

Predicting the risk of inpatient hypoglycemia with machine learning using electronic health records

^{1,2}Yue Ruan*, ^{3,4}Alexis Bellot*, ⁵Zuzana Moysova, ^{1,2}Garry D. Tan, ^{1,2}Alistair Lumb, ⁵Jim Davies, ^{3,4}Mihaela Van Der Schaar, ^{1,2}Rustam Rea

¹Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford University Hospitals NHS Foundation Trust, UK

²Oxford NIHR Biomedical Research Centre, UK

³Department of Mathematics, University of Cambridge, Cambridge, UK

⁴Alan Turing Institute, London, UK

⁵Big Data Institute, University of Oxford Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK.

Running Title

Predicting inpatient hypoglycaemia

Corresponding author:

Dr Rustam Rea

OCDEM, Churchill Hospital, Oxford, OX3 7LE, UK

Tel ++44 7803047223

Email: Rustam.rea@nhs.net

*YR and AB contribute equally to data analysis

Word count:	3725 (<4000)
Abstract:	250
Figures:	1
Tables:	4
Supplemental online material:	1

Abstract

Objective

We analyzed data from inpatients with diabetes admitted to a large university hospital to predict the risk of hypoglycemia through the use of machine learning algorithms.

Research Design and Methods

Four years of data was extracted from a hospital electronic health record system. This included laboratory and point-of-care blood glucose (BG) values to identify biochemical and clinically significant hypoglycaemic episodes ($BG \leq 3.9$ and ≤ 2.9 mmol/L respectively). We used patient demographics, administered medications, vital signs, laboratory results and procedures performed during the hospital stays to inform the model. Two iterations of the dataset included the doses of insulin administered and the past history of inpatient hypoglycaemia. Eighteen different prediction models were compared using the area under curve of the receiver operating characteristics (AUC_ROC) through a ten-fold cross validation.

Results

We analyzed data obtained from 17,658 inpatients with diabetes who underwent 32,758 admissions between July 2014 and August 2018. The predictive factors from the logistic regression model included people undergoing procedures, weight, type of diabetes, oxygen saturation level, use of medications (insulin, sulfonylurea, metformin) and albumin levels. The machine learning model with the best performance was the XGBoost model (AUC_ROC 0.96. This outperformed the logistic regression model which had an AUC_ROC of 0.75 for the estimation of the risk of clinically significant hypoglycaemia.

Conclusions

Advanced machine learning models are superior to logistic regression models in predicting the risk of hypoglycemia in inpatients with diabetes. Trials of such models should be conducted in real time to evaluate their utility to reduce inpatient hypoglycaemia.

Introduction

Hypoglycemia is a common and serious complication affecting people with diabetes [1]. It is an inappropriately low blood glucose which results in significant morbidity in people with Type 1 diabetes and in many people with Type 2 diabetes [2]. A blood glucose level of ≤ 3.9 mmol/l is defined as Level 1 hypoglycemia. A blood glucose level of 2.9 mmol/l and lower is defined as Level 2 hypoglycemia, requiring immediate action as at that level neurogenic and neuroglycopenic symptoms begin to occur [3]. Hypoglycemia can lead to permanent neurological damage if not treated promptly and can be ultimately fatal [1].

Hypoglycemia is an important and common clinical problem under inpatient settings. Retrospective analysis of a US electronic medical records database showed a 20% incidence of hypoglycaemia and 7% incidence of severe hypoglycaemia [4]. There was an associated 66% increase in adjusted inpatient mortality risk and greater than 50% increase in length of hospitalization stay. In a cross-sectional national audit of over 200 hospitals in the UK, the 2017 National Diabetes Inpatient Audit showed that almost one in five people with diabetes experience hypoglycemia during their hospital stay. Although only 7% experience a severe (Level 2) hypoglycemic episode, this rises to 26.9% of all patients with Type 1 diabetes with 185 people over the course of 1 week requiring injectable rescue treatment for their hypoglycemia. Inpatient hypoglycemia has been implicated in the development of adverse clinical and economic outcomes, including increased mortality [5-7], adverse cardiovascular outcomes [8, 9] and increased duration of hospital stay [6, 10, 11].

A recent review article has highlighted the urgent need for evidence-based methodologies to reduce inpatient hypoglycemia. Several strategies have already been developed to predict and prevent the occurrence of inpatient hypoglycemia [12]. One approach to reducing inpatient hypoglycemia is to retrospectively analyze historical clinical data and develop a prediction tool to determine the individualized risk of hypoglycemia during an inpatient

admission. With such a prediction tool, prevention measures can be tested in inpatients with high hypoglycemia risk. The possibility of developing such a prediction tool lies in the growing availability of rich clinical datasets stored in a hospital's electronic patient records (EPR) system.

Previous studies have used clinical information from local healthcare systems to develop inpatient hypoglycemia risk prediction tools [13-15]. In one study, the prediction model developed by the researchers was tested in a clinical trial; this demonstrated the feasibility of using such a model to reduce severe hypoglycemia (glucose less than 40 mg/dl or 2.2 mmol/l) in inpatients with diabetes [14].

