Review Article

narra j

Review Article

Predicting the risks of stroke, cardiovascular disease, and peripheral vascular disease among people with type 2 diabetes with artificial intelligence models: A systematic review and meta-analysis

Aqsha Nur[1,2], Sydney Tjandra[3*], Defin A. Yumnanisha[3], Arnold Keane[3] and Adang Bachtiar[1]

[1]Faculty of Public Health, Universitas Indonesia, Depok, Indonesia; [2]Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom; [3]Faculty of Medicine, Universitas Indonesia, Depok, Indonesia

*Corresponding author: sydney.tjandra@ui.ac.id

Abstract

Macrovascular complications, including stroke, cardiovascular disease (CVD), and peripheral vascular disease (PVD), significantly contribute to morbidity and mortality in individuals with type 2 diabetes mellitus (T2DM). The aim of this study was to evaluate the performance of artificial intelligence (AI) models in predicting these complications, emphasizing applicability in diverse healthcare settings. Following PRISMA guidelines, a systematic search of six databases was conducted, yielding 46 eligible studies with 184 AI models. Predictive performance was assessed using the area under the receiver operating characteristic curve (AUROC). Subgroup analyses examined model performance by outcome type, predictor data (lab-only, non-lab, mixed), and algorithm type. Heterogeneity was evaluated using I^2 statistics, and sensitivity analyses addressed outliers and study biases. The pooled AUROC for all AI models was 0.753 (95%CI: 0.740–0.766; I^2=99.99%). Models predicting PVD achieved the highest AUROC (0.794), followed by cerebrovascular diseases (0.770) and CVD (0.741). Gradient-boosting algorithms outperformed others (AUROC: 0.789). Models with lab-only predictors had superior performance (AUROC: 0.837) compared to mixed (0.759) and non-lab predictors (0.714). External validations reported reduced AUROC (0.725), underscoring limitations in generalizability. AI models show moderate predictive accuracy for T2DM macrovascular complications, with laboratory-based predictors being key to performance. However, the limited external validation and reliance on high-resource data restrict implementation in low-resource settings. Future efforts should focus on non-lab predictors, external validation, and context-appropriate AI solutions to enhance global applicability.

Keywords: Artificial intelligence, cardiovascular disease, type 2 diabetes mellitus, stroke, diabetic nephropathy and vascular disease

Introduction

At least 800 million people were estimated to live with diabetes in 2022, of which more than 90% were type 2 diabetes mellitus (T2DM) [1]. T2DM complications, such as stroke, cardiovascular diseases (CVDs), and peripheral vascular diseases (PVDs), increase the 5-year mortality, particularly for people living in low- and middle-income countries (LMICs) [2].

https://narraj.org/

Received: January 15, 2025 - Accepted: March 17, 2025 - Published online: March 19, 2025

According to World Health Organization (WHO), 75% of CVD deaths occur in LMICs [3]. As the global burden continues to rise, there is an urgent need for precise and early risk stratification methods to enable timely preventive measures for T2DM complications [4]. In this context, the use of artificial intelligence (AI) and machine learning models has garnered significant interest for their potential to enhance predictive accuracy in the management of T2DM complications [5,6]. These technologies promise to transform traditional healthcare approaches by leveraging vast amounts of data to uncover complex patterns and relationships that may not be readily apparent through conventional statistical methods [7].

Previous systematic reviews have primarily focused on the potential of AI in various aspects of T2DM care, particularly in predicting the onset of the disease [8,9]. For instance, recent meta-analyses have demonstrated the utility of AI in forecasting T2DM-related outcomes [8-10], yet none have comprehensively addressed the prediction of macrovascular complications associated explicitly with T2DM. This gap highlights the necessity of a focused investigation into how AI can be harnessed to predict complications like stroke, CVD, and PVD in patients already diagnosed with T2DM, including its deployment in low-resource settings [13].

The aim of this study was to explore the performance of machine learning algorithms in predicting the risk of macrovascular complications among individuals with T2DM, specifically the predictive capabilities of AI models in forecasting stroke, CVD, and PVD. This meta-analysis provided an in-depth analysis of subgroup performances, comparing models with various predictor types, including lab-only and mixed predictors, and examining the implications of these differences. We also highlight the challenges and limitations of current AI models, particularly their applicability in low-resource settings.

Methods

Protocol registration

This review was systematically developed, conducted, and reported following the preferred reporting items for systematic review and meta-analysis (PRISMA) checklist [14] and the filled PRISMA checklist is presented in **Underlying data**. The protocol has been registered at The International Prospective Register of Systematic Reviews (PROSPERO) under the reference ID CRD42023489167.

Databases and search strategy

Six databases (Scopus, PubMed, Embase, Wiley Online Library, IEEE Xplore, and Google Scholar) were searched for articles published between January 1, 2000, and November 30, 2023. Keywords employed were “type 2 diabetes,” “artificial intelligence,” “prediction,” “complication,” “stroke,” “cardiovascular disease,” and “peripheral vascular disease,” as well as their MeSH terms and subsets, combined with Boolean operators (**Underlying data**). Search results were exported to Rayyan (www.rayyan.ai) and duplicates were removed, followed by manual deduplication and screening decisions.

Eligibility criteria

Each article was screened for the participants, intervention, comparison, outcomes, and timeframe (PICOT) inclusion criteria [15], as presented in **Table 1,** by at least two investigators independently (AN, ST, RH, SW). The articles should include: (1) adults aged 18 years or above with T2DM; (2) intervention developed or implemented with AI, such as machine learning and deep learning; (3) outcome of prediction performances for stroke, CVD, or PVD; (4) diagnostic or prognostic studies with a cohort or case-control design capable of exhibiting temporality; (5) use of any actual medical dataset; and (6) written in English. We excluded studies that (1) had mixed populations with type 1 and/or prediabetes patients; (2) mainly explained theoretical models not tested on human subjects; (3) involved drugs as the intervention; (4) were reviews, framework developments, conference abstracts, proposals, editorials, commentaries, and qualitative studies; and (5) had irretrievable full-text. After titles and abstracts were screened on Rayyan, full-text screening was conducted to reconfirm eligibility. Discrepancies were resolved through consensus.

Table 1. PICOT inclusion and exclusion criteria

	Participants (P)	Intervention (I)	Comparison (C)	Outcomes (O)	Timeframe (T)	Others
Inclusion criteria	Adults with type 2 diabetes mellitus; actual medical records/datasets	Artificial intelligence (AI) development and implementation, including machine learning and deep learning	Standard or best-practice diagnostic modalities	Prediction performances for stroke, cardiovascular disease, or peripheral vascular disease	Between January 1, 2000, to November 30, 2023	Diagnostic or prognostic studies; English language
Exclusion criteria	Mixed populations with other types of diabetes or prediabetes; non-human subjects	Classical statistical models such as logistic regression, without specific mention of AI; drug involvement		Prediction performances for diabetes; qualitative studies		Irretrievable full-text; theoretical models; reviews; framework developments; proposals

Data extraction

A data extraction instrument was developed to tabulate several characteristics and details from all included studies, including (1) author and year, (2) country of origin, (3) study design, (4) data source, (5) single or multi-centered, (6) population profile (including number of patients, age, and proportion of males), (7) predictors, (8) whether external validation was employed, (9) AI/machine learning algorithm used, (10) outcome (stroke, CVD, or PVD), (11) data period and follow-up, (12) data pre-processing details, and (13) internal validation setup. We also extracted the main outcome(s) model performances in metrics such as F-measures, the area under the receiving operating curve (AUROC), c-statistics, sensitivity/recall, specificity, accuracy, and precision/positive predictive values.

Risk of bias assessment

Included studies were divided among AN, ST, DY, and AK, with each study being independently assessed by two investigators for risk of bias and applicability using the signaling questions on the Prediction model Risk of Bias Assessment Tool (PROBAST) [6]. Any discrepancies were discussed to reach a consensus.

