Received: 28 November 2022 | Revised: 30 January 2023 | Accepted: 12 February 2023

DOI: 10.1111/dom.15021

ORIGINAL ARTICLE

WILEY

Machine learning for non-invasive sensing of hypoglycaemia while driving in people with diabetes

Vera Lehmann MD[1] | Thomas Zueger MD[1,2] | Martin Maritsch PhD[2] | Mathias Kraus PhD[2,3] | Caroline Albrecht BSc[1] | Caterina Bérubé MSc[2] | Stefan Feuerriegel PhD[4] | Felix Wortmann PhD[5] | Tobias Kowatsch PhD[2,6,7] | Naïma Styger BSc[1] | Sophie Lagger MSc[1] | Markus Laimer MD[1] | Elgar Fleisch PhD[2,5] | Christoph Stettler MD[1]

[1]Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[2]Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

[3]School of Business, Economics and Society, Friedrich-Alexander University Erlangen-Nürnberg, Nürnberg, Germany

[4]Institute of AI in Management, LMU Munich, Munich, Germany

[5]Institute of Technology Management, University of St. Gallen, St. Gallen, Switzerland

[6]Institute for Implementation Science in Health Care, University of Zurich, Zurich, Switzerland

[7]School of Medicine, University of St. Gallen, St. Gallen, Switzerland

Correspondence
Christoph Stettler, MD, Department of Diabetes, Endocrinology, Nutritional Medicine and Metabolism, Inselspital, Bern University Hospital, University of Bern, CH-3010, Bern, Switzerland.
Email: christoph.stettler@insel.ch

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: Sinergia grant SNF CRSII5_183569

Abstract

Aim: To develop and evaluate the concept of a non-invasive machine learning (ML) approach for detecting hypoglycaemia based exclusively on combined driving (CAN) and eye tracking (ET) data.

Materials and Methods: We first developed and tested our ML approach in pronounced hypoglycaemia, and then we applied it to mild hypoglycaemia to evaluate its early warning potential. For this, we conducted two consecutive, interventional studies in individuals with type 1 diabetes. In study 1 (n = 18), we collected CAN and ET data in a driving simulator during euglycaemia and pronounced hypoglycaemia (blood glucose [BG] 2.0-2.5 mmol L^{-1}). In study 2 (n = 9), we collected CAN and ET data in the same simulator but in euglycaemia and mild hypoglycaemia (BG 3.0-3.5 mmol L^{-1}).

Results: Here, we show that our ML approach detects pronounced and mild hypoglycaemia with high accuracy (area under the receiver operating characteristics curve 0.88 ± 0.10 and 0.83 ± 0.11, respectively).

Conclusions: Our findings suggest that an ML approach based on CAN and ET data, exclusively, enables detection of hypoglycaemia while driving. This provides a promising concept for alternative and non-invasive detection of hypoglycaemia.

KEYWORDS
diabetes complications, hypoglycaemia, type 1 diabetes, glycaemic control

* Vera Lehmann, Thomas Zueger, Martin Maritsch, and Mathias Kraus have equal first authorship.

† Elgar Fleisch and Christoph Stettler have equal last authorship.

1 | INTRODUCTION

Hypoglycaemia is a dangerous acute complication of diabetes[1,2] associated with impairments of cognitive, executive and psychomotor functions,[3–5] thereby interfering with the performance of many everyday activities, including driving. Despite ongoing and important developments in diabetes treatment, hypoglycaemia is responsible for a substantial and increasing number of driving accidents.[6–9] While intermittent self-monitoring of capillary blood glucose (SMBG) is still the standard in many countries, continuous glucose monitoring (CGM) offers the advantage of permanent glucose control. However, CGM is limited by invasiveness, availability and costs, and is subject to an inherent time lag in hypoglycaemia.[10] Of note, in a recent prospective study, individuals with type 1 diabetes spent a considerable amount of time in hypoglycaemia while driving,[11] corroborating the need for alternative and complementary methods to detect hypoglycaemia while driving. Here, we develop a machine learning (ML) approach to detect hypoglycaemia exclusively from driving and gaze behaviour. There is a growing body of evidence examining hypoglycaemia prediction algorithms based on physiological, nutritional, insulin and/or CGM data.[12,13] However, to the best of our knowledge, no study has so far aimed to detect hypoglycaemia using ML methodology based on driving and gaze behaviour data.

Cars permanently generate a broad spectrum of granular real-time information on various driving features, transmitted via the Controller Area Network (CAN) bus. Additionally, cameras are increasingly installed in modern vehicles[14] to monitor driver behaviour and vigilance, also in (semi-)autonomous driving situations. A hypoglycaemia warning system based on CAN and eye tracking (ET) data could provide a non-invasive, complementary and scalable approach to reduce accidents in people with diabetes. In this article, we present the concept of a ML approach using CAN and ET data to detect hypoglycaemia during driving.

2 | MATERIALS AND METHODS

2.1 | Study design and population

We conducted two non-randomized, interventional studies in individuals with type 1 diabetes from October 2019 to July 2020 (study 1), and from November 2021 to March 2022 (study 2). We included active drivers aged 21-50 years (up to 60 years for study 2). Key exclusion criteria included motion sickness, pregnancy or breastfeeding, severe organ dysfunction, alcohol or drug abuse, and medication known to interfere with driving performance (e.g. sedatives, opioids). The eligibility criteria are listed in the supporting information (Appendix S1, Supplementary Methods). The studies were conducted at the University Hospital of Bern in collaboration with the ETH Zurich, and the University of St. Gallen, following the Declaration of Helsinki, the guidelines of good clinical practice, Swiss health laws, and the ordinance on clinical research. Each participant gave informed written consent. Both studies were approved by the local ethics committee Bern, Switzerland (2019-00579, 2021-002018) and were registered on ClinicalTrials.gov (NCT04035993 and NCT05183191).

