	Age at initial diagnosis of type-1 diabetes	N/A
Prior use of CGM	Whether person with T1D had previous experience of using CGM before the study	N/A
Age	Age at study commencement	N/A
Years since original diagnosis	Year since initial diagnosis of type-1 diabetes	N/A
Most recent HbA1c	Most recent recorded test result of HbA1c	N/A
Device usage fraction	Fraction of time (as specified by the time-period) of which the CGM device was used	(1 hour, 1 day, 1 week)
Fraction of time high	Fraction of time (as specified by the time-period) of which CGM readings were above 270 mg/dL	(1 hour, 1 day, 1 week)
Fraction of time low	Fraction of time (as specified by the time-period) of which CGM readings were below 70 mg/dL	(1 hour, 1 day, 1 week)
Average	Mean of blood glucose readings over specified time-period	(1 hour, 1 day, 1 week)
Standard deviation	Standard deviation of blood glucose readings over specified time-period	(1 hour, 1 day, 1 week)

		1 week)
Largest increase between readings	Largest increase in blood glucose level between consecutive readings within specified time-period	(1 hour, 1 day, 1 week)
Largest decrease between readings	Largest decrease in blood glucose level between consecutive readings within specified time-period	(1 hour, 1 day, 1 week)
Maximum number of consecutive increases	Most consecutive readings where blood glucose levels increase over defined time-period	(1 hour, 1 day, 1 week)
Maximum number of consecutive decreases	Most consecutive readings where blood glucose levels decrease over defined time-period	(1 hour, 1 day, 1 week)

Table 1: Summary of input features used by the models to make predictions. A sub-set of features are computed for various time-ranges (i.e. 1 hour, 1 day, 1 week) and considered as independent features.

Targets

To generate targets for our model predictions, we generate two binary variables referring to hypoglycaemic (< 70 mg/dL) and hyperglycaemic (> 270 mg/dL) events. A feature set is generated for each unique CGM reading, at which point we check if the CGM user's blood glucose level falls within these regions in the following 60-minutes (i.e. positive prediction). Blood glucose readings already within the hypoglycaemic or hyperglycaemic regions are removed from the modelling dataset to avoid artificially boosting model performance metrics. Figure 1 shows a schematic of blood glucose levels through a given day, regions of hypoglycaemia and hyperglycaemia and timestamps of model predictions prior (i.e. target).

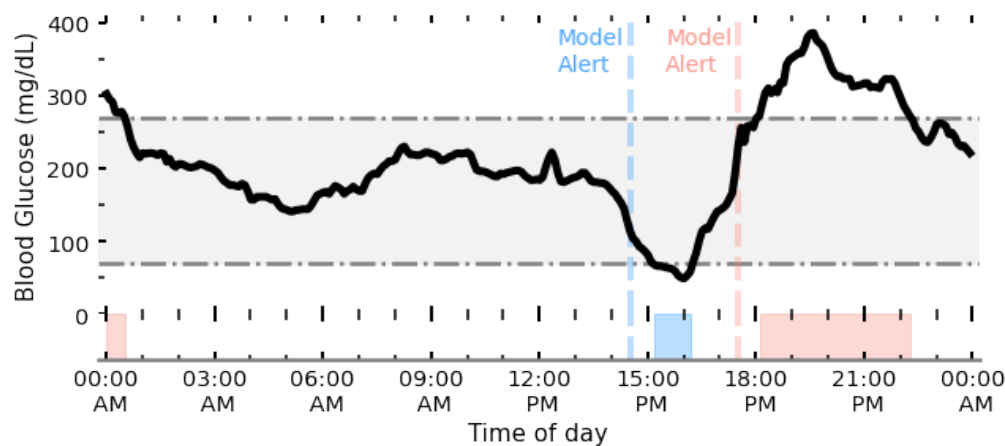


Figure 1: Schematic of blood glucose levels (black line) for a young adult with T1D tracked by CGM. The grey shaded region shows the desired range to keep blood glucose levels between (70mg/dL < BG < 270mg/dL). Our algorithm aims to predict (ahead-of-time) when a person with T1D will go below (hypoglycaemia) and above (hyperglycaemia) this range. Regions of low and high blood glucose are shaded blue and red respectively, with the corresponding first prediction event horizon (i.e. when our model first made a positive prediction of hypo/hyper) shown by the dashed line.

