Maritsch Martin (Orcid ID: 0000-0001-9920-0587)

Stettler Christoph (Orcid ID: 0000-0003-1691-6059)

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

**Running title:** Hypoglycemia detection during driving

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

[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

[*] equal first authorship; [#] equal last authorship.

**Corresponding author:** Prof Christoph Stettler, MD,
E-Mail: christoph.stettler@insel.ch, Phone: +41 31 632 40 70

**Word count:** Abstract: 222, Manuscript: 3500; **Display items:** 3 Figures, 2 Tables; **References**: 37

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/dom.15021

## Abstract

**Aims:** Hypoglycemia is one of the most dangerous acute complications of diabetes mellitus and is associated with an increased risk of driving mishaps. Current approaches to detect hypoglycemia are limited by invasiveness, availability, costs, and technical restrictions. In this work, we developed and evaluated the concept of a non-invasive machine learning (ML) approach detecting hypoglycemia based exclusively on combined driving (CAN) and eye tracking (ET) data.

**Materials and Methods:** We first developed and tested our ML approach in pronounced hypoglycemia, and, then, we applied it to mild hypoglycemia to evaluate its early warning potential. For this, we conducted two consecutive, interventional studies in individuals with type 1 diabetes mellitus. In study 1 (*n*=18), we collected CAN and ET data in a driving simulator during eu- and pronounced hypoglycemia (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 eu- and mild hypoglycemia (BG 3.0 – 3.5 mmol $L^{-1}$).

**Results:** Here, we show that our ML approach detects pronounced and mild hypoglycemia with high accuracy (area under the receiver operating characteristics curve [AUROC] 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, allows for detection of hypoglycemia while driving. This provides a promising concept for alternative and non-invasive detection of hypoglycemia.

**Study registration:** ClinicalTrials.gov (NCT04035993 and NCT05183191).

## Introduction

Hypoglycemia is a dangerous acute complication of diabetes mellitus[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, hypoglycemia 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, costs, and is subject to an inherent time lag in hypoglycemia[10]. Of note, in a recent prospective study, individuals with type 1 diabetes spent a considerable amount of time in hypoglycemia while driving[11], corroborating the need for alternative and complementary methods to detect hypoglycemia while driving. Here, we develop a machine learning (ML) approach to detect hypoglycemia exclusively from driving and gaze behavior. There is a growing body of evidence examining hypoglycemia 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 hypoglycemia using ML methodology based on driving and gaze behavior 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 behavior and vigilance, also in (semi-)autonomous driving situations. A hypoglycemia 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 paper, we present the concept of a machine learning (ML) approach using CAN and ET data to detect hypoglycemia during driving.

## Material and Methods

Study design and population

We conducted two non-randomized, interventional studies in individuals with type 1 diabetes from 10/2019 to 07/2020 (study 1), and from 11/2021 to 03/2022 (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 breast-feeding, 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 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, the 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, NCT05183191).

Study procedure

Supplementary Figure 1a 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., due to motion sickness) were excluded. Participants were fitted with the Dexcom G6 continuous glucose monitoring (CGM) system. Participants were instructed to refrain from alcohol, caffeine, and strenuous

physical activity 24 hours before the main visit. The main visit was postponed if sensor glucose was <3.0mmol $L^{-1}$ for >30min in the preceding 24 hours.