2.2 | Study procedure

Figure S1A depicts the visit schedule. After screening, participants familiarized themselves with the driving simulator during a test drive. Participants not capable of driving with the simulator (e.g. because of motion sickness) were excluded. Participants were fitted with the Dexcom G6 CGM system. Participants were instructed to refrain from alcohol, caffeine and strenuous physical activity for 24 hours before the main visit. The main visit was postponed if sensor glucose was less than 3.0 mmol L^{-1} for longer than 30 minutes in the preceding 24 hours.

For the main visit, participants were admitted to our clinical research unit after an overnight fast. During a controlled hypoglycaemia procedure, participants drove in euglycaemia and hypoglycaemia (Figure 1A) on a designated circuit using the driving simulator (Figure 1B), while CAN and ET data were recorded. We used a well-established driving simulator (Carnetsoft BV, Groningen, the Netherlands) as in previous studies on driving behaviour.[15–17] Eye gaze was recorded with a consumer eye tracker (Tobii Eye Tracker 4C, Tobii AB, Danderyd, Sweden). The intended BG range in hypoglycaemia was 2.0-2.5 mmol L^{-1} (in study 1) and 3.0-3.5 mmol L^{-1} (in study 2) (Figure 1C,D). In both studies, each driving session in euglycaemia and hypoglycaemia consisted of three environments (highway, rural and urban) completed in random order. The driving lasted 5 minutes in each environment and was separated by 1-2–minute breaks for intermittent BG measurement using the Biosen C-Line glucose analyser (EKF Diagnostics Holdings PLC, Penarth, Cardiff, UK). Participants were informed that a hypoglycaemic state aiming at a BG level of 2.0-2.5 mmol L^{-1} (study 1) or 3.0-3.5 mmol L^{-1} (study 2) was to be induced, but they were blinded to the BG values throughout the experiment. In euglycaemia and hypoglycaemia, participants rated eight hypoglycaemic symptoms, 'need-to-treat right now' and 'difficulty driving' on a seven-point scale (0 = none, 6 = extreme).[18] In addition, participants guessed their BG level (in mmol L^{-1}, one decimal place). Participants could abandon study procedures at any time point if they felt that the situation was unacceptable to them. After data collection and restoration of euglycaemia, the procedure was terminated if deemed safe by the investigator. A detailed description of the controlled hypoglycaemic state, the driving simulator, the eye tracker and the driving environments is provided in in the supporting information (Appendix S1, Supplementary Methods).

One to 3 days after the main visit, participants were scheduled for the close-out visit, including a safety assessment.

2.3 | Outcome and sample size calculation

The main outcome was the diagnostic accuracy of our ML approach to detect hypoglycaemia, quantified as the area under the receiver operating characteristic curve (AUROC). Traditional null hypothesis