Modelling

To determine the added value of machine learning we evaluate a baseline heuristic model, a logistic regression model and a gradient boosted tree-based model for both hypoglycaemia and hyperglycaemia prediction. Our baseline heuristic model is equivalent to a blood glucose threshold alert (i.e. predicting hypoglycaemia and hyperglycaemia within 60-minutes if blood glucose levels fall below 110mg/dL or go above 240mg/dL respectively). Our logistic regression model is aimed to emulate basic CGM alerts which extrapolate linear trends along with thresholds to make hypoglycaemia or hyperglycaemia predictions.

Finally, we make use of the XGBoost framework to implement a tree-based machine learning algorithm [30]. XGBoost makes use of an ensemble of weak learners (i.e. small trees) that are trained stage-wise through gradient boosting. This reduces overfitting while preserving or lowering variance in the prediction error [31], which frequently leads to gradient boosted trees outperforming other tree-based methods. Additionally, XGBoost naturally deals with continuous, binary/discrete, and missing data consistently; all of which are represented in our dataset. Model hyperparameters for our XGBoost models were selected using five-fold cross-validation of the complete training set using a sampler (Tree-structured Parzen Estimator) implemented with the Optuna library[32].

We randomly separate our CGM data into a hold-out test set (25%) and a training set (75%). Our supervised models (i.e. logistic regression and XGBoost) learn from the training set, and all models are evaluated using the same test sample. Overall, model performance was evaluated using the Area Under the Receiver Operating Curve (AUROC) and average precision, along with fixed measures of specificity and sensitivity.

Model explainability

Historically, machine learning algorithms are considered 'black-boxes' with little understanding of how predictions have been made. However, recent advances in *explainability* have led to individual predictions of tree-based algorithms being readily explainable[33].

To attribute the relative importance of each feature in predicting both hypoglycaemia and hyperglycaemia risk for our XGBoost model, we make use of the TreeExplainer algorithm as implemented in the SHAP (SHapley Additive exPlanations) library[33-35]. TreeExplainer efficiently calculates Shapley (SHAP) values[36], which aim to attribute payout (i.e. the prize) between coalitional players of a game. In the context of machine learning, SHAP values amount to the marginal contribution (i.e. change to the model prediction) of a feature amongst all possible coalitions (i.e. combinations of features). Practically, this means that for every individual prediction (negative or positive), the relative importance of every feature can be evaluated.

There is a rich history of global interpretation for machine learning models which summarise the average overall importance of features on predictions as a whole[37]. In a medical setting, however, tailored explanations for individuals are paramount, maximising the ability to understand their own data and ensure every person is evaluated fairly[38]. Shapley values are *locally accurate*, meaning that they can explain which features were relatively most important for an individual prediction (i.e. a hypoglycaemic or hyperglycaemic event). In addition, Shapley values are consistent (the values add up to the actual

prediction of the model) meaning they can also be used to check the global importance of a feature. Feature importance can therefore be checked periodically by averaging over a fixed time-period. Practically, this means that for a CGM user over a given time-period, the most important features leading to a prediction of hypoglycaemia or hyperglycaemia can be automatically evaluated. This gives immediate insight about an individual's blood glucose control, and intuition about what may be increasing their risk. Presenting reliable predictions with intuitive explanations, would enable users to be proactive in their control. Insightful control recommendations could empower users to feel closer to being on 'auto-pilot' (i.e. minimising the cognitive load burden).

We choose to implement SHAP over other local explainer algorithms (e.g. Lime[39]) since SHAP offers mathematical guarantees of trustworthiness (local accuracy, missingness, and consistency) which adhere to strict medical governance guidelines[33], and offers consistency between local explanations meaning global importance can be computed as well.

Results

Model evaluation

In Figure 2, we compare the performance of our baseline heuristic model against the machine learning classifiers (i.e. logistic regression and XGBoost). Performance is evaluated by the AUROC characteristic by comparing the model predictions of hypoglycaemia (left) or hyperglycaemia (right) 60-minutes ahead-of-time to the actual future readings. For hypoglycaemia, the baseline model achieved an AUROC of 0.811, the logistic regression 0.930 (95% CI:0.929-0.931) and XGBoost 0.998 (95% CI:0.998-0.998) evaluated on our hold-out test set. All confidence intervals (CI) are estimated from bootstrapping (sampling with replacement) for 500 resamples per model.