However, previous studies have only applied multilinear or logistic regression models on these datasets resulting in only a modest predictive capability of the models. Over the last few years, a number of advanced machine learning techniques have been developed within the field of biomedical engineering [16]. These can be used to create predictive models which can be tested and compared to traditional logistic regression models in order to determine the model with the best predictive performance. This is the first study to assess the performance of novel machine learning models in predicting the risk of inpatient hypoglycemia.

Research Design and Methods

We compared the performance of 18 different machine learning algorithms in predicting the risk of hypoglycemia in inpatients with diabetes.

1) Dataset

The study was approved by the Oxford University Hospitals NHS Foundation Trust Clinical Data Warehouse Programme Board following completion of a Data Protection

Impact Assessment. Data from Oxford University Hospitals NHS Foundation Trust was used which included the Cerner electronic patient record system, the laboratory information management system (LIMS) and the point-of-care testing (POCT) system. All the data used was collected for routine patient management with no additional data input required for the modelling. The dataset contains hospital admission data from 1st September 2014 to 30th June 2018 for qualified patients with diabetes. The qualified patients are defined as meeting the following criteria: (1) being an inpatient as coded in the EPR; (2) having one diagnosis code among E10(insulin-dependent diabetes mellitus), E11(non-insulin-dependent diabetes mellitus), E13(other specified diabetes mellitus), E14(unspecified diabetes mellitus) or O24(diabetes mellitus in pregnancy) as defined in the World Health Organization International Classification of Diseases–10th Revision (ICD-10)[17]; (3) having at least one blood glucose test performed during the hospital admissions. Hospital admission data for qualified patients, including patient demographics, procedures undertaken, diagnosis, laboratory tests, medication administration details and vital signs, were extracted from different source data systems and pooled into a final data table for use by the machine learning prediction models. A schematic representation of the data flow from the source data systems to the final dataset used in the present study is depicted in Figure S1.

2) *Predictors and outcome*

The outcome of interest in the present study is the risk of inpatient hypoglycemia during a hospital admission. We prepared two binary outcome variables $\text{Hypo}^{<4.0}$ and $\text{Hypo}^{<3.0}$ for each hospital admission representing two different severities of the potential hypoglycemic episodes since the degree of hypoglycemia may be influenced by different clinical predictors. We put value 1 to $\text{Hypo}^{<4.0}$ for any Level 1 hypoglycemic episode (any blood glucose measurement $< 4\text{mmol/l}$) and value 1 to $\text{Hypo}^{<3.0}$ for any Level 2 hypoglycemic episode (any blood glucose measurement $< 3\text{mmol/l}$) detected during the hospital admission, and value 0

to the two variables if no hypoglycemic episode detected (all blood glucose measurements > 4mmol/l).

We preprocessed the integrated dataset from the EPR and prepared 42 candidate predictors of interest based on clinical knowledge and previous studies. The predictors cover patient demographics (age, sex etc.), procedures (value 1 for at least one procedure undertaken), laboratory test results (sodium, potassium levels etc.), medication administration details (names and doses of medication delivered including different types of insulin) and vital signs (temperature, heart rate etc.). Additional variables were added to the dataset in order to improve the performance of the prediction algorithm. These included an episode of hypoglycaemia in a previous admission within 6 months. Table 1 shows the full list of predictors and how they were represented in the source data systems and in the prediction models with units of the predictors provided.

3) Prediction models

We evaluated the prediction performance of 18 different machine learning models on the dataset. The models were used to predict the risk of hypoglycaemia (either blood glucose less than 4.0mmol/L (Hypo^{<4.0}) or less than 3.0mmol/L (Hypo^{<3.0})). Forty two different variables were used as inputs into the prediction model. The models cover a wide range of commonly used classification algorithms including logistic regression, random forests, and artificial neural networks which have been previously demonstrated to be robust and applicable to big datasets.

4) Model validation and comparison

For internal model validation, we used a ten-fold cross-validation. We randomly selected nine-tenths of the dataset to be the training dataset (developing the model) and the remaining one-tenth to be used to validate the model.

The outcome variables indicate whether or not a hypoglycaemic episode occurred during an admission. The model was constructed to predict the probability of at least one hypoglycaemic episode occurring. We measured the area under curve of the receiver operating characteristics (AUC-ROC), which shows the probability that the model correctly ranks the risk of hypoglycaemia higher than no hypoglycaemia. The AUC-ROC is not sufficient on its own to use as a hypoglycaemia prognostic model as it does not take into account the prevalence of hypoglycaemia in the population. It assumes that positive and negative predictions are equally important. A detailed technical analysis of the shortcomings of the AUC-ROC was recently conducted by Saito et al [21]. In order to ensure a comprehensive assessment of predictive performance we used additional metrics. We used the terminology TP to represent the number of true positive predictions; FP to represent the number of false positives; TN to represent the number of true negatives; and FN to represent the number of false negatives. We defined precision (positive predictive value) as the ratio $TP/(TP + FP)$. This is a measure of the ability of the model to correctly predict a patient as having hypoglycaemia.. We defined recall (sensitivity) as the ratio $TP/(TP+FN)$. This is a measure of the ability of the model to label as hypoglycemic all of patients who did indeed develop hypoglycaemia. The precision and the recall were calculated for each machine learning model (Table 3).