For the main visit, participants were admitted to our clinical research unit after an overnight fast. During a controlled hypoglycemia procedure, participants drove in eu- and hypoglycemia (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, Netherlands) as in previous studies on driving behavior[16-18]. Eye gaze was recorded with a consumer eye tracker (Tobii Eye Tracker 4C, Tobii AB, Danderyd, Sweden). The intended BG range in hypoglycemia was 2.0–2.5mmol $L^{-1}$ (in study 1) and 3.0–3.5mmol $L^{-1}$ (in study 2), respectively (Figure 1c and 1d). In both studies, each driving session in eu- and hypoglycemia consisted of three environments (highway, rural, and urban) completed in random order. The driving lasted 5 min in each environment and was separated by 1–2 min breaks for intermittent BG measurement using the Biosen C-Line glucose analyzer (EKF Diagnostics Holdings plc., Penarth, Cardiff, Great Britain). Participants were informed that a hypoglycemic state aiming at a BG level of 2.0–2.5mmol $L^{-1}$ (study 1) or 3.0–3.5mmol $L^{-1}$ (study 2) was to be induced but they were blinded to the BG values throughout the experiment. In eu- and hypoglycemia, participants rated eight hypoglycemic symptoms, 'need-to-treat right now' and 'difficulty driving' on a seven-point scale (0=none, 6=extreme)[19]. 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 restoral of euglycemia, the procedure was terminated if deemed save by the

investigator. A detailed description of the controlled hypoglycemic state, the driving simulator, the eye tracker, and the driving environments is provided in the Supplementary Methods.

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

Outcome and sample size calculation

The main outcome was the diagnostic accuracy of our ML approach to detect hypoglycemia quantified as the area under the receiver operating characteristic curve (AUROC). Traditional null hypothesis 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[20] to extrapolate the discriminatory power of ML with increasing sample size. Due to 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 hypoglycemia was projected for a sample size of $n$=18. After completion of study 1, we implemented a bootstrap procedure[21] 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.

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 hypoglycemia (vs. euglycemia) and evaluated the performance using cross-validation. Second, the three ML models trained on data from pronounced hypoglycemia (study 1) were applied to previously unseen data from study 2 and their performance in detecting mild hypoglycemia (vs. euglycemia) was evaluated. We chose this approach because training the models on data from pronounced hypoglycemia (study 1) allows them to associate clear behavioral changes with hypoglycemia. Second, using data from mild hypoglycemia to evaluate the models allows us to see how well the models perform when the behavioral effects of hypoglycemia are weaker and to provide early warnings. In addition, this also allows validating the models on a separate population (study 2).

To reflect different generations of vehicles, we evaluated the performance to detect hypoglycemia separately with three ML models: (1) The *CAN+ET model* incorporating driving and gaze data, representing the latest state of available technology in modern cars. (2) A *CAN model* solely based on driving data, since contemporary cars are not yet generally equipped with ET. (3) An *ET model* using only gaze data, anticipating that the availability of (semi-)autonomous driving[22] will limit the role of CAN data in the future.

We followed best-practice in ML and use the following procedure for training and evaluation to ensure that the ML approach generalizes well to unseen individuals and to unseen road segments (see 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 cross-validation (*n*=18). Hyperparameters were tuned against the AUROC. For study 2, there was no training and no hyperparameter tuning; instead, we used the trained hypoglycemia detection ML models from study 1 and applied it 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 was only used for assessing the prediction performance. Thereby, we assumed mild hypoglycemia to have the same but weaker effects on driving behavior than pronounced hypoglycemia. We also experimented with other training and evaluation procedures (Supplementary Table 7), 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 behavior and vehicle velocity ('brake pedal position', 'gas pedal position', 'steering wheel angle', 'vehicle velocity'). For ET, four features were derived from two eye tracker signals ('gaze fixations', '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 as (i) we included a comparably homogeneous population of well-controlled, young individuals with type 1 diabetes, and (ii) as currently implemented advanced driver assistance systems in

production cars work without additional information about the driver[23, 24]. This is due to various reasons including privacy concerns and usability. Details on the feature engineering are outlined in the Supplementary Methods. The output of the three ML models was the probability of the participant driving in hypoglycemia vs. euglycemia. Additional ML modelling specifications and robustness checks can be found in the Supplementary Methods.

Reporting and 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 analyzed as follows: according to previous research[25], 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 euglycemia and hypoglycemia using paired *t*-tests or paired Wilcoxon signed rank tests, respectively. A $p$-value $<0.05$ was considered statistically significant.

Descriptive statistical analyses were performed using STATA version 16.0 (StataCorp LLC, College Station, Texas, USA). 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). Software used for data collection and management are described in the Supplementary Methods.

