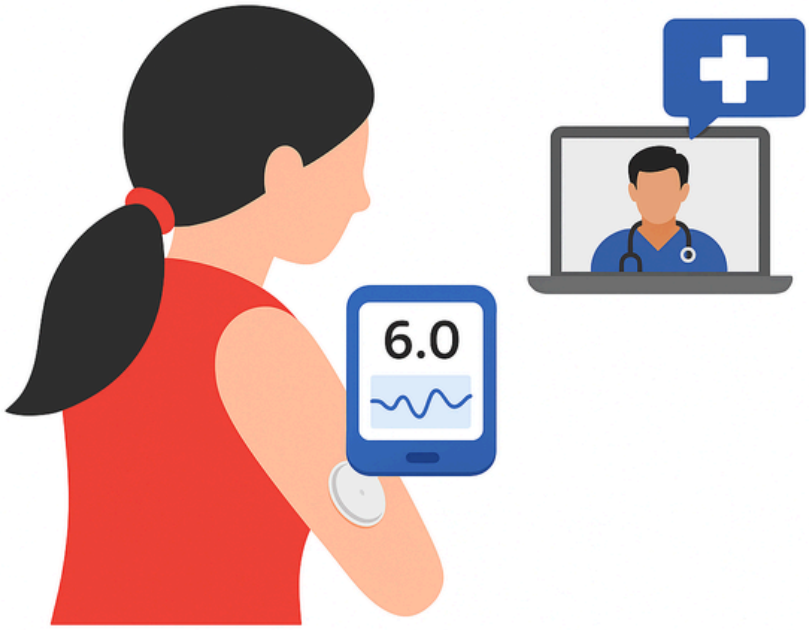
Clinical Science & Technology, 4th. semester master-project




# PREDICTING INDIVIDUALS DIAGNOSED WITH TYPE 2 DIABETES ACHIEVING HBA1C REDUCTION FROM TELEMONITORING - A POST HOC ANALYSIS USING MACHINE LEARNING

Tobias Fromreide & Lukas J. R. Nielsen

**Supervisor**

Stine Hangaard Casper, AAU

Thomas Kronborg Larsen, AAU

# *FORUDSIGELSE AF PERSONER MED TYPE 2-DIABETES, SOM OPNÅR HBA1C-REDUKTION VED TELEMONITORERING – EN POST HOC ANALYSE MED MASKINLÆRING*

AALBORG UNIVERSITY

**Titel:**
Forudsigelse af personer med type 2-diabetes, som opnår HbA1c-reduktion ved telemonitorering – en post hoc-analyse med maskinlæring

**Projektperiode:**
03.02.2025 - 02.06.2025

**Projektgruppe:**
10522

**Tema:**
Prædiktion af personer diagnosticeret med Type 2 diabetes som har bedst gavn af telemonitorering

**Forfattere:**
Lukas Nielsen
Tobias Fromreide

**Vejledere:**
Stine Hangaard Casper
Thomas Kronborg Larsen

**Referencesystem:**
Vancouver

**Antal tegn inkl. mellemrum, figurer og tabeller:**
56317

**Antal sider:**
23

**Bilag:**
2

**Figurer:**
7

**Tabeller:**
5

**Afleveringsdato:**
02.06.2025

**Baggrund og formål:** Diabetes mellitus (DM) påvirker cirka 537 millioner mennesker globalt, og antallet forventes at stige til 643 millioner i 2030, hvilket udgør en betydelig klinisk og økonomisk byrde. Telemonitorering er påvist at kunne forbedre glykæmisk kontrol hos personer med type 2-diabetes (T2DM), men effekten er variererende, og implementeringen er omkostningstung, hvilket gør målrettet patientudvælgelse afgørende. Maskinlæringsmodeller viser potentiale i diabetesbehandling, men i øjeblikket findes ingen modeller, til forudsigelsen af, hvem der har størst gavn af telemonitorering. Dette studie havde til formål at udvikle en prædiktiv model til at identificere personer med T2DM, der med størst sandsynlighed vil opnå effekt af telemonitorering.

**Metoder:** Denne post hoc-analyse anvendte data fra interventionsgruppen (n=163) i det randomiserede kontrollerede DiaMonT-studie. Et mål blev defineret som en ≥15 % reduktion i HbA1c. Manglende data blev imputeret, og datasættet blev stratificeret og derefter opdelt i trænings- og testdatasæt i et forhold på 80:20. Variabler blev udvalgt ved hjælp af fremadrettet sekventiel feature-selektion med fem-fold krydsvalidering og AUC. En logistisk regressionsmodel blev trænet på de udvalgte features og valideret på testdatasættet.

**Results:** Baseret på forward selection blev der i alt udvalgt seks variable: HbA1c_baseline, Betablokkere, opioider, DPP-4-hæmmere, GLP-1-receptoragonister og hypoglykæmi. Modellen opnåede en AUC på 0,72, en sensitivitet på 0,80 og en specificitet på 0,56.

**Konklusion:** En maskinlæringsmodel baseret på data fra DiaMonT-studiet viste moderat evne til at forudsige, hvilke personer med type 2-diabetes der opnåede >15 % reduktion i HbA1c som følge af telemonitorering. Resultaterne understøtter modellens potentiale for individuel behandlingsallokering, men yderligere forskning er nødvendig, før modellen kan anvendes selvstændigt i klinisk praksis.

# *PREDICTING INDIVIDUALS DIAGNOSED WITH TYPE 2 DIABETES ACHIEVING HBA1C REDUCTION FROM TELEMONITORING - A POST HOC ANALYSIS USING MACHINE LEARNING*

AALBORG UNIVERSITY

**Title**
Predicting individuals diagnosed with Type 2 Diabetes achieving HbA1c reduction from Telemonitoring - A post hoc analysis using machine learning

**Project period:**
03.02.2025 - 02.06.2025

**Project group:**
10522

**Theme:**
Prediction of individuals benefitting the most of telemonitoring

**Authors:**
Lukas Nielsen
Tobias Fromreide

**Supervisor:**
Stine Hangaard Casper

Thomas Kronborg Larsen

**System of reference:**
Vancouver

**Characters inclusive spaces, tables and figures:**
56317

**Pages:**
23

**Appendices:**
2

**Figures:**
7

**Tables:**
5

**Date of submission:**
02.06.2025

**Background and aim:** Diabetes mellitus affects approximately 537 million people globally, with an expected rise to 643 million by 2030, posing a significant clinical and economic burden. Telemonitoring can improve glycemic control in individuals with type 2 diabetes mellitus (T2DM), but its effectiveness varies and implementation is costly, making targeted patient selection crucial. While machine learning shows promise in diabetes care, no existing models predict who benefits most from telemonitoring. This study aimed to develop a predictive model to identify individuals with T2DM most likely to respond positively to telemonitoring

**Methods:** This post hoc analysis used data from the intervention group (n=163) of the DiaMonT randomized controlled trial. A binary outcome of ≥15% HbA1c reduction was defined. Missing data were imputed, and data stratified before being split into training and test sets using a 80:20 ratio. Features were selected using forward sequential feature selection with five-fold cross-validation and AUC. A logistic regression model was trained on the selected features and validated on the test set.

**Results:** Based on the forward selection a total of six features were included: HbA1c_baseline, Beta blockers, Opioids, DPP-4, GLP-1 receptor agonist and Hypoglycemia. The model achieved an AUC of 0.72, sensitivity of 0.80 and specificity of 0.56.

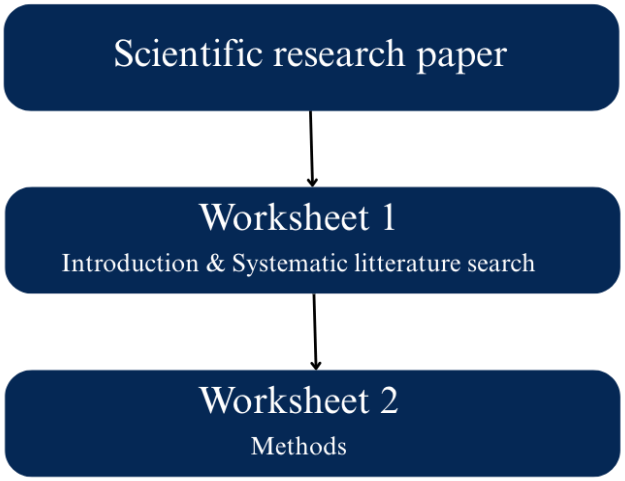
**Conclusion:** A machine learning model based on DiaMonT trial data moderately predicted which individuals with T2DM achieved >15% HbA1c reduction from telemonitoring. The results support its potential for personalized treatment allocation, though further research is needed before independent clinical application.

# Preface

This project was performed by Lukas Nielsen and Tobias Fromreide as part of the 4th. semester of the Master's program in Clinical Science and Technology at Aalborg University, during the period from February 2025 to June 2025. The results of the project will be of relevance for both public and private organizations as well as healthcare professionals working with Type 2 Diabetes Mellitus, digital health, and the development of machine learning models in clinical settings. We would like to extend our sincere thanks to our main supervisor, Stine Hangaard Casper, Professor at Aalborg University and Co-supervisor, Thomas Kronborg Larsen, Professor at Aalborg University, for their constructive guidance throughout the project.

# Readers guide

This thesis has been prepared in the form of a scientific article accompanied by two supplementary worksheets, as illustrated in Figure 1. These appendices provide an in-depth theoretical understanding of the thesis background, structured literature search, methodology, and perspectives. The scientific article is intended for submission to the *Journal of Diabetes Science and Technology*.




**Figur 1: Readers guide**

The research paper and its accompanying worksheets use the Vancouver reference, where references are numbered. All references are listed in the article and the corresponding appendices with the respective number. In both the research paper and worksheets, references to studies and other literature in the text are cited by stating the surname of the first author, followed by *et al.* if there are multiple authors. Each worksheet and its subsections are numbered. Supplementary material related to the worksheets is included as appendices at the end. Appendices are labeled with "Appendix" followed by a sequential number. The written work is supplemented throughout with figures and tables, which are numbered after chronological sequence and have all been prepared by the group members. Abbreviations are written out in full the first time they appear and shown in parentheses.

# Predicting individuals diagnosed with Type 2 Diabetes achieving HbA1c reduction from telemonitoring - A post hoc analysis using machine learning

T. Fromreide[a] and L.J.R Nielsen[a]

[a] Aalborg University, Department of Health Science and Technology, Selma Lagerloefvej 249 9260 Gistrup Denmark

## Abstract

**Background and aim:** Diabetes mellitus affects approximately 537 million people globally, with an expected rise to 643 million by 2030, posing a significant clinical and economic burden. Telemonitoring can improve glycemic control in individuals with type 2 diabetes mellitus (T2DM), but its effectiveness varies and implementation is costly, making targeted patient selection crucial. While machine learning shows promise in DM care, no existing models predict who benefits most from telemonitoring. This study aimed to develop a predictive model to identify individuals with T2DM most likely to respond positively to telemonitoring

**Methods:** This post hoc analysis used data from the intervention group (n=163) of the DiaMonT randomized controlled trial. A binary outcome of ≥15% HbA1c reduction was defined. Missing data were imputed, and data stratified before being split into training and test sets using a 80:20 ratio. Features were selected using forward sequential feature selection with five-fold cross-validation and area under the curve (AUC) as the performance metric. A logistic regression model was trained on the selected features and validated on the test set.