5) Model development

Following the development of the initial model (IH – Inpatient hypo), two further iterations of the dataset were carried out. In the first change to the dataset, the dose of insulin was added and the variable Intravenous insulin was indicated separately (IH+). In the second change, any previous admission in the last six months containing a low glucose (< 3 and <4 mmol/l) were added (PH – previous hypo). The 18 different machine learning models were

then re-run on each of these new datasets and the model outputs compared to the original dataset. The last column in Table 1 shows which variable is used in each dataset.

6) Variable ranking

We sought to understand how the different variables contributed to the predictions by the XGBoost model (the best predictive model among the 18 tested models). We evaluated the predictive power of each individual variable by providing XGBoost with one variable at a time, and assessing the diagnostic accuracy of the model that it constructed using only that variable. We evaluated the AUC-ROC (using 10-fold cross-validation) in order to get a full picture of each variable's predictive power.

7) Statistical analysis

All statistical analyses were performed using Python 3.6 and R version 3.3. All algorithms were implemented with the machine learning library 'sklearn'[18] that contains all the algorithms and data science utilities used in this report. Internal validation was obtained via 10-fold stratified cross-validation. Performance comparisons were made with t-tests from which we specified p-values less than 0.001 to be considered statistically significant.

Results

1) Baseline characteristics

We analyzed data obtained from 17,658 inpatients with diabetes [9,277 males, age 66(18) years, mean(SD)] who underwent 32,758 hospital admissions between July 2014 and August 2018. We identified all the biochemical (Level 1) and clinically significant (Level 2) hypoglycemic episodes during these admissions. The incidence of biochemical hypoglycemia during a hospital admission was 21.5% and that of clinically significant hypoglycemia was 9.6%. This is in keeping with data from the National Diabetes Inpatient Audit [19].

A selection of the baseline characteristics of the inpatient cohort and the glycemic outcomes are reported in Table 2.

2) *Model performance*

The performance metrics of the machine learning models tested on the PH dataset are presented in Table 3. The AUC_ROC varied between 0.62 and 0.96 for different machine learning models. The estimation performance was better when predicting the risk of $\text{Hypo}^{<3.0}$ compared to predicting that of $\text{Hypo}^{<4.0}$. The AUC_ROC for the logistic regression model was acceptable with 0.73 and 0.75 for estimation of the risk of $\text{Hypo}^{<4.0}$ and $\text{Hypo}^{<3.0}$ respectively. However, the best performing model for predicting the risk of $\text{Hypo}^{<4.0}$ and $\text{Hypo}^{<3.0}$ was the XGBoost model which had the highest AUC_ROC (0.96 for both), the highest precision value (0.88) as well as a high recall value (0.70) among all the models. Figure 1 shows the ROC curves contrasting the logistic regression, gradient boosting, decision tree and XGBoost models. Table S4 shows the normalized confusion matrix for these four models with the true positive and true negative rates on the upper left and lower right in the matrices, respectively. The XGBoost model was again the best performing model with a true positive rate of 0.98 and a true negative rate of 0.71 (Table S4).

3) *Logistic regression model*

Estimated regression coefficients with standard errors and P-values from the logistic regression model with the PH dataset are presented in Table S3. The variables which are significant predictors of hypoglycaemia are shown in Table 4. Similar predictors for $\text{Hypo}^{<4.0}$ and $\text{Hypo}^{<3.0}$ are found. Significant predictors for both $\text{Hypo}^{<4.0}$ and $\text{Hypo}^{<3.0}$ included weight, type of diabetes, oxygen saturation, albumin level, sulfonylurea use, metformin use, intravenous insulin titration, long acting human insulin use, procedures undertaken, and previous hypoglycemic episode.

4) Model development

Two iterations of the original dataset were carried out during the development of the prediction model. Table S1 shows that the addition of a variable to discriminate patients on intravenous insulin compared to those on subcutaneous insulin and a variable for the dose of subcutaneous insulin (IH+ dataset) increased the best performing model (XGBoost) ROC_AUC by 3 percentage points for Hypo^{<4.0}. When the PH dataset was used, all models showed higher AUC values while the XGBoost and gradient boosting models stood out with a significant increase of ROC_AUC by 15 percentage points.

5) Variable ranking

Table S2 demonstrates the relative importance of the variables with the top three most important variables being previous hypoglycemic episodes, albumin levels and type 2 diabetes. A number of novel predictive variables were identified from the machine learning method. These include several vital signs and medications which have logical clinical rationale underlying their importance to hypoglycaemia. Further studies will be required to confirm their importance in the development of hypoglycaemia.

Conclusions

To our knowledge, this is the first study comparing the performance of advanced machine learning models in predicting the risk of inpatient hypoglycemia. With the rich inpatient dataset collected from the EPR system in a large university hospital, the eighteen machine learning models showed high predicting power with an average ROC_AUC at 0.85 for the detection of clinically significant hypoglycemia. The model with the highest ROC_AUC (0.96) was the XGBoost model which outperformed the logistic regression model. This

model performed equally well in predicting both Level 1 hypoglycaemia (blood glucose < 4mmol/l) and Level 2 hypoglycaemia (blood glucose < 3mmol/l).

