



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/238069/>

Version: Published Version

Article:

Haris, Mohammad, Romer, Elizabeth, Younsi, Tanina et al. (2026) Prediction models for incident stroke in the community:a systematic review and meta-analysis of predictive performance. European heart journal. Digital health. ztaf147. ISSN: 2634-3916

<https://doi.org/10.1093/ehjdh/ztaf147>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Prediction models for incident stroke in the community: a systematic review and meta-analysis of predictive performance

Mohammad Haris ^{1,2,3}, Elizabeth Romer⁴, Tanina Younsi ³, Jianhua Wu ⁵, Harriet Larvin⁵, Chris Wilkinson^{6,7}, Alan Cameron ⁸, Giulio F. Romiti^{9,10}, Gregory Y. H. Lip^{10,11,12}, Ramesh Nadarajah ^{1,2,3,*†}, and Chris P. Gale ^{1,2,3,†}

¹Leeds Institute for Cardiovascular and Metabolic Medicine, University of Leeds, 6 Clarendon Way, Leeds, LS2 9DA, UK; ²Leeds Institute of Data Analytics, University of Leeds, Clarendon Way, Leeds, LS2 9NL, UK; ³Department of Cardiology, Leeds Teaching Hospitals NHS Trust, Great George Street, Leeds, LS1 3EX, UK; ⁴Department of Cardiology, Airedale NHS Foundation Trust, Keighley, UK; ⁵Wolfson Institute of Population Health, Queen Mary University of London, London, UK; ⁶Hull York Medical School, University of York, York, UK; ⁷Academic Cardiovascular Unit, South Tees NHS Foundation Trust, James Cook University Hospital, Middlesbrough, UK; ⁸School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, UK; ⁹Department of Translational and Precision Medicine, Sapienza - University of Rome, Rome, Italy; ¹⁰Liverpool Centre for Cardiovascular Science, at University of Liverpool, Liverpool John Moores University and Liverpool Heart & Chest Hospital, Liverpool, United Kingdom; ¹¹Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; and ¹²Department of Cardiology, Lipidology and Internal Medicine with Intensive Coronary Care Unit, Medical University of Białystok, Białystok, Poland

Received 25 August 2025; revised 8 November 2025; accepted 3 December 2025; online publish-ahead-of-print 5 February 2026

Aims

Stroke is the second leading cause of death and the third leading cause of disability worldwide. We performed a systematic review and meta-analysis of multivariable models applicable to the prediction of incident stroke in community cohorts.

Methods and results

Ovid Medline and Embase were searched for studies related to stroke and prediction models from inception to 3 November 2025. Measures of discrimination were extracted and pooled by Bayesian meta-analysis, with heterogeneity assessed through a 95% prediction interval (PI). Risk of bias was assessed using the Prediction model Risk Of Bias Assessment Tool and certainty in effect estimates by Grading of Recommendations, Assessment, Development and Evaluation. Forty-one studies met the inclusion criteria, describing 80 prediction models, with two (R-FSRS and Basic IS) eligible for meta-analysis, including 969 514 participants. Both R-FSRS (summary *c*-statistic 0.714, 95% CI 0.681–0.747) and Basic IS (0.709, 95% CI 0.647–0.769) showed acceptable discrimination performance. Risk of bias was high in 66% of models, and both models showed reduced performance when excluding development cohorts and studies at high risk of bias (R-FSRS, 0.667, 95% CI 0.604–0.727; Basic IS 0.701; 95% CI 0.583–0.807). Only 43% of studies reported calibration, and no model underwent clinical utility analysis or a clinical impact study.

Conclusion

Many models have been derived for stroke prediction, however, they are rarely externally validated, and studies are limited by a high risk of bias, poor reporting of calibration and a lack of clinical utility analysis or prospective validation. Thus, the evidence base is insufficient to translate these models to clinical practice.

Lay Summary

This study reviewed and combined results from multiple research studies to understand how well different equations can predict the risk of having a stroke in the general population. While some equations can reasonably estimate stroke risk, issues with data quality and a lack of real-world testing mean it is unclear how useful they are in everyday healthcare.

- Only two equations (Revised Framingham Stroke Risk Score and Basic Ischaemic Stroke) had acceptable accuracy, but performed worse when studies with a higher risk of errors were excluded.
- None of the equations have been tested for practical benefits, so their effectiveness in real-life medical settings remains unknown.