Quantitative data analysis

Studies reporting AUROCs as model performances were aggregated through a random-effects meta-analysis with R. When neither the standard error, range, nor standard deviation was available, we ran the Hanley and McNeil's approach [16] with R to approximate the standard error based on the AUROC, sample size, and number of complication cases [14-16]. To assess publication bias, funnel plots and Egger's regression were generated with MedCalc (MedCalc Software Ltd., Ostend, Belgium). Moreover, outliers, defined as models whose 95% confidence intervals did not overlap with meta-analysis result, were excluded to generate sensitivity analyses. As substantial heterogeneity remains, subgroup analyses were conducted for outcome types, external validation, algorithms, country income levels, risk of bias, missing data process details, cross-validation, and predictor data. All visualizations were generated with R 4.4.2 in RStudio (R Foundation for Statistical Computing, Vienna, Austria).

Results

Study characteristics

A total of 2,513 studies were found during the initial search across seven databases in addition to hand searching. After removing 512 duplicate records, 2,001 records were screened for their titles and abstracts. Subsequently, 1,895 records were excluded, leaving 106 reports to be retrieved. Studies without available full texts were excluded, resulting in 95 studies being assessed for

eligibility. Of these, 49 studies were excluded for various reasons: unsuitable population (17 studies), irrelevant outcome (29 studies), unsuitable study design (two studies), and text not in English (one study). Ultimately, 46 studies were included in the systematic review, with 30 included in the quantitative analysis. The selection process is depicted in **Figure 1**.

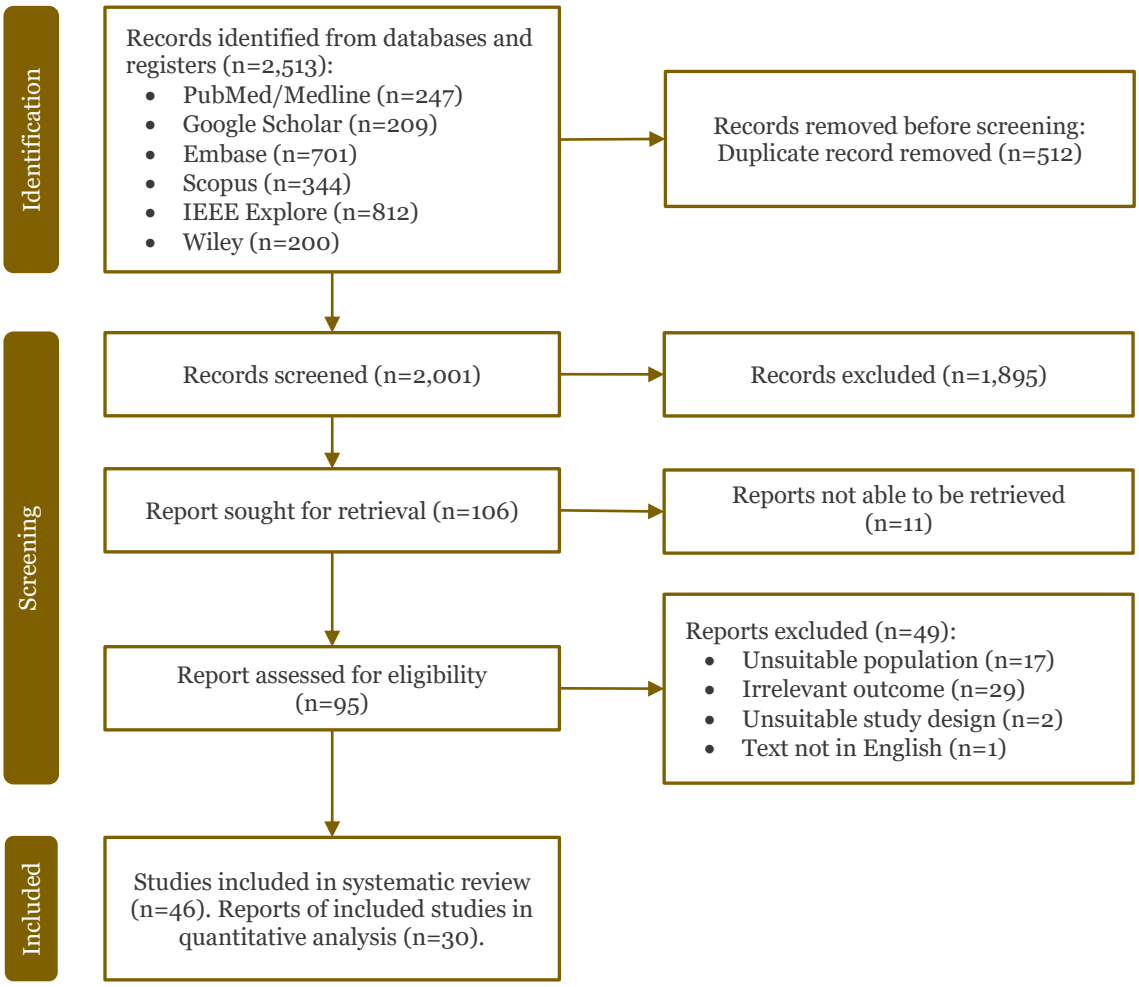

Figure 1. PRISMA flowchart of studies selection process.

The overall characteristics of the study, including the participants and outcomes of each study, are presented in **Table 2**. The systematic review included 46 studies from various countries, of which the majority of studies (29; 63%) were conducted in high-income countries (HICs) [19-40], followed by upper-middle-income countries (UMICs) (9; 19.6%) [36,41-48], and low- and middle-income countries (LMICs) (8; 17.4%) [49-56]. India contributed the most studies from LMICs, China dominated the UMICs, and the United States led in HICs [36,41-48]. Most data were sourced from hospital medical records in the respective countries [21,24-27,29,32,33,38,39,41,42,45-49,52-54,57-59] and national database centers [20,22,31,34-37,40,44,50,51,56,60-63], while some datasets were from specific trials or studies [19,23,28,30,43,64]. Sample sizes from the studies varied from 111 [47] to 1,910,674 [31]. Complete study characteristics are detailed in the **Underlying Data**.

A total of 250 predictors were identified across 46 studies and grouped into four categories: demographics (13 predictors; 5.2%), clinical (50 predictors; 20%), comorbidity (33 predictors; 13.2%), and laboratory (154 predictors; 61.6%) (**Figure 2**). The most commonly used demographic predictors were age, sex, and race, while clinical factors such as body mass index, blood pressure, and antidiabetic medication history were frequently used. Comorbidities such as hypertension, heart disease, and renal disease appeared in 50% of the studies, while laboratory parameters such as HbA1c, high-density lipoproteins, and cholesterol levels were the most commonly used laboratory predictors (**Figure 2**).

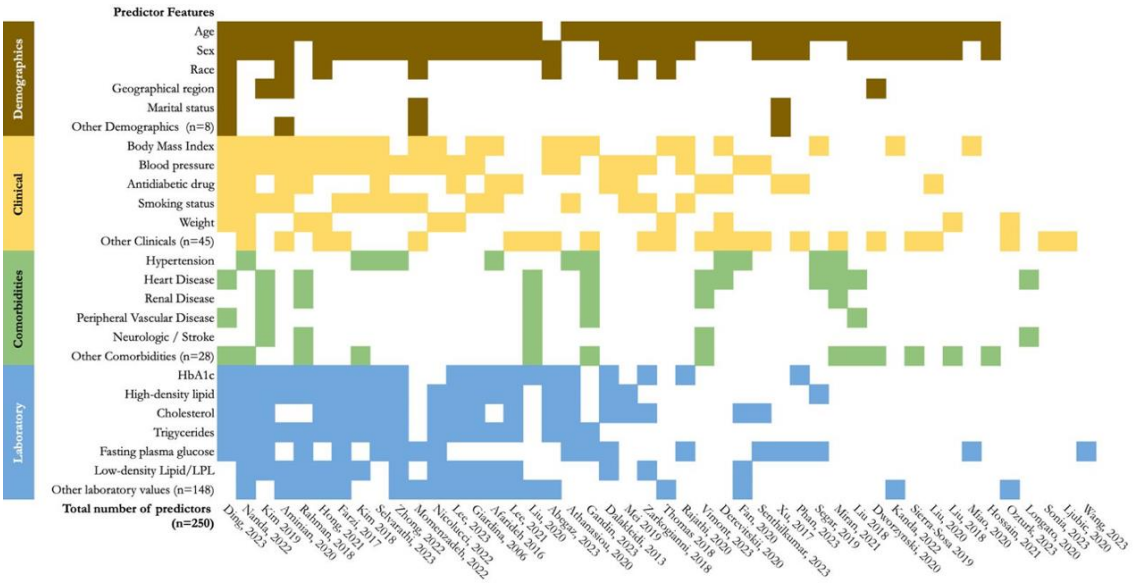

Figure 2. Heatmap of most common predictor features in included studies, ordered and categorized. The presence of color for specific predictors indicates that the corresponding reference study incorporated those predictors in the model.