Among the eighteen evaluated machine learning models, most machine learning prognostic models performed markedly better than the predictions of the linear regression model. XGBoost is a highly flexible nonparametric model which integrates a large number of other machine learning models (decision trees). It was consistently the best performer with the highest AUC_ROC, the highest precision and good recall. It is important to note that XGBoost performed significantly better than the linear regression model which has previously been used to predict inpatient hypoglycaemia. There was an improvement of over 20 percentage points in terms of AUC_ROC for prediction of both biochemical and clinically significant hypoglycemia. The high predictive capability of the XGBoost model also came with a high precision and recall showing low levels of overestimation of risk and low levels of missed events.

With the stepwise iteration of the dataset, the predictive ability of the XGBoost model and some other models improved significantly. This was in contrast to the logistic regression model which showed no significant improvement. This emphasizes the importance of developing a clinically relevant and comprehensive dataset on which to base any machine learning in order to optimize the capability of the learning algorithm.

Our dataset covers a wide range of potential predictors of inpatient hypoglycemia. The logistic regression model detected a number of significant predictors for clinically significant hypoglycemia including weight, type of diabetes, diastolic blood pressure, oxygen saturation, temperature, albumin levels, sulfonylurea use, metformin use, intravenous insulin use, high dose of long-acting human insulin, and people undertaking procedures (see Table S3). Previous studies also found similar predictors such as the albumin levels and glucose-lowering drugs[13, 15, 20]. In our dataset, undertaking any kind of procedure was also found

to be a significant predictor. This is clinically understandable as procedures may disrupt the daily in hospital routine of food intake and drug administration and thereby cause increased variability in glycemic levels.

Our machine learning models outperform other inpatient hypoglycaemia prediction models which have been published using logistic or multivariate regression techniques. These have shown a discrimination of between 0.70 and 0.80 [15, 20]. Our model development also includes a variable for previous hypoglycaemia which has not previously been included in other prediction models.

Machine learning models have been widely used within hospital information systems to predict the risk of emergency admission, sepsis in the intensive care unit and identifying type 2 diabetes using electronic health records[21-23]. The performance of the present study also compares favourably to these other predictive models which have an AUC_ROC of between 0.75 and 0.85.

There are several key strengths of the present study. Firstly, we evaluated a wide range of machine learning models and compared their predicting ability against the most commonly used statistical model which we used as a benchmark model. Secondly, we used an iterative approach to develop the model with additional variables which revealed how significant improvements to the model could be achieved. Thirdly, this was the largest dataset used to predict inpatient hypoglycaemia containing the data for 32,758 hospital admissions. It also integrated clinical information which previous studies have not considered before such as the medication dosage information and hypoglycaemia in previous admissions.

However, as with all modelling efforts, there are also limitations. One limitation with the dataset is the unavailability of carbohydrate intake/meal content information from the EPR during the hospital admissions. The carbohydrate intake has a direct impact on the postprandial glucose excursions and consequently the prandial insulin doses titrated to

individuals and could be an important predictor for hypoglycemic events. This information is unavailable as it is not routinely recorded electronically in hospitals. Secondly, the current electronic health record does not record the level of hypoglycaemia unawareness, prior continuous glucose monitoring data or prior self-monitoring of blood glucose. This data is likely to be an important factor in developing hypoglycaemia while in hospital, although for acutely unwell patients this data may not be directly applicable. Thirdly, our dataset is derived from a single organisation and the generalizability of our best performing machine learning model needs to be tested in other datasets and evaluated in other centres. Finally, although we have developed a highly predictive model for inpatient hypoglycaemia, the feasibility of using this model needs to be tested within a live electronic patient record in order to confirm the ability of the model to receive data in real time and ensure that the model performs as strongly as the current data suggests.

One of the advantages of the current prediction model is that it uses variables that are readily accessible within the EPR. As a result, the model can be integrated into a decision support system under the EPR framework. In practice, the decision support system would access the clinical information of a new inpatient, feed the required information to the prediction model which would then calculate the risk of the patient experiencing either biochemical or clinically significant hypoglycaemia during his or her hospital admission. This would enable the decision support system to suggest appropriate treatment options based on individual risk levels, thereby reducing hypoglycaemia and its consequent associated morbidity and potentially reduce the economic burden of prolonged hospital stay due to hypoglycemia. In a previous single centre study, a linear regression prediction model with a sensitivity of 50% and specificity of 71% was used to detect patients at risk of hypoglycaemia. Clinician education resulted in medication change in 40% of patients, and a reduction of 68% in the rate of severe hypoglycaemia in alerted high-risk patients versus non-

alerted high-risk patients [14]. The benefits of a significantly more powerful prediction model based on machine learning need to be evaluated in a large multi-centre randomised controlled trial.