Data and code availability statement

The code for independent replication is available on GitHub (*link will be inserted upon acceptance of this manuscript*). The datasets from the current study are available from the corresponding author upon reasonable request. All data shared will be de-identified.

## Results

The final analysis included 18 individuals with type 1 diabetes from study 1 (age 32.2±7.1y, 12 male, $HbA_{1c}$ 7.1±0.6%, 54±7mmol $mol^{-1}$), and 9 individuals from study 2 (age 47.5±10.5y, 7 male, $HbA_{1c}$ 7.3±0.8%, 56±9mmol $mol^{-}$, see Table 1). There was no overlap in these participants across the studies. The study flows are displayed in the Supplementary Figure 1.

Mean venous BG during hypoglycemia was 2.37±0.18mmol $L^{-1}$ in study 1 and 3.31±0.15mmol $L^{-1}$ in study 2 (Figure 1d). Corresponding mean CGM values were 3.30±0.44mmol $L^{-1}$ and 3.81±0.64mmol $L^{-1}$, respectively. CGM readings were significantly higher compared to BG values in both studies during hypoglycemia ($p$<0.001 for all comparisons). Individual BG values are shown in the Supplementary Figure 2, the self-rated symptoms by the participants are summarized in the Supplementary Table 1.

Overall, the feature engineering approach described in the Supplementary Methods led to 18,844 (9,881) observations for study 1 (study 2), out of which 9,101 (4,804) observations come from driving in euglycemia and 9,743 (5,077) from driving in hypoglycemia. The distribution of observations across the different BG levels is shown in Supplementary Figure 3. For detection of pronounced hypoglycemia (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 hypoglycemia (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 Supplementary Figures 4 and 5, and Supplementary Table 3.

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 Supplementary Figure 6. 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 (Supplementary Tables 5–7).

## Discussion

The main findings of our prospective, interventional studies in people with type 1 diabetes evaluating hypoglycemia detection while driving in a simulator are three-fold: First, a non-invasive ML approach purely based on driving and gaze behavior data and without measurement of glucose (CAN+ET model) detected pronounced hypoglycemia with high accuracy. Second, our ML approach was also applicable to mild hypoglycemia, thereby allowing for early warnings. Third, limiting the model to driving data (CAN model) or gaze data (ET model), exclusively, still resulted in an acceptable detection of both hypoglycemic 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 the 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 hypoglycemia[3-5]. While SMBG is a standard approach, it is not suitable for detecting hypoglycemia while driving. CGM offers continuous glucose readings but is limited by invasiveness, availability, and compromised accuracy particularly in hypoglycemia[26]. The costs for a CGM system are considered around thousand to several thousand dollars a year, depending on the country and the manufacturer. In addition, coverage of CGM by health insurances is limited and a majority of people living with diabetes still does not use or have access to this technology[27, 28]. In contrast, our approach leverages data that is 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[23, 24]. Of note, the accompanying CGM system significantly underestimated the degree of hypoglycemia in both of our studies, corroborating the potential of the ML approach to improve the accuracy of hypoglycemia detection. While manual calibration could mitigate this limitation of factory-calibrated CGM systems[30], 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 warning but is likely to worsen glycemic control[31], while repetitive adaptation before and after each ride may not be realistic in clinical practice.

The interpretation of the mean coefficients of the input features (Supplementary Figure 6) allowed for an analysis of the behavioral changes while driving in hypoglycemia. Driving behavior based on CAN data was characterized by a decrease in the standard deviation (SD) of vehicle controls (steering, brake and gas pedal) in hypoglycemia, 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 analyzing the ET data, the model feature coefficients revealed a less situational and wandering gaze behavior, 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 behavior in hypoglycemia 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 hypoglycemia[19, 33]. These changes indicate (near) mishaps and are thus unsuitable parameters for a preventive system. In contrast, the proposed ML approach relies upon driving features that describe more subtle changes in driving behavior, allowing for detection of changes in an earlier stage. This is corroborated by the fact that our ML approach still achieved an adequate performance when tested in mild hypoglycemia. In line with the literature[34-36], participants reported few symptoms and overestimated their BG levels during mild hypoglycemia (Supplementary Table 1), and a majority reported that they would continue driving in this state. Such findings, established in well-controlled individuals with preserved hypoglycemia awareness according to established criteria[15], further emphasizes the need for alternative hypoglycemia detection methods.