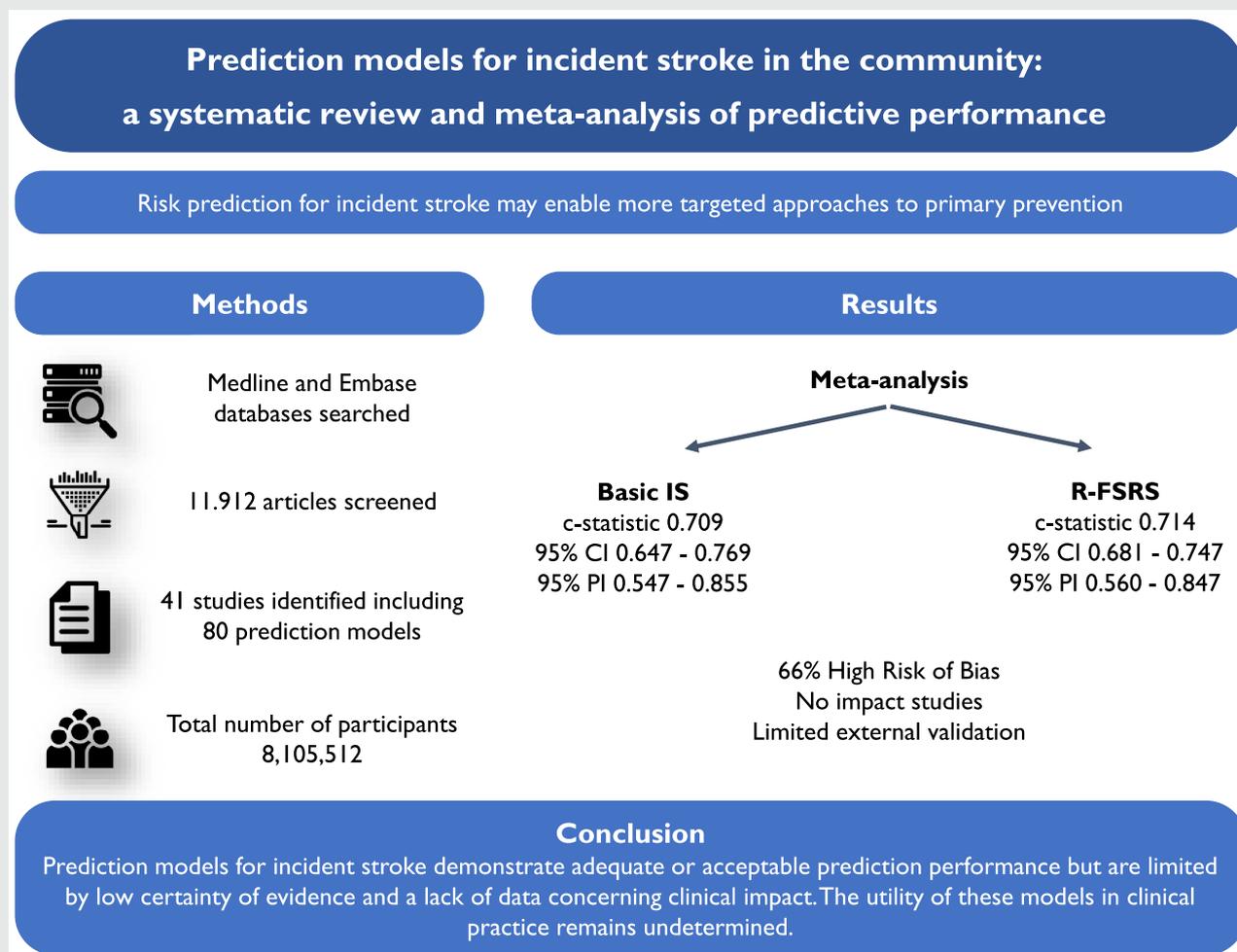
* Corresponding author. Tel: +44 (0) 113 343 3241, Email: r.nadarajah@leeds.ac.uk

†The last two authors are joint senior authors

© The Author(s) 2026. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical abstract



Keywords stroke • Cerebrovascular disease • Prediction • Prevention • Community

Key Learning Points

What is already known:

- Stroke is a leading global cause of death and disability, with prevention strategies emphasising risk prediction tools.
- Several multivariable models for stroke prediction exist, but many include variables not routinely collected or are developed in specific subpopulations.
- Previous systematic reviews have not clearly evaluated which models perform best in the general population.

What this study adds:

- This is a comprehensive review of stroke prediction models that use routinely available variables.
- Only two models (Revised Framingham Stroke Risk Score and Basic Ischaemic Stroke) were suitable for meta-analysis, showing acceptable performance, which reduced when studies at high risk of bias and derivation cohorts were excluded.
- No studies reported clinical utility or real-world impact, highlighting an urgent need for external validation and impact studies before clinical implementation.

Introduction

The absolute number of cases of stroke continues to rise due to an ageing population.¹ Stroke is already the second leading cause of death worldwide, responsible for almost seven million deaths annually,^{2,3} and the burden of stroke-related death is expected to increase by 50% in the coming decades.^{4,5} Moreover, stroke is the third leading cause of adult disability worldwide.^{2,3} Accordingly, it carries a significant healthcare and societal cost burden, accounting for approximately 34% of total global healthcare expenditure^{6,7} and costing in excess of US \$891 billion annually.⁵

While advancements in reperfusion therapies, dedicated stroke unit care, and early rehabilitation have conferred significant improvements in outcomes, prevention remains the optimal approach to reduce the burden of stroke.^{8,9} Risk prediction for incident stroke may enable more targeted approaches to primary prevention. Many risk prediction tools have been derived to predict the composite risk of cardiovascular disease overall,^{10,11} and stroke specifically.^{12–14} Guidelines for the primary prevention of stroke emphasize the use of these risk prediction tools to aid in employing effective strategies to prevent incident stroke.^{8,15–17}

However, the clinical utility of existing stroke prediction tools in the general population remains uncertain, as no prior evidence specifically addresses models applicable to this setting. Previous systematic reviews have included specific subpopulations not reflective of the general population, included composite scores which often do not give a specific risk of developing stroke, or included scores with variables not readily available in the community.^{18–24} To address this knowledge gap, we performed a systematic review of prediction models for incident stroke in the general population and conducted a quantitative synthesis of predictive performance.

Methods

This systematic review is registered on PROSPERO (CRD42024511343) and has been reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (see [supplementary material](#)).

