Intelligence-Based Medicine 6 (2022) 100072

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

Artificial intelligence models for predicting cardiovascular diseases in people with type 2 diabetes: A systematic review

Minhong Wang [a,*], Farah Francis [b], Holger Kunz [a], Xiang Zhang [d], Cheng Wan [c], Yun Liu [c], Paul Taylor [a], Sarah H. Wild [b], Honghan Wu [a,**]

[a] Institute of Health Informatics, University College London, London, UK
[b] Usher Institute, University of Edinburgh, Edinburgh, UK
[c] Department of Medical Informatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China
[d] School of Computer Science and Engineering, Southeast University, Nanjing, China

ARTICLE INFO

Keywords:
Cardiovascular disease
Type 2 diabetes
Machine learning
Systematic review

ABSTRACT

Background: People with type 2 diabetes have a higher risk of cardiovascular disease morbidity and mortality. We aim to distil the evidence, summarize the developments, and identify the gaps in relevant research on predicting cardiovascular disease in type 2 diabetes people using AI techniques in the last ten years.
Methods: A systematic search was carried out for literature published between 1st January 2010 and 30th May 2021 in five medical and scientific databases, including Medline, EMBASE, Global Health (CABI), IEEE Xplore and Web of Science Core Collection. All English language studies describing AI models for predicting cardiovascular diseases in adults with type 2 diabetes were included. The retrieved studies were screened and the data from included studies were extracted by two reviewers. The survey and synthesis of extracted data were conducted based on predefined research questions. IJMEDI checklist was used for quality assessment.
Results: From 176 articles identified by the search, 5 studies with sample sizes ranging from 560 to 203,517 met our inclusion criteria. The models predicted the risk of multiple cardiovascular diseases over 5 or 10 years. Ensemble learning, particularly random forest, is the most used algorithm in these models and consistently provided competitive performance. Commonly used features include age, body mass index, blood pressure measurements, and cholesterol measurements. Only one study carried out external validation. The area under the receiver operating characteristic curve for derivation cohorts varied from 0.69 to 0.77. AI models achieved better performance than conventional models in some specific scenarios.
Conclusions: AI technologies seem to show promising performance (AUROC in external validation: 0.75 compared to 0.69 from conventional risk scores) for cardiovascular disease prediction in type 2 diabetes people. However, only one of the reviewed models conducted an external validation. Quality of reporting was low in general, and all models lack reproducibility and reusability.

1. Introduction

Type 2 diabetes is a complex disease with increasing prevalence worldwide [1] that leads to continuing and comprehensive medical care requirements [2]. People with type 2 diabetes have risks of death and cardiovascular events that are 2–4 times higher than in the general population [3,4]. Risk calculators for CVD events tailored for people with type 2 diabetes would be helpful to inform treatment decisions.

A previous study [5] carried out a systematic review of published studies up to 2011 on prediction models for the risk of CVD in people with type 2 diabetes. They identified 12 models specifically developed for people with type 2 diabetes. All were traditional statistical models. A more recent review [6] identified 19 models for risk of CVD in people with type 2 diabetes and concluded that many have a high risk of bias, and few have been externally validated. A comparison of the performance of risk scores for CVD for people with type 2 diabetes has been performed by Ref. [7] and by Ref. [8]. Both studies carried out an external validation of existing Cox proportional hazards models and

* Corresponding author.
** Corresponding author.
E-mail addresses: minhong.wang@ucl.ac.uk (M. Wang), honghan.wu@ucl.ac.uk (H. Wu).

https://doi.org/10.1016/j.ibmed.2022.100072
Received 9 May 2022; Received in revised form 25 July 2022; Accepted 16 August 2022
Available online 24 August 2022

concluded that CVD risk scores developed by conventional methods could not accurately identify the risk of CVD events among people with type 2 diabetes.

The question therefore arises whether AI models can improve CVD risk prediction in people with type 2 diabetes. The aim of this paper is to systematically identify, describe and critically appraise studies that have used AI models to predict CVD in people with type 2 diabetes. The review will consider the data used for model development and validation, the AI algorithms adapted for modelling, and the performance of currently existing models.

2. Methods

To carry out a systematic review on existing AI models for predicting CVD in people with type 2 diabetes, we registered our protocol in PROSPERO (reg no: CRD42021255001) and have used the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Guidelines to report our findings [9].