All three ML models showed good performance in the highway environment, where the traffic context is more monotonous than in other settings. In 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 hypoglycemic ranges and driving environments. BG, the gold standard, was measured with high frequency, confirming that the glycemic 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 hypoglycemia,

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 to other proposed hypoglycemia 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 parameters (30 and 90 Hz, respectively) and BG values (5 – 10 min) 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, since hypoglycemia induction was ethically justifiable in this population. This limits generalization to multi-morbid individuals and other populations affected by hypoglycemia (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. Since 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 hypoglycemia 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 (euglycemia followed by hypoglycemia) may have introduced bias. This was chosen to avoid a carry-over effect since driving after hypoglycemia may be affected up to 75 min after restoration of euglycemia[37]. Lastly, the frequency of eu- and hypoglycemic values was balanced in the current study, not reflecting clinical reality. While this may increase 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 behavior data can detect hypoglycemia 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, and MM 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, 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.

## Acknowledgments

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 Center Berne, and the Department of Research of the University Hospital Berne.

The authors would like to thank all the volunteers for their participation and enthusiasm. Thanks go to the 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, California, USA) provided product support (CGM systems). Dexcom had no role in trial conduct nor in data analysis or manuscript drafting.

## Conflict of interest

The authors declare no competing interests.

## References


1. Frier BM. Hypoglycaemia in diabetes mellitus: epidemiology and clinical implications. *Nature reviews Endocrinology*. Dec 2014;10(12):711-22. 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*. Feb 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*. Mar 1989;12(3):193-7.

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*. Jan 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*. Oct 2013;36(10):3240-6. 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*. Aug 2003;26(8):2329-34.

8. Cox DJ, Ford D, Gonder-Frederick L, et al. Driving mishaps among individuals with type 1 diabetes: a prospective study. *Diabetes Care*. Dec 2009;32(12):2177-80. 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*. Apr 2009;26(4):404-8. 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*. Aug 04 2015;9(5):1006-15. 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*. 2022/7/21 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/metabolism research and reviews*. 2021;37(7):e3449-e3449. doi:10.1002/dmrr.3449

14. Koert Groeneveld DV, Steffen Schierholz. Meet the S-Class DIGITAL: "My MBUX" (Mercedes-Benz User Experience): At home on the road – luxurious and digital. Accessed Juli 19, 2021, https://media.daimler.com/marsMediaSite/en/instance/ko.xhtml?oid=46757031

15. 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*. Apr 1995;18(4):517-22.

16. 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 Reality*. 2019/03/01 2019;23(1):61-70. doi:10.1007/s10055-018-0342-7

17. van Winsum W. A threshold model for stimulus detection in the peripheral detection task. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2019/08/01/ 2019;65:485-502. doi:https://doi.org/10.1016/j.trf.2019.08.014

18. van Winsum W. Optic Flow and Tunnel Vision in the Detection Response Task. *Human Factors*. 2019;61(6):992-1003. doi:10.1177/0018720818825008

19. 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*. Feb 2000;23(2):163-70.

20. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification performance. OriginalPaper. *BMC Medical Informatics and Decision Making*. 2012-02-15 2012;12(1):8. doi:doi:10.1186/1472-6947-12-8

21. Efron B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*. 1979;7(1):1-26, 26.

22. committee O-RAD. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE International; 2021.