Another potential application of the prediction model is the selection of high risk patients who may benefit from an advanced treatment option such as continuous glucose monitoring (CGM) or closed-loop insulin delivery. The use of continuous glucose monitoring in the inpatient setting was considered at a recent symposium where the trials in both the ICU and non-ICU settings were reviewed [24]. While there was some evidence that CGM may reduce the rates of severe hypoglycaemia, it was recognised that there was limited data on clinical outcomes and that such technology may be most suitable for ‘populations.. at high risk for glucose variability and hypoglycaemia’. Closed-loop insulin delivery, or artificial pancreas, is a novel treatment option for people with diabetes who require exogenous insulin administration[25]. The system titrates insulin based on real-time glucose monitoring and a titration algorithm. Previous clinical studies have shown promise glycemic results of the closed-loop systems under inpatient settings[26, 27]. However, the system is costly and cannot be applied in every inpatient with diabetes. Prediction models such as the one described in this study could be used as a pre-selection tool to determine which patients would benefit the most from CGM or automated insulin delivery.

In conclusion, this study demonstrates for the first time, the utility of advanced machine learning models in predicting the risk of hypoglycemia in inpatients with diabetes. We have shown that these models are significantly better in predicting inpatient hypoglycaemia than the traditional logistic regression model. However further trials are needed to determine if this prediction model provides a significant clinical advantage over traditional logistic regression analysis or more simple risk factor prediction models eg insulin use alone. Such machine learning models need to be evaluated within a real-time clinical setting to

demonstrate their ability to predict hypoglycaemia following admission. The use of new technological methods such as machine learning and artificial intelligence are not a substitute for clinicians, but should be used to enhance clinical judgement and support everyday decisions. Multi-centre clinical trials are now needed to evaluate their utility within a clinical decision support system and reduce the burden of hypoglycaemia in hospital.

Acknowledgements

YR is supported by a Novo Nordisk Postdoctoral Fellowship run in partnership with the University of Oxford. The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Guarantor statement:

RR and YR are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis

Author Contributions:

YR and RR co-designed the study analysis. YR and AB carried out the data analysis. YR and RR drafted the manuscript. All authors contributed to the interpretation of the results and critical review of the paper.

Conflict of Interest

For all the authors listed in this paper – No potential conflicts of interest relevant to this article were reported

Funding Statement

YR salary was funded by Novo Nordisk Postdoctoral Fellowship run in partnership with the University of Oxford.

RR, JD, GT, AL are partly funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC).

References

1. Cryer PE, Davis SN, Shamon H. Hypoglycemia in diabetes. *Diabetes Care*. 2003;26(6):1902-12. PubMed PMID: 12766131.
2. Group UKHS. Risk of hypoglycaemia in types 1 and 2 diabetes: effects of treatment modalities and their duration. *Diabetologia*. 2007;50(6):1140-7. doi: 10.1007/s00125-007-0599-y. PubMed PMID: 17415551.
3. Agiostratidou G, Anhalt H, Ball D, Blonde L, Gourgari E, Harriman KN, et al. Standardizing Clinically Meaningful Outcome Measures Beyond HbA1c for Type 1 Diabetes: A Consensus Report of the American Association of Clinical Endocrinologists, the American Association of Diabetes Educators, the American Diabetes Association, the Endocrine Society, JDRF International, The Leona M. and Harry B. Helmsley Charitable Trust, the Pediatric Endocrine Society, and the T1D Exchange. *Diabetes Care*. 2017;40(12):1622-30. doi: 10.2337/dc17-1624. PubMed PMID: 29162582; PubMed Central PMCID: PMC5864122.
4. Brodovicz KG, Mehta V, Zhang Q, Zhao C, Davies MJ, Chen J, et al. Association between hypoglycemia and inpatient mortality and length of hospital stay in hospitalized, insulin-treated patients. *Curr Med Res Opin*. 2013;29(2):101-7. doi: 10.1185/03007995.2012.754744. PubMed PMID: 23198978.