**Results:** Based on the forward selection a total of six features were included: HbA1c_baseline, Beta blockers, Opioids, DPP-4, GLP-1 receptor agonist and Hypoglycemia. The model achieved an AUC of 0.72, sensitivity of 0.80 and specificity of 0.56.

**Conclusions:** A machine learning model based on DiaMonT trial data moderately predicted which individuals with T2DM achieved >15% HbA1c reduction from telemonitoring. The results support its potential for personalized treatment allocation, though further research is needed before independent clinical application.

**Keywords:** Type 2 diabetes - Telemonitoring - Prediction - Machine Learning - HbA1c

## Introduction

The prevalence of individuals living with diabetes mellitus (DM) worldwide was in 2021 estimated to be 537 million with an expected growth to 643 million in 2030 and 800 million by 2045 (1). The condition is a leading cause of mortality and is responsible for an estimated 6.7 million deaths worldwide per year (1). The condition is highly costly with an estimated amount in 2021 of 966000 USD million pr. year globally (1,2). DM is a group of metabolic disorders characterized by hyperglycemia, which is caused by defects in insulin action, insulin secretion or both (3). The most common form is Type 2 diabetes mellitus (T2DM), accounting for 90-95% of all individuals living with DM. In T2DM the elevated blood glucose levels are due to a combination of insulin resistance and dysfunction of insulin producing B-cells, which over time becomes less effective, leading to inadequate insulin production and hyperglycemia (4). In healthy individuals, blood glucose levels are maintained within a small range, typically between 3.9 and 7.8 mmol/L (5). Blood glucose levels are also measured as glycated hemoglobin (HbA1c), which reflects the average blood glucose levels over a period of two to three months (6). Hyperglycemia can have significant long-term consequences leading to further expenses and mortality (6,7). The International Diabetes Federation recommends a general target of HbA1c levels <53 mmol/mol for individuals diagnosed with T2DM (8). However, research shows that only approximately 50% achieve a level of ≤53 mmol/mol, showing a significant gap between guideline targets and clinical reality (9,10). A modern approach to glycemic control is telemonitoring, which enables individuals to continuously track their blood glucose levels using Continuous Glucose Monitoring (CGM) systems, with real-time data accessible via mobile devices or eHealth platforms (11). Telemonitoring has shown potential in improving glycemic control in people with T2DM (11–13). However, the clinical benefits vary across individuals and evidence suggests that the effect is more pronounced in studies with short duration and smaller samples (11,14). Telemonitoring comes with a high cost, whereas efficient patient selection is crucial for implementation (15–17). A systematic literature search across multiple scientific databases revealed that recent advances in machine learning (ML) have enabled predictive modelling in various diabetes-related aspects, such as hyperglycemia risk and retinopathy screening (18–22). However, to the best of the authors knowledge, no earlier evidence exists for predicting which individuals diagnosed with T2DM derive the greatest from telemonitoring. The aim of this study was to develop a predictive model capable of identifying individuals most likely to achieve an effect of telemonitoring.

## Methods

### *Data collection*

This study is a post hoc analysis based on data from the DiaMonT trial, a randomized controlled study (NCT04981808) investigating telemonitoring in adults with insulin-treated T2D (23). Participants were recruited from Steno Diabetes Center North Denmark (Aalborg University Hospital) and Steno

Diabetes Center Zealand (Nykøbing Falster Hospital) (23). Individuals were included if they had a confirmed insulin-treated T2DM for at least 12 months, were able to use a smartphone and trial-related devices, could read and understand Danish and were the age of minimum 18 years (23). The participants were randomized into either the intervention group or control group. The intervention group (n=200) were provided with a CGM, connected pen (NovoPen) for long-acting insulin (and an additional pen for short-acting insulin if applicable), activity tracker (Fitbit Charge 4), and a smart-phone (unless they prefer using their own) (23). Participants received training at inclusion and used the devices continuously to collect data on glucose levels, insulin management, physical activity and sleep (23). The participants received calls after one week, one month and two months, and additionally as needed (23). The control group (n=200) received standard care according to Danish guidelines, along with blinded smart insulin pens (for long-acting insulin, and short-acting if applicable) and a blinded CGM, worn during the first and last 20 days of the trial (23). The data were collected for later analysis (23). Both the intervention and control group were monitored for a total of three months (23).

## *Data preprocessing*

Only data from the intervention group were used for model development. All variables were converted to numerical format, and repeated observations were aggregated making each patient represented by a single row. The outcome was defined as a binary variable indicating ≥15% reduction in HbA1c from baseline to follow-up. Missing values were estimated using a multivariate imputation method (Python: *IterativeImputer,* Library: *scikit-learn*) (24). To preserve class distribution between training and test data, stratification was applied prior to data splitting (25). The data was stratified into the following groups: >20% decrease, 10-20% decrease, 0-10% decrease, 0-10% increase and 10-20% increase. Afterwards, the dataset was split into training and test sets using an 80:20 ratio. The training data were standardized (Python: *StandardScaler,* Library: *scikit-learn*) to improve numerical stability during logistic regression (26) .

## *Feature selection*

Sequential feature addition was performed (Python: *SequentialFeatureSelection,* Library*: mlxtend*) prior to training the model. Logistic regression was used as the classifier and the scoring of features was done by the area under the curve (AUC) using five-fold-cross validation. Floating was enabled to allow the algorithm to both add and remove features dynamically. Features were sequentially added if they improved the mean cross-validated AUC score by ≥ 0.005.

## *Model training and validation*

The training and test data were reduced including only the selected features. Subsequently, the training data were standardized (Python: *StandardScaler,* Library: *scikit-learn*) and the same transformation was applied to the test data ensuring consistent

scaling without risking data leakage. A logistic regression model was then trained with the final features, and finally evaluated on the test data using AUC as the performance metric.

### *Software*

All data preprocessing and modeling were conducted in Python. The Anaconda distribution was used to manage the programming environment and Jupyter Notebook as the development interface for code, analysis and data visualisation.

## Results

The preprocessing resulted in the total of 163 participants in the final dataset leaving 168 excluded due to being in the control group or not completing the study. In the final dataset a total of 331 missing values were imputed. The baseline characteristics for the included participants can be found in Table 1. A Kernel Density Plot illustrating the change in density distribution of HbA1c values from baseline to follow-up was made (Figure 1). As a result from the sequential feature selection a total of six features out of 75 potential features were selected for model training as their inclusion resulted in an mean increase of >0.005 in AUC (Figure 2). The six selected features were HbA1c_baseline, Beta blockers, Opioids, DPP-4, GLP-1 receptor agonists and Hypoglycemia. The best performing logistic regression model was able to differentiate between the participants having a 15% reduction in HbA1c from baseline to follow-up and those who did not, with an average AUC of 0.87 across five-fold cross-validation on the training data (Figure 3). When evaluated on the independent test data, the final model achieved an AUC of 0.72 (Figure 4). In total the model performs with a sensitivity of 0.8, specificity of 0.56, positive predictive value (PPV) of 0.6, negative predictive value (NPV) of 0.77 and an accuracy of 0.67.

**Table 1.** Baseline Characteristics for included patients

| Characteristic | Value |
|---|---|
| Age, years | 61.18 ± 11.33 |
| HbA1c, mmol | 62.09 ± 12.94 |
| BMI | 33.06 ± 6.63 |
| Years with diabetes | 19.43 ± 14.43 |
| Smoking | 2.25 ± 0.65 |
| Alcohol | 1.25 ± 0.67 |
| Exercise | 1.42 ± 0.70 |

*Continuous variables are presented by mean ± standard deviation.*

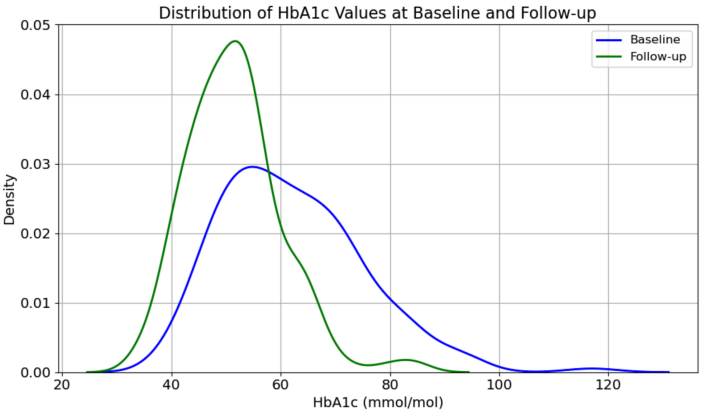



**Figure 1:** Kernel density plot showing the distribution of HbA1c values at baseline and at the end of follow-up.

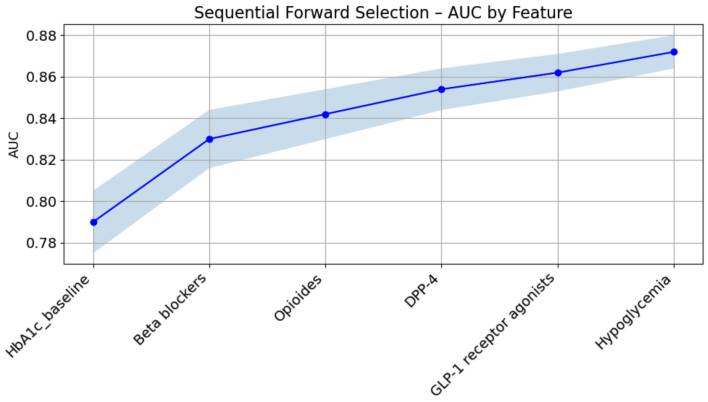



**Figure. 2:** Forward selection showing the number of features included against the averaged increase in AUC with standard deviation from five-fold cross-validation.