2.1. Search strategy

We comprehensively searched publications from 1st January 2010 to 30th May 2021 in Medline, EMBASE, Global Health (CABI), IEEE Xplore and Web of Science Core Collection. Keywords used for the search were "Cardiovascular disease", "CVD", "Type 2 diabetes", "Score", "Predictor"," Prediction", "Machine learning" and "Artificial intelligence". We delivered exploded search on the terms "Cardiovascular disease", "Type 2 diabetes", "Machine learning" and "Artificial intelligence" in Medline, EMBASE, and Global Health to capture all narrower terms in associated with the broader concepts. As all models reported in the systematic review carried out in 2012 in the field [5] were traditional statistical models, we included ten years studies from 2010 onward for analyzing AI models application in this field. The full search strategy was provided in the Supplementary table 3.

2.2. Inclusion and exclusion criteria

We included studies that 1) were carried out on T2D adults cohorts, 2) predicted the risk of T2D patients developing any type of CVD (over a time period), 3) developed AI models (machine learning or deep learning models), 4) developed prediction models over a time period, and 5) the publication is in English. Our exclusion criteria were: 1) studies not written in English, 2) non-AI predictive models (for example conventional statistical models using the proportional hazards approach), 3) models that are not re-implementable, 4) biomarker studies, 5) drug target studies, 6) classification models in cross-sectional studies.

2.3. Data extraction

Two reviewers (MW and FF) used Covidence (https://www.covidence.org/) for screening eligible studies and data extraction independently. According to the inclusion and exclusion criteria, we identified eligible studies based on the title, abstract, and full text. Any discrepancies between the two reviewers were discussed, and a third reviewer (HW) was consulted when no consensus was achieved.

2.4. Data analysis

Results from the data extraction were analyzed according to the questions predetermined before starting the review. The questions cover the aspects of the source of data, model development and model performance. Narrative (descriptive) summaries were provided on a qualitative attribute, such as prediction model development methods, performance measures used, patient selections, and the used dataset.

2.5. Quality assessment

Two reviewers (MW and FF) independently used the IJMEDI checklist [10] to evaluate the quality of the included studies. All divergence has been solved by discussion. The checklist contains 30 questions and covers six dimensions of the studies, including problem understanding, data understanding, data preparation, modeling, validation, and deployment. Each question can be answered as OK (adequately addressed), mR (sufficient but improvable), and MR (inadequately addressed). Based on previous studies [11], we attempted to assign the scores to the answers based on the priority of the items. We assigned 2,1, and 0 points for OK, mR and MR in high-priority items, and assigned 1, 0.5, and 0 points respectively in low-priority items. The maximum number of points is 50 and study quality can be divided into low (0–19.5), medium (20–34.5), and high (35–50).

The quality of the reports was also assessed by following the reporting guideline TRIPOD-AI checklist [12].

3. Results

Using our search strategies in five databases, we identified five studies from which we extracted and summarized data (see Fig. 1 for a flow chart describing study selection).

3.1. Data sources

Table 1 describes the characteristics of the populations used to derive AI models for CVD risk among people with type 2 diabetes. Ref. [14,17] were carried out on the same dataset. Only three out of the five studies provided information on demographic characteristics.

Only two [13,16] out of the five studies provided information on missing data and the method of handling missing data. Ref. [13] provided information on missing data for each variable. Ref. [16] excluded participants who had >10% missing data. Imputation was used for both studies to handle missing data.

Only two studies provided information on the method of selecting people with type 2 diabetes. Ref. [15] used ICD 10 code and Ref. [13] provided a comprehensive cohort derivation protocol.

For external validation, Ref. [16] used a subset of ALLHAT participants. Ref. [13] used other eligible non-surgical patients' data from the same clinic health system for model validation. The other three studies only used a development dataset for model evaluation. No study conducted external validation using data from another population.

3.2. Model development summary

Table 2 provides the summary of model development from the selected studies, four of which developed models for predicting CVD risk in the next 5 years, with one using a 10-year time period. All studies adapted machine learning data-driven approaches for model developments. Random forest modelling was used in three studies. Ref. [16] used the random forest approach for feature selection followed by generating a proportional hazards regression model.

Only three studies provided information on the data split ratio and hyper-parameter tuning for model development. Grid search was the most common approach for hyper-parameter tuning. All studies faced challenges from imbalanced data. Three provided information on how this was addressed using class weights, sub-sampling, and ensemble modelling.

It is worth noting that the reported models have not addressed reproducibility. Ref. [13] provided a web version and smartphone application of the risk score, Ref. [15] published the source code for model development online. However, none of the studies provided the parameters or code for their models. Hence, none of the models can be reproduced to allow use in other datasets.

Table 2 also includes the features used in each of the studies.