Both machine learning models demonstrated excellent predictive power for hypoglycaemia, with a clear advantage in using XGBoost. We note that despite its crudeness, our baseline heuristic model also performs well; demonstrating the use of threshold-based alerts on CGM devices in forward planning. Regardless, a more powerful predictive model means a lower false-alarm rate can be achieved, while maintaining the safety of the predictions. Reducing alarm-fatigue for CGM users is an important goal, and more skilful models help enable this. In Table 2, additional measures of model skill are given, including average precision, sensitivity, and specificity. Sensitivity and specificity are evaluated from dichotomising model predictions at probability $P=0.5$. Again, we find a clear performance increase for our XGBoost model, in-keeping with the high performance of decision tree based methods[40] and commercial hybrid loop systems[41].

High performance is also seen for hyperglycaemia, with the baseline model achieving an AUROC of 0.734, the logistic regression 0.862 (95% CI:0.861-0.862) and XGBoost 0.989 (95% CI:0.989-0.990). Average precision, sensitivity, and specificity demonstrate similar trends with XGBoost being the most skilful. For each modelling approach we note that the model skill is lower for hyperglycaemia prediction in comparison to hypoglycaemia, suggesting

prediction of lower blood glucose events is better suited to our modelling choices.

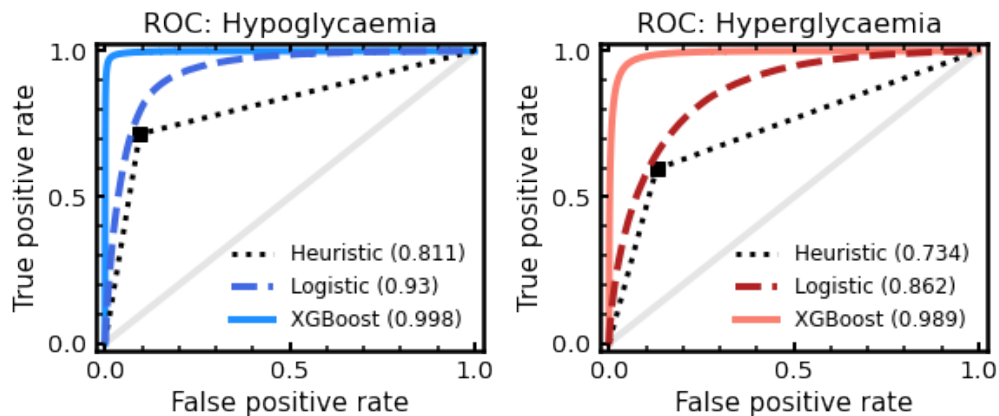


Figure 2: Receiving operator characteristic (ROC) for our models of hypoglycaemia (left) and hyperglycaemia (right) prediction. In each panel, a XGBoost model (solid line) and a logistic regression model (dashed line) are compared to a baseline heuristic (dotted line). A zero skill model is represented by the solid grey line. The total area under each curve (i.e. AUROC score) is given in the brackets.

MODEL	AUROC	AVERAGE PRECISION	SPECIFICITY (P _{THRES} = 0.5)	SENSITIVITY (P _{THRES} = 0.5)
Hypoglycaemia				
Heuristic	0.811	0.121	0.906	0.716
Logistic Reg.	0.930 [0.929-0.931]	0.244 [0.240-0.247]	0.827	0.905
XGBoost	0.998 [0.998-0.998]	0.953 [0.951-0.954]	0.994	0.945
Hyperglycaemia				
Heuristic	0.733	0.258	0.872	0.595
Logistic Reg.	0.862 [0.861-0.862]	0.453 [0.450-0.456]	0.752	0.817
XGBoost	0.989 [0.989-0.990]	0.931 [0.930-0.932]	0.931	0.970

Table 2: Summary of model performance metrics for both hypoglycaemia and hyperglycaemia prediction. A baseline heuristic, logistic regression and an XGBoost model are evaluated for each target. Summary statistics (AUROC and average precision) are shown with 95% CI in square brackets. Sensitivity and specificity are evaluated from dichotomising model predictions at probability P= 0.5.