23. Your Volvo car will understand you when you're not at your best – and step in when you need support. Accessed 13.12.2022, 2022. https://www.media.volvocars.com/global/en-gb/media/pressreleases/304020

24. Smart Eye's Automotive Driver Monitoring Technology Has Been Installed in more than 1,000,000 Cars on Roads Globally. Accessed 13.12.2022, 2022.

https://smarteye.se/news/smart-eye-announces-its-automotive-driver-monitoring-technology-has-been-installed-in-more-than-1000000-cars-on-roads-globally/

25. 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

26. 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*. Oct 2020;43(10):e142-e143. doi:10.2337/dc20-0882

27. 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. *Journal of diabetes science and technology*. 0(0):19322968211049783. doi:10.1177/19322968211049783

28. Kruger DF, Anderson JE. Continuous Glucose Monitoring (CGM) Is a Tool, Not a Reward: Unjustified Insurance Coverage Criteria Limit Access to CGM. *Diabetes technology & therapeutics*. 2021/09/01 2021;23(S3):S-45-S-55. doi:10.1089/dia.2021.0193

29. Matt T. Car Ownership Statistics in the U.S. https://www.valuepenguin.com/auto-insurance/car-ownership-statistics#national

30. 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*. Mar 2021;22(2):271-278. doi:10.1111/pedi.13160

31. 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*. Jan 1 2020;4(1):bvz005. doi:10.1210/jendso/bvz005

32. Kashevnik A, Shchedrin R, Kaiser C, Stocker A. Driver Distraction Detection Methods: A Literature Review and Framework. *IEEE Access*. 2021;9:60063-60076. doi:10.1109/ACCESS.2021.3073599

33. Cox DJ, Gonder-Frederick L, Clarke W. Driving decrements in type I diabetes during moderate hypoglycemia. *Diabetes*. Feb 1993;42(2):239-43.

34. Mitrakou A, Ryan C, Veneman T, et al. Hierarchy of glycemic thresholds for counterregulatory hormone secretion, symptoms, and cerebral dysfunction. *The American journal of physiology*. Jan 1991;260(1 Pt 1):E67-74. doi:10.1152/ajpendo.1991.260.1.E67

35. Schwartz NS, Clutter WE, Shah SD, Cryer PE. Glycemic thresholds for activation of glucose counterregulatory systems are higher than the threshold for

symptoms. *The Journal of clinical investigation*. Mar 1987;79(3):777-81. doi:10.1172/jci112884

36. 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*. Aug 1994;37(8):797-807. doi:10.1007/bf00404337

37. 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*. Mar 2008;57(3):732-6. doi:10.2337/db07-0695

## Figure legends

**Figure 1: Overview. (a)** Hypoglycemia induction procedures for study 1 and study 2 using variable insulin aspart and glucose administration with corresponding driving sessions in eu- and hypoglycemia. The intended range for blood glucose (BG) in hypoglycemia 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-minute drives in three different environments (highway, rural, and urban) while in-vehicle driving (CAN) and eye tracking (ET) data was collected. **(b)** Driving simulator, eye tracker, and glucose management setup in both studies. **(c)** Key characteristics of study 1 and study 2. **(d)** Venous BG in hypoglycemia for study 1 and study 2 shown as boxplots. Overall, BG in hypoglycemia 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 x IQR from the edge of the box. Values outsides the whisker range are illustrated by dots.

**Figure 2: Procedure for building and evaluating our machine learning models.**

**Figure 3: Machine learning (ML) detects pronounced and mild hypoglycemia based on driving and gaze data.** Reported is the area under the curve for the receiver operating characteristic (AUROC) to detect hypoglycemia. Here, we report the performance in detecting **(a)** pronounced hypoglycemia (study 1) and **(b)** mild hypoglycemia (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 gray dashed line shows the performance of a model that has no discriminatory power and decides at random (AUROC=0.50). ROC, receiver operating characteristic.