Search strategy and inclusion criteria

The CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) was used to frame the research question (see [supplementary material](#)). The Ovid Medline and Ovid Embase databases were searched from inception to 3 November 2025 using a combination of keywords and subject headings related to stroke and prediction models (see [supplementary material](#)), restricted to human studies. The full search strategy can be found in the [supplementary material](#). We performed forward and backward citation searches and reviewed previous systematic reviews. We used Endnote's duplicate identification tool and then manually removed all remaining duplicates.

Articles were included if they were an original study in adults (≥ 18 years of age), developed and/or validated a prediction model for incident stroke based on multivariable analysis, provided a prediction performance metric for discrimination performance for a stroke outcome and were written in English. Articles were excluded if they included patients with stroke at baseline, only reported measures of association between risk factors and incident stroke rather than a full prediction model, studied only a subset of the general population (for example, individuals with a particular morbidity), or incorporated variables that would not be routinely available with high completeness in community health records (e.g. Townsend score, waist circumference, exercise, diet, uric acid, left ventricular hypertrophy; [supplementary material](#)). Where studies reported multiple models, only those utilising variables readily obtainable in community settings were included.

We uploaded records to a systematic review web application (Rayyan, Qatar Computing Research Institute).²⁵ Three investigators (M.H., E.R. and T.Y.) independently screened them for inclusion by title, abstract, full text and [supplementary materials](#). Disagreements were resolved by consultation with a fourth investigator (R.N.).

Data extraction and quality assessment

Two investigators (M.H. and E.R.) independently extracted the data from the included studies based on the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS). All data came from the primary reference, unless otherwise stated.

To allow quantitative synthesis and assessment of the predictive performance of the models, we extracted measures of discrimination and calibration.²⁶ Discrimination assesses the model's ability to differentiate between individuals who will experience the outcome and those who will not. To assess discrimination, we extracted data on the c-statistic or the area under the receiver operating characteristic (AUROC), along with their corresponding 95% confidence intervals (CI). If the reported CI was missing, we computed it using the methods outlined by Debray *et al.*²⁶ Calibration evaluates the accuracy of the model's predicted probabilities, and we extracted all performance measures reported. Three investigators (M.H., E.R. and T.Y.) assessed the models for risk of bias and applicability to our review question using the Prediction model Risk Of Bias ASsessment Tool (PROBAST).²⁷ We also checked for reporting of the clinical utility of a model (net benefit in the form of decision curve analysis or decision analytical modelling, which can be used to integrate the benefits and harms of using a model for clinical decision support) and conducted forward citation searching for studies determining the impact (clinical and cost-effectiveness) of using models in real world clinical practice.

Data synthesis and statistical analysis

We reported continuous variables as means \pm standard deviation (SD) and categorical variables as percentages. We evaluated statistical significance in all analyses at the 0.05 level. When a study reported on multiple cohorts and presented separate data for each cohort, we assessed model performance separately for each cohort within that study. A funnel plot was created to check for publication bias.²⁸

We conducted a Bayesian meta-analysis of discrimination through a summary measure of c-statistic and corresponding 95% CI. We calculated the 95% prediction interval (PI) to depict the extent of between-study heterogeneity and to indicate a possible range for prediction model performance in a new validation.²⁹ A PI is a statistical measure to estimate a range for the predicted model performance in a new validation of the model with a certain level of confidence. Summary c-statistics of <0.60 , $0.60–0.69$, $0.70–0.80$, and >0.80 were defined *a priori* as inadequate, adequate, acceptable and good based on prior publications.^{30,31}

All Bayesian meta-analysis models assume random effects by default. Results are based on the posterior median. PIs are directly obtained from the corresponding posterior quartiles. The standard model for random effects meta-analysis assumes that the 'true' performance is normally distributed within and across studies.³² Within-study normality of performance estimates can be justified with this selection of included studies because they are all large. Snell *et al.* showed that the between-study distribution of the c-statistic on the original scale is not normally distributed when there is variability in the predictor effect across studies (which is likely in this selection of studies, as they include different populations and adopt slightly different definitions for predictors).³² They found that the logit scale is more appropriate for the estimation of the PI. Consequently, we used the 'valmeta' function of the 'metamisc' package in R software, which applies a logit transformation to the c-statistic prior to calculation of the summary c-statistic and PI.^{33–35}

For appropriate prior distributions, we borrowed from earlier work by Debray *et al.*, which recommended a half Student-t distribution with location m , scale σ , and ν degrees of freedom, where we set $m = 0$ and define σ equal to the largest empirical value of $\hat{\tau}$ (to allow for more extreme values of heterogeneity).²⁹ These hyperparameter values allow to penalize the extent of between-study heterogeneity when the number of included validation studies is low.²⁹ Further, we also used $\nu = 3$ to ensure that the variance $\sigma^2 - \nu/(\nu - 2)$ exists and samples of τ were truncated above 10 to rule out unreasonable values. Thus, the resulting priors are given as $\tau_{disc} \sim \text{Student-}t(0, 0.5^2, 3)T[0.10]$, which has been shown to allow for large but realistic values for between-study heterogeneity.²⁹

Our primary analysis assessed overall discrimination for models that had three or more cohorts with c-statistic data, excluding studies with a high risk of bias. We performed sensitivity analysis in which we assessed discrimination for each model, excluding results from derivation cohorts. We