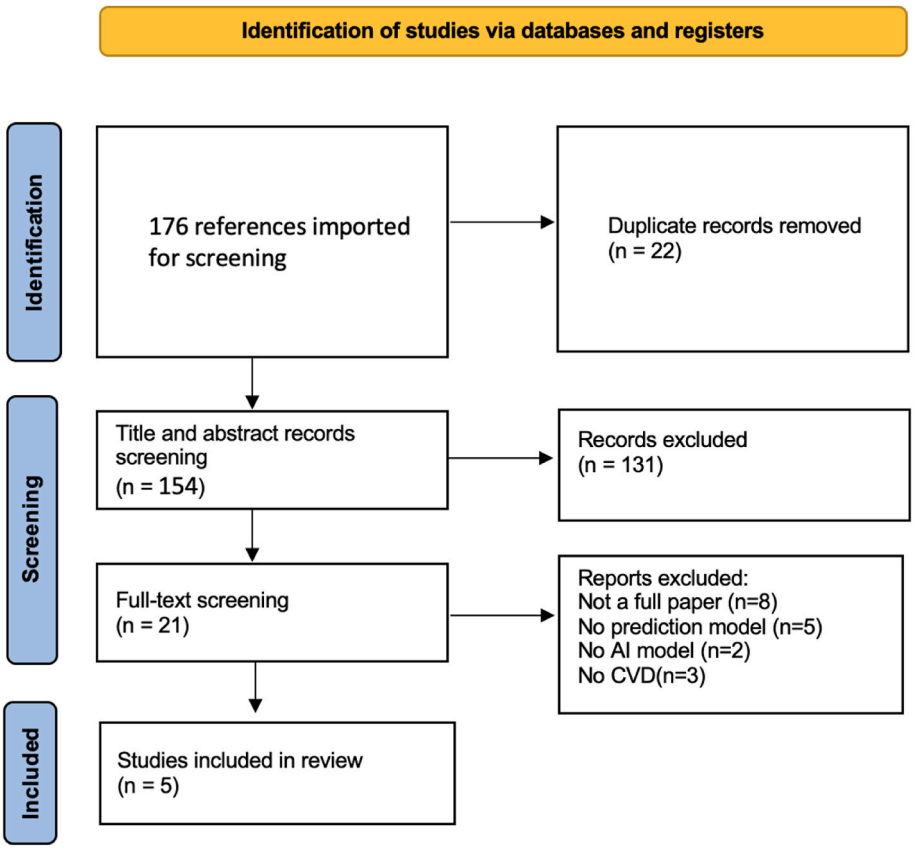

Fig. 1. Flow chart describing study identification for review of AI models for CVD risk prediction among people with type 2 diabetes.

Common features include age, body mass index, blood pressure measurements, and cholesterol measurements. In respect of the interpretability of the models, Ref. [14] demonstrated that the SHAP value can be used to explain machine learning models. The top 3 important features are BMI, HDL Cholesterol, and glycosylated hemoglobin. Ref. [13,15] analyzed the feature importance in random forest models and gradient boosting models. As Ref. [17] carried out models with higher complexity, they pointed out that making the representation of each risk factor's impact and interpreting the predicted outputs are more difficult.

3.3. Model performance summary

Multiple outcomes were used in the studies. Ref. [15] developed multiple models for predicting heart failure, myocardial infarction, stroke, and cardiovascular disease separately. Ref. [17] developed models for coronary heart disease and stroke. Ref. [14] developed models for fatal or non-fatal cardiovascular disease including stroke and coronary heart disease. Ref. [13] developed models for all-cause mortality, coronary artery disease events, heart failure, and nephropathy. Ref. [16] developed models for incident heart failure.

The area under the receiver operating characteristic (AUROC) was used for evaluating the model performance in all selected studies. Table 3 summarized the performance of the best performing models within each study based on AUROC (ranging from 0.69 to 0.77). The results from other metrics were also included based on the reports from each study. It is important to note that, although the AUROC is a widely used metric for comparing machine learning algorithms, it is not best suited for use with imbalanced datasets.

Refs. [13,15] compared the performance of machine learning models and regression models (including Cox proportional hazards and logistic regression). Although the difference in performance is not large, the results showed ML models outperformed regression models in many outcomes of interest, including predicting heart failure, myocardial infarction, cardiovascular disease, and coronary artery disease events. One of the potential causes of these results might be that logistic regression is not well suited for predicting long-term outcomes. However, the studies did not carry out a time-to-event analysis to compare different models, which could assist on explaining the performance discrepancy.