Quality of the studies

The overall risk of bias indicates that 78% of the articles were rated as having a high or uncertain risk of bias, with specific distributions of nine articles rated as low risk [21,23-26,39,41,42,60], 10 as unclear risk [19,33-38,45,48,59], and 27 as high risk [20,22,27-32,36,40,43,44,46,47,49,50, 52-57,59,62,64] (**Figure 3**). The high risk of bias predominantly originated from the outcome domain due to the uncertainty in determining outcomes without knowledge of predictor information. Additionally, the analysis domain contributed significantly to the high risk, primarily due to inadequate handling of missing data or improper imputation methods and the low number of participants with the outcome. On the other hand, nearly all studies (91%) showed no concerns regarding applicability concerns.

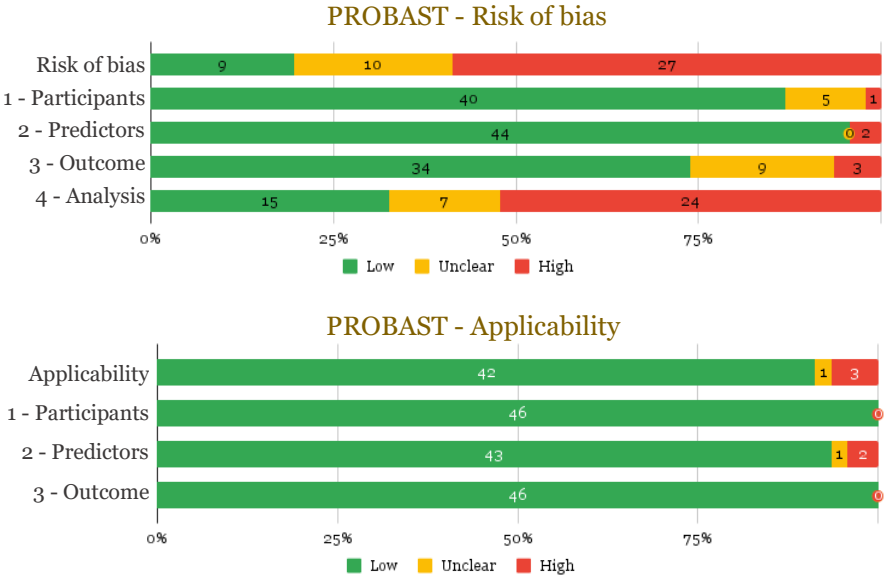

Figure 3. PROBAST summary of 46 studies included.

Table 2. Characteristics of the included studies in the systematic review

Author (Year)	Country	Types of predictors				Algorithm	Outcome
		Demographic information	Clinical data	Comorbidities	Laboratory parameters		
Nanda (2022) [49]	India	+	+	+	+	RF, KNN, AdaBoost, SVM-PolyK, SVM-RBF, Naive-Bayes, Bagging, Stacking (RF + KNN)	Diabetic foot ulcer type
Senthilkumar (2023) [50]	India	+	+	N/A	+	LR, RF, AdaBoost, Multilayer perceptron.	CVD
Sonia (2023) [51]	India	N/A	+	N/A	N/A	NN (DNN, AlexNet, LeNet-5, DNHRV), Inception, VGG-16, LSTM	CVD
Ding (2023) [41]	China	+	+	+	+	LR, Gradient Boosting (AdaBoost, GBDT), RF, SVM	CAD,CeVD, PVD
Abegaz (2023) [19]	United States	+	+	N/A	+	RF, XGBoost, LR, Weighted ensemble model	MACE, MI, HF, stroke
Selvarathi (2023) [52]	India	+	+	+	+	HWNN, CCGLSTM,Traditional LSTM, Convolution LSTM, Convolution GLSTM	CVD
Vimont (2023) [20]	France	+	+	N/A	+	LR, RF, NN	CVD (HF, PAD, MI, UA, TIA, CV-related death), stroke
Gandin (2023) [21]	Italy	+	+	+	+	Cox proportional hazards regression, Non-linear PH-DNN	HF
Zhong (2022) [42]	China	+	+	+	+	NN (KNN, ANN), SVM linear, SVM radial, Decision tree, RF, XGBoost, LR (LR with Lasso)	ACS
Miran (2021) [22]	United States	+	N/A	+	N/A	LR, NN, RF, XGBoost	HF
Momenzadeh (2022) [23]	United States	+	+	N/A	+	GB (AdaBoost, XGBoost), SVC, ET, RF, LR	CVD
Nicolucci (2022) [24]	Italy	+	+	N/A	+	XGBoost	CVD
Hong (2021) [25]	United States	+	+	N/A	+	Cox proportional hazard, LASSO regression	CHD, HF, stroke
Aminian (2020) [26]	United States	+	+	N/A	+	Cox-based (Cox proportional hazards, exponential, Fine-Gray), RF	Coronary artery events
Athanasiou (2020) [27]	Greece	+	+	+	+	XGBoost	CVD
Miao (2020) [43]	China	+	+	N/A	+	KNN, SVM	CVD
Hossain (2021) [28]	Australia	+		+		LR, SVM, Decision Tree, NB, RF KNN.	CVD
Zarkogianni (2018) [29]	Greece	+	+	N/A	+	RF, Hybrid ensemble, SOM classifier, BLR model, CART, NB, FFN	CHD, stroke
Segar (2019) [30]	United States	+	+	+	+	RSF, Cox-based method	HF
Derevitskii (2020) [44]	Russia	+	+	N/A	N/A	XGBoost	CHF, AF
Ljubic (2020) [31]	United States	N/A	+	N/A	N/A	RNN, RNN GRU	AP, atherosclerosis, IHD, MI, PVD
Fan (2020) [45]	China	+	+	N/A	+	RF	CHD
Dalakleidi (2013) [32]	Greece	+	+	N/A	+	GA	CVD
Xu (2017) [46]	China	+	N/A	N/A	+	LR	CeVD
Lee (2023) [33]	South Korea	+	+	N/A	+	GRU-ODE-Bayes-based	CVD
Wang (2023) [47]	China	N/A	N/A	N/A	+	RF	Diabetic foot