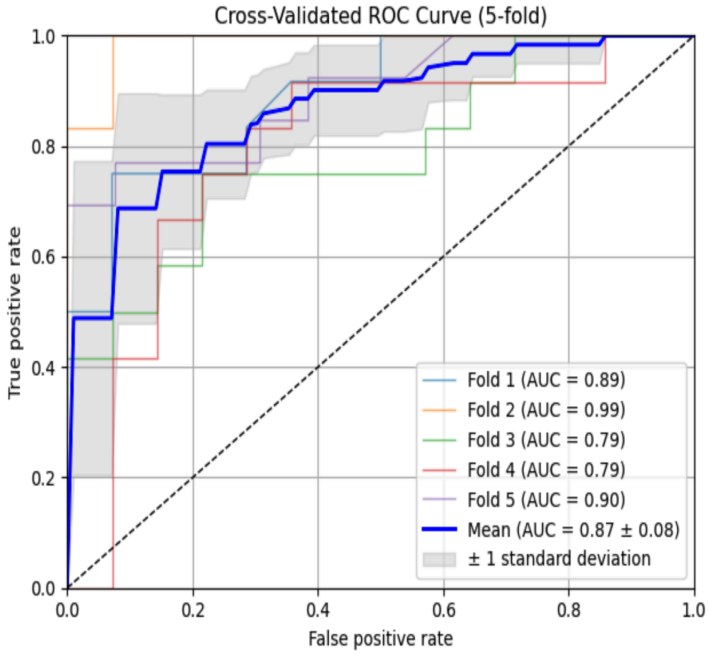



**Figure 3:** Receiver operating characteristics for each of the five-fold cross-validation along with the mean curve and standard deviation.

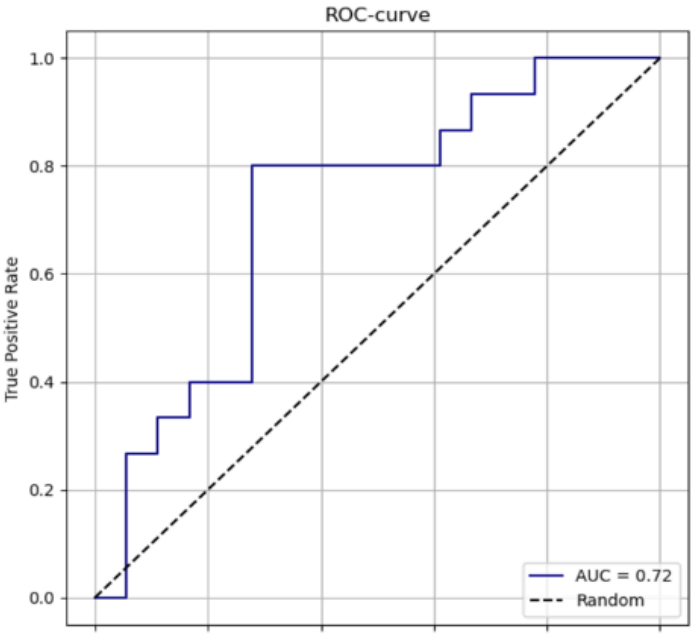



**Figure 4:** Receiver operating characteristics curve for the test set.

## Discussion

The aim of this study was to develop a predictive model capable of identifying individuals with T2DM most likely to benefit from telemonitoring, defined with an outcome of achieving ≥15% reduction in HbA1c. The final logistic regression model achieved an AUC of 0.87 on the training set, 0.72 on the test set, with a sensitivity of 0.80 and a specificity of 0.56. To the best of the authors knowledge, no previous studies have attempted to predict which individuals diagnosed with T2DM derive the greatest from telemonitoring.

Literature highlights the importance of relevant features in diabetic treatment. Lugner et al., identified the top ten predictors of T2DM and found that using only the top 10 features, including HbA1c as the most predictive feature, achieved an AUC of 0,88 (27). Similarly, Widiarti et al., showed that cardiovascular medications such as beta blockers can significantly

impact blood glucose levels in individuals diagnosed with T2DM (28). Our sequential feature selection process yielded six final features that each increased the average AUC ≥0.005. These findings align with the literature as the highest increase in AUC was seen after adding HbA1c at baseline and individuals treated with beta blockers, suggesting strong predictive value from glycemic control and comorbidity features. Additionally, all six features represent clinically relevant markers known from the literature supporting the validity of our results (29,30).

The level of performance indicates a moderate discriminative ability. The Receiver Operating Characteristics (ROC) curve (figure 4) shows a reasonable trade-off between sensitivity and specificity with a clear separation from the discrimination line. A rise in the early stage indicates that high sensitivity can be achieved while maintaining a relatively low false positive rate. Given our aim to identify individuals with T2DM with high benefit from telemonitoring, a sensitivity of 0.8 indicates that the model is reasonably effective in detecting the intended target group. Wild et al. supports this in a RCT study, highlighting that telemonitoring improves glycemic control in individuals with T2DM, underscoring the importance of sensitivity in clinical support (31).

While the model does not reach high performance (AUC>0.8), it shows consistent ability above chance level to identify patients with a high benefit from telemonitoring which supports its potential use as a tool for clinical decision support in healthcare (32). Our model performance is comparable to similar models developed within the scientific field of DM. Nagaraj et al. developed a ML model to predict short- and long-term HbA1c response insulin initiation in individuals with T2DM (30). The model achieved AUCs of 0.8 and 0.81 for short- and long-term prediction, demonstrating the feasibility of such models in long-term risk stratification (30). Furthermore, Fu et al. compared several ML models including logistic regression, to predict glycemic control over a 52-week period in individuals with T2DM (33). The logistic regression model achieved an AUC-score of 0.68 which indicated that the model could support individualized treatment planning and improve care making in clinically relevant even with a moderate AUC-score (33). These findings support that predictive models with an AUC of 0.70-0.80 may be valuable tools in clinical decision-making. However, Zale et al. argue that despite a moderate performance, such models may offer limited benefits if not integrated into real-time clinical workflows (34). Based on the ROC characteristics and an AUC of 0.72, the model appears clinically useful as a support tool for early
risk stratification, particularly in settings where sensitivity is prioritized.

The drop in AUC-performance from training to test indicates limited generalizability to new unseen data, described in the literature as a possible result of overfitting (35,36). Martin et al. recommends larger, more heterogeneous datasets to improve generalizability and reduce the risk of overfitting (37). However, the risk of overfitting is not solely

dependent on sample size; factors such as model complexity and the feature selection strategy also have substantial impact (37). Subtle shifts in population characteristics, including HbA1c outliers, may have influenced our results, as shown by the Kernel Density plot with a few extreme values (Figure 1). Moreover, a notable observation from the plot is that a substantial proportion of patients exhibit relatively low HbA1c levels at baseline, indicating no need for telemonitoring to lower their HbA1c level and are unlikely to achieve a 15% reduction (8). This indicates a potential risk of distributional imbalance between the training and test set, which may have resulted in overfitting the model and contributed to the drop in AUC performance. To potentially prevent overfitting it could be relevant to modify the outcome to a lower percentage in HbA1c reduction. A meta-analysis by Giugliano et al. demonstrated that even small reductions in HbA1c are clinically relevant, particularly in individuals with lower baseline levels (38). This supports the idea that lowering the outcome, potentially could include more individuals who still have a smaller, yet clinically significant effect from a better glycemic control, potentially enhancing both model sensitivity, specificity and risk of overfitting (38). However, maintaining an outcome of 15%, while reaching an acceptable sensitivity of 0.8 it was considered acceptable in order to identify those most likely to benefit from telemonitoring.

In this study logistic regression was chosen as the supervised learning algorithm. However, with our classification problem it could also be interesting to apply other ensemble models such as XGBoost or Random Forest (39,40). These models can be more difficult to interpret than logistic regression, but their performance can be superior. Noticeably, a recent DM prediction study demonstrated that ensemble models outperformed the method used in this study in terms of AUC, even when trained on a limited dataset (41). Nevertheless, it is still considered that using logistic regression was the appropriate choice for this project due to its simplicity, suitability for binary classification and ability to perform reliably on small datasets (42). These characteristics make logistic regression particularly relevant in clinical research where model transparency is essential (42).

The outcome was set as a binary variable. However, changing this to a continuous measure such as a significant change in HbA1c over time could provide a more clinically relevant prediction for individuals with a moderate improvement who did not achieve 15% reduction. Using such a regression outcome could potentially preserve more information, allowing the model to detect more subtle patterns and variations in the data (43). However, this would require changing the methodological approach of using logistic regression as this model is not appropriate for continuous outcomes. In that case, regression based models such as XGBoost and random forest would be more suitable.

A limitation of this study was a relatively small sample size of individuals (n=163), which may lead to an overoptimistic model performance (44). Furthermore, this could affect the models external validity, raising questions about its performance on

larger and more heterogeneous populations, whereas this can be interesting to explore in future research.

Another limitation was the necessity of imputing data, which could potentially lead to bias and hereby reduce the generalizability of the results potentially affect the model performance (45).

## Conclusion

A predictive ML model based on data from the DiaMonT trial was successfully developed achieving a moderate performance in predicting which individuals diagnosed with T2DM achieves >15% in HbA1c reduction from telemonitoring. These findings suggest potential for supporting individualized targeted treatment strategies in T2DM care. However, the results should be interpreted with caution, and further research is required to explore whether ML can become a reliable and independent tool for clinical decision-making.

## Abbreviations

BMI, body mass index; T2D, type 2 diabetes; CGM, continuous glucose monitoring; ROC, receiver operating characteristic;

AUC, area under the receiver operating characteristics curve.

## Acknowledgment

The authors would like to show their gratitude towards Stine Hangaard and Thomas Kronborg at Aalborg University, The Department of Health Science and Technology for providing the data on which this study was based.

## Declaration of Conflicting Interests

With respect to the research, authorship and/or publication of this article the authors report no conflicts of interest related to this.

## Funding

The authors received no financial support during the research period.

## References

1. Global diabetes data report 2000 — 2045 [Internett]. [sitert 7. mars 2025]. Tilgjengelig på: https://diabetesatlas.org/data/

2. Global Increase in Diabetes Prevalence Imposes a Substantial Health and Economic Burden | Published by Journal of Health Economics and Outcomes Research [Internett]. [sitert 30. mai 2025]. Tilgjengelig på: https://jheor.org/post/1265-global-increase-in-diabetes-prevalence-imposes-a-substantial-health-and-economic-burden

3. Schuster DP, Duvuuri V. Diabetes mellitus. Clinics in Podiatric Medicine and Surgery. 1. januar 2002;19(1):79–107.

4. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, mfl. Type 2 diabetes mellitus. Nat Rev Dis Primers. 23. juli 2015;1(1):1–22.

5. Norton L, Shannon C, Gastaldelli A, DeFronzo RA. Insulin: The master regulator of glucose metabolism. Metabolism. 1. april 2022;129:155142.

6. Diabetesforeningen [Internett]. [sitert 7. april 2025]. Langtidsblodsukker og type 2-diabetes. Tilgjengelig på: https://diabetes.dk/din-diabetes/type-2-diabetes/blodsukker-og-maling/langtidsblodsukker-hba1c/

7. Rawshani A, Rawshani A, Franzén S, Sattar N, Eliasson B, Svensson AM, mfl. Risk Factors, Mortality, and Cardiovascular Outcomes in Patients with Type 2 Diabetes. N Engl J Med. 16. august 2018;379(7):633–44.