5. Borzi V, Fontanella A. The clinical impact of hypoglycemia in hospitalized patients. *Italian Journal of Medicine*. 2015;9(1):11-9. doi: <http://dx.doi.org/10.4081/itjm.2015.549>. PubMed PMID: 603449345.
6. Gomez-Huelgas R, Guijarro-Merino R, Zapatero A, Barba R, Guijarro-Contreras A, Tinahones F, et al. The frequency and impact of hypoglycemia among hospitalized patients with diabetes: A population-based study. *J Diabetes Complications*. 2015;29(8):1050-5. doi: <https://dx.doi.org/10.1016/j.jdiacomp.2015.07.018>. PubMed PMID: 26279321.
7. Turchin A, Matheny ME, Shubina M, Scanlon SV, Greenwood B, Pendergrass ML. Hypoglycemia and clinical outcomes in patients with diabetes hospitalized in the general ward. *Diabetes Care*. 2009;32(7):1153-7. doi: <http://dx.doi.org/10.2337/dc08-2127>. PubMed PMID: 354871408.
8. Akhavan P, Aghili R, Malek M, Ebrahim Valojerdi A, Khamseh ME. Hypoglycemia: Adverse Cardiovascular Outcomes in Non-Critically Ill People with Type 2 Diabetes. *Arch Iran Med*. 2016;19(2):82-6. doi: 0161902/AIM.003. PubMed PMID: 26838076.
9. Carey M, Boucai L, Zonszein J. Impact of hypoglycemia in hospitalized patients. *Current Diabetes Reports*. 2013;13(1):107-13. doi: <https://dx.doi.org/10.1007/s11892-012-0336-x>. PubMed PMID: 23065370.
10. Evans M, Wolden ML, Thorsted BL, McEwan PC, Jacobsen JL. Inpatient hypoglycaemia increases length of hospital stay and all-cause mortality risk. *Diabetic Medicine*. 2015;32:23. doi: <http://dx.doi.org/10.1111/dme.12665> 12. PubMed PMID: 71820881.
11. Nirantharakumar K, Marshall T, Kennedy A, Narendran P, Hemming K, Coleman JJ. Hypoglycaemia is associated with increased length of stay and mortality in people with diabetes who are hospitalized. *Diabet Med*. 2012;29(12):e445-8. doi: 10.1111/dme.12002. PubMed PMID: 22937877.
12. Ruan Y, Tan GD, Lumb A, Rea RD. Importance of inpatient hypoglycaemia: impact, prediction and prevention. *Diabet Med*. 2019;36(4):434-43. doi: 10.1111/dme.13897. PubMed PMID: 30653706.
13. Stuart K, Adderley NJ, Marshall T, Rayman G, Sitch A, Manley S, et al. Predicting inpatient hypoglycaemia in hospitalized patients with diabetes: a retrospective analysis of 9584 admissions with diabetes. *Diabet Med*. 2017;34(10):1385-91. doi: 10.1111/dme.13409. PubMed PMID: 28632918.
14. Kilpatrick CR, Elliott MB, Pratt E, Schafers SJ, Blackburn MC, Heard K, et al. Prevention of inpatient hypoglycemia with a real-time informatics alert. *J Hosp Med*. 2014;9(10):621-6. doi: 10.1002/jhm.2221. PubMed PMID: 24898687.
15. Mathioudakis NN, Everett E, Routh S, Pronovost PJ, Yeh HC, Golden SH, et al. Development and validation of a prediction model for insulin-associated hypoglycemia in non-critically ill hospitalized adults. *BMJ Open Diabetes Res Care*. 2018;6(1):e000499. doi: 10.1136/bmjdr-2017-000499. PubMed PMID: 29527311; PubMed Central PMCID: PMC5841507.
16. Park C, Took CC, Seong JK. Machine learning in biomedical engineering. *Biomed Eng Lett*. 2018;8(1):1-3. doi: 10.1007/s13534-018-0058-3. PubMed PMID: 30603186; PubMed Central PMCID: PMC6208556.
17. WHO. World Health Organization. ICD-10 version:2010. 2010 [cited 2019 Mar 20] <https://icdwho.int/browse10/2010/en-!G35>. 2010.
18. Pedregosa. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-30.

19. NHS. <https://files.digital.nhs.uk/pdf/s/7/nadia-17-rep.pdf> 2017.
20. Elliott MB, Schafers SJ, McGill JB, Tobin GS. Prediction and prevention of treatment-related inpatient hypoglycemia. *Journal of Diabetes Science & Technology*. 2012;6(2):302-9. doi: <https://dx.doi.org/10.1177/193229681200600213>. PubMed PMID: 22538139.
21. Rahimian F, Salimi-Khorshidi G, Payberah AH, Tran J, Ayala Solares R, Raimondi F, et al. Predicting the risk of emergency admission with machine learning: Development and validation using linked electronic health records. *PLoS Med*. 2018;15(11):e1002695. doi: 10.1371/journal.pmed.1002695. PubMed PMID: 30458006; PubMed Central PMCID: PMCPMC6245681 following competing interests: JT receives funding for DPhil provided by Rhodes Trust and Clarendon Fund, is Chair on board of CHASE (incorporated association), travel grant from European Society of Hypertension, British Research Council training grant, Special Consultant for Bendelta. KR receives a stipend as a specialty consulting editor for PLOS Medicine and serves on the journal's editorial board.
22. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach. *JMIR Med Inform*. 2016;4(3):e28. doi: 10.2196/medinform.5909. PubMed PMID: 27694098; PubMed Central PMCID: PMCPMC5065680.
23. Zheng T, Xie W, Xu L, He X, Zhang Y, You M, et al. A machine learning-based framework to identify type 2 diabetes through electronic health records. *Int J Med Inform*. 2017;97:120-7. doi: 10.1016/j.ijmedinf.2016.09.014. PubMed PMID: 27919371; PubMed Central PMCID: PMCPMC5144921.
24. Wallia A, Umpierrez GE, Rushakoff RJ, Klonoff DC, Rubin DJ, Hill Golden S, et al. Consensus Statement on Inpatient Use of Continuous Glucose Monitoring. *J Diabetes Sci Technol*. 2017;11(5):1036-44. doi: 10.1177/1932296817706151. PubMed PMID: 28429611; PubMed Central PMCID: PMCPMC5950996.
25. Boughton CK, Hovorka R. Advances in artificial pancreas systems. *Sci Transl Med*. 2019;11(484). doi: 10.1126/scitranslmed.aaw4949. PubMed PMID: 30894501.
26. Bally L, Thabit H, Hartnell S, Andereggen E, Ruan Y, Wilinska ME, et al. Closed-Loop Insulin Delivery for Glycemic Control in Noncritical Care. *N Engl J Med*. 2018;379(6):547-56. doi: 10.1056/NEJMoa1805233. PubMed PMID: 29940126.
27. Thabit H, Hartnell S, Allen JM, Lake A, Wilinska ME, Ruan Y, et al. Closed-loop insulin delivery in inpatients with type 2 diabetes: a randomised, parallel-group trial. *Lancet Diabetes Endocrinol*. 2017;5(2):117-24. doi: 10.1016/S2213-8587(16)30280-7. PubMed PMID: 27836235.