Model explanation

In addition to increased predictive power, the added value of machine learning models can be demonstrated through explanations. Using SHAP we can evaluate the relative importance of features for a given positive prediction of hypoglycaemia or hyperglycaemia. SHAP is applied post model construction and therefore has no negative implications for performance. Figure 3 shows the overall relative importance of every input feature for predicting hypoglycaemic (left-panel) and hyperglycaemic (right-panel) events. The relative importance of a feature is quantified by the absolute average SHAP value. Since SHAP values are consistent across predictions, they can be averaged for individual CGM users, across any time-range, to provide immediate insight.

Here we provide the average relative importance for all CGM users in the study, but this diagram is trivially made for individual users. Unsurprisingly, the user's current blood glucose reading is most important for the model to make

predictions of both hypoglycaemia and hyperglycaemia. Time of day is also important, providing insight into the sleep and eating, physical activity and stress level habits of the CGM user and their relationship with blood glucose. Sudden drops (or increases) in blood glucose are important for predicting hypoglycaemia (hyperglycaemia) as shown by the short-term largest decrease (increase) between readings. Interestingly the long-term fraction of time low is found to be reasonably predictive of hypoglycaemic events, providing immediate insight into certain user's control habits.

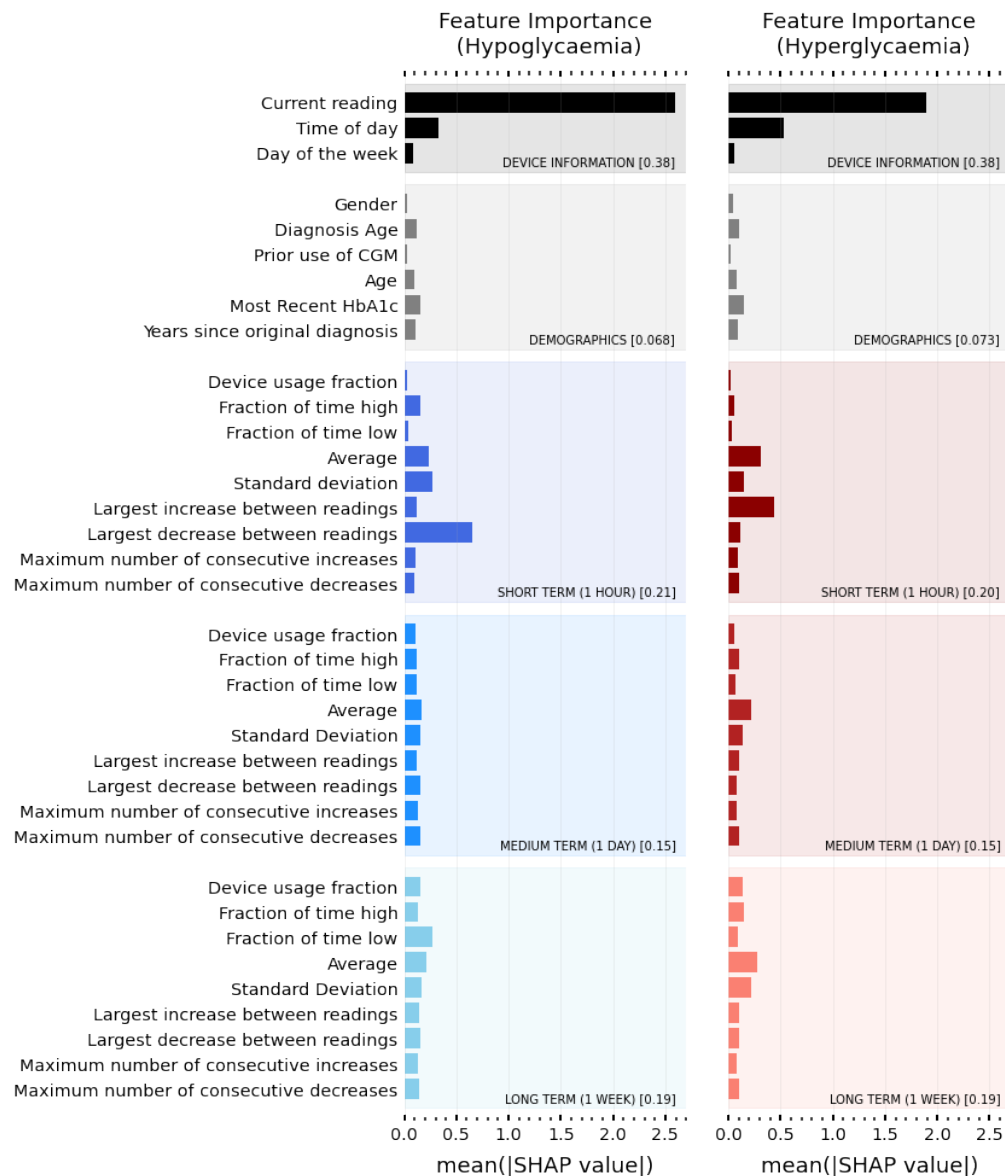
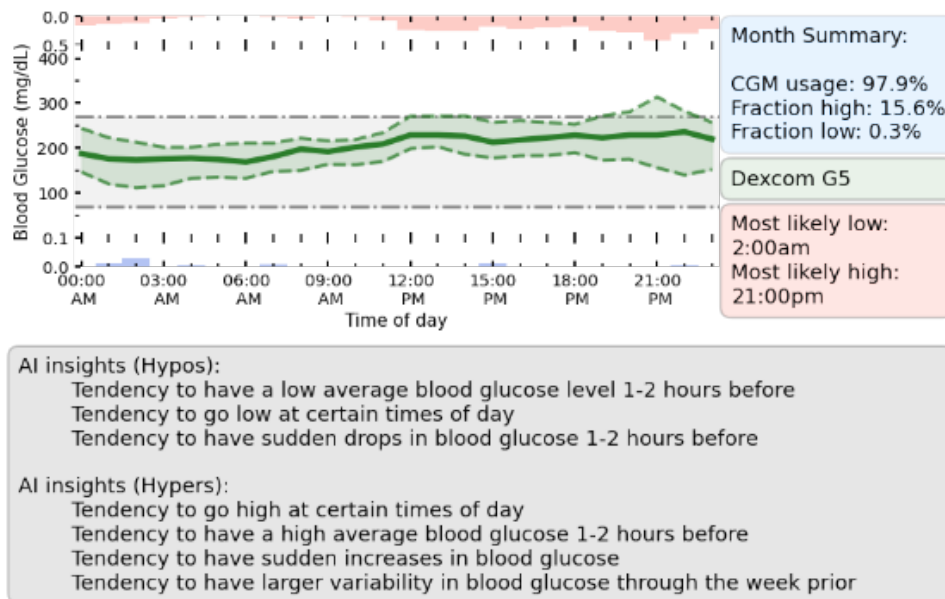