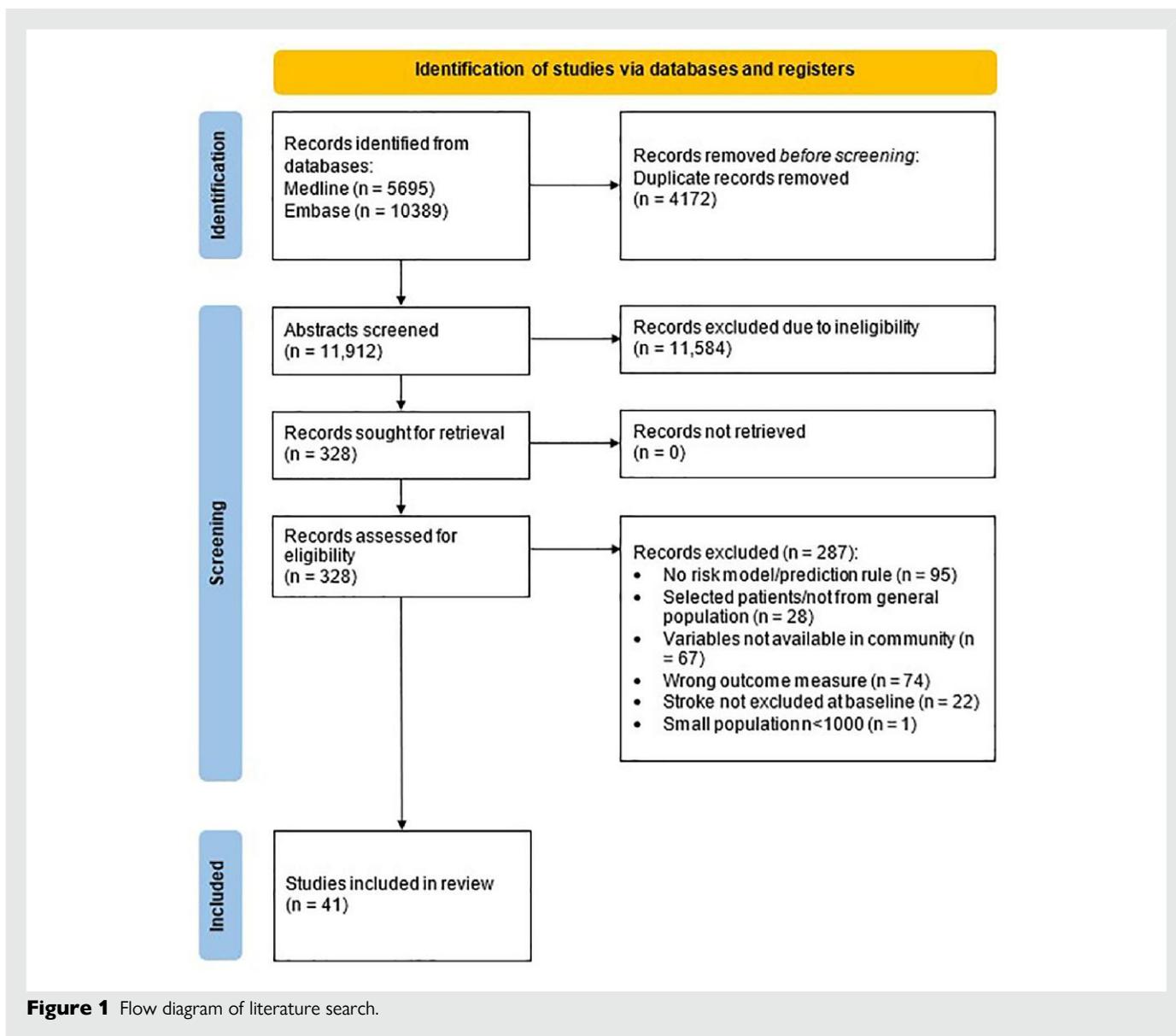


Figure 1 Flow diagram of literature search.

performed a further analysis comparing models employing regression techniques vs. machine learning techniques.

Results

Study selection

We identified 11 912 unique records, reviewed 328 full-text reports and included 41 studies (Figure 1). A list of excluded studies that met a number of the inclusion criteria is available in the [supplementary material](#).

Characteristics of included studies

In the 41 studies, 80 models were derived and/or validated across 48 cohorts from a range of countries located in Asia Pacific (n = 21), the United States (n = 14) and Europe (n = 13) (Table 1). The total number of participants included in the studies was 8 105 512, with cohort sizes ranging from 1131 to 1 730 828. The mean age ranged from 35.5 to

68.6 years, and the proportion of women ranged from 0.0% to 66.1% (see [Supplementary material online, Table S1](#)).

Characteristics of included prediction models

The 80 included models consisted of 51 multivariable regression models and 29 machine learning models. All studies reported the predictors in the models and prediction horizons ranged from one to 20 years. The range of reported AUROC was between 0.54 and 0.99, and 15 of the models were externally validated (see [Supplementary material online, Table S2](#)). Twenty-one models only predicted ischaemic strokes, and 59 models predicted both ischaemic and haemorrhagic strokes (see [Supplementary material online, Table S3](#)). The machine learning models were all supervised machine learning and included gradient boosting, logistic regression, random forest, support vector machine, neural networks and XGBoost (see [Supplementary material online, Table S4](#)).