8. Adler A, Bailey C, Day C, Colagiuri S. International Diabetes Federation. 2017. IDF Clinical Practice Recommendations for managing Type 2 Diabetes in Primary Care. Tilgjengelig på: https://idf.org/media/uploads/2023/05/attachments-63.pdf

9. Ashraf H, Faraz A, Ahmad J, Kohkan G. Achievement of guideline targets of glycemic and non-glycemic parameters in North Indian type 2 diabetes mellitus patients: A retrospective analysis. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 1. januar 2021;15(1):425–31.

10. Edelman SV, Polonsky WH. Type 2 Diabetes in the Real World: The Elusive Nature of Glycemic Control. Diabetes Care. 11. august 2017;40(11):1425–32.

11. Andrès E, Meyer L, Zulfiqar AA, Hajjam M, Talha S, Bahougne T, mfl. Telemonitoring in diabetes: evolution of concepts and technologies, with a focus on results of the more recent studies. J Med Life. 2019;12(3):203–14.

12. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, mfl. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. august 2019;42(8):1593–603.

13. Beck RW, Riddlesworth T, Ruedy K, Ahmann A, Bergenstal R, Haller S, mfl. Effect of Continuous Glucose Monitoring on Glycemic Control in Adults With Type 1 Diabetes Using Insulin Injections: The DIAMOND Randomized Clinical Trial. JAMA. 24. januar 2017;317(4):371–8.

14. Lau D, Manca DP, Singh P, Perry T, Olu-Jordan I, Ryan Zhang J, mfl. The effectiveness of continuous glucose monitoring with remote telemonitoring-enabled virtual educator visits in adults with non-insulin dependent type 2 diabetes: A randomized trial. Diabetes Research and Clinical Practice. 1. november 2024;217:111899.

15. Pettus JH, Zhou FL, Shepherd L, Preblick R, Hunt PR, Paranjape S, mfl. Incidences of Severe Hypoglycemia and Diabetic Ketoacidosis and Prevalence of Microvascular Complications Stratified by Age and Glycemic Control in U.S. Adult Patients With Type 1 Diabetes: A Real-World Study. Diabetes Care. desember 2019;42(12):2220–7.

16. Mengesha ME. ResearchGate. 2025 [sitert 12. mars 2025]. (PDF) The Cost-effectiveness of a Continuous Glucose Monitoring Device for Adult Diabetes Patients in Ethiopia: A Semi-Markov Modelling Study. Tilgjengelig på: https://www.researchgate.net/publication/387743947_The_Cost-effectiveness_of_a_Continuous_Glucose_Monitoring_Device_for_Adult_Diabetes_Patients_in_Ethiopia_A_Semi-Markov_Modelling_Study

17. Wright EE, Miller E, Bindal A, Poon Y. Addition of continuous glucose monitoring to glucagon-like peptide 1 receptor agonist treatment for type 2 diabetes mellitus - An economic evaluation. J Manag Care Spec Pharm. 1. februar 2025;31(2):127–36.

18. Giammarino F, Senanayake R, Prahalad P, Maahs DM, Scheinker D. A Machine Learning Model for Week-Ahead Hypoglycemia Prediction From Continuous Glucose Monitoring Data. J Diabetes Sci Technol [Internett]. 2024;((Giammarino F.) PE, Spoltore, Italy). Tilgjengelig på: https://www.embase.com/search/results?subaction=viewrecord&id=L2028909970&from=export

19. Alexiadis A, Tsanas A, Shtika L, Efopoulos V, Votis K, Tzovaras D, mfl. Next-Day Prediction of Hypoglycaemic Episodes Based on the Use of a Mobile App for Diabetes Self-Management. IEEE Access. 2024;12:7469–78.

20. Luong A, Cheung J, McMurtry S, Nelson C, Najac T, Ortiz P, mfl. Comparison of Machine Learning Models to a Novel Score in the Identification of Patients at Low Risk for Diabetic Retinopathy. Ophthalmol Sci [Internett]. 2025;5(1). Tilgjengelig på: https://www.embase.com/search/results?subaction=viewrecord&id=L2034756057&from=export

21. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, mfl. Predicting the risk of developing diabetic retinopathy using deep learning. Lancet Digit Heal. 2021;3(1):e10–9.

22. Kronborg T, Hangaard S, Hejlesen O, Vestergaard P, Jensen MH. The potential of predicting nocturnal hypoglycaemia for insulintreated individuals with type 2 diabetes in telemonitoring. Diabetologia. 2023;66((Kronborg T.; Hangaard S.; Hejlesen O.; Jensen M.H.) Department of Health Science and Technology, Aalborg University, Aalborg, Denmark):S379–80.

23. Hangaard S, Kronborg T, Hejlesen O, Aradóttir TB,

Kaas A, Bengtsson H, mfl. The Diabetes teleMonitoring of patients in insulin Therapy (DiaMonT) trial: study protocol for a randomized controlled trial. Trials. 7. desember 2022;23:985.

24. scikit-learn [Internett]. [sitert 9. mai 2025]. IterativeImputer. Tilgjengelig på: https://scikit-learn/stable/modules/generated/sklearn.impute.IterativeImputer.html

25. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology. 1. august 2001;54(8):774–81.

26. Kotsiantis SB. Supervised Machine Learning: A Review of Classification Techniques. Tilgjengelig på: https://datajobs.com/data-science-repo/Supervised-Learning-%5BSB-Kotsiantis%5D.pdf

27. Lugner M, Rawshani A, Helleryd E, Eliasson B. Identifying top ten predictors of type 2 diabetes through machine learning analysis of UK Biobank data. Sci Rep. 24. januar 2024;14(1):2102.

28. Widiarti W, Saputra PBT, Savitri CG, Putranto JNE, Alkaff FF. The impact of cardiovascular drugs on hyperglycemia and diabetes: a review of 'unspoken' side effects. Hellenic Journal of Cardiology. 1. mai 2025;83:71–7.

29. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A Systematic Review of Reviews Evaluating Technology-Enabled Diabetes Self-Management Education and Support. J Diabetes Sci Technol. september 2017;11(5):1015–27.

30. Nagaraj SB, Sidorenkov G, van Boven JFM, Denig P. Predicting short‐ and long‐term glycated haemoglobin response after insulin initiation in patients with type 2 diabetes mellitus using machine‐learning algorithms. Diabetes Obes Metab. desember 2019;21(12):2704–11.

31. Wild SH, Hanley J, Lewis SC, McKnight JA, McCloughan LB, Padfield PL, mfl. Supported Telemonitoring and Glycemic Control in People with Type 2 Diabetes: The Telescot Diabetes Pragmatic Multicenter Randomized Controlled Trial. PLOS Medicine. 26. juli 2016;13(7):e1002098.

32. Zimmerman RM, Hernandez EJ, Tristani-Firouzi M, Yandell M, Steinberg BA. Explainable artificial intelligence for stroke risk stratification in atrial fibrillation. European Heart Journal - Digital Health. 20. mai 2025;6(3):317–25.

33. Fu X, Wang Y, Cates RS, Li N, Liu J, Ke D, mfl. Implementation of five machine learning methods to predict the 52-week blood glucose level in patients with type 2 diabetes. Front Endocrinol [Internett]. 20. januar 2023 [sitert 23. mai 2025];13. Tilgjengelig på: https://www.frontiersin.org/articles/10.3389/fendo.2022.1061507/full

34. Zale A, Mathioudakis N. Machine Learning Models for Inpatient Glucose Prediction. Curr Diab Rep. 1. august 2022;22(8):353–64.

35. Hu Z, Li X, Yuan Y, Xu Q, Zhang W, Lei H. Development and validation of machine learning models for predicting venous thromboembolism in colorectal cancer patients: A cohort study in China. International Journal of Medical Informatics. 1. mars 2025;195:105770.

36. Mbah C, Thierens H, Thas O, Neve JD, Chang-Claude J, Seibold P, mfl. Pitfalls in Prediction Modeling for Normal Tissue Toxicity in Radiation Therapy: An Illustration With the Individual Radiation Sensitivity and Mammary Carcinoma Risk Factor Investigation Cohorts. International Journal of Radiation Oncology, Biology, Physics. 1. august 2016;95(5):1466–76.

37. Martin PP, Graulich N. Navigating the data frontier in science assessment: Advancing data augmentation strategies for machine learning applications with generative artificial intelligence. Computers and Education: Artificial Intelligence. 1. desember 2024;7:100265.

38. Giugliano D, Maiorino M, Bellastella G, Chiodini P, Esposito K. Relationship of baseline HbA1c, HbA1c change and HbA1c target of < 7% with insulin analogues in type 2 diabetes: a meta-analysis of randomised controlled trials. International Journal of Clinical Practice. 2011;65(5):602–12.

39. Lev A. XGBoost versus Random Forest | JFrog ML [Internett]. 2022 [sitert 23. mai 2025]. Tilgjengelig på: https://www.qwak.com/post/xgboost-versus-random-forest

40. Aguilera-Venegas G, López-Molina A, Rojo-Martínez G, Galán-García JL. Comparing and tuning machine learning algorithms to predict type 2 diabetes mellitus. Journal of Computational and Applied Mathematics. 1. august 2023;427:115115.

41. Sampath P, Elangovan G, Ravichandran K, Shanmuganathan V, Pasupathi S, Chakrabarti T, mfl. Robust diabetic prediction using ensemble machine learning models with synthetic minority over-sampling technique. Sci Rep. 22. november 2024;14(1):28984.

42. I.Olufemi, C.Obunadike, A. Adefabi, D. Abimbola. Application of Logistic Regression Model in Prediction of Early Diabetes Across United States. IJSMR. 2023;06(05):34–48.

43. Hoskin T. Data Types [Internett]. Mayo Clinic Department of Health Sciences Research; [sitert 30. juni 2025]. Tilgjengelig på: https://www.mayo.edu/research/documents/data-types/doc-20408956

44. Flint C, Cearns M, Opel N, Redlich R, Mehler DMA, Emden D, mfl. Systematic misestimation of machine learning performance in neuroimaging studies of depression. Neuropsychopharmacology. juli 2021;46(8):1510–7.

45. Brownlee J. Iterative Imputation for Missing Values in Machine Learning [Internett]. [sitert 23. mai 2025]. Tilgjengelig på: https://machinelearningmastery.com/iterative-imputation-for-missing-values-in-machine-learning/

# Table of contents

# Worksheet 1

The elaboration of the introduction includes a description of diabetes mellitus (DM) and an in-depth explanation of the underlying mechanisms. Finally, there will be a description of the treatment for type 2 diabetes including telemonitoring and artificial intelligence.

## 1.0 Introduction

### 1.1 Incidence and prevalence

According to the International Diabetes Federation (IDF), in 2021 537 million people were diagnosed with DM worldwide. This number is predicted to rise exponentially and is estimated to be 643 million by 2030 and almost 800 million by 2045. In 2021, over 6.7 million people died from DM, making this disease the leading cause of death in the world (1). In Denmark approximately 310.000 people were diagnosed with DMs in 2021. Furthermore, it is estimated that around 100,000 people live with DM but are not yet diagnosed. Similar to global diabetes-related death rates, this disease is also a leading cause of death in Denmark, claiming approximately 8,000 lives in 2021 (2).