Category	Predictor	Data in EPR	Data in models	Unit	Completeness (%)	Datasets
Demographics	Age	Year of birth	Computed based on the year of admission	years	100	IH
	Sex	Male/female	Binary variable (1/0)	NA	100	IH
	Ethnicity	Ethnicity (categorical value)	Categorical variable (white British, African etc.)	NA	100	IH
	Weight	Weight measured at time of admission	Weight value	kg	71	IH
	Height	Height measured at time of admission	Height value	cm	59	IH
	Type of diabetes	Type of diabetes (categorical value)	Categorical variable (T1D/T2D/other)	NA	100	IH
Vital signs	Diastolic blood pressure	Multiple measurements	Average value throughout the admission	mm Hg	73	IH
	Systolic blood pressure	Multiple measurements	Average value throughout the admission	mm Hg	73	IH
	Heart rate	Multiple measurements	Average value throughout the admission	/min	71	IH
	Oxygen saturation	Multiple measurements	Average value throughout the admission	%	73	IH
	Temperature	Multiple measurements	Average value throughout the admission	Celsius	72	IH
Laboratory tests	Albumin	Multiple measurements	Average value throughout the admission	g/L	81	IH
	Amylase	Multiple measurements	Average value throughout the admission	IU/L	15	IH
	C-peptide	Multiple measurements	Average value throughout the admission	pmol/L	17	IH
	Cortisol	Multiple measurements	Average value throughout the admission	nmol/L	26	IH
	Creatinine	Multiple measurements	Average value throughout the admission	umol/L	80	IH
	C-Reactive Protein	Multiple measurements	Average value throughout the admission	mg/L	78	IH
	eGFR	Multiple measurements	Average value throughout the admission	ml/min/1.73m ²	80	IH
	Hemoglobin	Multiple measurements	Average value throughout the admission	g/L	80	IH
	HbA1c	Multiple measurements	Average value throughout the admission	%	42	IH
	Potassium	Multiple measurements	Average value throughout the admission	mmol/L	80	IH
	Sodium	Multiple measurements	Average value throughout the admission	mmol/L	80	IH
	White cells	Multiple measurements	Average value throughout the admission	x10 ⁹ /L	79	IH
	Sulfonylurea	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	DPP-4	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
Medications	GLP-1	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Metformin	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Morphine	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Pioglitazone	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Bisoprolol	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Amitriptyline	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Pregabalin	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Dexamethasone	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Prednisolone	Drug dose and time	Binary variable (1 for on drug and 0 for not)	NA	100	IH
	Intravenous insulin	Multiple rates of insulin infusion	Binary variable (1 for on IV insulin and 0 for not)	NA	100	IH+
	Insulin (rapid-acting analogue)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
	Insulin (mixed analogue)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
	Insulin (long-acting analogue)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
	Insulin (short-acting human)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
	Insulin (mixed human)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
	Insulin (intermediate-acting human)	Multiple doses of different amount	Average total daily insulin dose	unit	100	IH+
Procedures	Procedure indication	Procedure name and time	Binary variable (1 for had at least one procedure during the admission and 0 for not)	NA	100	IH+
Previous hypoglycemia	Previous biochemical hypoglycemia	Blood glucose measurements	Binary variable (1 for had at least one blood glucose < 4mmol/l)	NA	63	PH
	Previous clinically significant hypoglycemia	Blood glucose measurements	Binary variable (1 for had at least one blood glucose < 3mmol/l)	NA	63	PH

Insulin (rapid acting analogue): "Insulin aspart", "Insulin lispro", "Insulin glulisine", "Insulin faster acting aspart"

Insulin (mixed analogue): "Insulin aspart biphasic (Novomix 30)", "Insulin lispro biphasic (Humalog Mix 25 and Humalog Mix 50)"

Insulin (long acting analogue): "Insulin glargine", "Insulin detemir", "Insulin degludec"

Insulin (short acting human): "Insulin Actrapid", "Insulin Humulin S"

Insulin (mixed human): "Insulin Humulin M3"

Insulin (intermediate acting human): "Insulin Insulatard", "Insulin Humulin I"