Table 1 Overview of included studies

Study	Cohort (country)	Study aim	Stroke cases (n)/total patients (n) (%)	Outcome definition	Enrolment period (mean F/U in years)	Exclusion criteria	Overall ROB
Arafa <i>et al.</i> 2021 ³⁶	Suita (Japan)	D, IV	372/6641 (5.6)	Cerebral infarction, ICH, SAH	1989–1999 (17.1)	History of cardiovascular diseases Age >79 or <30 y, Missing data	High
Assman <i>et al.</i> 2007 ³⁷	PROCAM (Germany)	D	85/8130 (1)	Focal neurological deficit >24 h due to vascular event	1978–1995 (12)	Haemorrhagic or undetermined stroke	High
Borglykke <i>et al.</i> 2010 ³⁸	MORGAM (Denmark)	D, IV	2928/88290 (3.3)	Fatal and non-fatal stroke	1982–1997	History of MI/stroke Missing data	High
Camen <i>et al.</i> 2020 ³⁹	MORGAM (Denmark)	EV	3033/82881 (3.7)	Ischaemic/haemorrhagic/Indeterminate	1982–2010 (12.7)	History of MI/stroke/CAD	High
Chambless <i>et al.</i> 2004 ⁴⁰	ARIC (US)	D	434/14685 (3)	Ischaemic	1987–1989 (12.3)	History of MI/stroke/CAD Race other than black or white Missing data	High
Chien <i>et al.</i> 2010 ⁴¹	Chin-Shan (Taiwan)	D, IV, EV	240/3512 (6.8)	Ischaemic/haemorrhagic	1990–2007 (15.9 median)	History of stroke/TIA	High
Chun <i>et al.</i> 2022 ⁴²	CKB (China)	D, EV	43234/503842 (8.6)	Ischaemic/haemorrhagic	2004–2008 (9)	—	High
Di Castelnuovo <i>et al.</i> 2019 ⁴³	BiomarCaRE (Finland, Sweden, Ireland, France, Italy)	D	1550/58173 (2.7)	Ischaemic/haemorrhagic	—	History of stroke	High
Ferret <i>et al.</i> 2014 ⁴⁴	ARIC (US) Rotterdam (Netherlands) CHS (US)	D	2884/27493 (10.5)	Ischaemic/haemorrhagic	—	History of stroke/AF Current anticoagulation Not Caucasian/African American	Low
Flueckiger <i>et al.</i> 2018 ⁴⁵	MESA (US)	EV	231/6712 (3.4)	Ischaemic or haemorrhagic	—	ABI >1.4	High
Foraker <i>et al.</i> 2016 ⁴⁶	Jackson Heart Study (US)	EV	112/4140 (2.7)	Ischaemic or haemorrhagic	—	History of stroke Missing data	Low
Harada <i>et al.</i> 2018 ⁴⁷	JALS (Japan)	D	1351/67969 (2)	Ischaemic or haemorrhagic	2002–2004 (6.9)	History of stroke or heart disease Age >90 or <40 y, Missing data	High
Hilvo <i>et al.</i> 2022 ⁴⁸	FINRISK 2002 (Finland)	D	249/7810 (3.2)	Ischaemic or haemorrhagic	—	SAH	High
Hong <i>et al.</i> 2023 ⁴⁹	FOS, ARIC, MESA, REGARDS (US)	EV	2199/62482 (3.5)	Ischaemic, haemorrhagic or other stroke	1971–2008 (10)	—	Low
Howard <i>et al.</i> 2017 ⁵⁰	REGARDS (US)	D, EV	939/23983 (3.9)	Stroke	2003–2007 (8.2)	—	High
Hung <i>et al.</i> 2019 ⁵¹	NHIRD (Taiwan)	D, IV	2544/840487 (0.3)	Ischaemic stroke	2003–2003 (8)	—	Low
Hung <i>et al.</i> 2018 ⁵²	NHIRD (Taiwan)	D, IV	4795/552898 (0.9)	Ischaemic stroke	2003–2003 (5)	—	High
Hunter <i>et al.</i> 2022 ⁵³	FHS, FOS, FHS-3rd gen, FHS-OMNI2, FHS new offspring (US)	D, IV	14983/113714 (13.2)	Ischaemic stroke/TIA	1948–2018 (5)	Haemorrhagic stroke	Low
Jung <i>et al.</i> 2018 ⁵⁴	KCPS-II (South Korea)	D, IV	823/144594 (5.7)	Ischaemic or haemorrhagic	2004–2005	—	High