### 1.2 Health economics

IDF claims that globally DM related health expenditure is approximately 966.000 USD million. This amounts to around 1,838 USD per person (1). Put into perspective, this is equivalent to approximately 50% of an average European citizen's monthly salary. In Denmark, the total expenditure per year was in 2021 estimated to be 2,426 USD million. This amounts to around 7,844 USD per person which is 4,3 times higher than the global average individual (2).

### 1.3 Definition of diabetes mellitus

DM, is a group of metabolic disorders characterized by hyperglycemia, which is caused by defects in insulin action, insulin secretion or both. The chronic hyperglycemia of DM is usually associated with long-term dysfunction, damage, and failure of different organs especially the eyes, nerves, kidneys, heart, and blood vessels (3).

In healthy individuals, blood glucose levels are maintained within a small range, typically between 3.9 and 7.8 mmol/L, depending on fasting or postprandial states (4) (5). This balance is regulated through the interplay between insulin and glucagon. Insulin lowers blood glucose by promoting glucose uptake in cells, whereas glucagon increases blood glucose by stimulating glycogenolysis and gluconeogenesis in the liver. Norton et al. (2022) highlights insulin as the primary regulator of glucose metabolism, emphasizing that an imbalance between insulin and glucagon, as seen in individuals with

DM, leads to persistent hyperglycemia and metabolic complications (4). Hyperglycemia can have significant long-term consequences (6). Glycated hemoglobin (HbA1c) works as a central marker for assessing blood glucose control over a longer period. It is often measured using the average blood glucose concentration over approximately three-four months. In individuals with DM, HbA1c is a critical indicator, which not only confirms poor glycemic regulation when elevated, but higher values of HbA1c is also strongly associated with serious complications such as cardiovascular disease, retinopathy and nephropathy. According to the Danish Diabetes Association, the national target for long term HbA1c levels is set between 48 mmol/mol for those with type 2 diabetes mellitus (T2DM) and 53 mmol/mol for those with type 1 diabetes mellitus (T1DM) (6). The International Diabetes Federation recommends a general target of HbA1c levels <53 mmol/mol for individuals diagnosed with T2DM (7)

## 1.4 Diabetes classifications

DM consists of various subgroups based on the pathogenesis of the disease rather than its treatment. These subgroups have been continuously revised, most recently in 2024 by the American Diabetes Association (ADA) (8). DM is categorized into the following groups: T1DM, T2DM, Other Specific Types of Diabetes, and Gestational Diabetes (9).

### 1.4.1 Type 1 Diabetes Mellitus

T1DM is a chronic autoimmune disease characterized by an progressive destruction of insulin-producing b-cells in the pancreas, leading to insulin deficiency and hyperglycemia (10). In healthy individuals, these cells are responsible for producing insulin, which regulates the blood glucose levels (11). The disease is primarily mediated by self-reactive lymphocytes, particularly T-cells, which infiltrate the pancreatic islet and contribute to B-cell destruction (11) (12).

Most cases of T1DM are autoimmune (Type 1a), while a smaller subgroup of individuals is present with idiopathic T1DM (Type 1b), where autoantibodies are detectable (10).

The exact cause of T1DM remains unclear, but it is believed to arise from a combination of genetic predisposition and environmental factors. The autoimmune process often begins years before any clinical symptoms appear, typically with the presence of autoantibodies, subsequent B-cell dysfunction which leads to hyperglycemia (10) (12).

### 1.4.2 Type 2 Diabetes Mellitus

T2DM is a chronic metabolic disorder characterized by elevated blood glucose levels. The increase in glucose levels is due to a combination of insulin resistance and dysfunction of insulin producing B-cells (13), which over time becomes less effective, leading to inadequate insulin production and hyperglycemia (13). In the early stages of T2DM, peripheral tissues such as skeletal muscle and the

liver become less responsive to insulin, impairing the uptake of glucose and increasing hepatic glucose production (14). The pancreatic B-cells enhance the insulation secretion to maintain normal blood glucose levels. However, over time, this compensatory response fails due to the progressive loss of B-cells (14). In addition to insulin resistance and B-cell dysfunction, individuals also experience an increase in lipolysis in adipose tissue contributing to a metabolic dysregulation due to the release of fatty acids leading to an exacerbate insulin resistance and impairment of B-cells (14). Chronic inflammation and oxidative stress also plays a role in the blood glucose regulation as pro-inflammatory cytokines and reactive oxygen species contribute to both insulin resistance and B-cell reduction, further reducing insulin production (14).

Another factor in T2DM is the dysregulation of incretin hormones as glucagon-like peptide-1 (GLP-1), which normally enhances insulin secretion in response to food intake. This leads to a reduced insulin secretion and inadequate suppression of glucagon (14).

### 1.4.3 Gestational diabetes and other types

This classification of DM includes several distinct subcategories that do not fit within the requirements of T1DM and T2DM. These forms often arise from genetic mutations, diseases affecting the pancreas, or hormonal disorders (9). A well-known example is Maturity-Onset Diabetes of the Young (MODY), which is a monogenic form caused by mutations in specific genes that regulate insulin secretion (9). Another category includes secondary DM caused by conditions such as pancreatitis, Cushing's syndrome, or the prolonged use of certain medications (9). DM can also occur during pregnancy known as gestational diabetes (9). This typically resolves postpartum and is associated with an increased risk of developing T2DM in the future (9).

Due to the numerous forms of DM, this project cannot encompass all diagnoses. Therefore, the study focuses specifically on individuals with T2DM. Consequently, the following sections will exclusively address this diagnosis.

## 1.5 Risk factors of developing T2DM

According to a systematic review performed by Dendup et al, the development of T2DM can be influenced by various risk factors (15). The review states that one of the most significant predictors of T2DM is obesity and overweight, as excess body fat, visceral fat in particular, contributes to insulin resistance and impaired glucose metabolism (15). Closely related to this comes physical activity whereas an inactive lifestyle reduces the body's ability to utilize glucose effectively, leading to a potential metabolic dysregulation (15). Dietary habits also play a crucial role, with diets high in refined carbohydrates, processed foods and unhealthy fats strongly associated with an increased risk (15). Additionally, other behavioural factors such as smoking and alcohol consumption, have been

linked to a higher risk of DM as these contribute with an exacerbating inflammatory response that can impair insulin signaling (15). Other socioeconomic disparities can contribute in the development such as low income, low education levels, limited access to healthy food options and healthcare services (15). The above-mentioned risk factors are further stated in a review by Bellou et al. which also expands on these findings by identifying additional non-genetic factors (16). The review highlights the significance of different serum biomarkers, personality traits such as low conscientiousness and exposure to air pollution and certain medical conditions, including hypertension, metabolic syndrome, gestational DM, and preterm birth (16).

## 1.6 Complications of T2DM

T2DM can lead to dangerous complications, which are categorized into micro- and macrovascular complications (17). Microvascular complications are defined by affecting the small blood vessels and can lead to diabetic retinopathy, nephropathy, and neuropathy (18). Macrovascular complications impact the larger blood vessels, which increases the risk of ischemic heart disease, potentially leading to angina pectoris, heart attacks, and heart failure (18). Additionally, individuals with T2DM have a higher risk of stroke and peripheral atherosclerosis, potentially causing claudication, chronic foot ulcers, and amputations (18).

The progression of these complications depends on the level of risk factors such as elevated blood glucose, blood pressure, and cholesterol (18). In the prevention of complications, it is highly relevant for individuals to maintain optimal blood glucose levels, blood pressure, and lipid levels through lifestyle modifications, including healthy diet, physical activity, and smoking cessation (18).

## 1.7 Treatment of T2DM

The treatment of T2DM focuses on improving insulin resistance and preserving or restoring beta-cell function in the pancreas. Modifications in lifestyle, such as weight loss and increased physical activity, are crucial for enhancing insulin sensitivity and reducing the burden on beta cells (19). T2DM can also be treated through pharmacological interventions, including metformin and other antidiabetic agents, used to improve insulin resistance and stimulate insulin secretion (19). In special cases, bariatric surgery may be considered to achieve significant weight loss and potentially induce DM remission (19).

A growing and modern approach in T2DM management is telemonitoring, which allows the opportunity for individuals to track their blood glucose levels in real time. Telemonitoring uses technology to remotely track health, enhancing DM management and reducing complications (20) . A systematic review performed by Zhu et al. revealed that telemonitoring led to a significant reduction in HbA1C levels over a period of 180 days. Methods of telemonitoring in individuals with T2DM are

various whereas SMS-based telemonitoring, mobile applications, integrated eHealth platforms and automated remote monitoring are among the most used tools (20). One of the most popular tools is Continuous Glucose Monitoring (CGM), which includes a subcutaneous sensor to continuously measure glucose levels (20). This enables individuals to detect hyper- and hypoglycemia and research has shown that CGM may significantly lower HbA1c levels (21) (22).

Despite its many advantages, challenges still remain on the documented effect. The review by Zhu concluded that telemonitoring had a significant reduction in blood pressure but not of clinical relevance, no effect on BMI/weight, no effect on fasting blood glucose (23). Even though the review documents an effect on HbA1c, the effect was most pronounced in studies with shorter duration (< 6 months) and smaller sample sizes, suggesting that the impact may diminish over time (23). Furthermore there remains complications with the integration of telemonitoring in individuals (23) with DM, particularly regarding cost. The devices are expensive, especially for individuals without insurance or government coverage (20). In relation to the high costs, several studies emphasize that it is important to identify which subgroups benefit the most from telemonitoring (24) (25). Pettus et al. highlights the difficulty of identifying which individuals truly benefit from telemonitoring, making its widespread use hard to justify given the high costs (26). Variability in adherence, lifestyle, and healthcare access complicates selection, risking inefficiency and inequitable access. Therefore, innovative technologies, such as AI, have been identified as potential tools to improve prediction and optimize the selection of telemonitoring, making it easier to determine who will benefit the most (27) (25).

## 1.8 The potential of machine Learning in T2DM

Machine learning (ML) has been increasingly applied in DM research, primarily in predictive models, patient management and risk assessment. A study performed by Giammarino et al. built a ML-model to predict the likelihood of hypoglycemic events occurring within a week based on data from CGM which demonstrated an area under the curve (AUC) of 0.74 across multiple cohorts, which correlates with a moderate predictive accuracy (28). Another approach was seen in a study performed by Alexiadis et al. which used ML in predicting next-day hypoglycemia in individuals with T2DM using data collected from a self-managed mobile app (29). The results showed an accuracy of 81.4% based on a Random Forest test, highlighting the potential of mobile health integration in DM management (29).