Figure 3: Overall importance ranking of input features for predicting hypo (left panel) and hyper (right panel) risk. Average (absolute) SHAP value for predictive features over all study participants. A higher value corresponds to a more important feature in decision making. Features are grouped into categories (Device information, Demographics, Short term (1 hour), Medium Term (1 day), Long term (1 week)). The fractional contribution (i.e. sum over all features in that category) of a given category is given in the square brackets.

User interface

Despite CGM providing a wealth of information to both users and clinicians, the sheer volume of data makes it hard to quickly draw conclusions about blood

glucose control. Quick summary metrics such as the fraction of time-in-range (e.g. $70\text{mg/dL} < \text{BG} < 270\text{mg/dL}$) are the baseline for assessing control. By considering the most predictive model features that led to predictions of hypoglycaemic or hyperglycaemic events, we can draw further personalised insights into an individual's blood glucose control. In Figure 4, we present a prototype dashboard which summarises a randomly selected user's CGM data over a given month, along with potential insights derived from explainable machine learning. In addition to metrics such as time above or below range, we provide the user's average blood glucose through the day, along with the most likely times for our model to predict hypoglycaemia (red, above green line) or hyperglycaemia (blue, below green line) for the individual. We select the top features for predicting both hypoglycaemia and hyperglycaemia for the user and summarise this information as control recommendations in the grey box. This provides a quick glance into the specifics of the user's blood glucose control; enabling the user to be better informed to avoid potential events in the future. One AI insight (grey box) for this user is that they tend to go high at specific times of day. Looking at the fraction of time spent high on the dashboard through the day (red box and histogram), this peaks around 21:00pm, hence the user should consider insulin dosages around their evening meal.