## Tables

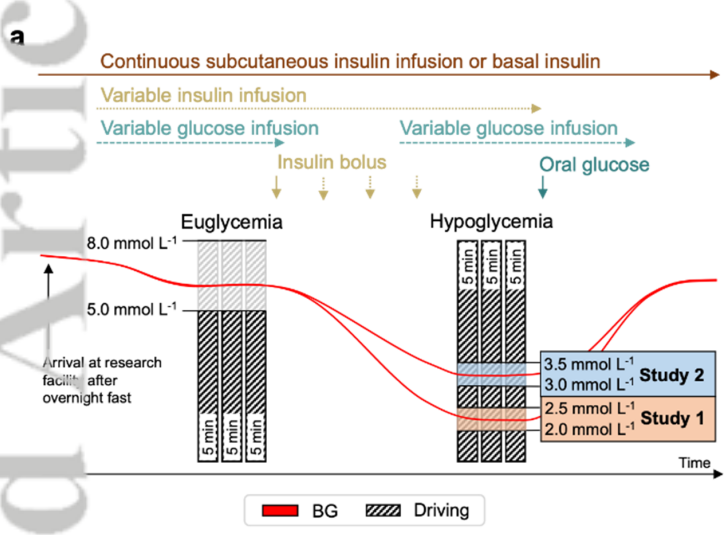
| Variable | Study 1 (*n*=18) | Study 2 (*n*=9) |
|---|---|---|
| Age [years] | 32.2±7.1 | 47.5±10.5 |
| Sex | 6 female, 12 male | 2 female, 7 male |
| Insulin treatment | 12 CSII, 6 MDI | 4 CSII, 5 MDI |
| Weight [kg] | 85.0±22.5 | 84.6±21.5 |
| Height [m] | 1.76±0.10 | 1.76±0.08 |
| BMI [kg $m^{-2}$] | 27.1±5.0 | 27.2±5.5 |
| TDD [IU $day^{-1}$ $kg^{-1}$] | 0.69±0.17 | 0.59±0.13 |
| $HbA_{1c}$ [%] | 7.1±0.6 | 7.3±0.8 |
| $HbA_{1c}$ [mmol $mol^{-1}$] | 54±7 | 56±9 |
| Clarke score >3 | 0 / 18 | 2 / 9 * |
| Diabetes duration [years] | 19.5±11.0 | 20.8±10.9 |
| Driving experience [years] | 14.1±7.6 | 25.8±13.3 |
| Kilometers driven per year [km $year^{-1}$] | 9,356±7,837 | 12,944±9,625 |

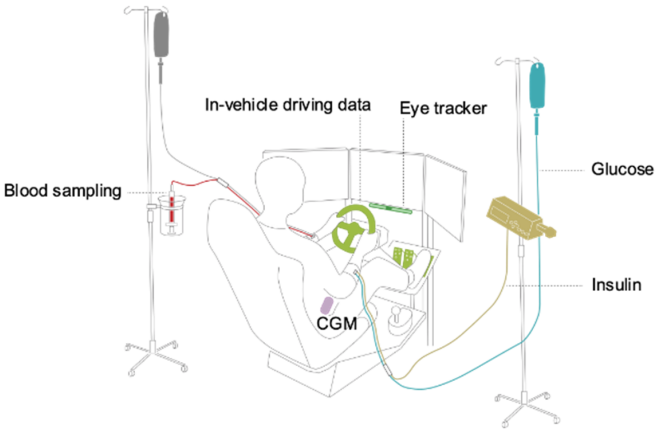
*two participants reported a Clarke Score of 4 points

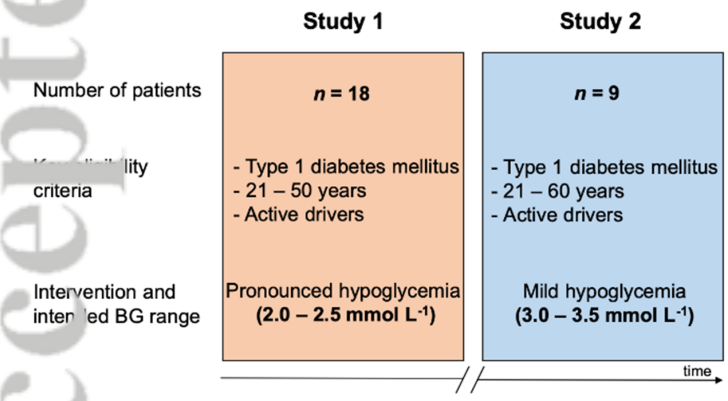
**Table 1: Baseline characteristics of the participants.** Shown are the mean values±standard deviation for continuous variables. A Clarke score of higher than 3 points indicates impaired awareness of hypoglycemia. BMI, body mass index; CSII, continuous subcutaneous insulin infusion; $HbA_{1c}$, glycated hemoglobin; IU, insulin units; MDI, multiple daily injections; TDD, total daily insulin dose.