Author (Year)	Country	Types of predictors				Algorithm	Outcome
		Demographic information	Clinical data	Comorbidities	Laboratory parameters		
Ozturk (2023) [35]	United Kingdom	N/A	+	N/A	+	NN, RF, NB, SVM, Ensemble	Hypertension
Kanda (2022) [60]	Japan	+	+	N/A	+	XGBoost, NN, LR, Cox proportional hazard	CVD
Lee (2021) [34]	Hong Kong	+	+	N/A	+	RSF, Multivariate Cox model, CISF	AMI
Liu (2020) [36]	United States	+	+	N/A	N/A	MTFL, OS-MTL, Private-shared MTL, Single-task learning with Lasso	PVD, CVD, CeVD
Farzi (2017) [53]	Iran	+	+	N/A	+	RF, J48 (Decision Tree), LMT, NBTree, SMO, MLP, Naïve Bayes, Bayes Net, RBF	CVD
Longato (2020) [37]	Italy	+	+	N/A	N/A	NN	CVD, stroke
Giardina (2006) [38]	United Kingdom	+	+	N/A	+	WkNN, kNN, GA, RI	CHD
Phan (2023) [39]	Taiwan	+	+	N/A	+	LR, GB (GBM, LGBM, AdaBoost, XGB), RF, Voting ensemble, LDA	Ischemic stroke
Rahman (2018) [54]	Bangladesh	+	+	+	+	LR, RF, Decision Tree with AdaBoost, SVM, NB, Decision Tree	CVD
Liu (2018) [40]	United States	+	+	N/A	N/A	MTL, STL	Vascular disease
Rajathi (2020) [55]	India	+	+	N/A	+	HWNN, SOM, MTLBO	CVD
Liu (2020) [61]	China	+	N/A	N/A	+	RF, BN, NB, C5.0	Macrovascular complications, diabetic foot
Dworzynski (2020) [62]	Denmark	+	+	N/A	N/A	LR (Reference LR, Logistic ridge regression), RF, GB	HF, MI, stroke, CVD
Afarideh (2016) [56]	Iran	+	+	N/A	+	Cox proportional hazards, ANN	CVD
Liu (2018) [63]	United States	+	+	N/A		LR, MTL (STL, MTFL, MTRL, FETR, TREFLES)	Vascular disease
Mei (2019) [48]	China	+	+	N/A	+	LR, NN (TSNN, KENN), Decision fusion, Pooled cohort equations	CVD
Sierra-Sosa (2019) [57]	Spain	+	N/A	+	N/A	LR (LR Ridge, LR Lasso), LDA, SVM	MI
Thomas (2018) [64]	United States	+	+	N/A	+	Patient network	Stroke, MI, HF
Kim (2019) [58]	United States	+	+	+	+	MTL, GBM, LASSO	IHD, CHF, CVD, PVD
Kim (2018) [59]	United States	+	+	+	+	MTL, Lasso-penalized Cox-regression	IHD, PVD, CHF, CVD

AMI: acute myocardial infarction; ANN: artificial neural network; BN: Bayesian network; CAD: coronary artery disease; CeVD: cerebrovascular disease; CHF: congestive heart failure; CISF: conditional inference survival forest; CVD: cardiovascular disease; FETR: feature and task relationship learning; GA: genetic algorithm; GB: gradient boosting; GBDT: gradient boosting decision tree; GRU: gated recurrent unit; HF: heart failure; HWNN: hybrid wavelet neural network; IHD: ischemic heart disease; KENN: knowledge-enhanced neural network; LDA: linear discriminant analysis; LGBM: light gradient boosting machine; LR: logistic regression; MACE: major adverse cardiopulmonary events; MI: myocardial infarction; MLP: multi-layer perception; MTLBO: modified teaching learning based optimization; MTLR: multi-task relationship learning; MTFL: multi-task feature learning; N/A: not applicable; OS-MTL: outcome-specific multi-task learning; PAD: peripheral artery disease; PH-DNN: proportional hazards deep neural network; PVD: peripheral vascular disease; RBF: radial base function; RF: random forest; RI: random initialization; RSF: random survival forest; SOM: self-organized mapping; STL: single task learning; SVM: support vector machine; TIA: transient ischemic attack; TSNN: teacher-student network; UA: unstable angina; WkNN: weighted k-nearest neighbors; XGB: extreme gradient boosting.

Predictive performance

The most common machine learning algorithm used was gradient boosting, which led to the highest AUROC performance (0.789), followed by random forests (AUROC of 0.776) (**Table 3**). Neural networks, despite their ability to capture complex relationships, were less frequently used due to their high computational demands and the need for large datasets, achieving an AUROC of 0.759. A majority of studies (29; 63.04%) utilized k-fold cross-validation for internal validation. The outcomes assessed included CVD (coronary artery disease, myocardial infarction, heart failure, atrial fibrillation), cerebrovascular disease (stroke), and PVD.

All 184 models were pooled with a random effects model, obtaining an AUROC of 0.753 (95%CI: 0.740–0.766; I^2=99.99%; p<0.001), with significant overall publication bias (p-Egger of 0.026). For 80 models of CVDs, an AUROC of 0.741 (95%CI: 0.721–0.760; I^2=99.78%; p<0.001) was obtained. Meanwhile, 25 models of PVD and 38 models of cerebrovascular diseases obtained AUROCs of 0.794 (95%CI: 0.758–0.831; I^2=97.23%; p<0.001) and 0.770 (95%CI: 0.743–0.797; I^2=99.73%; p<0.001) respectively. Subgroup analysis results are detailed in **Table 3**.

Table 3. Subgroup analyses summary of areas under the receiver operating characteristics (AUROCs) of machine learning prediction models for type 2 diabetes macrovascular complications based on various characteristics

Subgroups	Number of prediction model	Random effect AUROC	Lower 95%CI	Higher 95%CI	Heterogeneity (%)	p-value
All studies	184	0.753	0.740	0.766	99.99	<0.001
Outcome types						
Cardiovascular	80	0.741	0.721	0.76	99.78	<0.001
Peripheral vascular/diabetic foot	25	0.794	0.758	0.831	97.23	<0.001
Stroke/cerebrovascular	38	0.770	0.743	0.797	99.73	<0.001
Mixed	41	0.741	0.716	0.765	100	<0.001
External validation data						
Yes	56	0.725	0.708	0.742	97.85	<0.001
No	128	0.765	0.749	0.782	99.99	<0.001
Machine learning algorithm						
Cox-based	14	0.712	0.664	0.760	98.53	<0.001
Gradient boosting	37	0.789	0.761	0.817	99.56	<0.001
Logistic regression	23	0.731	0.711	0.752	99.55	<0.001
Multi-task learning	18	0.699	0.665	0.733	99.99	<0.001
Neural network	11	0.759	0.722	0.797	98.55	<0.001
Random forest	30	0.776	0.742	0.810	99.73	<0.001
Others	51	0.752	0.726	0.777	99.99	<0.001
Country income						
HIC	134	0.737	0.723	0.751	99.99	<0.001
LIC/LMIC/UMIC	50	0.800	0.774	0.825	97.31	<0.001
Risk of bias						
Low/medium	110	0.780	0.765	0.794	99.76	<0.001
High	74	0.711	0.691	0.731	99.99	<0.001
Missing data process detailed						
Yes	114	0.775	0.760	0.790	99.66	<0.001
No	70	0.717	0.696	0.738	99.99	<0.001
Cross-validation						
Yes	127	0.759	0.743	0.775	99.99	<0.001
No	57	0.739	0.717	0.761	98.88	<0.001
Predictor data						
No lab	29	0.714	0.696	0.731	100	<0.001
Lab only	3	0.837	0.784	0.890	0	<0.001
Mixed	152	0.759	0.745	0.774	99.98	<0.001

CI: confidence interval; HIC: high-income countries; LIC: low-income countries; LMIC: lower-middle-income countries; UMIC: upper-middle-income countries

We excluded outliers and retrieved 83 models with an overall AUROC of 0.746 (95%CI: 0.742–0.75; I^2=99.86%; p<0.001). This is comparable to the initial meta-analysis, showing robustness despite outliers. Similarly, outcome and predictor subgroup sensitivity analyses were conducted, with results in **Figure 4** and **Table 4**. Most notably, the PVD outcome subgroup had an AUROC of 0.820 (95%CI: 0.798–0.842; p<0.001) with a heterogeneity of I^2=0%.