Discussion

The key contributions of our work are as follows:

- 1) Machine learning models with state-of-the-art performance for predicting hypoglycaemia (AUROC:0.998) and hyperglycaemia (AUROC: 0.989) 60-minutes in advance. This performance is high relative to simple algorithms[42-44] and comparable machine learning approaches[23, 45].
- 2) With careful feature engineering, we have demonstrated how machine learning explanations (SHAP) can be utilised to understand specifics about an individual's control. SHAP also adds transparency to model predictions, aiding assurance that all individuals are evaluated fairly.
- 3) Provided a prototype dashboard to help young adults with T1D and clinicians make use of CGM data and the insight from machine learning explanations.

Technological advances represent a significant opportunity to help reduce self-care burden on an individual with T1D, and reduce the risk of health complications arising from poor glycaemic control. In particular, for young adults, automated feedback from CGM may be an important tool for reducing risk, at times of transition (from paediatric to adult care units) and where glycaemic control can be at a minimum.

Ahead-of-time machine learning predictions are of personal and clinical value as they give the CGM user more time to adjust self-care and reduce risk. Our tree-based model demonstrated a significant performance increase relative to threshold based and linear models. This performance increase is vital for reducing alert burden on the user, since more certain predictions require less total alerts while maintaining safety of the device.

Despite the wealth of information provided by CGM devices, part of the problem is deriving quick insight that is useful for people with T1D, their family carers, and clinicians[46, 47]. Machine learning explanations can help summarise what specifics in an individual's glycaemic control led to increased risk of either hypoglycaemia or hyperglycaemia. Used in combination with directly derived metrics (e.g. time-in-range), their utility can be in providing quick-glance specific recommendations about how to reduce risk.

Limitations

Limitations of this work include the reliance on the user to comply in using the CGM device. For our results, we only generate predictions when the user has used the device for 80% of the prior week. While predictions can still be generated with a lower usage compliance, this will inevitably decrease prediction performance, and care must be taken about when machine learning enhancement can be implemented safely. Furthermore, while current CGM devices are generally accurate, they are not infallible and considerations must be made for the safety of systems reliant on their accuracy[48].

Another limitation of this study is the lack of insulin and carbohydrate data. Including this information could enable specific recommendations about insulin and carbohydrate dosages through the day. Including information tracked by smart watches, such as physical activity and stress levels, would not only improve predictions, but provide far more powerful intuitive recommendations. Having contextual information (e.g. high stress levels or even self-reported event markers such as drinking, sickness or exercise) would be critical for empathetic recommendations and reducing burden for the user.

In this work we chose to train hypoglycaemia and hyperglycaemia models using data from all CGM users in our cohort. In practice it may be more suitable to train *individual* models per CGM user, which may be better tailored to the individual. However, it would be more complex to make direct comparisons between relative feature importance for different CGM users, and hence left outside the scope of this paper.

Conclusion

We introduced a framework for high-performance prediction and explanation of hypoglycaemia and hyperglycaemia for young adults. Careful feature selection enables both accurate short-term risk prediction, and intuitive feedback about an individual's blood glucose control. The key benefit of adopting a machine learning framework lies in the ability to provide more accurate ahead-of-time predictions (in comparison to more simplistic derived alerts), potentially reducing burden on the young adult potentially going through transition with their care practices. Combining these models with explanations enables both users and clinicians to gain immediate insight into an individual's blood glucose control, automatically highlighting what specific trends lead to increased risk.

Acknowledgements

We acknowledge funding from UKRI Trustworthy Autonomous Systems Hub (Grant code: RITM0372366).