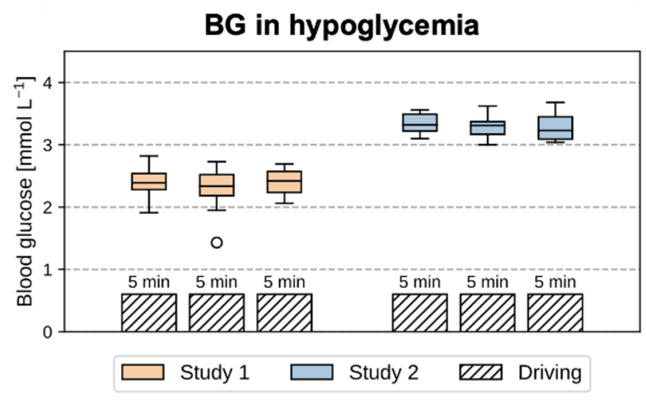
| | | AUROC | AUPRC | BACC | F1 | MCC | Sensitivity | Specificity |
|---|---|---|---|---|---|---|---|---|
| Study 1 | CAN+ET | 0.88±0.10 | 0.90±0.10 | 0.85±0.10 | 0.87±0.10 | 0.70±0.19 | 0.86±0.14 | 0.83±0.14 |
| | CAN | 0.81±0.13 | 0.86±0.11 | 0.80±0.09 | 0.81±0.13 | 0.61±0.16 | 0.79±0.19 | 0.81±0.15 |
| | 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 |

**Table 2: Machine learning (ML) detects pronounced and mild hypoglycemia based on driving and gaze data.** Reported is the performance in detecting pronounced (study 1) and mild hypoglycemia (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. AUROC, area under the curve for the receiver operating characteristic, AUPRC, area under the precision-recall curve; BACC, balanced accuracy; F1, F1-score; MCC, Matthews correlation coefficient.