3.4. Reporting quality

Table 4 summarizes the quality report based on the IJMEDI checklist for assessing AI model development and validation. The detailed results of the quality assessment are provided in the Supplementary table 1. While there is no study identified as high quality based on the checklist, all studies fell into medium quality (20–34.5 points) based on the checklist. While all studies addressed the aspects of problem understanding and modeling with high quality, the descriptions of data understanding, data preparation, validation and deployment are insufficient. Fig. 2 shows the proportion of the answers in different categories in the high- and low-priority items.

The quality report based on TRIPOD checklist is provided in Supplement Table 2.

4. Discussion

This review describes five AI models for CVD risk prediction that have been specifically derived for people with type 2 diabetes. This is a smaller number of models compared to the models that have been developed using traditional proportional-hazard models.

All models were developed in European and North American populations indicating restricted geographic and socio-demographic diversity. This includes a lack of models developed for populations of low-

Table 1
Summary of data sources used for model development among studies describing AI models for CVD risk prediction among people with type 2 diabetes.

Study	Source of data	Practice setting	Country	Study population size	Outcome (proportion affected)	Cohort characteristics					
						Age, years, (mean ± SD)	Sex (% women)	BMI, kg/m2 (mean ± SD)	Race/ethnicity	Diabetes duration (mean ± SD)	HbA1c (%) (mean ± SD)
[13]	Electronic health records	–	The United States	13,722 Surgical group: 2287	Coronary artery events: 3.67% Heart failure: 3.32%	52.5 (43.7, 60.5)[a]	65.5	45.1 (40, 51.8)[a]	White 75.8% Black 19.3% Other 2.4% Missing 2.5%	–	7.1 (6.3, 8.2)[a]
				Nonsurgical group:11,435	Coronary artery events:5.91% Heart failure:9.23%	54.8 (46.2, 62.5)[a]	64.2	42.6 (39.4, 47.2)[a]	White 69.9% Black 24.5% Other 2.1% Missing 3.5%	–	7.1 (6.4, 8.4)[a]
[14]	Electronic health records	Hospital data	Greece	560	CVD 7.32%	–	–	–	–	–	–
[15]	Patient register	Hospital data	Denmark	203,517	CVD 27.52%	–	47.03	–	–	–	–
[16]	Accord trial and ALLHAT	Clinical Trial	The United State and Canada	8,756	Incident heart failure: 3.6%	62.7 ± 6.6	38.5	32.1 ± 5.4	Black 18.5% Hispanic 7.5% Other 11.3% White 62.7%	9.0 (10.0)[a]	8.3 (1.1)[b]
[17]	Electronic health records	Hospital data	Greece	560	CVD 7.32%	58.56 ± 10.70	53.04	29.49 ± 5.54	–	7.67 ± 7.37	7.43 ± 1.81

[a] Median (interquartile range).
[b] mg/dL.

or middle-income countries that have a different mix of lifestyles and distribution of ethnic groups, which affect the prevalence of diabetes and its link with CVD risk. The average age of included populations is more than 50 years, reflecting the importance of age as a risk factor for diabetes and CVD.

All studies included an imbalanced dataset where as little as 3.3% of people develop the outcome of interest. While this accurately represents the real prevalence of the outcome in the population of interest, there is a risk that prediction will be biased towards the larger subgroup as some studies did not report their approach to handle data imbalance. Variation in the degree of imbalance between studies may also mean that the AUROC is not the best metric on which to compare performance. While all studies used AUROC in model evaluation, there is an adequacy of measures used in model evaluation. By using metrics that focus on one class, for example sensitivity-specificity metrics and precision-recall metrics, can help us to understand the performance from imbalanced datasets. However, most of the studies did not report the results from these metrics.

While multiple studies proved that machine learning models achieved good performance within the development data, there are still many challenges in deploying AI models in clinical settings. High-quality large datasets are needed to develop the models that should be externally validated. None of our selected studies published the models online or elsewhere, which makes it impossible to reproduce and externally validate the models.

Refs. [7,8] reported that several existing Cox models only achieved AUROC ranging from 0.66 to 0.67 and 0.62–0.69 separately in external validation carried out in separate UK populations. One of the five reviewed studies carried out external validation and reported an AUROC of 0.74, which is a promising result from AI-assisted models. However, there are no studies that carried out external validation in a different population from the development data. Even though reported AI models achieved good performance internally, most models still lack external validation. The generalizability to low- or middle-income countries also needs to be assessed. Hence, the robustness of the models still needs to be further investigated.