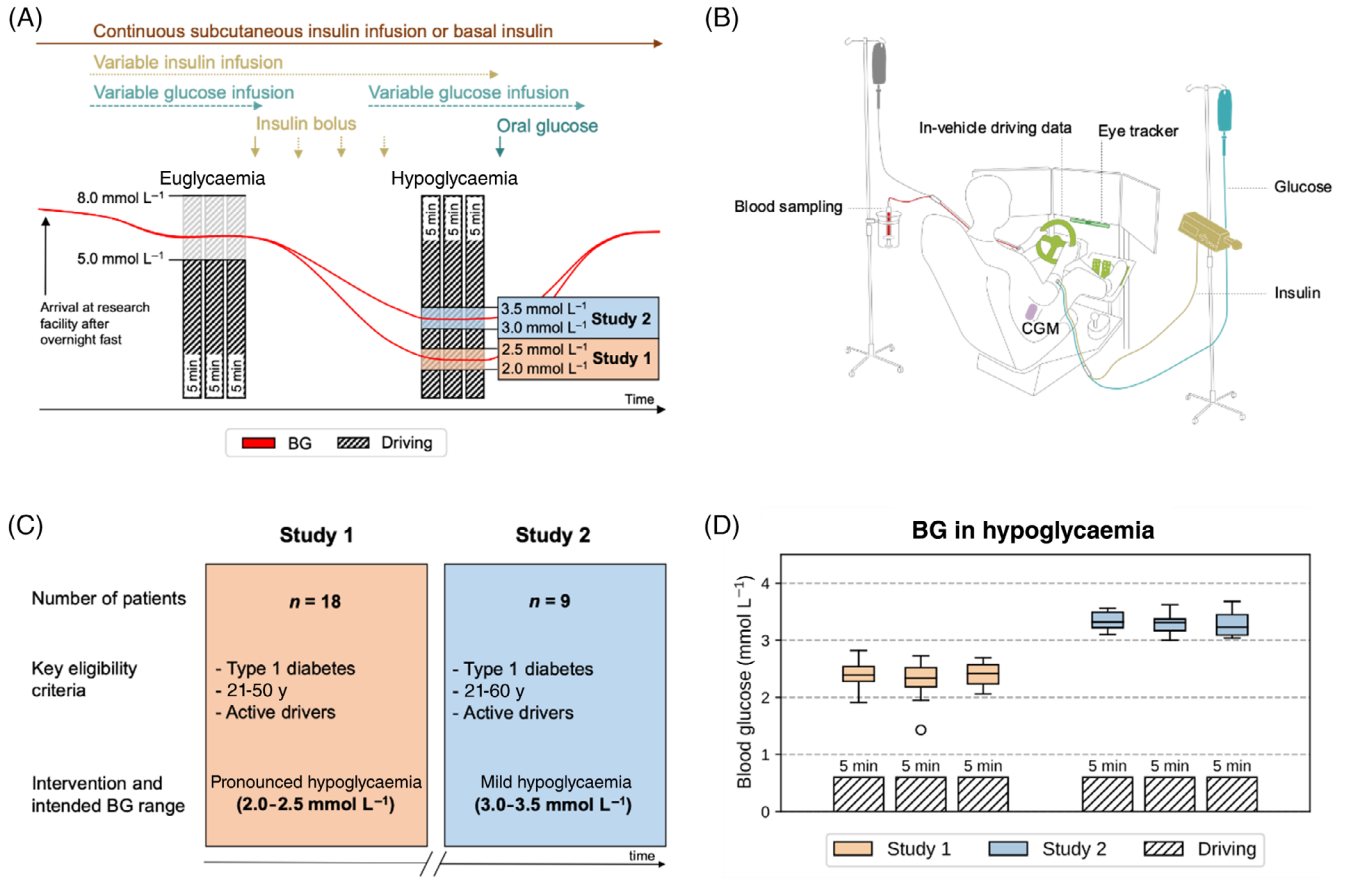

FIGURE 1 Overview. A, Hypoglycaemia induction procedures for study 1 and study 2 using variable insulin aspart and glucose administration with corresponding driving sessions in euglycaemia and hypoglycaemia. The intended range for blood glucose (BG) in hypoglycaemia was 2.0-2.5 mmol L^{-1} in study 1, and 3.0-3.5 mmol L^{-1} in study 2. Driving sessions consisted of three 5-min drives in three different environments (highway, rural and urban), while in-vehicle driving (CAN) and eye tracking (ET) data were collected. B, Driving simulator, ET and glucose management set-up in both studies. C, Key characteristics of study 1 and study 2. D, Venous BG in hypoglycaemia for study 1 and study 2 shown as boxplots. Overall, BG in hypoglycaemia was stable across both studies. The line within the box of the boxplot shows the median, the inner bounds of the box correspond to the interquartile range (IQR = 25th to 75th percentiles) and the outer bounds (i.e. whiskers) correspond to the most extreme data points no more than 1.5 × IQR from the edge of the box. Values outsides the whisker range are illustrated by dots. CGM, continuous glucose monitoring

testing that lends itself to power calculation was not applicable to our study (i.e. there is no null hypothesis for the development of ML models). Therefore, we implemented an established methodology from a previous study[19] to extrapolate the discriminatory power of ML with increasing sample size. Because of the lack of pre-existing literature in the field, this method was applied to preliminary data that we retrieved in a pilot study (n = 3) to calculate the sample size for study 1. Based on this approach, an AUROC of 0.85 to detect pronounced hypoglycaemia was projected for a sample size of n = 18. After completion of study 1, we implemented a bootstrap procedure[20] to suggest a sample size for study 2. Specifically, after training our ML models, we computed 10 000 random samples with replacement for the out-of-sample AUROC of n patients and then inspected the bootstrapped distribution. For a sample size of n = 9, we registered a mean AUROC of 0.88 with a standard deviation of 0.03. We thus aimed for n = 9 completing study 2, which was expected to give precise estimates of the diagnostic accuracy with good confidence.

2.4 | ML approach

We developed and tested our ML approach in a two-step manner (Figure 2): first, based on data from study 1 (n = 18), we built three ML models (named CAN+ET, CAN and ET) to detect pronounced hypoglycaemia (vs. euglycaemia) and evaluated the performance using crossvalidation. Second, the three ML models trained on data from pronounced hypoglycaemia (study 1) were applied to previously unseen data from study 2 and their performance in detecting mild hypoglycaemia (vs. euglycaemia) was evaluated. We chose this approach because training the models on data from pronounced hypoglycaemia (study 1) allows them to associate clear behavioural changes with hypoglycaemia. Also, using data from mild hypoglycaemia to evaluate the models allows us to see how well the models perform when the behavioural effects of hypoglycaemia are weaker and to provide early warnings. In addition, this also allows validating the models on a separate population (study 2).

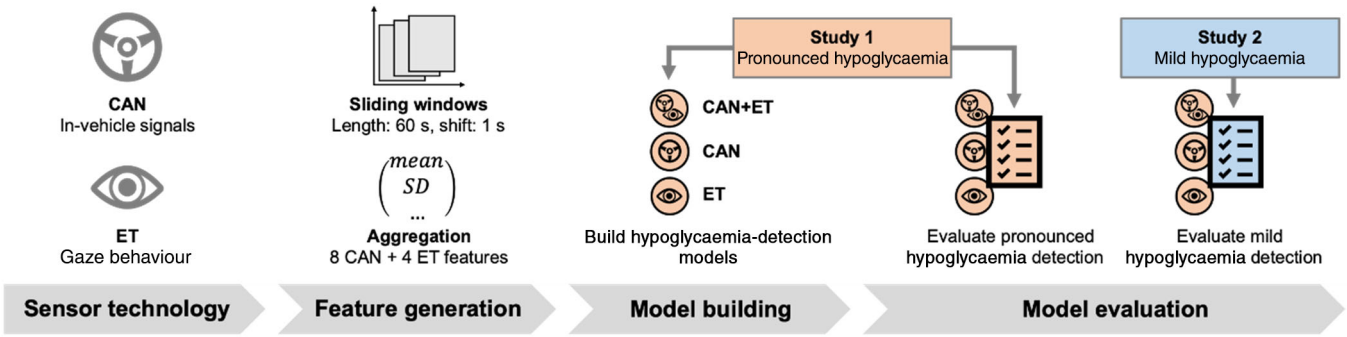

FIGURE 2 Procedure for building and evaluating our machine learning models. CAN, controller area network; ET, eye tracking; SD, standard deviation

To reflect different generations of vehicles, we evaluated the performance to detect hypoglycaemia separately with three ML models: (a) the CAN + ET model incorporating driving and gaze data, representing the latest state of available technology in modern cars; (b) the CAN model solely based on driving data, because contemporary cars are not yet generally equipped with ET; and (c) the ET model using only gaze data, anticipating that the availability of (semi-)autonomous driving[21] will limit the role of CAN data in the future.

We followed best practice in ML and used the following procedure for training and evaluation to ensure that the ML approach generalizes well to unseen individuals and to unseen road segments (see the supporting information, Supplementary Methods). To this end, all evaluations were performed using out-of-sample data to assess the ML models on previously unseen road segments and unseen individuals. For study 1, we used leave-one-subject-out crossvalidation (n = 18). Hyperparameters were tuned against the AUROC. For study 2, there was no training and no hyperparameter tuning; instead, we used the trained hypoglycaemia detection ML models from study 1 and applied them to each participant from study 2 (n = 9). That is, there was no additional training with data from study 2; instead, data from study 2 were only used for assessing the prediction performance. Thereby, we assumed mild hypoglycaemia to have the same but weaker effects on driving behaviour than pronounced hypoglycaemia. We also experimented with other training and evaluation procedures (Table S7), where we arrived at consistent conclusions. Eventually, results are reported as the out-of-sample prediction performance averaged across study participants (i.e. macro-average). To quantify the variation in the prediction performance across participants, we further report the standard deviation of the performance at participant level in both studies.