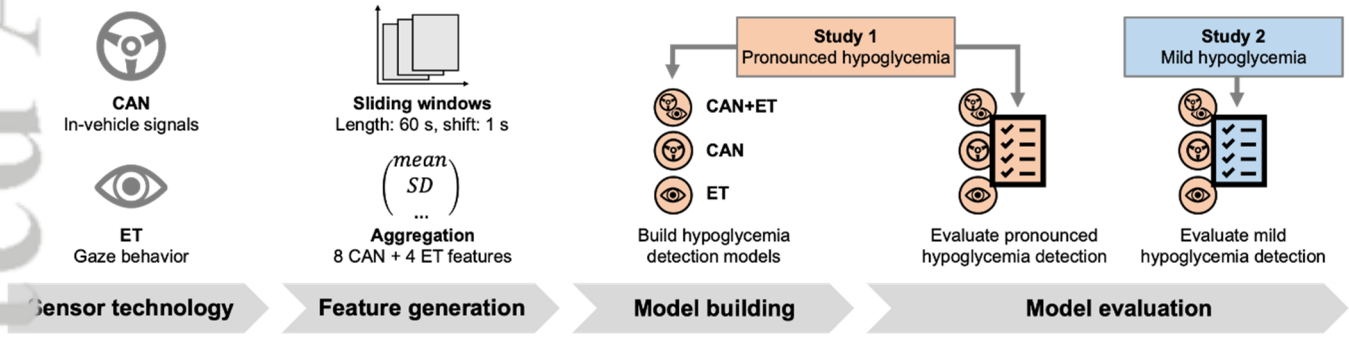


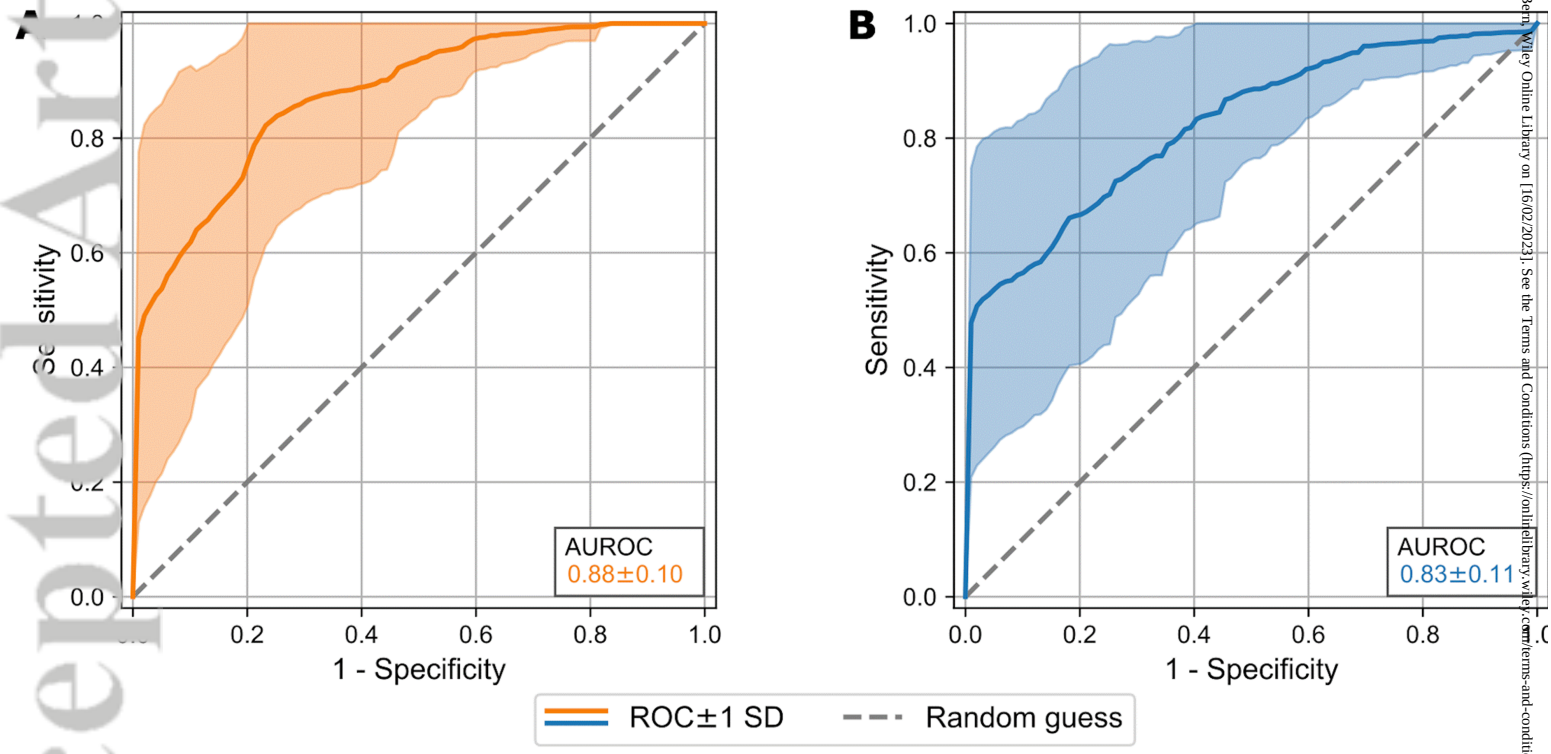




DOM_15021_dom-22-1319-op-File002.tif




DOM_15021_dom-22-1319-op-File003.tif




DOM_15021_dom-22-1319-op-File004.tif