In addition, while all studies focused on predicting CVD events, there was a wide variety of outcomes, including but not limit to death, coronary artery events, heart failure, myocardial infarction, and stroke. The performance of the best prediction model might be heavily dependent on the type of outcome. A further analysis on comparing performance based on specific outcome could provide us a better understanding on the best performance models.

Compared to traditional proportional hazards models, AI models also face challenges in interpretability. Having an explainable AI model, one that is less of a 'black box', would encourage clinical adoption [18]. Within the selected studies, reporting feature importance was the only approach to increase the interpretability of the models. Ref. [13,15] analyzed the feature importance and Ref. [14] reported the SHAP value of the features.

It is established that people with metabolic diseases such as type 2 diabetes have comorbidities that provide challenges when developing prediction models [19,20]. Prediction studies need to adapt to changes in risk factors such as high cholesterol, high blood pressure and kidney disease [21]. For instance, none of the studies considers interventions such as medications, treatments, and lifestyle changes, which may affect

Table 2
Summary of model development processes and used treatments among studies describing AI models for CVD risk prediction among people with type 2 diabetes.

Study	Data pre-processing		Feature selection	Imbalance treatment	Hyper-parameter tuning	Modelling algorithm	Time period	Features
	Handling missing data	Data split ratio (train:test)						
[13]	Imputation	80:20	Not provided	Not provided	Not provided	Random forest	10 years	Age, BMI, Heart failure, Insulin, Smoking status, non-insulin diabetes medications, Other antihypertensive medications, Sex, Nephropathy, COPD, eGFR, Dyslipidemia, Race, Systolic blood pressure, Renin-angiotensin system inhibitors, Cerebrovascular disease, Hypertension, HbA1c, Lipid-lowering medications, Peripheral arterial disease, Aspirin, Warfarin, Diastolic blood pressure, Coronary artery disease, Diabetic neuropathy, Triglycerides
[14]	Not provided	90:10	Incremental feature selection	Sub-sampling approach	Grid search	XGBoost	5 years	BMI, HbA1c, HDL Cholesterol, Total Cholesterol, Triglycerides, Age, Fasting Glucose, Diabetes duration, Pulse Pressure, Smoking Habit, Hypertension, Lipid-lowering therapy, Parental History of Diabetes
[15]	Not provided	70:30 (test set split for model selection and final validation)	Not provided	Class weights	Grid search with three-fold cross-validation on the training data.	Logistic regression, random forest, gradient boosting	5 years	26 canonical features, 3,423 hospital diagnoses, 2,015 hospital procedures, 670 prescriptions, and 47 primary care interactions
[16]	Imputation	50:50	Stepwise backward selection, stepwise forward selection, and permutation-based random survival forest (RSF) selection	Not provided	Not provided	Random survival forest (RSF) based Cox PH relationship modelling [a]	5 years	Age, BMI, SBP, FPG, QRS, Serum Cr, DBP, HDL-C, Prior MI, Prior CABG
[17]	Not provided	90:10	Not provided	Different ensemble methods	Theoretical justifications given for parameter values	Hybrid of ensembles of wavelet neural network, and self-organizing maps	5 years	Age, diabetes duration, BMI, HbA1c, pulse pressure, fasting glucose, total cholesterol, triglycerides, HDL cholesterol, smoking, sex, hypertension, lipid lowering therapy, aspirin, insulin therapy and parental history of diabetes

Abbreviation and explanation: BMI: Body mass index; COPD: chronic obstructive pulmonary disease; eGFR: estimated glomerular filtration rate; HbA1c: glycosylated hemoglobin; HDL/HDL-C: high-density lipoprotein cholesterol; SBP: systolic blood pressure; FPG: fast plasma glucose; CR: creatinine; DBP: diastolic blood pressure; MI: myocardial infarction; CABG: coronary artery bypass grafting.

[a] Used random forest for feature selection following by generating proportional hazards regression model.

Table 3
Summary of performance from the best performing model among studies describing AI models for CVD risk prediction among people with type 2 diabetes.