The input to the ML models consisted of eight features for CAN data, derived from four in-vehicle data signals reflecting the driver behaviour and vehicle velocity ('brake pedal position', 'steering wheel angle' and 'vehicle velocity'). For ET, four features were derived from two eye tracker signals ('gaze fixations' and 'gaze velocity'). All features were standardized by subtracting the mean and scaling to unit variance. Each feature was computed within a sliding window of 60 seconds. We did not use driver characteristics (e.g. age) as inputs to our models because (a) we included a comparably homogeneous population of well-controlled, young individuals with type 1 diabetes, and (b) currently implemented advanced driver assistance systems in production cars work without additional information about the driver.[22,23] This is attributed to various reasons, including privacy concerns and usability. Details of the feature engineering are outlined in in the supporting information (Supplementary Methods). The output of the three ML models was the probability of the participant driving in hypoglycaemia versus euglycaemia. Additional ML modelling specifications and robustness checks can be found in the in the supporting information (Supplementary Methods).

2.5 | Reporting and the software used

Unless otherwise specified, results are reported as mean ± standard deviation (SD). Paired BG and CGM values were checked for normal distribution using the Shapiro–Wilk test and compared using paired *t*-tests or Wilcoxon signed rank tests. Self-rated symptoms were analysed as follows: according to previous research,[24] neurogenic and neuroglycopenic scores were calculated by averaging scores of the four neurogenic and neuroglycopenic symptoms, respectively. The overall symptom score was calculated by averaging scores of all eight symptoms. Symptom scores, single symptoms and self-estimated BG levels were checked for normal distribution, and compared between euglycaemia and hypoglycaemia using paired *t*-tests or paired Wilcoxon signed rank tests, respectively. A *P* value of less than .05 was considered statistically significant.

Descriptive statistical analyses were performed using STATA version 16.0 (StataCorp LLC, College Station, TX). All ML models were implemented using Python 3.8 with the Python packages scikit-learn (version 0.24.2). The package XGBoost (version 1.3.3) was used additionally for the robustness checks. Input features to the ML models were computed using numpy (version 1.20.1) and scipy (version 1.6.2). Evaluation metrics were computed using scikit-learn (version 0.24.2). The software used for data collection and management are described in the in the supporting information (Supplementary Methods).

TABLE 1 Baseline characteristics of the participants

Variable	Study 1 (n = 18)	Study 2 (n = 9)
Age (y)	32.2 ± 7.1	47.6 ± 10.5
Sex	6 female, 12 male	2 female, 7 male
Insulin treatment	12 CSII, 6 MDI	4 CSII, 5 MDI
Weight (kg)	85.0 ± 22.5	84.6 ± 21.5
Height (m)	1.76 ± 0.10	1.76 ± 0.08
BMI (kg m^{-2})	27.1 ± 5.0	27.2 ± 5.5
TDD (IU day^{-1} kg^{-1})	0.69 ± 0.16	0.59 ± 0.13
HbA1c (%)	7.1 ± 0.6	7.3 ± 0.8
HbA1c (mmol mol^{-1})	54 ± 7	56 ± 9
Clarke score > 3	0 / 18	2 / 9[a]
Diabetes duration (y)	19.5 ± 11.0	20.8 ± 10.9
Driving experience (y)	14.1 ± 7.6	25.8 ± 13.3
Kilometres driven per year (km $year^{-1}$)	9356 ± 7837	12 944 ± 9625

Note: Shown are the mean values ± standard deviation for continuous variables. A Clarke score of higher than 3 points indicates impaired awareness of hypoglycaemia.
Abbreviations: BMI, body mass index; CSII, continuous subcutaneous insulin infusion; IU, insulin units; MDI, multiple daily injections; TDD, total daily insulin dose.
[a]Two participants reported a Clarke score of 4 points.

2.6 | Data and code availability statement

The code for independent replication is available on GitHub (https://github.com/im-ethz/ML-For-Hypoglycemia-Detection-While-Driving-In-Simulator). The datasets from the current study are available from the corresponding author upon reasonable request. All data shared will be deidentified.

3 | RESULTS

The final analysis included 18 individuals with type 1 diabetes from study 1 (age 32.2 ± 7.1 years, 12 males, HbA1c 7.1% ± 0.6% [54 ± 7 mmol mol^{-1}]) and nine individuals from study 2 (age 47.6 ± 10.5 years, seven males, HbA1c 7.3% ± 0.8% [56 ± 9 mmol mol^{-1}]; Table 1). There was no overlap in these participants across the studies. The study flows are displayed in the Figure S1.

Mean venous BG during hypoglycaemia was 2.37 ± 0.18 mmol L^{-1} in study 1 and 3.31 ± 0.15 mmol L^{-1} in study 2 (Figure 1D). Corresponding mean CGM values were 3.30 ± 0.44 and 3.81 ± 0.64 mmol L^{-1}, respectively. CGM readings were significantly higher compared with BG values in both studies during hypoglycaemia ($P < .001$ for all comparisons). Individual BG values are shown in the Figure S2; the self-rated symptoms by the participants are summarized in the Table S1.