Continued

Table 1 Continued

Study	Cohort (country)	Study aim	Stroke cases (n)/total patients (n) (%)	Outcome definition	Enrolment period (mean F/U in years)	Exclusion criteria	Overall ROB
Lee et al. 2020 ⁵⁵	NHIS, NHIS-NSC (South Korea)	D, IV	3484/973055 (0.4)	Ischaemic, haemorrhagic or other stroke	2010–2011	—	High
Li X et al. 2022 ⁵⁶	Jilin University Cohorts 1–3 (China)	D, IV	15833/32366 (49)	Ischaemic stroke	2018–2021 (5)	—	Low
Li Y et al. 2022 (1) ⁵⁷	CPRD (UK)	D, IV	224442/1730828 (12.9)	Ischaemic, haemorrhagic or any stroke	—	—	High
Lolak et al. 2023 ⁵⁸	Ramathibodi Hospital (Thailand)	D, IV	9659/275247 (3.5)	Ischaemic or haemorrhagic	2010–2020	History of stroke <2 visits during period	High
Majed et al. 2013 ⁵⁹	PRIME (combined, France, Ireland)	D, IV	138/9638 (1.4)	First stroke event	1991–1993 (10)	History of stroke or CAD Missing data	Low
Marrugat et al. 2014 ⁶⁰	FRESCO Study (Spain)	D, EV	786/50408 (1.6)	Ischaemic or haemorrhagic	1992–2005 (9.3)	History of cardiovascular disease	Low
Teoh et al. 2018 ⁶¹	TJTC Hospital (Japan)	D, IV	2725/8175 (33.3)	First stroke	2001–2015	History of stroke/MI/HF Age <45 or >94 years	High
Wannamethee et al. 2005 ⁶²	BRHS (UK)	D, IV	291/5128 (5.7)	Stroke	1978–2000 (21.3)	History of stroke/coronary heart disease/diabetes	High
Wu et al. 2020 ⁶³	CLHLS (China)	D, IV	56/1131 (5)	Stroke	2012–2012 (2)	History of stroke Age <60 years Missing data	Low
Yang et al. 2023 ⁶⁴	CHARLS (China)	D, IV	312/5844 (5.3)	Stroke	2011–2011 (7)	Age <45 years	Low
Yatsuya et al. 2016 ⁶⁵	JPHC I, JPHC II (Japan)	D, IV, EV	863/27270 (3.2)	Ischaemic or haemorrhagic	1990–1994 (16.4)	History of stroke/MI Missing data	Low
Yatsuya et al. 2013 ⁶⁶	JPHC I, JPHC II (Japan)	D, IV, EV	790/15672 (5)	Ischaemic or haemorrhagic	1990–1994 (14)	History of cardiovascular disease Missing data	Low
Zhang X et al. 2005 ⁶⁷	Beijing steelworker cohort (China)	D, IV	118/4400 (2.7)	Ischaemic or haemorrhagic	1974–1980 (13.5)	Age <35 years	High
Zhang Y et al. 2020 ⁶⁸	BLSA (China)	D, EV	106/1203 (8.8)	Stroke	2009–2009 (4.8)	History of stroke Age >84 years Missing data	Low
Dufouil et al. 2017 ⁶⁹	FHS, REGARDS, 3C (US)	D, EV	974/36364 (2.7)	Ischaemic or haemorrhagic	1948–2007	History of stroke Age <55 or >84 years	Low
Hung et al. 2017 ⁷⁰	NHIRD (Taiwan)	D, IV	4944/798611 (0.6)	Ischaemic stroke	2003–2003 (5)	History of stroke Missing data	High
Sun et al. 2021 ⁷¹	UKB, CPRD (UK)	D, IV	2347/306654 (0.8)	Ischaemic or haemorrhagic	2006–2010 (8.1)	History of stroke/TIA/heart disease/ peripheral vascular disease, taking lipid-lowering drugs, and Missing data	High

Continued

Table 1 Continued

Study	Cohort (country)	Study aim	Stroke cases (n)/total patients (n) (%)	Outcome definition	Enrolment period (mean F/U in years)	Exclusion criteria	Overall ROB
Xing <i>et al.</i> 2019 ⁷²	China MUCA, CIMIC (China)	EV	2690/84961 (3.2)	Ischaemic or haemorrhagic	MUCA 1992–1994 (17.1) CIMIC 2007–2008 (5.9)	History of cardiovascular disease Age <35 or >74 years	High
Bos <i>et al.</i> 2017 ⁷³	Rotterdam (Netherlands)	EV	502/7966 (6.3)	Ischaemic or haemorrhagic	2000–2006 (7)	History of cardiovascular disease/AF	High
D'Agostino <i>et al.</i> 2008 ⁷⁴	FHS, FOS (US)	D, IV, EV	177/8491 (2)	Ischaemic/haemorrhagic/ TIA	1968–1987 (12)	History of cardiovascular disease Age <30 or >74 y, Missing data	Low
Li Y <i>et al.</i> 2022 (2) ⁷⁵	CPRD (UK)	D, IV, EV	74547/1096275 (6.8)	Ischaemic/haemorrhagic/ TIA	1985–2015 (5)	Missing data	Low
Vu <i>et al.</i> 2024 ⁷⁶	Suita (Japan)	D	438/7389 (5.9)	Ischaemic or haemorrhagic	1989–1999 (15)	History of cardiovascular diseases Age >79 or <30 y, Missing data	High

ACVD, atherosclerotic cardiovascular disease; ADA, Active Data Augmenter; AF, atrial fibrillation; ARIC, Atherosclerosis Risk In Communities; BLSA, Beijing Longitudinal Study of Aging; BN, Bayesian Network; BRHS, British Regional Heart Study; CAD, coronary artery disease; CHARLS, China Health and Retirement Longitudinal Study; CHS, Cardiovascular Health Study; CIMIC, Community Intervention of Metabolic Syndrome in China; CKB, China Kadoorie Biobank; CLHLS, Chinese Longitudinal Healthy Longevity Survey; CNN, Convolutional Neural Network; CORSAIB, COR Sa Illes Balears; CPRD, Clinical Practice Research Datalink; CVD, cardiovascular disease; D, derivation; DNN, Deep Neural Network; EV, external validation; FHS, Framingham Heart Study; FOS, Framingham Offspring Study; FRESCO, (Función de Riesgo Española de acontecimientos Coronarios y Otros; FSRS, Framingham Stroke Risk Score; GBDT, Gradient-Boosted Decision Tree; GRU, Gated Recurrent Unit; HF, heart failure; ICH, intracerebral haemorrhage; IHD, ischaemic heart disease; IS, ischaemic stroke; IV, internal validation; JALS, Japan Arteriosclerosis Longitudinal Study; JPHC, Japan Public Health Center; LR, Logistic Regression; LVH, left ventricular hypertrophy; MESA, Multi-Ethnic Study of Atherosclerosis; MGL, Multiple Group Learning; MI, myocardial infarction; MORGAM, MONica Risk, Genetics, Archiving and Monograph; NHIRD, National Health Insurance Research Database; NHIS, National Health Insurance Service; PCE, Pooled Cohort Equation; PRIME, Prospective Epidemiological Study of Myocardial Infarction; PROCAM, Prospective Cardiovascular Münster; R-FSRS, Revised Framingham Stroke Risk Score; REGARDS, Reasons for Geographic and Racial Differences in Stroke; RF, Random Forest; RLR, Relational Logistic Regression; RNN, Recurrent Neural Network; SAH, subarachnoid haemorrhage; SDA, Straightforward Data Augmentation; SRSRF, Self-Reported Stroke Risk Stratification; SVM, Support Vector Machine; TAN, Tree-Augmented Naïve Bayes; TIA, transient ischaemic attack; TJTC, Tsuyama Jifukai Tsuyama Chuo; TRS, Traditional Risk Score; XG Boost, extreme gradient boosting