Study	Internal validation setup		Model performance (Internal)			Model performance (External)
	Cross-validation	Stratified data split or not	AUROC	Other metrics		
				Metric	Result	
[13]	5-fold	Stratified	0.75	Index of prediction accuracy	0.14	–
[14]	10-fold	Stratified	0.7113 ± 0.1169 (Mean ± SD)	Sensitivity (Mean ± SD)	0.71 ± 0.2385	–
[15]	3-fold[a]	–	0.69(95% CI 0.68–0.70)	–	–	–
[16]	–	Random	0.77[95% CI 0.75–0.80]	Hosmer-Lemeshow statistic	$\chi^2 = 9.63$, P = 0.29	0.74[95% CI 0.72–0.76]
[17]	10-fold	DUPLEX data splitting	0.7148 ± 0.1573 (Mean ± SD)	Accuracy (Mean ± SD)	0.7179 ± 0.0906	–
				Specificity (Mean ± SD)	0.7264 ± 0.0944	
				Sensitivity (Mean ± SD)	0.61 ± 0.2665	
				Brier score (Mean ± SD)	0.0007 ± 0.0002	

[a] 3-fold cross-validation carried out in training data for hyperparameter tuning and model type selection.

risks and that therefore need to be taken into consideration when developing prediction models [22–24]. From this review, there were no indications on how the prediction models might be able to incorporate these factors. Only one study [13] attempted to predict when a CVD event would occur. Being able to estimate when CVD is likely to develop may be helpful in managing type 2 diabetes, allowing clinicians and patients to work towards a realistic target to prevent a CVD event. As the onset of CVD is a time-dependent event and might not be adequately

Table 4
Quality assessment scores of the 5 included studies according to the IJMEDI checklist.

	Problem Understanding (10)	Data Understanding (6)	Data Preparation (8)	Modeling (6)	Validation (12)	Deployment (8)	Total (50)
[13]	10	3	2	6	6	1.5	28.5
[14]	9	1	2	6	8	2	28
[15]	10	3	2	6	9	1.5	31.5
[16]	10	3	2	6	7	1	29
[17]	9	2	2	6	10	0.5	29.5

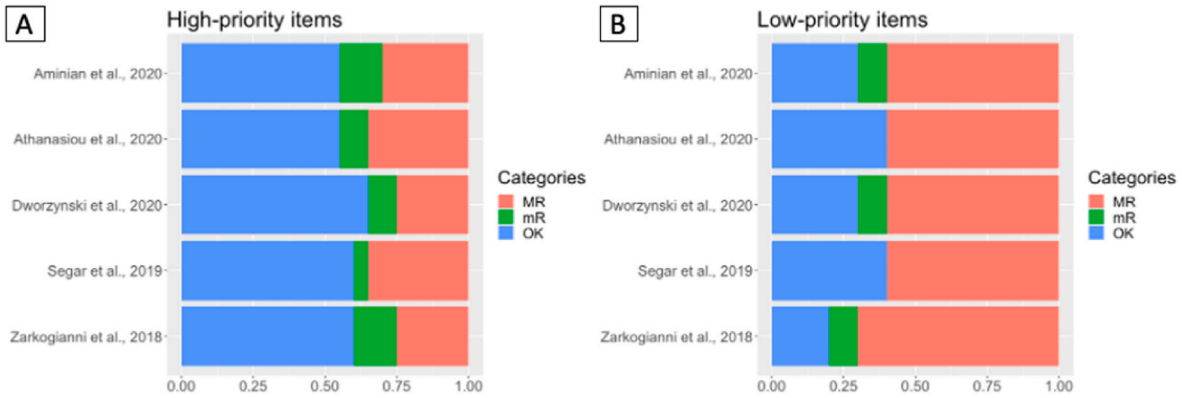

Fig. 2. Proportion of the answers in different categories in the A) high- and B) low-priority items. OK = adequately addressed; mR = sufficient but improvable; MR = inadequately addressed.

captured by a binary classifier, classification AI models might not be sufficient for predicting CVD. Currently, there is still a lack of applying regression AI models in this field.

Another point to consider is using clinically appropriate predictors for the model depending on whether incident or recurrent CVD events are the outcome of interest. For example, Ref. [13] used prevalent CVD outcomes such as heart failure and coronary artery disease to predict further CVD events that could be helpful for supporting secondary prevention.

The main limitation of this review is the small number of studies describing CVD prediction models developed for type 2 diabetes using AI, one of which had a particularly small sample size. Next, this study is limited to English journal publications which exclude studies in other languages.

In conclusion, we have identified five studies that have developed AI risk prediction models for CVD in people with type 2 diabetes. AI models have the potential to achieve better performance than traditional models, but also need to be reproducible, externally validated, and found to be of value in clinical practice. Ref. [25] concluded in their systematic review on more generic CVD predictions: "*The usefulness of most of the models remains unclear owing to methodological shortcomings, incomplete presentation, and lack of external validation and model impact studies*." Six years on, our review draws the same conclusion.