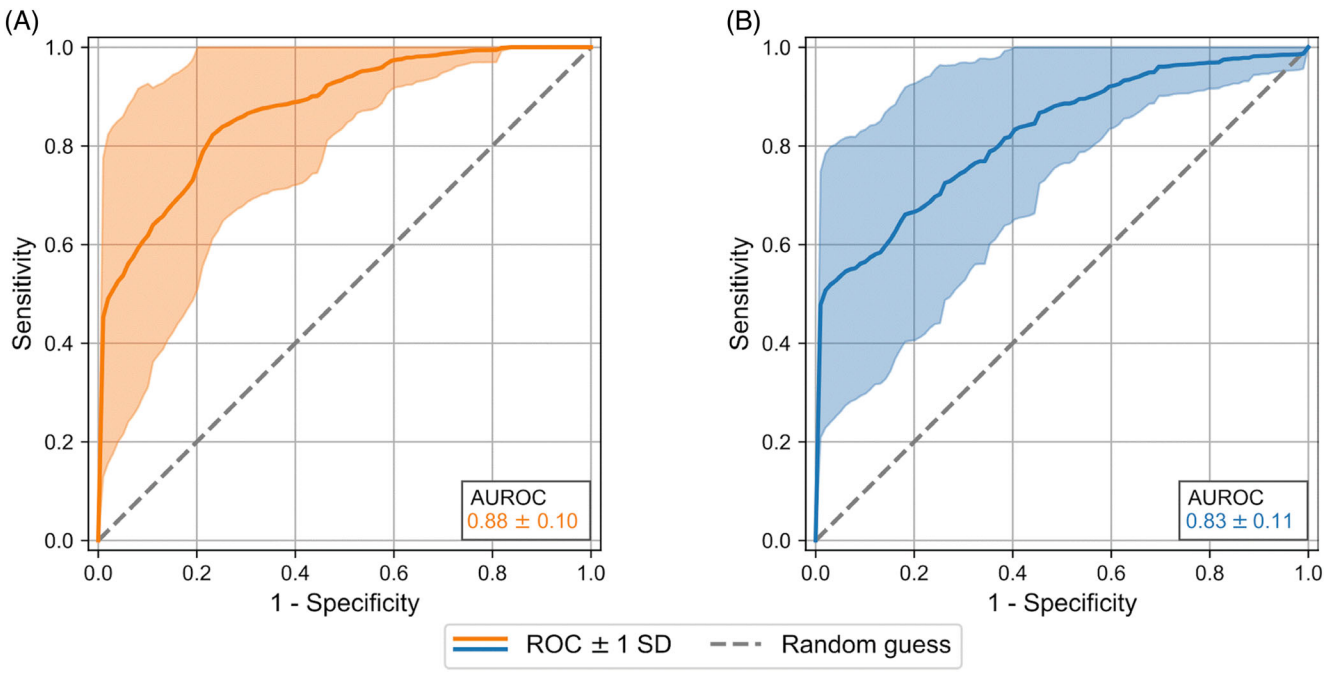

FIGURE 3 Machine learning detects pronounced and mild hypoglycaemia based on driving and gaze data. Reported is the area under the curve for the receiver operating characteristic (AUROC) to detect hypoglycaemia. Here, we report the performance in detecting A, Pronounced hypoglycaemia (study 1), and B, Mild hypoglycaemia (study 2) using combined in-vehicle driving and eye tracking data (CAN+ET). The AUROC illustrates the mean true positive rate (= sensitivity) against the false positive rate (= 1 − specificity). The shaded areas illustrate the standard deviation (SD) at various thresholds across the participants. The grey dashed line shows the performance of a model that has no discriminatory power and decides at random (AUROC = 0.50). ROC, receiver operating characteristic

TABLE 2 Machine learning detects pronounced and mild hypoglycaemia based on driving and gaze data

		AUROC	AUPRC	BACC	F1	MCC	Sensitivity	Specificity
Study 1	CAN+ET	0.88 ± 0.10	0.90 ± 0.10	0.85 ± 0.10	0.87 ± 0.10	0.70 ± 0.19	0.86 ± 0.14	0.83 ± 0.14
	CAN	0.81 ± 0.13	0.86 ± 0.11	0.80 ± 0.09	0.81 ± 0.13	0.61 ± 0.16	0.79 ± 0.19	0.81 ± 0.15
	ET	0.81 ± 0.15	0.87 ± 0.12	0.81 ± 0.10	0.82 ± 0.13	0.63 ± 0.20	0.79 ± 0.20	0.83 ± 0.19
Study 2	CAN+ET	0.83 ± 0.11	0.92 ± 0.06	0.80 ± 0.08	0.80 ± 0.13	0.57 ± 0.16	0.71 ± 0.19	0.88 ± 0.13
	CAN	0.75 ± 0.05	0.88 ± 0.04	0.74 ± 0.05	0.85 ± 0.09	0.53 ± 0.10	0.88 ± 0.16	0.59 ± 0.18
	ET	0.75 ± 0.19	0.86 ± 0.12	0.76 ± 0.12	0.86 ± 0.07	0.52 ± 0.23	0.86 ± 0.11	0.65 ± 0.25

Note: Reported is the performance in detecting pronounced (study 1) and mild hypoglycaemia (study 2) as mean ± standard deviation. Across both studies, we report the performance metrics using combined in-vehicle driving and eye tracking data (CAN+ET), and driving (CAN) or gaze (ET) data exclusively. Abbreviations: AUROC, area under the curve for the receiver operating characteristic; AUPRC, area under the precision-recall curve; BACC, balanced accuracy; F1, F1-score; MCC, Matthew's correlation coefficient.

Overall, the feature engineering approach described in the supporting information (Appendix S1, Supplementary Methods) led to 18 844 (9881) observations for study 1 (study 2), of which 9101 (4804) came from driving in euglycaemia and 9743 (5077) from driving in hypoglycaemia. The distribution of observations across the different BG levels is shown in Figure S3. For detection of pronounced hypoglycaemia (study 1), the CAN+ET model showed an overall area under the receiver operating characteristics curve (AUROC) of 0.88 ± 0.10 (Figure 3A). The corresponding area under the precision-recall curve (AUPRC) was 0.90 ± 0.10. The CAN model achieved an AUROC of 0.81 ± 0.13 and the ET model showed an AUROC of 0.81 ± 0.15 (Table 2).