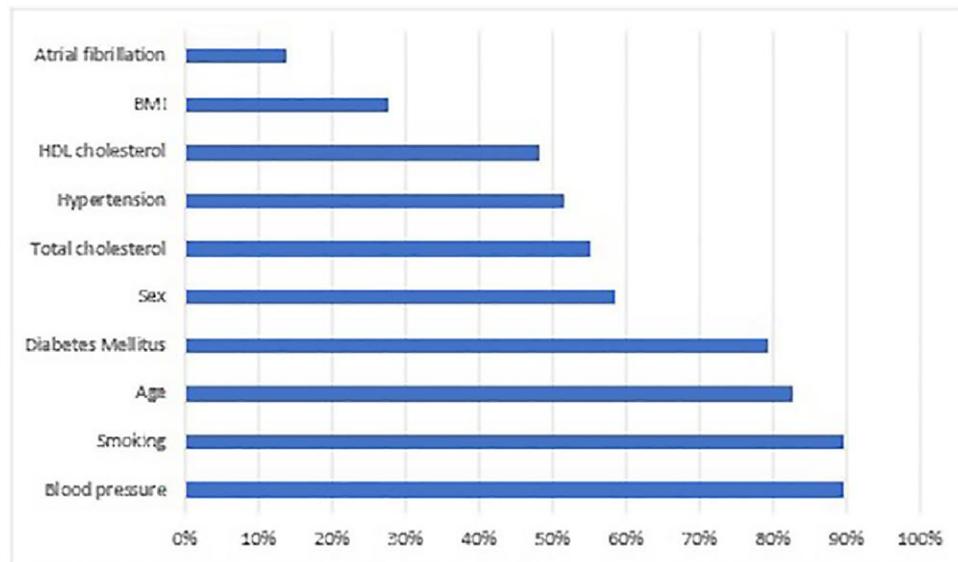


Figure 2 An overview of the 10 predictors most frequently incorporated in the prediction models in this study. HDL, high-density lipoprotein; BMI, body mass index.

Only 18 studies reported calibration outcomes. Of these, 10 reported Hosmer-Lemeshow statistic, 4 reported O:E ratio and 11 reported calibration plots.

The most common predictors in the regression models were blood pressure (BP) (89.7%), smoking (89.7%), age (82.8%) and diabetes mellitus (79.3%), as shown in [Figure 2](#). Other variables included physical measurements such as BP and body mass index (BMI), medical conditions such as hypertension and atrial fibrillation, investigations such as blood glucose and cholesterol, and medications such as antihypertensives. Machine learning and regression-based models used broadly similar sets of predictors, and their discrimination performance was comparable (see [Supplementary material online, Tables S2, S5 and S6](#)).

Clinical utility and clinical impact of included models

No studies conducted a clinical utility analysis, and forward citation searching did not identify studies of clinical impact for the included risk prediction models.

Risk of bias assessment

Overall, 66% of model results were at high risk of bias solely driven by high risk of bias in the analysis domain, mainly due to the handling of missing data ([Figure 3, Supplementary material online, Table S7](#)).

Meta-analysis

The Revised Framingham Stroke Risk Score (R-FRS) and Basic Ischaemic Stroke (Basic IS) models were eligible for the primary meta-analysis, which included the derivation and validation cohorts and incorporated a total of 969 514 participants. Both models had acceptable discrimination performance with a summary c-statistic of 0.714; 95% CI 0.681–0.747 for R-FRS and 0.709; 95% CI 0.647–0.769 for Basic IS ([Figure 4](#)). After excluding results from the development cohorts, both models showed reduced discriminatory performance with a summary c-statistic of 0.667; 95% CI 0.604–0.727 for R-FRS and 0.701; 95% CI 0.583–0.807 for Basic IS (see [Supplementary material online, Figure S1](#)). Most studies

evaluated the R-FRS model separately in men and women. Analysis according to sex revealed similar discriminatory performance in men (summary c-statistic 0.706, 95% CI 0.657–0.755) and women (0.716; 95% CI 0.667–0.761) (see [Supplementary material online, Figure S2](#)).