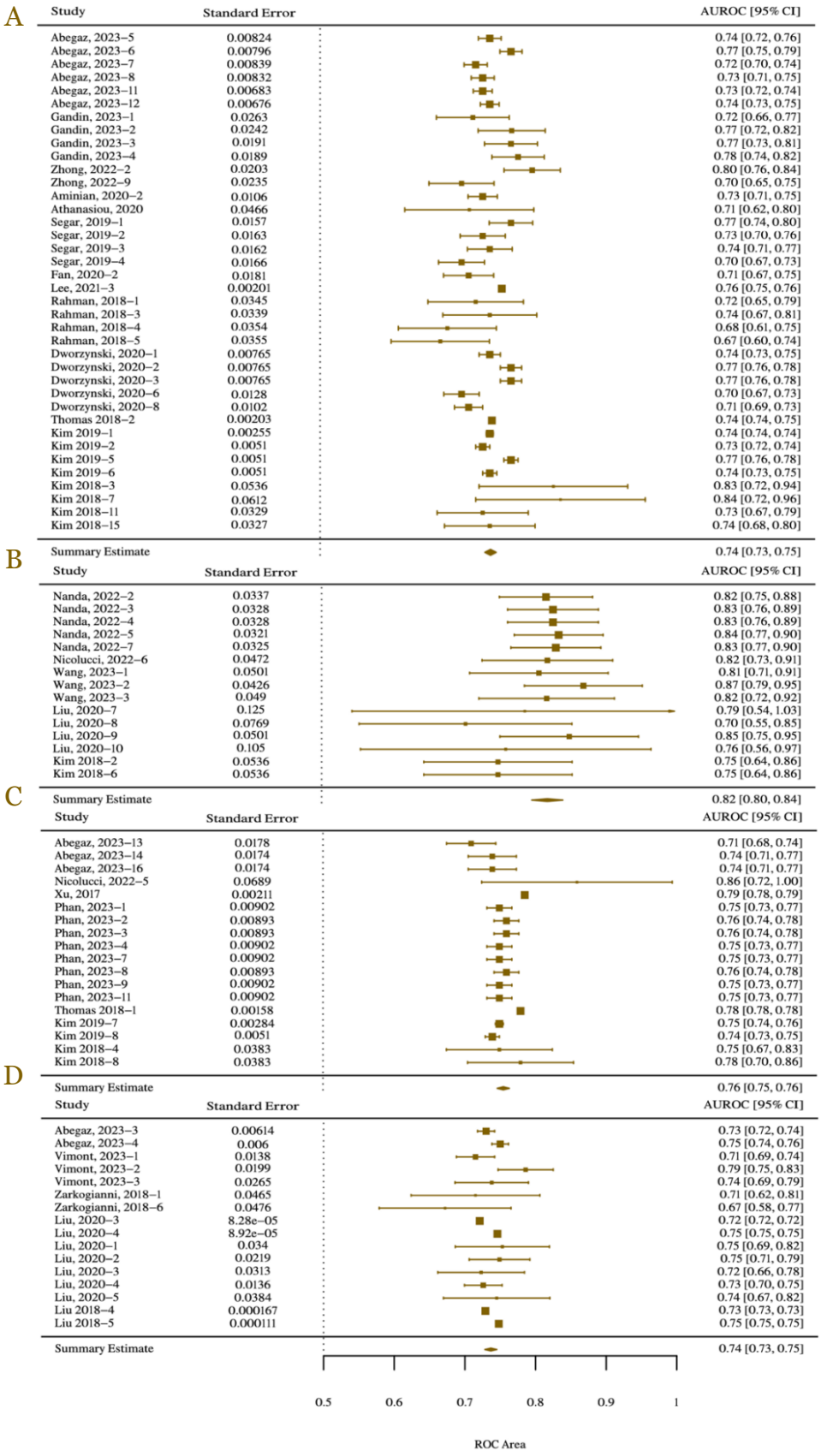

Figure 4. Forest plots of artificial intelligence areas under the operating curve (AUROCs), excluding outliers, in predicting T2DM complications: (A) cardiovascular (AUROC: 0.741; 95%CI: 0.733–0.749; *p*-Het<0.001; overall *p*<0.001); (B) peripheral vascular (AUROC: 0.820; 95%CI: 0.798–0.842; *p*-Het: 0.864; overall *p*<0.001); (C) cerebrovascular (AUROC: 0.756; 95%CI: 0.747–0.764; *p*-Het<0.001; overall *p*<0.001); and (D) mixed (AUROC: 0.737; 95%CI: 0.729–0.745; *p*-Het: 0.653; overall *p*<0.001).

Table 4. Sensitivity analyses summary of AUROCs of machine learning prediction models for type 2 diabetes macrovascular complications based on outcomes and predictors

Subgroups	Number of prediction model	Random effect AUROC	Lower 95%CI	Higher 95%CI	Heterogeneity (%)	*p*-value
All studies	83	0.746	0.742	0.750	99.86	<0.001
Outcome types						
Cardiovascular	38	0.741	0.733	0.749	80.99	<0.001
Peripheral vascular/diabetic foot	15	0.820	0.798	0.842	0.00	<0.001
Stroke/cerebrovascular	18	0.756	0.747	0.764	92.42	<0.001
Mixed	16	0.737	0.729	0.745	99.97	<0.001
External validation data						
No lab	15	0.710	0.703	0.718	99.90	<0.001
Mixed	73	0.753	0.748	0.758	92.31	<0.001

AUROC: area under the receiver operating characteristic; CI: confidence interval

In the lab-only predictors subgroup, no outliers were identified. We observed significant publication bias for peripheral vascular (*p*-Egger of 0.028) and cerebrovascular complications (*p*-Egger of 0.018). Funnel plots of AUROCs against standard errors are presented in **Figure 5**.

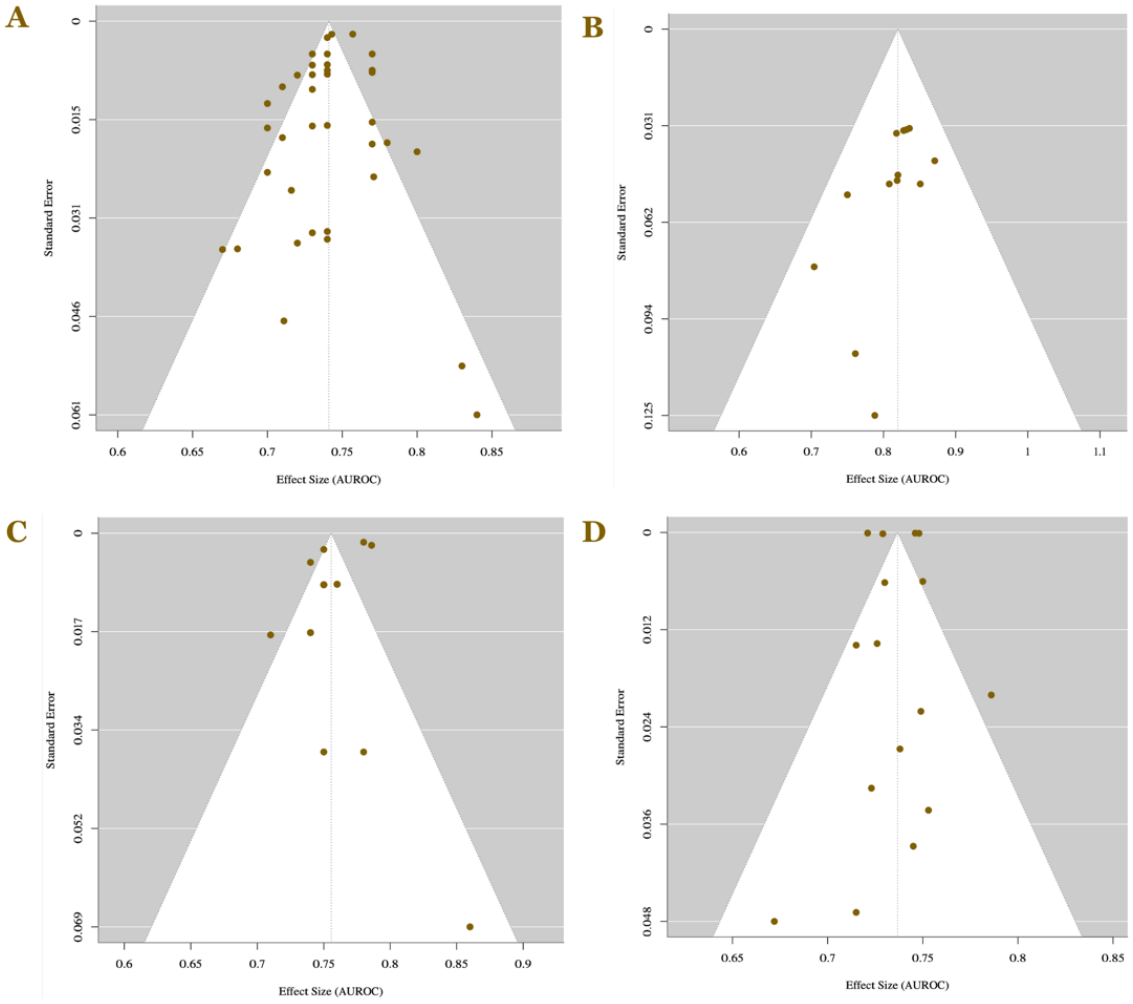

Figure 5. Funnel plots of artificial intelligence areas under the operating curve (AUROCs), excluding outliers, in predicting type 2 diabetes complications: (A) cardiovascular (*p*-Egger: 0.236); (B) peripheral vascular (*p*-Egger: 0.028); (C) cerebrovascular (*p*-Egger: 0.018); and (D) mixed (*p*-Egger: 0.938).