Table 2. Baseline characteristics and glycemc outcomes of the inpatients cohorts

Predictors	Inpatients with diabetes (N = 17,658) Inpatient hospital admissions (n = 32,758)	
Sex, N(%)		
	Female	8,381(47)
	Male	9,277(53)
Age, mean(SD)		66(18)
Ethnicity, N(%)		
	White British	12,511(70.8)
	African	116(0.7)
	Pakistani	331(1.9)
	Chinese	53(0.3)
	Indian	254(1.4)
	Not stated	2869(16.2)
	Other	1524(8.6)
Type of diabetes, N(%)		
	Insulin-dependent diabetes	1696(9.6)
	Non-insulin-dependent diabetes	14006(79.3)
	Other forms	1956(11.1)
Systolic blood pressure, mean(SD)		132.5(18.2)
eGFR, mean(SD)		29.8(6.4)
Hemoglobin, mean(SD)		29.9(6.4)
Medication use		
Sulfonylurea, n(%)		6,435(19.6)
DPP-4, n(%)		1,415(4.3)
GLP-1, n(%)		349(1.1)
Metformin, n(%)		10,756(32.8)
Insulin, n(%)		
	Intravenous insulin	4,678(14.3)
	Rapid-acting analogue	3,954(12.1)
	Mixed-acting analogue	1,553(4.7)
	Long-acting analogue	5,118(15.6)
	Short-acting human	3,561(10.9)
	Mixed-acting human	1,388(4.2)
	Intermediate-acting human	2,394(7.3)
Procedures, n(%)		22,931(70.0)
Glycemc outcomes		
Hypoglycemia, n(%)		
	Biochemical hypoglycemia	7,030(21.5)
	Clinically significant hypoglycemia	3,154(9.6)
Blood glucose level, mean(SD)		10.1(4.7)

N(%), number of patients and percentage over the total number of patients; n(%), number of admissions and percentage over the total number of admissions

Table 3. Performance metrics of the machine learning models based on the PH dataset

Machine learning algorithm	Biochemical hypoglycemia (blood glucose < 4 mmol/l)			Clinically significant hypoglycemia (blood glucose <3 mmol/l)		
	AUC_ROC	Precision	Recall	AUC_ROC	Precision	Recall
Logistic regression	0.73	0.48	0.10	0.75	0.39	0.10
SGD	0.74	0.12	0.10	0.77	0.10	0.10
kNN	0.62	0.40	0.18	0.62	0.30	0.15
Decision tree	0.81	0.70	0.71	0.84	0.68	0.73
Gaussian naïve Bayes	0.81	0.47	0.68	0.86	0.33	0.81
Bernoulli naïve Bayes	0.82	0.60	0.60	0.86	0.47	0.67
Multinomial naïve Bayes	0.75	0.10	0.10	0.79	0.10	0.10
SVM	0.79	0.73	0.10	0.83	0.41	0.10
QDA	0.77	0.23	0.96	0.89	0.15	0.97
Random forest	0.94	0.86	0.67	0.93	0.96	0.66
Extra trees	0.93	0.85	0.68	0.93	0.94	0.66
LDA	0.88	0.69	0.75	0.90	0.72	0.72
Passive aggressive	0.76	0.46	0.25	0.77	0.33	0.10
AdaBoost	0.89	0.68	0.60	0.93	0.63	0.46
Bagging	0.93	0.84	0.70	0.92	0.93	0.67
Gradient boosting	0.96	0.87	0.70	0.96	0.96	0.67
XGBoost	0.96	0.88	0.70	0.96	0.97	0.67
MLP	0.74	0.57	0.17	0.78	0.47	0.14
Mean (SD)	0.82(0.10)	0.59(0.25)	0.49(0.29)	0.85(0.10)	0.55(0.31)	0.48(0.31)

Abbreviations: SGD: stochastic gradient descent, kNN: k nearest neighbor, SVM: support vector machine, QDA: quadratic discriminant analysis, LDA: linear discriminant analysis, MLP: multilayer perceptron (artificial neural network).

Table 4. Most significant predictors from the logistic regression model

Predictors	Biochemical hypoglycemia (blood glucose < 4 mmol/l)			Clinically significant hypoglycemia (blood glucose <3 mmol/l)		
	Coefficient	p-value	z-score	Coefficient	p-value	z-score
PrevLowGlucose3 (+)	3.842	<0.001	28.42	4.021	<0.001	20.39
Albumin level (-)	-0.078	<0.001	-27.22	-0.074	<0.001	-19.51
Intravenous insulin (+)	0.639	<0.001	15.43	0.501	<0.001	9.82
Procedure indication (+)	0.485	<0.001	14.87	0.339	<0.001	6.81
Sulfonylurea (+)	0.572	<0.001	14.24	0.311	<0.001	5.35
Type 2 diabetes (-)	-0.820	<0.001	-13.68	-0.656	<0.001	-7.88
Weight (-)	-0.010	<0.001	-7.42	-0.012	<0.001	-6.38
Oxygen saturation (+)	0.059	<0.001	6.31	0.067	<0.001	5.30
Metformin (-)	-0.212	<0.001	-6.02	-0.258	<0.001	-5.02
Long-acting human insulin (+)	0.011	<0.001	4.77	0.010	<0.001	5.35
Rapid-acting human insulin (+)	0.023	<0.001	4.11		NS	
Mixed insulin analogue (+)	0.007	<0.001	3.67		NS	

NS: not significant

Factors with a positive Coefficient value increase the risk of hypoglycaemia and factors with a negative Coefficient value decrease the risk of hypoglycaemia. The factors are listed in order of effect size on the logistic regression model. Eg an increase in albumin value reduces the risk of hypoglycaemia, people with Type 2 diabetes have an decreased risk of hypoglycaemia. A (+) or (-) sign is given to each of the factors to indicate the effect direction.

Figure 1. ROC curves for logistic regression, XGBoost and decision tree model when predicting biochemical hypoglycemia.