Authors' contributions

MW, FF and HW designed the study. MW and FF conducted the literature screening. HW was consulted for screening. MW and FF reviewed the papers and extracted the data. MW analyzed the data. MW, FF and HW interpreted the data. MW and FF wrote the first draft of the manuscript. SHW, HK, PT, XZ, CW, YL and HW critically appraised the manuscript. All authors approved the submission for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Summary table

- An extensive number of models have been developed to predict the risk of cardiovascular diseases in people with type 2 diabetes.
- Most previous studies developed and externally validated convention methods of risk prediction such as Cox proportional hazards models.
- Previous external validation showed that cardiovascular disease risk scores developed by conventional methods could not accurately identify the risk of cardiovascular disease among people with type 2 diabetes.
- This study identified and reviewed existing AI models for predicting risks of cardiovascular diseases in people with type 2 diabetes.
- Within the limited number of studies, AI models showed the potential to achieve more accurate predictions compared to conventional risk prediction methods.
- This study pointed out the limitations of the current AI modelling studies, including the absence of reproducible models, omission of preprocessing details, limited description of data provenance, and the lack of external validation.

Acknowledgements

This study was supported by Medical Research Council and Health Data Research UK (MR/S004149/1, MR/S004149/2); British Council (UCL-NMU-SEU international collaboration on Artificial Intelligence in Medicine: tackling challenges of low generalizability and health inequality); National Institute for Health Research (NIHR202639); Advanced Care Research Centre at the University of Edinburgh.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmed.2022.100072.

References

[1] Einarson TR, Acs A, Ludwig C, Panton UH. Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007–2017. Cardiovasc Diabetol 2018;17(1). https://doi.org/10.1186/s12933-018-0728-6.

[2] Rawshani A, Rawshani A, Franzén S, Sattar N, Eliasson B, Svensson A-M, Zethelius B, Miftaraj M, Mcguire DK, Rosengren A, Gudbjörnsdottir S. Risk factors, mortality, and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med 2018;379(7):633–44. https://doi.org/10.1056/nejmoa1800256.

[3] Rawshani A, Rawshani A, Franzén S, Eliasson B, Svensson A-M, Miftaraj M, Mcguire DK, Sattar N, Rosengren A, Gudbjörnsdottir S. Mortality and cardiovascular disease in type 1 and type 2 diabetes. N Engl J Med 2017;376(15):1407–18. https://doi.org/10.1056/nejmoa1608664.

[4] Larsson SC, Wallin A, Håkansson N, Stackelberg O, Bäck M, Wolk A. Type 1 and type 2 diabetes mellitus and incidence of seven cardiovascular diseases. Int J Cardiol 2018;262:66–70. https://doi.org/10.1016/j.ijcard.2018.03.099.

[5] van Dieren S, Beulens JW, Kengne AP, Peelen LM, Rutten GE, Woodward M, van der Schouw YT, Moons KG. Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review. Heart 2012;98(5):360–9. https://doi.org/10.1136/heartjnl-2011-300734.

[6] Galbete A, Tamayo I, Librero J, Enguita-Germán M, Cambra K, Ibáñez-Beroiz B. Cardiovascular risk in patients with type 2 diabetes: a systematic review of prediction models. Diabetes Res Clin Pract 2021:109089. https://doi.org/10.1016/j.diabres.2021.109089.

[7] Read SH, van Diepen M, Colhoun HM, Halbesma N, Lindsay RS, McKnight JA, McAllister DA, Pearson ER, Petrie JR, Philip S, Sattar N, Woodward M, Wild SH, Scottish Diabetes Research Network Epidemiology G. Performance of cardiovascular disease risk scores in people diagnosed with type 2 diabetes: external validation using data from the national scottish diabetes register. Diabetes Care 2018;41(9):2010–8. https://doi.org/10.2337/dc18-0578.

[8] Dziopa K, Asselbergs FW, Gratton J, Chaturvedi N, Schmidt AF. Cardiovascular risk prediction in type 2 diabetes: a comparison of 22 risk scores in primary care settings. Diabetologia 2022. https://doi.org/10.1007/s00125-021-05640-y.

[9] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, Mcdonald S, Mckenzie JE. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021;n160. https://doi.org/10.1136/bmj.n160.

[10] Campagner FCA. The need to separate the wheat from the chaff in medical informatics. Int J Med Inf 2021;153:104510. https://doi.org/10.1016/j.ijmedinf.2021.104510.

[11] Zhou Y, et al. Machine learning predictive models for acute pancreatitis: a systematic review. Int J Med Inf 2022;157:104641. https://doi.org/10.1016/j.ijmedinf.2021.104641.

[12] Heus P, Damen JAAG, Pajouheshnia R, Scholten RJPM, Reitsma JB, Collins GS, Altman DG, Moons KGM, Hooft L. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. BMJ Open 2019;9(4):e025611. https://doi.org/10.1136/bmjopen-2018-025611.