Discussions

The pooled analysis of 184 models demonstrated a moderate overall performance with an AUROC of 0.753. The models exhibited varying levels of performance based on the specific outcome types. For CVD outcomes, the AUROC was 0.741, while models predicting PVD and cerebrovascular disease achieved higher AUROCs of 0.794 and 0.770, respectively. A meta-analysis developed a random forest model with the highest AUROC of 0.918 for predicting diabetic foot ulcers in individuals with T2DM [65]. These findings suggest that while the models generally exhibit robustness, their effectiveness may differ depending on the type of predicted macrovascular complication.

High heterogeneity was observed across the included studies, with I^2 values approaching 100% in most cases. This substantial heterogeneity emphasizes the variability in model performance, which could be attributed to differences in study populations, data sources, predictor variables, and machine learning algorithms. The high heterogeneity also underscores the significance of context-specific factors and suggests that predictive accuracy might improve with models tailored to specific populations and settings. Sensitivity analysis further supports the robustness of the findings, with AUROC slightly decreasing to 0.746 after excluding outliers, while heterogeneity remained high (I^2=99.86%). This consistency indicates that extreme values did not unduly influence the overall conclusions.

The application of these machine learning models in predicting the risk of T2DM complications might help overcome some limitations of existing conventional scoring systems. For instance, the Framingham risk score, a well-established heart disease risk assessment tool, was developed for the general population rather than individuals with T2DM [18]. Consequently, risk scores designed for the general population may not provide optimal discriminatory ability for individuals with T2DM [67,68].

Distribution of AI model development suggests that a country's income level may influence the extent of AI research and publications. A previous bibliometric study mapped the publication of AI in healthcare, noting that the United States and China were among the top nine countries, with the United States leading (41.84%) and China second (14.70%) [69]. Our study also highlighted that the most prominent institutions funding AI research were from the United States. The disparity in the number of studies from non-high-income countries might be due to limited healthcare or AI infrastructure and resources, despite the fact that 80% of the global population resides in these countries where public health challenges are growing [70,71]. Therefore, research in developing countries, including on AI, is essential to ensure that findings are applicable to their specific contexts.

Given the varying availability of predictor data, testing an existing machine learning model in different settings requires consideration of data availability, especially in low-resource settings where laboratory parameters may not be readily accessible. In our subgroup analysis, models without laboratory data (using only demographic, clinical, or comorbidity data) had an AUROC of 0.714, while models with laboratory data achieved an AUROC of 0.837. This demonstrates that non-lab-based models can perform comparably to lab-based models, and suggests that further improvement in non-lab models is feasible [72,73].

To improve performance, several strategies can be employed, such as hyperparameter tuning and exploring different algorithms that can optimize the models [74]. In addition, models that described their missing data handling tended to perform better, suggesting that appropriate imputation techniques, such as those utilizing autoencoders, would be beneficial in improving data pre-processing for AI model development [25-27]. With regards to algorithm used, boosting algorithms and random forests, being ensemble learning algorithms, offer advantages such as reduced risk of overfitting and effective handling of both categorical and numerical predictors [78,79].

External validation led to a consistent decrease in model performance, with an AUROC of 0.725 compared to the AUROC of 0.765 observed during internal validation. This result suggests that development models tend to overestimate performance [30-32]. Only 11 (23.91%) studies performed external validation, an important step in assessing the generalizability of prediction models [83]. For certain studies with small or non-representative datasets, external validation

may not be necessary [84]. However, for comprehensive T2DM complication prediction models, external validation is crucial to ensure applicability across diverse clinical settings [85].

Furthermore, our analysis revealed significant risks of bias in the studies included in this review. Many studies exhibited a high or unclear risk of bias due to issues like incomplete data, insufficient population sampling, and inadequate consideration of key predictors. The reliance on internal validation and the underrepresentation of diverse patient groups further contribute to these biases [86]. Multi-center studies or those based on national databases tend to have lower bias [87], and future studies should prioritize diverse, high-quality training data and effective data handling to improve model accuracy and reliability [88].

The application of AI and machine learning in predicting complications is still in its early stages but holds significant promise. Early and accurate diagnosis could enable timely interventions, but this requires rigorous validation and scrutiny. Future research should focus on enhancing the predictive power of non-lab-based models and conducting extensive external validations. Additionally, to ensure the ethical and practical use of AI or machine learning in healthcare, a secure framework focusing on data protection, patient consent, and algorithmic transparency must be established. Collaboration among policymakers, bioethicists, and researchers is crucial to overcoming the challenges in AI model implementation.

Our study provides valuable insights into the capabilities and limitations of AI models in predicting T2DM complications. The inclusion of studies from multiple countries with varying income levels offers a broad perspective. However, high heterogeneity in the included studies is a key limitation, albeit commonly observed in published AI model performance meta-analyses [81,82]. Moreover, external validation is limited, and superior predictive accuracy relies on laboratory data. Future studies should focus on improving non-lab-based models to enhance their applicability in low-resource settings.

Conclusions

This review highlights the potential of machine learning models in predicting macrovascular complications in T2DM. Despite moderate performance and high heterogeneity, the findings underscore the need for context-specific models tailored to specific populations. Future research should aim to improve the performance of non-lab-based models and expand external validation to enhance their applicability in diverse clinical settings. Collaboration and ethical considerations will be critical to the successful integration of AI in healthcare.

Ethics approval

Not required.

Acknowledgments

Stevano Wijoyo (Universitas Indonesia) supported the initial screening of this review. Dante Harbuwono (Universitas Indonesia) and Sri Laksmiastuti (Universitas Trisakti) suggested the clinical impact of AI on diabetes, which guided the protocol of this review.

Competing interests

All the authors declare that there are no conflicts of interest.

Funding

This research was funded by the Indonesia Endowment Fund for Education, Ministry of Finance, Republic of Indonesia.

Underlying data

Data from this study are publicly available on the Open Science Framework, a public, open-access repository, at https://osf.io/7gh9m/. Please contact the corresponding author for any further inquiries.

Declaration of artificial intelligence use

Artificial intelligence (AI) (rayyan.ai) was employed in identifying potential duplicates; the following decision of deduplication and screening processes were executed manually. AI-assisted improvements in the paper's language and readability. AI was not utilized during data collection, analysis, or visualization.

How to cite

Nur A, Tjandra S, Yumnanisha DA, *et al.* Predicting the risks of stroke, cardiovascular disease, and peripheral vascular disease among people with type 2 diabetes with artificial intelligence models: A systematic review and meta-analysis. Narra J 2025; 5 (1): e2116 - http://doi.org/10.52225/narra.v5i1.2116.