ML has also been used in the screening of diabetic retinopathy. Luong et al. conducted a study comparing ML models with a Retinopathy Risk Score (RSS) designed to identify individuals at low risk for diabetic retinopathy (30). Findings showed that the RSS performed comparably to four out of five ML models. Another study by Bora et al. developed a learning system that used color fundus photographs to predict the risk of developing retinopathy within two years (31). Results showed that

the model achieved an AUC of 0.79, improving risk stratification when combined with existing clinical factors (31).

Another area of research focused on predicting nocturnal hypoglycemia in individuals with insulin-treated T2DM (32). The study developed by Kronborg et al. used a ML model, which used CGM data to predict nocturnal hypoglycemia (32). The model, trained with data from 67 individuals, achieved an AUC of 0.82, suggesting potential for real-world application (32).

Despite the increasing use of ML in DM treatment, several limitations may hinder a broader applicability. A bibliometric analysis performed by Garcia-Jaramillo et al. showed that most studies focused on detection, whereas management and treatment optimization were significantly underexplored (33). Similarly, a review by Eghbali-Zarch & Masoud investigated applications in insulin management, identifying challenges related to affordability and accessibility (34).

Although a substantial amount of literature was found on ML and prediction in DM, it is primarily focused on hypoglycemia risk assessment and retinopathy detection. Based on the current knowledge of the project, it has not been possible to identify any literature investigating the prediction of which subgroup of individuals with T2DM benefits the greatest from telemonitoring. Given that telemonitoring, in particular CGM, has shown to improve glycemic control and reduce complications, its implementation should be targeted towards individuals, who derive the greatest clinical and economic benefit (21) (20). The high cost of CGM and variability in individual adherence makes it crucial to develop predictive models that can identify the most suitable individuals for telemonitoring (24).

This gap is crucial to address as it could enhance healthcare resource allocation, ensuring that telemonitoring is used where it is most effective, maximizing clinical outcomes and cost-efficiency.

# 2.0 Systematic literature search

The structured literature search has formed the basis for the development of section 1.8. The execution and structure of the search are described below.

## 2.1 purpose of the literature search

The purpose of the structured literature search was to identify scientific research and explore the state of art in prediction models and ML within DM.

## 2.2 Scientific databases

The structured literature search was conducted in three selected databases: 1) PubMed, 2) Embase, 3) Scopus. These databases were chosen based on their content related to the project, which is visualized

in Table 1. A combination of results from these databases was considered sufficient in uncovering existing literature from the field.

| SCIENTIFIC DATABASES | |
|---|---|
| **DATABASE** | **DESCRIPTION OF THE DATABASE'S CONTENT** |
| **PUBMED** | The database contains over 37 million citations and abstracts related to biomedical literature. |
| **EMBASE** | The database contains references to studies related to medicine and biomedicine. |
| **SCOPUS** | Large scientific database covering multidisciplinary fields, including medicine and natural sciences. |

**Table 1:** Database overview and description

## 2.3 Inclusion and exclusion criteria

Predefined inclusion and exclusion criteria were developed to ensure the relevance of the literature related to the research topic. The criteria used in the structured literature search are illustrated in Table 2

| INCLUSION CRITERIA | EXCLUSION CRITERIA |
|---|---|
| Language: Danish, Norwegian, Swedish & English<br><br>Timespan: 2000-2025<br><br>Peer-reviewed articles<br><br>Studies evaluating the use of ML for prediction in the context of type 2 diabetes. | Studies missing abstracts<br><br>No acces to full text studies |

**Table 2:** Showing in- and exclusion criteria for the systematic literature research

## 2.4 Search block

The search strategy and selected databases were designed to increase the reliability of the results by maintaining a standardized approach (35). Variations in the database's thesauri resulted in structural differences in the searches. In accordance with the predefined inclusion criteria, restrictions of language and timespan were applied as a filter for each database. The search terms were structured using a block search approach, categorized into Population, Intervention, and Outcome. See Table 3.

| POPULATION | INTERVENTION | OUTCOME |
|---|---|---|
| Diabetes mellitus, type 2 [Thesaurus]<br>Diabetes (free text)<br>Type 2 diabetes (free text)<br>t2dm (free text)<br>Diabet* | Machine Learning [Thesaurus]<br>Forecasting [Thesaurus]<br>Predict* modeling (free text) | Telemedicine [Thesaurus]<br>Telehealth [Thesaurus]<br>Telemonitoring (free text) |

**Table 3:** Systematic block search approach

## 2.5 Search strategy

To ensure relevance and a strong foundation to our systematic literature search, relevant free-text and thesaurus terms were extracted through an initial unstructured literature search process (36). Afterwards, the identified terms were then used as the basis for the structured literature search which was designed as a block search with three blocks. The block search model was inspired by the PICO model but was simplified to PIO, excluding the comparison element, to ensure relevance to this project (36).

In each search block, both thesaurus and free-text terms were used in the selected databases. Additionally, truncations (*) were applied to ensure the inclusion of different variations of the specific search terms. Whenever relevant thesaurus terms, such as MeSH, were relevant, these were also incorporated into the search blocks to enhance relevant variations of the included thesaurus term (36).

## 2.6 Identified literature

A total of 1192 studies were identified through the systematic literature search in the three selected databases. Table 4 presents the number of studies identified from each database. Each database gave a different search string. Appendix 1 shows an example of a search string used in Scopus.

| IDENTIFIED STUDIES IN THE SELECTED DATABASES | |
|---|---|
| DATABASE | IDENTIFIED STUDIES |
| PUBMED | 30 |
| EMBASE | 252 |
| SCOPUS | 910 |
| **TOTAL** | **1192** |

**Table 4:** Total number of identified studies from selected databases

## 2.7 The screening process

A reference list was downloaded from each database and uploaded into the reference management tool, Zotero (37). The software initially identified and removed duplicates. The remaining studies were screened at the abstract/title level based on the in/exclusion criteria established by the group.

The total number of studies was divided into two groups in Zotero and distributed among the two group members, who conducted the screening independently. In cases of disagreement or uncertainty regarding a source, a note was added for further discussion in the group. The literature was next screened at full-text level to enhance the reliability and validity of the selection process. Any uncertainties regarding a study were resolved following the same procedure as mentioned above. The final studies were critically appraised using the respective "Joanna Briggs Institute Critical Appraisal Checklists (JBI) (Appendix 2) (38).

A flow diagram of the screening process was made according to the "Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (39). The PRISMA flow diagram visualizes the identification, screening, and inclusion of the literature. In total 8 studies were included in the project. PRISMA flow chart can be seen in Figure 2.

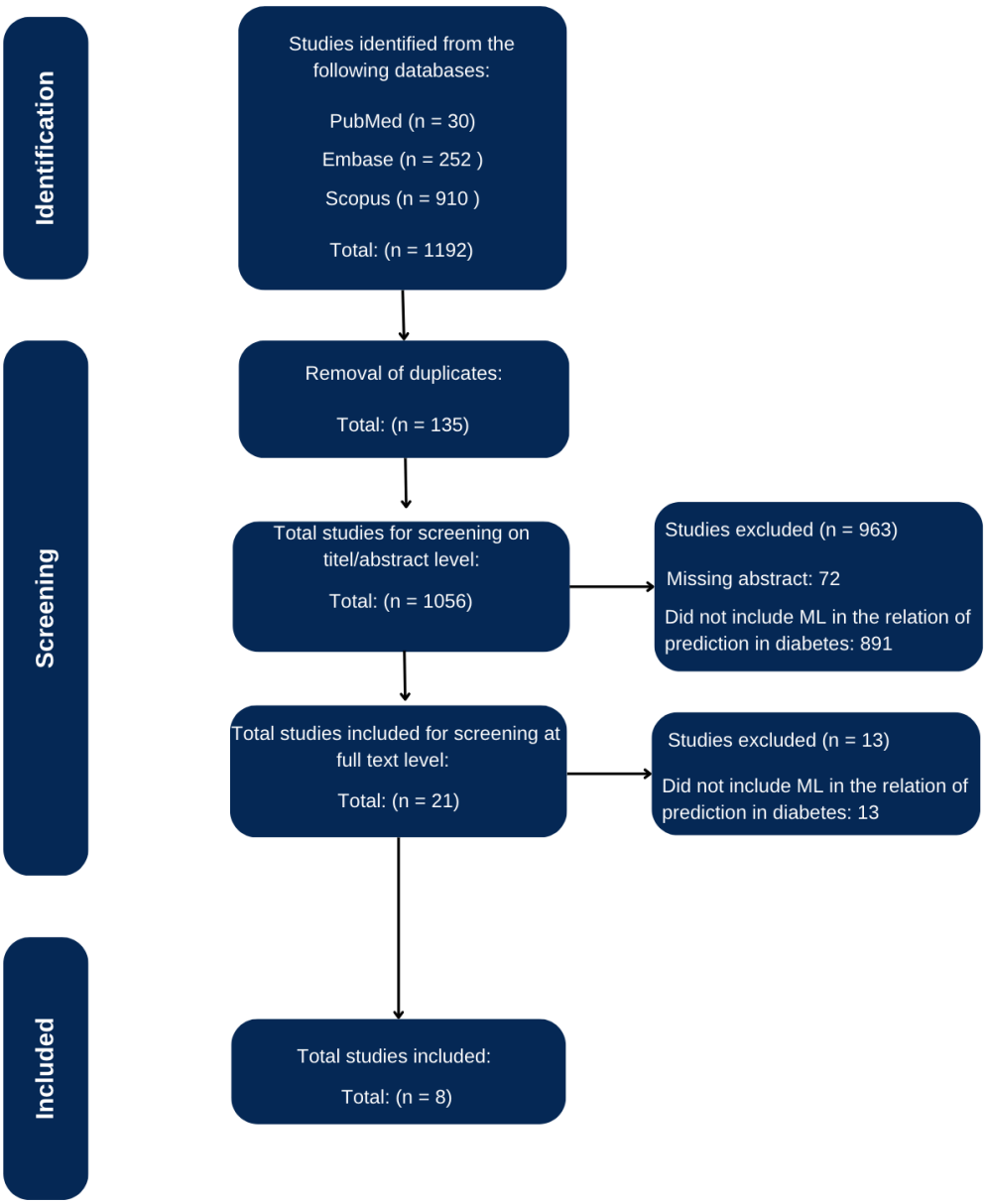



**Figure 2**: PRISMA flow chart

# Worksheet 2

## 3.0 Method overview

ML is a subfield within artificial intelligence focused on addressing the questions of developing and testing algorithms that can detect patterns in data and generate predictions using given information (40). Given the strong potential of ML to predict given outcomes it is particularly suitable and relevant for clinical prediction problems, such as estimating HbA1c in people living with T2DM. In this project the outcome variable, HbA1c_change, was predefined and a supervised learning approach was applied, allowing the model to learn the relationship between our data variables and the target

outcome (40). Logistic regression was used as the supervised learning algorithm to classify individuals' effect of telemonitoring based on their HbA1c_change.