1. Inchiostro, S., R. Candido, and F. Cavalot, *How can we monitor glycaemic variability in the clinical setting?* Diabetes, Obesity and Metabolism, 2013. **15**(s2): p. 13-16.
2. Leese, G.P., et al., *Frequency of severe hypoglycemia requiring emergency treatment in type 1 and type 2 diabetes: a population-based study of health service resource use.* Diabetes care, 2003. **26**(4): p. 1176-1180.
3. Skriverhaug, T., et al., *Long-term mortality in a nationwide cohort of childhood-onset type 1 diabetic patients in Norway.* Diabetologia, 2006. **49**(2): p. 298-305.
4. Musen, G., et al., *Impact of diabetes and its treatment on cognitive function among adolescents who participated in the Diabetes Control and Complications Trial.* Diabetes care, 2008. **31**(10): p. 1933-1938.
5. Lauretti, E., et al., *Glucose deficit triggers tau pathology and synaptic dysfunction in a tauopathy mouse model.* Translational Psychiatry, 2017. **7**(1): p. e1020-e1020.
6. Collaboration, E.R.F., *Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies.* The Lancet, 2010. **375**(9733): p. 2215-2222.
7. Forbes, J.M. and M.E. Cooper, *Mechanisms of diabetic complications.* Physiological reviews, 2013. **93**(1): p. 137-188.
8. Zhou, B., et al., *Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants.* The Lancet, 2016. **387**(10027): p. 1513-1530.
9. Borus, J.S. and L. Laffel, *Adherence challenges in the management of type 1 diabetes in adolescents: prevention and intervention.* Current opinion in pediatrics, 2010. **22**(4): p. 405.

10. Datye, K.A., et al., *A review of adolescent adherence in type 1 diabetes and the untapped potential of diabetes providers to improve outcomes*. Current Diabetes Reports, 2015. **15**(8): p. 1-9.
11. Clarke, W.L., et al., *Maternal fear of hypoglycemia in their children with insulin dependent diabetes mellitus*. Journal of Pediatric Endocrinology and Metabolism, 1998. **11**(Supplement): p. 189-194.
12. Patton, S.R., et al., *Fear of hypoglycemia in parents of young children with type 1 diabetes mellitus*. Journal of Clinical Psychology in medical settings, 2008. **15**(3): p. 252-259.
13. Haugstvedt, A., et al., *Fear of hypoglycaemia in mothers and fathers of children with Type 1 diabetes is associated with poor glycaemic control and parental emotional distress: a population - based study*. Diabetic Medicine, 2010. **27**(1): p. 72-78.
14. Group, J.D.R.F.C.G.M.S., *Continuous glucose monitoring and intensive treatment of type 1 diabetes*. New England Journal of Medicine, 2008. **359**(14): p. 1464-1476.
15. Group, J.D.R.F.C.G.M.S., *Effectiveness of continuous glucose monitoring in a clinical care environment: evidence from the Juvenile Diabetes Research Foundation continuous glucose monitoring (JDRF-CGM) trial*. Diabetes Care, 2010. **33**(1): p. 17-22.
16. Rodbard, D., *Continuous glucose monitoring: a review of successes, challenges, and opportunities*. Diabetes technology & therapeutics, 2016. **18**(S2): p. S2-3-S2-13.
17. Langendam, M., et al., *Continuous glucose monitoring systems for type 1 diabetes mellitus*. Cochrane Database of Systematic Reviews, 2012(1).
18. Liebl, A., et al., *Continuous glucose monitoring: evidence and consensus statement for clinical use*. Journal of diabetes science and technology, 2013. **7**(2): p. 500-519.
19. Laffel, L.M., et al., *Effect of continuous glucose monitoring on glycemic control in adolescents and young adults with type 1 diabetes: a randomized clinical trial*. JAMA, 2020. **323**(23): p. 2388-2396.
20. Buckingham, B.A., et al., *Evaluation of a predictive low-glucose management system in-clinic*. Diabetes technology & therapeutics, 2017. **19**(5): p. 288-292.
21. Cichosz, S.L., et al., *A novel algorithm for prediction and detection of hypoglycemia based on continuous glucose monitoring and heart rate variability in patients with type 1 diabetes*. Journal of diabetes science and technology, 2014. **8**(4): p. 731-737.
22. van Doorn, W.P., et al., *Machine learning-based glucose prediction with use of continuous glucose and physical activity monitoring data: The Maastricht Study*. PloS one, 2021. **16**(6): p. e0253125.
23. Dave, D., et al., *Feature-based machine learning model for real-time hypoglycemia prediction*. Journal of Diabetes Science and Technology, 2021. **15**(4): p. 842-855.
24. Jensen, M.H., et al., *Prediction of nocturnal hypoglycemia from continuous glucose monitoring data in people with type 1 diabetes: a proof-of-concept study*. Journal of diabetes science and technology, 2020. **14**(2): p. 250-256.
25. Pérez-Gandía, C., et al., *Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring*. Diabetes technology & therapeutics, 2010. **12**(1): p. 81-88.