[13] Aminian A, Zajichek A, Arterburn DE, Wolski KE, Brethauer SA, Schauer PR, Nissen SE, Kattan MW. 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):852–9. https://doi.org/10.2337/dc19-2057.

[14] Athanasiou M, Sfrintzeri K, Zarkogianni K, Thanopoulou AC, Nikita KS. An explainable XGBoost–based approach towards assessing the risk of cardiovascular disease in patients with Type 2 Diabetes Mellitus. 2020-10-01. In: 2020 IEEE 20th international conference on bioinformatics and bioengineering (BIBE); 2020.

[15] Dworzynski P, Aasbrenn M, Rostgaard K, Melbye M, Gerds TA, Hjalgrim H, Pers TH. Nationwide prediction of type 2 diabetes comorbidities. Sci Rep 2020;10(1). https://doi.org/10.1038/s41598-020-58601-7.

[16] Segar MW, Vaduganathan M, Patel KV, Mcguire DK, Butler J, Fonarow GC, Basit M, Kannan V, Grodin JL, Everett B, Willett D, Berry J, Pandey A. 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):2298–306. https://doi.org/10.2337/dc19-0587.

[17] Zarkogianni K, Athanasiou M, Thanopoulou AC, Nikita KS. Comparison of machine learning approaches toward assessing the risk of developing cardiovascular disease as a long-term diabetes complication. IEEE J Biomed Health Inf 2018;22(5):1637–47. https://doi.org/10.1109/jbhi.2017.2765639.

[18] Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inf Decis Making 2020;20(1). https://doi.org/10.1186/s12911-020-01332-6.

[19] Nowakowska M, Zghebi SS, Ashcroft DM, Buchan I, Chew-Graham C, Holt T, Mallen C, Van Marwijk H, Peek N, Perera-Salazar R, Reeves D, Rutter MK, Weng SF, Qureshi N, Mamas MA, Kontopantelis E. The comorbidity burden of type 2 diabetes mellitus: patterns, clusters and predictions from a large English primary care cohort. BMC Med 2019;17(1). https://doi.org/10.1186/s12916-019-1373-y.

[20] Guerrero-Fernandez de Alba I, Orlando V, Monetti VM, Mucherino S, Gimeno-Miguel A, Vaccaro O, Forjaz MJ, Poblador Plou B, Prados-Torres A, Riccardi G, Menditto E. Comorbidity in an older population with type-2 diabetes mellitus: identification of the characteristics and healthcare utilization of high-cost patients. Front Pharmacol 2020;11:586187. https://doi.org/10.3389/fphar.2020.586187.

[21] Iglay K, Hannachi H, Joseph Howie P, Xu J, Li X, Engel SS, Moore LM, Rajpathak S. Prevalence and co-prevalence of comorbidities among patients with type 2 diabetes mellitus. Curr Med Res Opin 2016;32(7):1243–52. https://doi.org/10.1185/03007995.2016.1168291.

[22] Nissen SE. Effect of muraglitazar on death and major adverse cardiovascular events in patients with type 2 diabetes mellitus. JAMA 2005;294(20):2581. https://doi.org/10.1001/jama.294.20.joc50147.

[23] Lincoff AM, Tardif J-C, Schwartz GG, Nicholls SJ, Rydén L, Neal B, Malmberg K, Wedel H, Buse JB, Henry RR, Weichert A, Cannata R, Svensson A, Volz D, Grobbee DE. Effect of aleglitazar on cardiovascular outcomes after acute coronary syndrome in patients with type 2 diabetes mellitus. JAMA 2014;311(15):1515. https://doi.org/10.1001/jama.2014.3321.

[24] Scirica BM, Braunwald E, Raz I, Cavender MA, Morrow DA, Jarolim P, Udell JA, Mosenzon O, Im K, Umez-Eronini AA, Pollack PS, Hirshberg B, Frederich R, Lewis BS, Mcguire DK, Davidson J, Steg PG, Bhatt DL. Heart failure, saxagliptin, and diabetes mellitus: observations from the SAVOR-TIMI 53 randomized trial. Circulation 2014;130(18):1579–88. https://doi.org/10.1161/circulationaha.114.010389.

[25] Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, Lassale CM, Siontis GCM, Chiocchia V, Roberts C, Schlüssel MM, Gerry S, Black JA, Heus P, Van Der Schouw YT, Peelen LM, Moons KGM. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ 2016. https://doi.org/10.1136/bmj.i2416. i2416.