When transferring the three ML models to mild hypoglycaemia (study 2), the CAN+ET model showed an overall AUROC of 0.83 ± 0.11 (Figure 3B), and AUPRC of 0.92 ± 0.06. The CAN model achieved an AUROC of 0.75 ± 0.05 and the ET model showed an AUROC of 0.75 ± 0.19 (Table 2).

Additional performance metrics are displayed in Table 2. The AUPRC plots and the performance across different environments (highway, rural and urban) are shown in the Figures S4 and S5, and Table S3.

To explain the decision-making of the ML models, we interpret the coefficients of the input features for CAN+ET, CAN and ET in the, Figure S6. Robustness checks include the evaluation of other (non-)linear ML models, a sensitivity analysis of the detection performance across different window lengths, and with different training and evaluation procedures (Tables S5–S7).

4 | DISCUSSION

The main findings of our prospective, interventional studies in people with type 1 diabetes evaluating hypoglycaemia detection while driving in a simulator are 3-fold: first, a non-invasive ML approach, solely based on driving and gaze behaviour data and without measurement of glucose (i.e. the CAN+ET model), detected pronounced hypoglycaemia with high accuracy. Second, our ML approach was also applicable to mild hypoglycaemia, thereby allowing for early warnings. Third, limiting the model to driving data (the CAN model) or gaze data (the ET model), exclusively, still resulted in acceptable detection of both hypoglycaemic levels. This corroborates the potential of our ML approach to be applied in widely available cars without eye-tracking cameras (CAN only), as well as expanding its use to future cars with (semi-)automated driving (ET only).

Driving a vehicle involves the complex management of speed, braking and steering. High levels of cognitive, executive and psychomotor functions are required, all of which are affected negatively by hypoglycaemia.[3–5] Although SMBG is a standard approach, it is not suitable for detecting hypoglycaemia while driving. CGM offers continuous glucose readings but is limited by invasiveness, availability and compromised accuracy, particularly in hypoglycaemia.[25] The cost of a CGM system is estimated to be approximately one thousand to several thousand dollars a year, depending on the country and the manufacturer. In addition, coverage of CGM by health insurance is limited and most people living with diabetes still do not use or have access to this technology.[26,27] By contrast, our approach leverages data that are already being recorded by vehicles, making it a scalable and cost-effective solution not requiring additional sensors installed in the car or attached to the body. Moreover, there is a growing economic interest in in-vehicle warning systems, as car manufacturers are increasingly integrating health-related features into their vehicles.[22,23] Of note, the accompanying CGM system significantly underestimated the degree of hypoglycaemia in both of our studies, corroborating the potential of the ML approach to improve the accuracy of hypoglycaemia detection. While manual calibration could mitigate this limitation of factory-calibrated CGM systems,[28] it would not eliminate the delay of CGM as described previously.[10] Conversely, setting CGM alarm thresholds to a higher level may translate into earlier warnings, but will probably worsen glycaemic control,[29] while repetitive adaptation before and after each journey may not be realistic in clinical practice.

The interpretation of the mean coefficients of the input features (Figure S6) allowed for an analysis of the behavioural changes while driving in hypoglycaemia. Driving behaviour based on CAN data was characterized by a decrease in the standard deviation (SD) of vehicle controls (steering, brake and gas pedal) in hypoglycaemia, indicating a less proactive driving style with reduced fine motor control. Drivers intervened more abruptly, which was reflected in higher energy (i.e. sum of squares) in vehicle control signals. When analysing the ET

data, the model feature coefficients revealed less situational and wandering gaze behaviour, which was reflected in a lower number of gaze fixations, as well as a higher mean and a lower SD in gaze velocity. Observations in CAN and ET were consistent in that they both depicted behaviour in hypoglycaemia as more monotonous, less situational and less fine control driven.

Earlier simulator studies in individuals with type 1 diabetes have reported more time off-road and across the midline in hypoglycaemia.[18,30] These changes indicate (near) mishaps and are thus unsuitable variables for a preventive system. By contrast, the proposed ML approach relies upon driving features that describe more subtle changes in driving behaviour, allowing for the detection of changes at an earlier stage. This is corroborated by the fact that our ML approach still achieved an adequate performance when tested in mild hypoglycaemia. In line with the literature,[31–33] participants reported few symptoms and overestimated their BG levels during mild hypoglycaemia (Table S1), and a majority reported that they would continue driving in this state. Such findings, established in well-controlled individuals with preserved hypoglycaemia awareness according to established criteria,[34] further emphasizes the need for alternative hypoglycaemia detection methods.

All three ML models showed good performance in the highway environment, where the traffic context is more monotonous than in other settings. By contrast, the urban and rural environments appeared more challenging. In urban and rural settings, drivers have to operate the steering wheel and pedals more frequently and significantly, as well as shift their gaze more often (traffic lights, pedestrians, junctions, etc.).

The strength of our study is its prospective and interventional design using a standardized protocol, providing data from different hypoglycaemic ranges and driving environments. BG, the gold standard, was measured with high frequency, confirming that the glycaemic target ranges during the experiments were reliably met and maintained within narrow ranges. In a two-step manner, we developed and tested our ML models in independent populations and across different ranges of hypoglycaemia, irrespectively of individual thresholds for cognitive decline. Our dataset was collected in a well-established driving simulator, using CAN and ET data of contemporary car systems, thus providing a base for widespread applicability in the automotive sector. Compared with other proposed hypoglycaemia detection methods,[13] our approach allows for implementation without the need for additional sensors installed in the vehicle or attached to the body. All ML models were evaluated on unseen road segments and unseen individuals, which eliminates learning bias. While the current study focuses on people with diabetes, the concept may be applicable to other critical driver states caused by drowsiness and/or other medical conditions. However, this hypothesis needs validation in future studies.