26. Howsmon, D. and B.W. Bequette, *Hypo-and hyperglycemic alarms: devices and algorithms*. Journal of diabetes science and technology, 2015. **9**(5): p. 1126-1137.
27. Gani, A., et al., *Universal glucose models for predicting subcutaneous glucose concentration in humans*. IEEE Transactions on Information Technology in Biomedicine, 2009. **14**(1): p. 157-165.
28. Vehí, J., et al., *Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning*. Health informatics journal, 2020. **26**(1): p. 703-718.
29. (CDC), C.f.D.C.a.P. *Type 1 Diabetes*. 2021 [cited 2022; Available from: <https://www.cdc.gov/diabetes/basics/what-is-type-1-diabetes.html>].
30. Chen, T. and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016, Association for Computing Machinery: San Francisco, California, USA. p. 785–794.
31. Breiman, L., *Bias, variance, and arcing classifiers*. 1996, Tech. Rep. 460, Statistics Department, University of California, Berkeley
32. Akiba, T., et al. *Optuna: A next-generation hyperparameter optimization framework*. in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*. 2019.
33. Lundberg, S.M., et al., *Explainable AI for trees: From local explanations to global understanding*. arXiv preprint arXiv:1905.04610, 2019.
34. Lundberg, S.M. and S.-I. Lee. *A unified approach to interpreting model predictions*. in *Proceedings of the 31st international conference on neural information processing systems*. 2017.
35. Lundberg, S.M., et al., *Explainable machine learning predictions to help anesthesiologists prevent hypoxemia during surgery*. bioRxiv, 2017: p. 206540.
36. Shapley, L.S., *17. A value for n-person games*. 2016: Princeton University Press.
37. Kuhn, M. and K. Johnson, *Applied predictive modeling*. Vol. 26. 2013: Springer.
38. Rajkomar, A., et al., *Ensuring fairness in machine learning to advance health equity*. Annals of internal medicine, 2018. **169**(12): p. 866-872.
39. Ribeiro, M.T., S. Singh, and C. Guestrin. " *Why should i trust you?*" *Explaining the predictions of any classifier*. in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 2016.
40. Dave, D., et al., *Improved low-glucose predictive alerts based on sustained hypoglycemia: Model development and validation study*. JMIR diabetes, 2021. **6**(2): p. e26909.
41. Forlenza, G.P., et al., *Successful at-home use of the tandem control-IQ artificial pancreas system in young children during a randomized controlled trial*. Diabetes technology & therapeutics, 2019. **21**(4): p. 159-169.
42. Biester, T., et al., " *Let the algorithm do the work*": *reduction of hypoglycemia using sensor-augmented pump therapy with predictive insulin suspension (SmartGuard) in pediatric type 1 diabetes patients*. Diabetes technology & therapeutics, 2017. **19**(3): p. 173-182.
43. Vu, L., et al. *Predicting nocturnal hypoglycemia from continuous glucose monitoring data with extended prediction horizon*. in *AMIA Annual Symposium Proceedings*. 2019. American Medical Informatics Association.

44. Kodama, S., et al., *Ability of current machine learning algorithms to predict and detect hypoglycemia in patients with diabetes mellitus: meta-analysis*. JMIR diabetes, 2021. **6**(1): p. e22458.
45. Deng, Y., et al., *Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients*. NPJ Digital Medicine, 2021. **4**(1): p. 1-13.
46. Polonsky, W.H. and A.L. Fortmann, *Impact of real-time CGM data sharing on quality of life in the caregivers of adults and children with type 1 diabetes*. Journal of Diabetes Science and Technology, 2022. **16**(1): p. 97-105.
47. Polonsky, W.H. and D. Hessler, *What are the quality of life-related benefits and losses associated with real-time continuous glucose monitoring? A survey of current users*. Diabetes technology & therapeutics, 2013. **15**(4): p. 295-301.
48. Schrangl, P., et al., *Limits to the evaluation of the accuracy of continuous glucose monitoring systems by clinical trials*. Biosensors, 2018. **8**(2): p. 50.