References

1. Zhou B, Rayner AW, Gregg EW, *et al.* Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: A pooled analysis of 1108 population-representative studies with 141 million participants. The Lancet 2024;404(10467):2077-2093.
2. Raghavan S, Vassy JL, Ho Y, *et al.* Diabetes mellitus-related all-cause and cardiovascular mortality in a national cohort of adults. J Am Heart Assoc 2019;8(4):e011295.
3. World Health Organization. Cardiovascular diseases (CVDs) fact sheet 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed: 2 August 2024.
4. Tomic D, Shaw JE, Magliano DJ. The burden and risks of emerging complications of diabetes mellitus. Nat Rev Endocrinol 2022;18(9):525-539.
5. Moor M, Banerjee O, Abad ZSH, *et al.* Foundation models for generalist medical artificial intelligence. Nature 2023;616(7956):259-265.
6. Moons KGM, Wolff RF, Riley RD, *et al.* PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. Ann Intern Med 2019;170(1):W1-W33.
7. Hunter DJ, Holmes C. Where medical statistics meets artificial intelligence. N Engl J Med 2023;389(13):1211-1219.
8. Sheng B, Pushpanathan K, Guan Z, *et al.* Artificial intelligence for diabetes care: Current and future prospects. Lancet Diabetes Endocrinol 2024;12(8):569-595.
9. Mohsen F, Al-Absi HRH, Yousri NA, *et al.* A scoping review of artificial intelligence-based methods for diabetes risk prediction. NPJ Digit Med 2023;6(1):197.
10. Kee OT, Harun H, Mustafa N, *et al.* Cardiovascular complications in a diabetes prediction model using machine learning: A systematic review. Cardiovasc Diabetol 2023;22(1):13.
11. Chowdhury MZI, Yeasmin F, Rabi DM, *et al.* Predicting the risk of stroke among patients with type 2 diabetes: A systematic review and meta-analysis of C-statistics. BMJ Open 2019;9(8):e025579.
12. Usman TM, Saheed YK, Nsang A, *et al.* A systematic literature review of machine learning based risk prediction models for diabetic retinopathy progression. Artif Intell Med 2023;143:102617.
13. Ciecierski-Holmes T, Singh R, Axt M, *et al.* Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: A systematic scoping review. NPJ Digit Med 2022;5(1):162.
14. Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
15. Riva JJ, Malik KMP, Burnie SJ, *et al.* What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc 2012;56(3):167-171.
16. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29-36.
17. The R Foundation. The R project for statistical computing. Available from: https://www.R-project.org/. Accessed: 1 May 2024.
18. Gardner J. Package "auctestr": Statistical testing for AUC data. Available from: https://cran.r-project.org/web/packages/auctestr/auctestr.pdf. Accessed: 1 May 2024.
19. Abegaz TM, Baljoon A, Kilanko O, *et al.* Machine learning algorithms to predict major adverse cardiovascular events in patients with diabetes. Comput Biol Med 2023;164:107289.

20. Vimont A, Béliard S, Valéro R, *et al.* Prognostic models for short-term annual risk of severe complications and mortality in patients living with type 2 diabetes using a national medical claim database. Diabetol Metab Syndr 2023;15(1):128.
21. Gandin I, Saccani S, Coser A, *et al.* Deep-learning-based prognostic modeling for incident heart failure in patients with diabetes using electronic health records: A retrospective cohort study. PLoS One 2023;18(2):e0281878.
22. Miran SM, Nelson SJ, Zeng-Treitler Q. A model-agnostic approach for understanding heart failure risk factors. BMC Res Notes 2021;14(1):184.
23. Momenzadeh A, Shamsa A, Meyer JG. Bias or biology? Importance of model interpretation in machine learning studies from electronic health records. JAMIA Open 2022;5(3):ooac063.
24. Nicolucci A, Romeo L, Bernardini M, *et al.* Prediction of complications of type 2 diabetes: A machine learning approach. Diabetes Res Clin Pract 2022;190:110013.
25. Hong D, Fort D, Shi L, *et al.* Electronic medical record risk modeling of cardiovascular outcomes among patients with type 2 diabetes. Diabetes Ther 2021;12(7):2007-2017.
26. Aminian A, Zajichek A, Arterburn DE, *et al.* Predicting 10-year risk of end-organ complications of type 2 diabetes with and without metabolic surgery: A machine learning approach. Diabetes Care 2020;43(4):852-859.
27. Athanasiou M, Thanopoulou AC, Sfrintzeri K, *et al.* An explainable XGBoost-based approach towards assessing the risk of cardiovascular disease in patients with type 2 diabetes mellitus. arXiv 2020:2009.06629.
28. Hossain ME, Uddin S, Khan A. Network analytics and machine learning for predictive risk modelling of cardiovascular disease in patients with type 2 diabetes. Expert Syst Appl 2021;164:113918.
29. Zarkogianni K, Athanasiou M, Thanopoulou AC. Comparison of machine learning approaches toward assessing the risk of developing cardiovascular disease as a long-term diabetes complication. IEEE J Biomed Health Inform 2018;22(5):1637-1647.
30. Segar MW, Vaduganathan M, Patel KV, *et al.* Machine learning to predict the risk of incident heart failure hospitalization among patients with diabetes: The WATCH-DM risk score. Diabetes Care 2019;42(12):2298-2306.
31. Ljubic B, Hai AA, Stanojevic M, *et al.* Predicting complications of diabetes mellitus using advanced machine learning algorithms. J Am Med Inform Assoc 2020;27(9):1343-1351.
32. Dalakleidi K, Zarkogianni K, Karamanos K, *et al.* A hybrid genetic algorithm for the selection of the critical features for risk prediction of cardiovascular complications in Type 2 Diabetes patients. 13th IEEE International Conference on BioInformatics and BioEngineering. Chania: IEEE; 2013.
33. Lee J, Choi Y, Ko T, *et al.* Prediction of cardiovascular complication in patients with newly diagnosed type 2 diabetes using an XGBoost/ GRU-ODE-Bayes-based machine-learning algorithm. Endocrinol Metab 2024;39(1):176-185.
34. Lee S, Zhou J, Guo CL, *et al.* Predictive scores for identifying patients with type 2 diabetes mellitus at risk of acute myocardial infarction and sudden cardiac death. Endocrinol Diabetes Metab 2021;4(3):e00240.
35. Ozturk B, Lawton T, Smith S, Habli I. Predicting progression of type 2 diabetes using primary care data with the help of machine learning. Stud Health Technol Inform 2023;302:38-42.
36. Liu B, Li Y, Ng K. Prognostication and outcome-specific risk factor identification for diabetes care via private-shared multi-task learning. 2020 IEEE International Conference on Big Data (Big Data). Atlanta: IEEE; 2020.
37. Longato E, Fadini GP, Sparacino G, *et al.* Prediction of cardiovascular complications in diabetes from pharmacy administrative claims. 2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON). Palermo: IEEE; 2020.
38. Giardina M, Azuaje F, Mccullagh P, Harper R. A supervised learning approach to predicting coronary heart disease complications in type 2 diabetes mellitus patients. Sixth IEEE Symposium on BioInformatics and BioEngineering (BIBE'06). Arlington, VA: IEEE; 2006.
39. Phan TP, Nguyen PA, Vo TH, *et al.* Artificial intelligent-based personalized predictive ischemic stroke among type 2 diabetes mellitus complication patients 2023. Available from: https://www.researchsquare.com/article/rs-3311009/v1. Accessed: 1 May 2024
40. Liu B, Li Y, Sun Z, *et al.* Early prediction of diabetes complications from electronic health records: A multi-task survival analysis approach. The Thirty-Second AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press; 2018.
41. Ding J, Luo Y, Shi H, *et al.* Machine learning for the prediction of atherosclerotic cardiovascular disease during 3-year follow up in Chinese type 2 diabetes mellitus patients. J Diabetes Investig 2023;14(11):1289-1302.
42. Zhong Z, Sun S, Weng J, *et al.* Machine learning algorithms identifying the risk of new-onset ACS in patients with type 2 diabetes mellitus: A retrospective cohort study. Front Public Health 2022;10:947204.
43. Miao L, Guo X, Abbas HT, *et al.* Using machine learning to predict the future development of disease. 2020 International Conference on UK-China Emerging Technologies (UCET). Glasgow: IEEE; 2020.