Limitations include a restricted sample size, owing to the complex and laborious study procedures. Conversely, the high resolution of driving and gaze variables (30 and 90 Hz, respectively) and BG values (5-10 minutes) provided a solid basis for the ML modelling process. The model was built on data of well-controlled and generally healthy individuals with type 1 diabetes, as hypoglycaemia induction was ethically justifiable in this population. This limits generalization to multimorbid individuals and other populations affected by hypoglycaemia (e.g. type 2 diabetes), where the approach needs separate validation. Currently, the detection capacity of the ML approach is limited to the specific glucose ranges of these studies and the performance in additional glucose ranges requires future research. As the study was performed in a simulator and not in real cars, we acknowledge the proof-of-concept character of our experiments. Given the potential risks of inducing hypoglycaemia while driving, this may however be an acceptable first step. In this study, we used CAN data analogous to the data collected in real cars. This does not include environmental data, which precludes conclusions on the performance of our model on predicting mishaps (e.g. crossing the midline). We acknowledge that the sequence of driving (euglycaemia followed by hypoglycaemia) may have introduced bias. This was chosen to avoid a carry-over effect because driving after hypoglycaemia may be affected for up to 75 minutes after restoration of euglycaemia.[35] Lastly, the frequency of euglycaemia and hypoglycaemic values was balanced in the current study, not reflecting clinical reality. While this may increase the probability of false positive alarms, this may again be acceptable at the current conceptual stage.

In conclusion, we provide proof-of-concept that a machine learning approach based on driving and gaze behaviour data can detect hypoglycaemia while driving. The approach may empower self-management and care of people with diabetes, and may be applicable to contemporary cars, while anticipating future developments in automotive technology.

AUTHOR CONTRIBUTIONS

VL, TZ, MM, and MK share first authorship. EF and CS share last authorship. The following authors contributed to the conception and design of the study: CS, EF, TZ, ML, FW, TK and SF; acquisition of data: VL, TZ, MK, MM, CB, CA, NS and SL; analysis of data: VL, TZ, MK, MM and SF; interpretation of data: VL, TZ, MK, MM, CB, SF, FW, TK and CS; writing the manuscript: VL, TZ, MM, MK, SF, FW and CS; critical review of the manuscript: CB, TK, ML and EF. CS is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final draft of the manuscript for submission.

ACKNOWLEDGEMENTS

The authors would like to thank all the volunteers for their participation and enthusiasm. Thanks go to Constance Bischoff and Joana Filipa Rodrigues Cunha Freitas (study nurses), Mona Spoerri (medical student) and Vincent Vorburger (trainee) for their valuable support during the clinical visits. We are grateful to Mirjam Zueger for her support in the graphics design. Dexcom (DexCom, Inc., San Diego, CA) provided product support (CGM systems). Dexcom had no role in trial conduct, nor in data analysis, nor manuscript drafting. The study was funded by the Swiss National Science Foundation (Sinergia grant, SNF CRSII5_183569), with additional grants provided by the Clinical Trials

Unit Berne, the Swiss Diabetes Foundation, the Diabetes Centre Berne, and the Department of Research of the University Hospital Berne. Open access funding provided by Inselspital Universitatsspital Bern.

CONFLICT OF INTEREST

The authors declare no competing interests.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/dom.15021.

DATA AVAILABILITY STATEMENT

The code for independent replication is available on GitHub (https://github.com/im-ethz/ML-For-Hypoglycemia-Detection-While-Driving-In-Simulator). The datasets from the current study are available from the corresponding author upon reasonable request. All data shared will be de-identified.

ORCID

Martin Maritsch https://orcid.org/0000-0001-9920-0587
Christoph Stettler https://orcid.org/0000-0003-1691-6059