The ML model developed in this project follows the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (41). In figure 3, a visualization of the methodological approach applied.

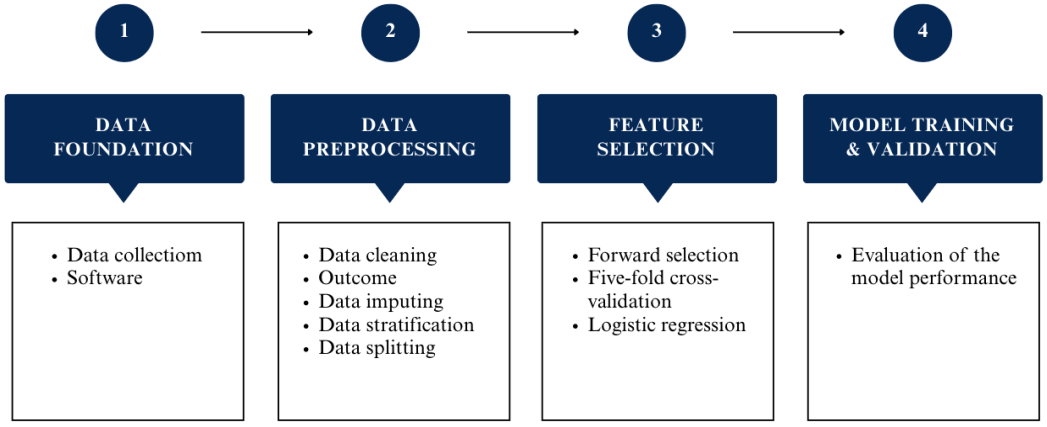



**Figure 3**: Overview of the machine learning pipeline illustrating the four main steps from data foundation to model validation.

## 3.1 Data foundation

Data were collected as a part of the DiaMonT trial, a randomized controlled study investigating telemonitoring in adults with insulin-treated T2D (42). Participants were recruited from Steno Diabetes Center North (Aalborg University Hospital) and Zealand (Nykøbing Falster Hospital) through endocrinology clinics, informational material, and diabetes-related organizations (42). Individuals were included if they had a confirmed T2D diagnosis for at least 12 months, were able to use a smartphone and trial-related devices, could read and understand Danish and were the age of minimum 18 years (42).

### 3.1.1 Software

For modelling and data preprocessing, the platform Anaconda was used to manage the Python environment and to launch JupyterNotebook. Anaconda is an open-source distributor used widely in the field of managing data in Python (43). The platform enables management of packages and environments, making it suitable for ML projects and data analysis. JupyterNotebook provides a flexible and interactive development environment where code, visualizations, and documentation can be combined in a single workflow (44). Pandas were imported for data handling, scikit-learn for ML, and matplotlib for visualisation in the same environment.

## 3.2 Data preprocessing

### 3.2.1 Data cleaning

Prior to analysis, several preprocessing steps were undertaken. Data from individuals enrolled at Aalborg and Falster centers were merged into one dataset containing all relevant clinical and questionnaire-based information collected at the beginning of the trial. Individuals from the control group were excluded in order to isolate the individuals from the intervention group. Furthermore, individuals who did not complete the full clinical trial of the intervention were also removed. All data were converted into numerical format allowing consistent interpretation. Repeated observations were merged so that each individual was represented by a single, structured row to allow the use of ML algorithms.

### 3.2.2 Definition of outcome

The outcome for the ML prediction model was defined as the percentage change in HbA1c levels from baseline testing to the final testing. The outcome was defined as a binary variable indicating either significant reduction or no significant reduction in HbA1c levels. Previous research on telemonitoring in individuals with T2DM has reported HbA1c reductions ranging from 6–7% up to 26% (45) (46), with several studies reporting reductions between 14% and 21% (47) (48) (49). Given this variation from clinical evidence, together with recommendations from national treatment guidelines for T2DM in Denmark, a conservative and evidence-based threshold in HbA1c reduction was set to be $\geq$15 % (6).

### 3.2.3 Data imputing

All missing values (NaN) in the final dataset were removed, as logistic regression cannot be performed on datasets containing missing values. NaN were located and imputed by using the “IterativeImputer" class from the “scikit-learn” library in Python. Iterative imputation was chosen because of its ability to model each feature with missing values as a function of the others. This method discovers multivariate and natural relationships within the data, typically resulting in more accurate and consistent imputations (50).

### 3.2.4 Data stratification and data splitting

To ensure that the data was proportionally split between the training and test sets, stratification was applied before splitting the data (51). A “change_group” column was created by dividing the percentage change in HbA1c levels into five bins, ensuring that all individuals were assigned to a defined group. The bins were: >20% decrease, 10–20% decrease, 0–10% decrease, 0–10% increase, and 10–20% increase. Dividing the data by using stratification helps preserve the distribution of these

groups across both training and test, preventing data imbalance, improving reliability and generalisability (51).

As the final step in preparing the data for ML, the dataset was split into training and test sets using an 80:20 ratio. The dataset was split after importing “train_test_split” from sklearn.model_selection, with “stratify=y” to preserve class distribution and “random_state=42” to ensure reproducibility. The distribution between the two subsets was verified by displaying the group proportions and dimensions directly in the Jupyter Notebook to ensure consistency in the split. “StandardScaler” was implemented and performed on training data to improve numerical stability during logistic regression (52).

## 3.3 Feature selection

Forward feature selection was performed on the training data. From the “sklearn.linear_model” pack, LogisticRegression was imported to ensure that the model was valid. Subsequently, the SequentialFeatureSelector was imported as the class to run the feature selection. The model evaluated combinations of features based on the Area Under the Curve (AUC) using five-fold cross-validation. At each iteration, a logistic regression model was evaluated using the currently selected features plus one additional candidate feature. Floating allowed previously included features to be removed if performance improved. The model’s AUC was calculated on the scaled training set, and the features leading to the highest performance improvement were retained. This process continued until no further gain in AUC over 0.005 was observed. For illustration a Receiver Operating Characteristics - Curve (ROC-Curve) was plotted.

### 3.3.1 Clinical relevance

To ensure that the final features were clinically relevant they were evaluated based on two criterias; their relevance to the aim of this study as documented in clinical studies and objective judgment regarding their relevance in the management of HbA1c in individuals with T2DM (53) (54). The final features are described in the literature as important for understanding and influencing HbA1c in individuals with T2DM (53) (54).

## 3.4 Model training & validation

The training and test datasets were reduced to include only the subset of features identified through sequential feature selection. To ensure consistent scaling while preventing data leakage, the training features were standardized using the “StandardScaler” from scikit-learn package, and to ensure that the test data was scaled similar, the test data was transformed the same way. Subsequently, a logistic regression model was trained on the training data including the right features. The model’s

performance was then evaluated on the independent test data using AUC as the primary metric for validation.

Model performance was assessed using sensitivity and specificity. A ROC-Curve was used to visualize the trade-off between sensitivity and 1-specificity across different thresholds. Sensitivity reflects the amount of individuals with high benefit from telemonitoring correctly identified, while specificity refers to correctly excluding those without benefit. The AUC shows the model´s ability to distinguish between those individuals who benefit from telemonitoring and those who do not. Higher AUC score indicates better overall model performance. In addition, a figure illustrating baseline and follow up HbA1c-distribution were analyzed.

# References


1. Global diabetes data report 2000 — 2045 [Internett]. [sitert 7. mars 2025]. Tilgjengelig på: https://diabetesatlas.org/data/

2. Denmark diabetes report 2000 — 2045 [Internett]. [sitert 7. mars 2025]. Tilgjengelig på: https://diabetesatlas.org/data/

3. Schuster DP, Duvuuri V. Diabetes mellitus. Clinics in Podiatric Medicine and Surgery. 1. januar 2002;19(1):79–107.

4. Norton L, Shannon C, Gastaldelli A, DeFronzo RA. Insulin: The master regulator of glucose metabolism. Metabolism. 1. april 2022;129:155142.

5. Katsarou A, Gudbjörnsdottir S, Rawshani A, Dabelea D, Bonifacio E, Anderson BJ, mfl. Type 1 diabetes mellitus. Nat Rev Dis Primers. 30. mars 2017;3(1):1–17.

6. Diabetesforeningen [Internett]. [sitert 7. april 2025]. Langtidsblodsukker og type 2-diabetes. Tilgjengelig på: https://diabetes.dk/din-diabetes/type-2-diabetes/blodsukker-og-maling/langtidsblodsukker-hba1c/

7. Adler A, Bailey C, Day C, Colagiuri S. International Diabetes Federation. 2017. IDF Clinical Practice Recommendations for managing Type 2 Diabetes in Primary Care. Tilgjengelig på: https://idf.org/media/uploads/2023/05/attachments-63.pdf

8. American Diabetes Association Professional Practice Committee, ElSayed NA, Aleppo G, Bannuru RR, Bruemmer D, Collins BS, mfl. 2. Diagnosis and Classification of Diabetes: *Standards of Care in Diabetes—2024*. Diabetes Care. 1. januar 2024;47(Supplement_1):S20–42.

9. Maraschin J de F. Classification of Diabetes. I: Diabetes [Internett]. Springer, New York, NY; 2013 [sitert 7. mars 2025]. s. 12–9. Tilgjengelig på: https://link-springer-com.zorac.aub.aau.dk/chapter/10.1007/978-1-4614-5441-0_2

10. American Diabetes Association. Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 1. januar 2010;33(Supplement_1):S62–9.

11. B cells in type 1 diabetes mellitus and diabetic kidney disease | Nature Reviews Nephrology [Internett]. [sitert 7. mars 2025]. Tilgjengelig på: https://www.nature.com/articles/nrneph.2017.138

12. Pugliese A. Pathogenesis of Type 1 Diabetes. I: Diabetes Epidemiology, Genetics, Pathogenesis, Diagnosis, Prevention, and Treatment [Internett]. Springer, Cham; 2018 [sitert 7. mars 2025]. s. 1–40. Tilgjengelig på: https://link-springer-com.zorac.aub.aau.dk/referenceworkentry/10.1007/978-3-319-27317-4_7-1

13. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, mfl. Type 2 diabetes mellitus. Nat Rev Dis Primers. 23. juli 2015;1(1):1–22.

14. Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, mfl. Pathophysiology of Type 2 Diabetes Mellitus. International Journal of Molecular Sciences. januar 2020;21(17):6275.

15. Dendup T, Feng X, Clingan S, Astell-Burt T. Environmental Risk Factors for Developing Type 2 Diabetes Mellitus: A Systematic Review. International Journal of Environmental Research and Public Health. januar 2018;15(1):78.

16. Risk factors for type 2 diabetes mellitus: An exposure-wide umbrella review of meta-analyses | PLOS One [Internett]. [sitert 7. mars 2025]. Tilgjengelig på: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0194127

17. Farmaki P, Damaskos C, Garmpis N, Garmpi A, Savvanis S, Diamantis E. Complications of the Type 2 Diabetes Mellitus. Current Cardiology Reviews. 1. november 2020;16(4):249–51.