44. Derevitskii IV, Kovalchuk SV. Machine learning-based predictive modeling of complications of chronic diabetes. Procedia Comput Sci 2020;178:274-283.
45. Fan R, Zhang N, Yang L, *et al.* AI-based prediction for the risk of coronary heart disease among patients with type 2 diabetes mellitus. Sci Rep 2020;10(1):14457.
46. Xu E, Li X, Mei J, *et al.* Applying risk models on patients with unknown predictor values: An incremental learning approach. Stud Health Technol Inform 2017;245:639-643.
47. Wang T, Wang M, Liu L, *et al.* Lower serum branched-chain amino acid catabolic intermediates are predictive signatures specific to patients with diabetic foot. Nutr Res 2023;119:33-42.
48. Mei J, Xia E. Knowledge learning symbiosis for developing risk prediction models from regional EHR repositories. Stud Health Technol Inform 2019;264:258-262.
49. Nanda R, Nath A, Patel S, *et al.* Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity. Med Biol Eng Comput 2022;60(8):2349-2357.
50. Ramakrishnan J, Kumar R, Senthilkumar G, *et al.* Diagnosing cardiovascular disease via intelligence in healthcare multimedia: A novel approach. Int J Nanotechnol 2023;20:182.
51. Evangelin Sonia SV, Nedunchezhian R, Rajalakshmi M. A multi-modal integrated deep neural networks for the prediction of cardiovascular disease in type-2 diabetic males. Automatika 2023;64(4):1315-1327.
52. Selvarathi C, Varadhaganapathy S. Deep learning based cardiovascular disease risk factor prediction among type 2 diabetes mellitus patients. Inform Technol Control 2023;52(1):215-227.
53. Farzi S, Kianian S, Rastkhadive I. Predicting serious diabetic complications using hidden pattern detection. 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI). Tehran: IEEE; 2017.
54. Rahman T, Farzana SM, Khanom AZ. Prediction of diabetes induced complications using different machine learning algorithms. Dhaka: BRAC University; 2018.
55. Rajathi K, Jeyarani RA, Christiana KB, Vahini TV. Predicting cardiovascular disease as a long-term diabetes complication using SOM. Int J Recent Technol Eng 2020;8(6):3288-3292.
56. Afarideh M, Aryan Z, Ghajar A, *et al.* Complex association of serum alanine aminotransferase with the risk of future cardiovascular disease in type 2 diabetes. Atherosclerosis 2016;254:42-51.
57. Sierra-Sosa D, Garcia-Zapirain B, Castillo C, *et al.* Scalable healthcare assessment for diabetic patients using deep learning on multiple GPUS. IEEE Trans Industr Inform 2019;15(10):5682-5689.
58. Kim E, Caraballo PJ, Castro MR, *et al.* Towards more accessible precision medicine: Building a more transferable machine learning model to support prognostic decisions for micro- and macrovascular complications of type 2 diabetes mellitus. J Med Syst 2019;43(7):185.
59. Kim E, Pieczkiewicz DS, Castro MR, *et al.* Multi-task learning to identify outcome-specific risk factors that distinguish individual micro and macrovascular complications of type 2 diabetes. AMIA Jt Summits Transl Sci Proc 2018;2017:122-131.
60. Kanda E, Suzuki A, Makino M, *et al.* Machine learning models for prediction of HF and CKD development in early-stage type 2 diabetes patients. Sci Rep 2022;12(1):20012.
61. Liu S, Zhang R, Shang X, *et al.* Analysis for warning factors of type 2 diabetes mellitus complications with Markov blanket based on a Bayesian network model. Comput Methods Programs Biomed 2020;188:105302.
62. Dworzynski P, Aasbrenn M, Rostgaard K, *et al.* Nationwide prediction of type 2 diabetes comorbidities. Sci Rep 2020;10(1):1-13.
63. Liu B, Li Y, Ghosh S, *et al.* Simultaneous modeling of multiple complications for risk profiling in diabetes care. arXiv 2018: 1802.06476.
64. Thomas PB, Robertson DH, Chawla NV. Predicting onset of complications from diabetes: A graph based approach. Appl Netw Sci 2018;3(1):48.
65. Nanda R, Nath A, Patel S, *et al.* Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity. Med Biol Eng Comput 2022;60(8):2349-2357.
66. Lloyd-Jones DM, Wilson PWF, Larson MG, *et al.* Framingham risk score and prediction of lifetime risk for coronary heart disease. Am J Cardiol 2004;94(1):20-24.
67. Chamnan P, Simmons RK, Sharp SJ, *et al.* Cardiovascular risk assessment scores for people with diabetes: A systematic review. Diabetologia 2009;52(10):2001-2014.
68. Wang J, Wang S, Zhu MX, *et al.* Risk prediction of major adverse cardiovascular events occurrence within 6 months after coronary revascularization: Machine learning study. JMIR Med Inform 2022;10(4):e33395.
69. Jimma BL. Artificial intelligence in healthcare: A bibliometric analysis. Telemat Inform Rep 2023;9:100041.

70. Franzen SRP, Chandler C, Lang T. Health research capacity development in low and middle income countries: Reality or rhetoric? A systematic meta-narrative review of the qualitative literature. BMJ Open 2017;7(1):e012332.
71. Yegros-Yegros A, van de Klippe W, Abad-Garcia MF, *et al.* Exploring why global health needs are unmet by research efforts: the potential influences of geography, industry and publication incentives. Health Res Policy Syst 2020;18(1):47.
72. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol 2022;75(1):25-36.
73. de Hond AAH, Steyerberg EW, van Calster B. Interpreting area under the receiver operating characteristic curve. Lancet Digit Health 2022;4(12):e853-e855.
74. Nematzadeh S, Kiani F, Torkamanian-Afshar M, *et al.* Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases. Comput Biol Chem 2022;97:107619.
75. Pavithrakannan R, Fenn NB, Raman S, *et al.* Imputation analysis of central tendencies for classification. 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). Toronto: IEEE; 2021.
76. Kim JC, Chung K. Multi-modal stacked denoising autoencoder for handling missing data in healthcare big data. IEEE Access 2020;8:104933-104943.
77. Mausor FH, Jaafar J, Taib SM. Missing values imputation using Fuzzy C means based on correlation of variable. 2020 International Conference on Computational Intelligence (ICCI). Bandar Seri Iskandar: IEEE; 2020.
78. Wassan S, Suhail B, Mubeen R, *et al.* Gradient boosting for health IoT federated learning. Sustainability 2022;14(24):16842.
79. Padmanabhan S, Tran TQB, Dominiczak AF. Artificial intelligence in hypertension. Circ Res 2021;128(7):1100-1118.
80. Siontis GCM, Tzoulaki I, Castaldi PJ, *et al.* External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol 2015;68(1):25-34.
81. Smith LA, Oakden-Rayner L, Bird A, *et al.* Machine learning and deep learning predictive models for long-term prognosis in patients with chronic obstructive pulmonary disease: A systematic review and meta-analysis. Lancet Digit Health 2023;5(12):e872-e881.
82. Silva K De, Lee WK, Forbes A, *et al.* Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis. Int J Med Inform 2020;143:104268.
83. Steyerberg EW, Moons KGM, van der Windt DA, *et al.* Prognosis research strategy (PROGRESS) 3: Prognostic model research. PLoS Med 2013;10(2):e1001381.
84. Martens FK, Kers JG, Janssens ACJW. External validation is only needed when prediction models are worth it (Letter commenting on: J Clin Epidemiol. 2015;68:25-34). J Clin Epidemiol 2016;69:249-250.
85. Wessler BS, Nelson J, Park JG, *et al.* External validations of cardiovascular clinical prediction models: A large-scale review of the literature. Circ Cardiovasc Qual Outcomes 2021;14(8):e007858.
86. Mora T, Roche D, Rodríguez-Sánchez B. Predicting the onset of diabetes-related complications after a diabetes diagnosis with machine learning algorithms. Diabetes Res Clin Pract 2023;204:110910.
87. Du Y, Rafferty AR, McAuliffe FM, *et al.* An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. Sci Rep 2022;12(1):1170.
88. Yousefi L, Tucker A. Predicting type 2 diabetes complications and personalising patient using artificial intelligence methodology. In: Stoian AP, editor. Type 2 diabetes - from pathophysiology to cyber systems. London: IntechOpen; 2021.