REFERENCES

1. Frier BM. Hypoglycaemia in diabetes mellitus: epidemiology and clinical implications. *Nat Rev Endocrinol*. 2014;10(12):711-722. doi:10.1038/nrendo.2014.170
2. Ishtiak-Ahmed K, Carstensen B, Pedersen-Bjergaard U, Jørgensen ME. Incidence trends and predictors of hospitalization for hypoglycemia in 17,230 adult patients with type 1 diabetes: a Danish register linkage cohort study. *Diabetes Care*. 2017;40(2):226-232. doi:10.2337/dc16-0862
3. Hoffman RG, Speelman DJ, Hinnen DA, Conley KL, Guthrie RA, Knapp RK. Changes in cortical functioning with acute hypoglycemia and hyperglycemia in type I diabetes. *Diabetes Care*. 1989;12(3):193-197.
4. Strachan MW, Ewing FM, Frier BM, McCrimmon RJ, Deary IJ. Effects of acute hypoglycaemia on auditory information processing in adults with type I diabetes. *Diabetologia*. 2003;46(1):97-105. doi:10.1007/s00125-002-0950-2
5. Graveling AJ, Deary IJ, Frier BM. Acute hypoglycemia impairs executive cognitive function in adults with and without type 1 diabetes. *Diabetes Care*. 2013;36(10):3240-3246. doi:10.2337/dc13-0194
6. ECRI. Diabetes and Commercial Motor Vehicle Safety (Federal Motor Carrier Safety Administration). June 2011 Update: Plymouth Meeting, Pennsylvania, ECRI. 2011.
7. Cox DJ, Penberthy JK, Zrebiec J, et al. Diabetes and driving mishaps: frequency and correlations from a multinational survey. *Diabetes Care*. 2003;26(8):2329-2334.
8. Cox DJ, Ford D, Gonder-Frederick L, et al. Driving mishaps among individuals with type 1 diabetes: a prospective study. *Diabetes Care*. 2009;32(12):2177-2180. doi:10.2337/dc08-1510
9. Skurtveit S, Strom H, Skrivarhaug T, Morland J, Bramness JG, Engeland A. Road traffic accident risk in patients with diabetes mellitus receiving blood glucose-lowering drugs. Prospective follow-up study. *Diabet Med*. 2009;26(4):404-408. doi:10.1111/j.1464-5491.2009.02699.x
10. Schmelzeisen-Redeker G, Schoemaker M, Kirchsteiger H, Freckmann G, Heinemann L, Del Re L. Time delay of CGM sensors: relevance, causes, and countermeasures. *J Diabetes Sci Technol*. 2015;9(5):1006-1015. doi:10.1177/1932296815590154
11. Merickel J, High R, Smith L, et al. Driving safety and real-time glucose monitoring in insulin-dependent diabetes. *Int J Automot Eng*. 2019;10(1):34-40. doi:10.20485/jsaeijae.10.1_34
12. Tsichlaki S, Koumakis L, Tsiknakis M. Type 1 diabetes hypoglycemia prediction algorithms: systematic review. *JMIR Diabetes*. 2022;7(3):e34699. doi:10.2196/34699
13. Diouri O, Cigler M, Vettoretti M, et al. Hypoglycaemia detection and prediction techniques: a systematic review on the latest developments. *Diabetes Metab Res Rev*. 2021;37(7):e3449. doi:10.1002/dmrr.3449
14. Koert Groeneveld DV, Schierholz S. Meet the S-Class DIGITAL: "My MBUX" (Mercedes-Benz User Experience): At home on the road – luxurious and digital. Accessed July 19, 2021 https://media.daimler.com/marsMediaSite/en/instance/ko.xhtml?oid=46757031
15. Bock O, Drescher U, Janouch C, Haeger M, van Winsum W, Voelcker-Rehage C. An experimental paradigm for the assessment of realistic human multitasking. *Virtual Real*. 2019;23(1):61-70. doi:10.1007/s10055-018-0342-7
16. van Winsum W. A threshold model for stimulus detection in the peripheral detection task. *Transp Res F: Traffic Psychol Behav*. 2019;65:485-502. doi:10.1016/j.trf.2019.08.014
17. van Winsum W. Optic flow and tunnel vision in the detection response task. *Hum Factors*. 2019;61(6):992-1003. doi:10.1177/0018720818825008
18. Cox DJ, Gonder-Frederick LA, Kovatchev BP, Julian DM, Clarke WL. Progressive hypoglycemia's impact on driving simulation performance. Occurrence, awareness and correction. *Diabetes Care*. 2000;23(2):163-170.
19. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification performance. *BMC Med Inform Decis Mak*. 2012;12(1):8. doi:10.1186/1472-6947-12-8
20. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat*. 1979;7(1):1-26.
21. Committee ORAD. *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Warrendale, PA: SAE International; 2021. https://doi.org/10.4271/J3016_202104
22. Your Volvo car will understand you when you're not at your best – and step in when you need support. Accessed December 13, 2022 2022 https://www.media.volvocars.com/global/en-gb/media/pressreleases/304020
23. Smart eye's automotive driver monitoring technology has been installed in more than 1,000,000 cars on roads globally. Accessed December 13, 2022 2022 https://smarteye.se/news/smart-eye-announces-its-automotive-driver-monitoring-technology-has-been-installed-in-more-than-1000000-cars-on-roads-globally/
24. Potashner D, Brown RE, Li A, Riddell MC, Aronson R. Paradoxical rise in hypoglycemia symptoms with development of hyperglycemia during High-intensity interval training in type 1 diabetes. *Diabetes Care*. 2019;42(10):2011-2014. doi:10.2337/dc19-0609
25. Farrell CM, McNeilly AD, Hapca SM, McCrimmon RJ. Real-time continuous glucose monitoring during a hyperinsulinemic-hypoglycemic clamp significantly underestimates the degree of hypoglycemia. *Diabetes Care*. 2020;43(10):e142-e143. doi:10.2337/dc20-0882
26. DeSalvo DJ, Noor N, Xie C, et al. Patient demographics and clinical outcomes among type 1 diabetes patients using continuous glucose monitors: data from T1D exchange real-world observational study. *J Diabetes Sci Technol*. 2021:19322968211049783. doi:10.1177/19322968211049783
27. Kruger DF, Anderson JE. Continuous glucose monitoring (CGM) is a tool, not a reward: unjustified insurance coverage criteria limit access to CGM. *Diabetes Technol Ther*. 2021;23(S3):S-45-S-55. doi:10.1089/dia.2021.0193
28. Nagl K, Berger G, Aberer F, et al. Performance of three different continuous glucose monitoring systems in children with type 1 diabetes

during a diabetes summer camp. *Pediatr Diabetes*. 2021;22(2):271-278. doi:10.1111/pedi.13160
29. Lin YK, Groat D, Chan O, et al. Alarm settings of continuous glucose monitoring systems and associations to glucose outcomes in type 1 diabetes. *J Endocr Soc*. 2020;4(1):bvz005. doi:10.1210/jendso/bvz005
30. Cox DJ, Gonder-Frederick L, Clarke W. Driving decrements in type I diabetes during moderate hypoglycemia. *Diabetes*. 1993;42(2):239-243.
31. Mitrakou A, Ryan C, Veneman T, et al. Hierarchy of glycemic thresholds for counterregulatory hormone secretion, symptoms, and cerebral dysfunction. *Am J Physiol*. 1991;260(1 Pt 1):E67-E74. doi:10.1152/ajpendo.1991.260.1.E67
32. Schwartz NS, Clutter WE, Shah SD, Cryer PE. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for symptoms. *J Clin Invest*. 1987;79(3):777-781. doi:10.1172/jci112884
33. Fanelli C, Pampanelli S, Epifano L, et al. Relative roles of insulin and hypoglycaemia on induction of neuroendocrine responses to, symptoms of, and deterioration of cognitive function in hypoglycaemia in male and female humans. *Diabetologia*. 1994;37(8):797-807. doi:10.1007/bf00404337
34. Clarke WL, Cox DJ, Gonder-Frederick LA, Julian D, Schlundt D, Polonsky W. Reduced awareness of hypoglycemia in adults with IDDM. A prospective study of hypoglycemic frequency and associated symptoms. *Diabetes Care*. 1995;18(4):517-522.
35. Zammitt NN, Warren RE, Deary IJ, Frier BM. Delayed recovery of cognitive function following hypoglycemia in adults with type 1 diabetes: effect of impaired awareness of hypoglycemia. *Diabetes*. 2008;57(3):732-736. doi:10.2337/db07-0695

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lehmann V, Zueger T, Maritsch M, et al. Machine learning for non-invasive sensing of hypoglycaemia while driving in people with diabetes. *Diabetes Obes Metab*. 2023;25(6):1668-1676. doi:10.1111/dom.15021