18. Stolar M. Glycemic Control and Complications in Type 2 Diabetes Mellitus. The American Journal of Medicine. 1. mars 2010;123(3, Supplement):S3–11.

19. Taylor R. Type 2 Diabetes: Etiology and reversibility. Diabetes Care. 14. mars 2013;36(4):1047–55.

20. Lau D, Manca DP, Singh P, Perry T, Olu-Jordan I, Ryan Zhang J, mfl. The effectiveness of continuous glucose monitoring with remote telemonitoring-enabled virtual educator visits in adults with non-insulin dependent type 2 diabetes: A randomized trial. Diabetes Research and Clinical Practice. 1. november 2024;217:111899.

21. Beck RW, Riddlesworth T, Ruedy K, Ahmann A, Bergenstal R, Haller S, mfl. Effect of Continuous Glucose Monitoring on Glycemic Control in Adults With Type 1 Diabetes Using Insulin Injections: The DIAMOND Randomized Clinical Trial. JAMA. 24. januar 2017;317(4):371–8.

22. Battelino T, Danne T, Bergenstal RM, Amiel SA, Beck R, Biester T, mfl. Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations From the International Consensus on Time in Range. Diabetes Care. august 2019;42(8):1593–603.

23. Zhu X, Williams M, Finuf K, Patel V, Sinvani L, Wolf-Klein G, mfl. Home Telemonitoring of Patients With Type 2 Diabetes: A Meta-Analysis and Systematic Review. Diabetes Spectr. februar 2022;35(1):118–28.

24. Wright EE, Miller E, Bindal A, Poon Y. Addition of continuous glucose monitoring to glucagon-like peptide 1 receptor agonist treatment for type 2 diabetes mellitus - An economic evaluation. J Manag Care Spec Pharm. 1. februar 2025;31(2):127–36.

25. Mengesha ME. ResearchGate. 2025 [sitert 12. mars 2025]. (PDF) The Cost-effectiveness of a Continuous Glucose Monitoring Device for Adult Diabetes Patients in Ethiopia: A Semi-Markov Modelling Study. Tilgjengelig på: https://www.researchgate.net/publication/387743947_The_Cost-effectiveness_of_a_Continuous_Glucose_Monitoring_Device_for_Adult_Diabetes_Patients_in_Ethiopia_A_Semi-Markov_Modelling_Study

26. Pettus JH, Zhou FL, Shepherd L, Preblick R, Hunt PR, Paranjape S, mfl. Incidences of Severe Hypoglycemia and Diabetic Ketoacidosis and Prevalence of Microvascular Complications Stratified by Age and Glycemic Control in U.S. Adult Patients With Type 1 Diabetes: A Real-World Study. Diabetes Care. desember 2019;42(12):2220–7.

27. Martens T, Beck RW, Bailey R, Ruedy KJ, Calhoun P, Peters AL, mfl. Effect of Continuous Glucose Monitoring on Glycemic Control in Patients With Type 2 Diabetes Treated With Basal Insulin: A Randomized Clinical Trial. JAMA. 8. juni 2021;325(22):2262–72.

28. Giammarino F, Senanayake R, Prahalad P, Maahs DM, Scheinker D. A Machine Learning Model for Week-Ahead Hypoglycemia Prediction From Continuous Glucose Monitoring Data. J Diabetes Sci Technol [Internett]. 2024;((Giammarino F.) PE, Spoltore, Italy). Tilgjengelig på: https://www.embase.com/search/results?subaction=viewrecord&id=L2028909970&from=export

29. Alexiadis A, Tsanas A, Shtika L, Efopoulos V, Votis K, Tzovaras D, mfl. Next-Day Prediction of Hypoglycaemic Episodes Based on the Use of a Mobile App for Diabetes Self-Management. IEEE Access. 2024;12:7469–78.

30. Luong A, Cheung J, McMurtry S, Nelson C, Najac T, Ortiz P, mfl. Comparison of Machine Learning Models to a Novel Score in the Identification of Patients at Low Risk for Diabetic Retinopathy. Ophthalmol Sci [Internett]. 2025;5(1). Tilgjengelig på: https://www.embase.com/search/results?subaction=viewrecord&id=L2034756057&from=export

31. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, mfl. Predicting the risk of developing diabetic retinopathy using deep learning. Lancet Digit Heal. 2021;3(1):e10–9.

32. Kronborg T, Hangaard S, Hejlesen O, Vestergaard P, Jensen MH. The potential of predicting nocturnal hypoglycaemia for insulintreated individuals with type 2 diabetes in telemonitoring. Diabetologia. 2023;66((Kronborg T.; Hangaard S.; Hejlesen O.; Jensen M.H.) Department of Health Science and Technology, Aalborg University, Aalborg, Denmark):S379–80.

33. García-Jaramillo M, Luque C, León-Vargas F. Machine Learning and Deep Learning Techniques Applied to Diabetes Research: A Bibliometric Analysis. Journal of Diabetes Science and Technology. 2024;18(2):287–301.

34. Eghbali-Zarch M, Masoud S. Application of machine learning in affordable and accessible insulin management for type 1 and 2 diabetes: A comprehensive review. Artif Intell Med [Internett]. 2024;151((Eghbali-Zarch M.; Masoud S., saramasoud@wayne.edu) Department of Industrial and Systems Engineering, Wayne State University, Detroit, MI, United States). Tilgjengelig på: https://www.embase.com/search/results?subaction=viewrecord&id=L2031695252&from=export

35. Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Research Synthesis Methods. 2020;11(2):181–217.

36. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc. oktober 2018;106(4):531–41.

37. Zotero | Your personal research assistant [Internett]. [sitert 9. mai 2025]. Tilgjengelig på: https://www.zotero.org/

38. Martin J. © Joanna Briggs Institute 2017 Critical Appraisal Checklist for Systematic Reviews and Research Syntheses. 2017;

39. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, mfl. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 29. mars 2021;372:n71.

40. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 17. juli 2015;349(6245):255–60.

41. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Medicine. 6. januar 2015;13(1):1.

42. Hangaard S, Kronborg T, Hejlesen O, Aradóttir TB, Kaas A, Bengtsson H, mfl. The Diabetes teleMonitoring of patients in insulin Therapy (DiaMonT) trial: study protocol for a randomized

controlled trial. Trials. 7. desember 2022;23:985.

43. Anaconda | Built to Advance Open Source AI [Internett]. [sitert 9. mai 2025]. Tilgjengelig på: https://www.anaconda.com/

44. Project Jupyter [Internett]. [sitert 9. mai 2025]. Tilgjengelig på: https://jupyter.org

45. Alanzi T, Alanazi NR, Istepanian R, Philip N. Evaluation of the effectiveness of mobile diabetes management system with social networking and cognitive behavioural therapy (CBT) for T2D. Mhealth. 2018;4:35.

46. Zhou P, Xu L, Liu X, Huang J, Xu W, Chen W. Web-based telemedicine for management of type 2 diabetes through glucose uploads: a randomized controlled trial. Int J Clin Exp Pathol. 2014;7(12):8848–54.

47. Carter EL, Nunlee-Bland G, Callender C. A Patient-Centric, Provider-Assisted Diabetes Telehealth Self-management Intervention for Urban Minorities. Perspect Health Inf Manag. 1. januar 2011;8(Winter):1b.

48. Warren R, Carlisle K, Mihala G, Scuffham PA. Effects of telemonitoring on glycaemic control and healthcare costs in type 2 diabetes: A randomised controlled trial. J Telemed Telecare. oktober 2018;24(9):586–95.

49. Yoon KH, Kim HS. A short message service by cellular phone in type 2 diabetic patients for 12 months. Diabetes Research and Clinical Practice. 1. februar 2008;79(2):256–61.

50. scikit-learn [Internett]. [sitert 9. mai 2025]. IterativeImputer. Tilgjengelig på: https://scikit-learn/stable/modules/generated/sklearn.impute.IterativeImputer.html

51. Steyerberg EW, Harrell FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology. 1. august 2001;54(8):774–81.

52. Kotsiantis SB. Supervised Machine Learning: A Review of Classification Techniques. Tilgjengelig på: https://datajobs.com/data-science-repo/Supervised-Learning-%5BSB-Kotsiantis%5D.pdf

53. Davies MJ, D'Alessio DA, Fradkin J, Kernan WN, Mathieu C, Mingrone G, mfl. Management of Hyperglycemia in Type 2 Diabetes, 2018. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care. desember 2018;41(12):2669–701.

54. Gaede P, Vedel P, Larsen N, Jensen GVH, Parving HH, Pedersen O. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. N Engl J Med. 30. januar 2003;348(5):383–93.

# Appendices



## Appendix 1:

( ALL ( diabetes ) OR TITLE-ABS-KEY ( diabetes AND mellitus AND type 2 ) OR ALL ( t2dm ) OR ALL ( type 2 diabetes ) AND TITLE-ABS-KEY ( machine AND learning ) OR TITLE-ABS-KEY ( forecasting ) OR ALL ( predicitive AND modelling ) AND TITLE-ABS-KEY ( telemedicine ) OR ALL ( telemonitoring ) OR TITLE-ABS-KEY ( telehealth ) )

## Appendix 2:

# JBI CRITICAL APPRAISAL CHECKLIST FOR COHORT STUDIES

Reviewer Tobias Fromreide & Lukas Nielsen

Date 20.03.2025

Author Flavia Giammarino, Ransalu Senanayake, Priya Prahalad, David M. Maahs, and David Scheinker Year 2024

| | Yes | No | Unclear | Not applicable |
|---|---|---|---|---|
| 1. Were the two groups similar and recruited from the same population? | ☐ | X | ☐ | ☐ |
| 2. Were the exposures measured similarly to assign people to both exposed and unexposed groups? | ☐ | X | ☐ | ☐ |
| 3. Was the exposure measured in a valid and reliable way? | X | ☐ | ☐ | ☐ |
| 4. Were confounding factors identified? | ☐ | X | ☐ | ☐ |
| 5. Were strategies to deal with confounding factors stated? | ☐ | X | ☐ | ☐ |

| | | | | |
|---|---|---|---|---|
| 6. Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)? | ☐ | X | ☐ | ☐ |
| 7. Were the outcomes measured in a valid and reliable way? | X | ☐ | ☐ | ☐ |
| 8. Was the follow up time reported and sufficient to be long enough for outcomes to occur? | X | ☐ | ☐ | ☐ |
| 9. Was follow up complete, and if not, were the reasons to loss to follow up described and explored? | ☐ | X | ☐ | ☐ |
| 10. Were strategies to address incomplete follow up utilized? | ☐ | X | ☐ | ☐ |
| 11. Was appropriate statistical analysis used? | X | ☐ | ☐ | ☐ |

Overall appraisal: Include **X** Exclude ☐ Seek further info ☐