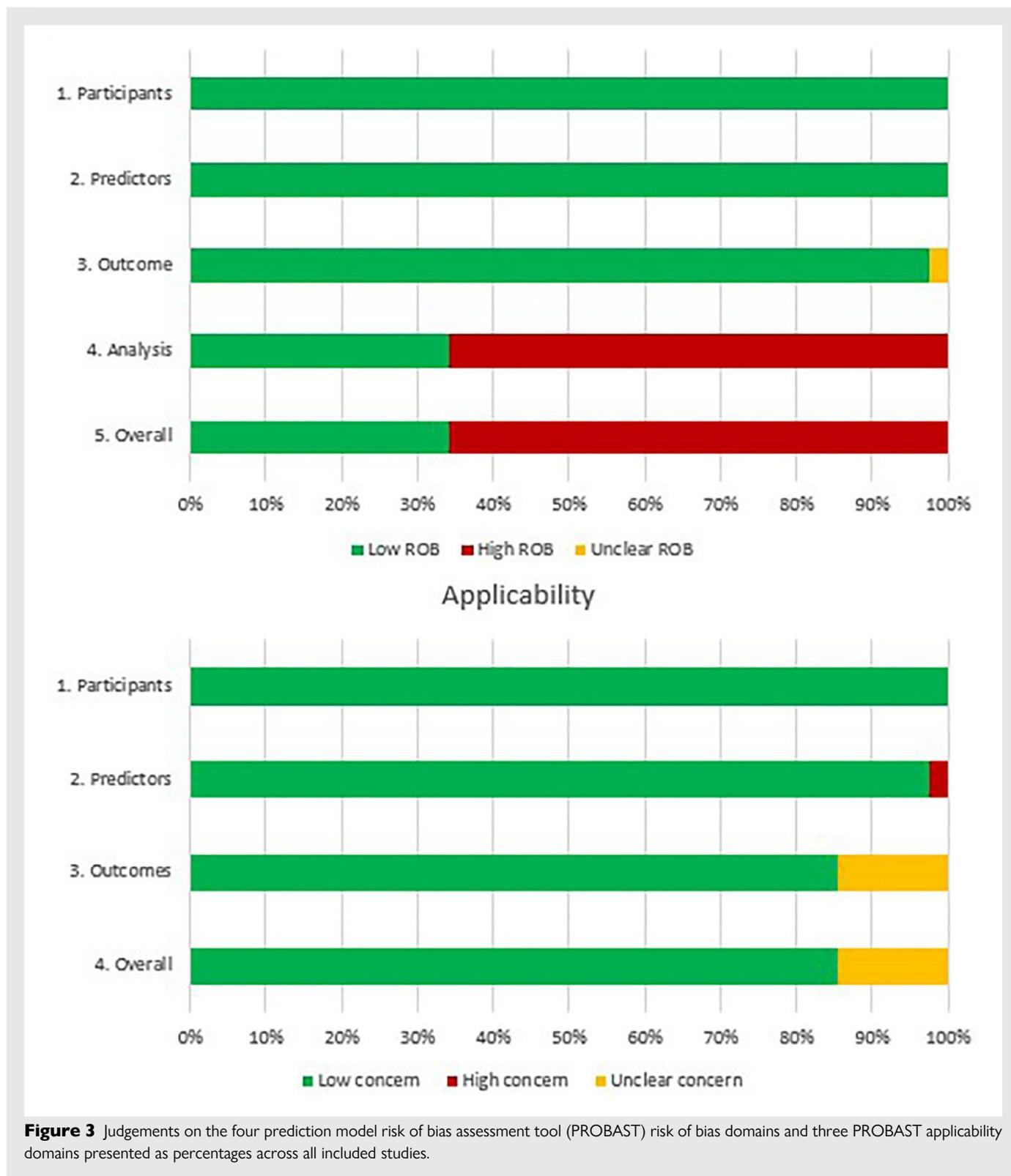
Analysis comparing regression models with machine learning models showed better performance in machine learning models, with a summary c-statistic of 0.853; 95% CI 0.823–0.882 for machine learning models and a summary c-statistic of 0.732; 95% CI 0.718–0.745 for regression models (see [Supplementary material online, Figure S3](#)). Funnel plots were symmetrical but with additional horizontal scatter, consistent with the presence of between-study heterogeneity (see [Supplementary material online, Figure S4](#)).

Discussion

This systematic review and meta-analysis included 80 models developed and/or validated in 48 cohorts representative of, and including variables readily available in, the general population for estimating an individual's average risk of incident stroke. The majority of models showed good discrimination performance, but only two models (R-FRS and Basic IS) were eligible for primary meta-analysis and demonstrated acceptable summary discrimination performance measures. However, after excluding the derivation cohorts and studies at high risk of bias, the performance of both models reduced to the point that they may not offer good discrimination of stroke. Additionally, clinical utility remains uncertain as neither of the models underwent a prospective investigation of clinical or cost-effectiveness.

Previous work

A recent review of prediction models for incident stroke included 17 studies until February 2022.¹⁸ However, it included models that have been derived and/or validated in specific populations and models with variables not readily available in the general population, limiting the clinical translatability of the findings to the primary prevention setting in the community. Furthermore, the authors performed a meta-analysis of all



the included models together, rather than delineating whether a particular model performed particularly well in practice.¹⁸

Other reviews have focused on the prediction of stroke in specific groups, not reflective of the general population, for example, in patients

on dialysis.²⁴ Some reviews explored composite scores for vascular disease, but these did not specifically address stroke,^{21,23} and others included scores with variables that are not usually readily available in the general population.^{20,22}

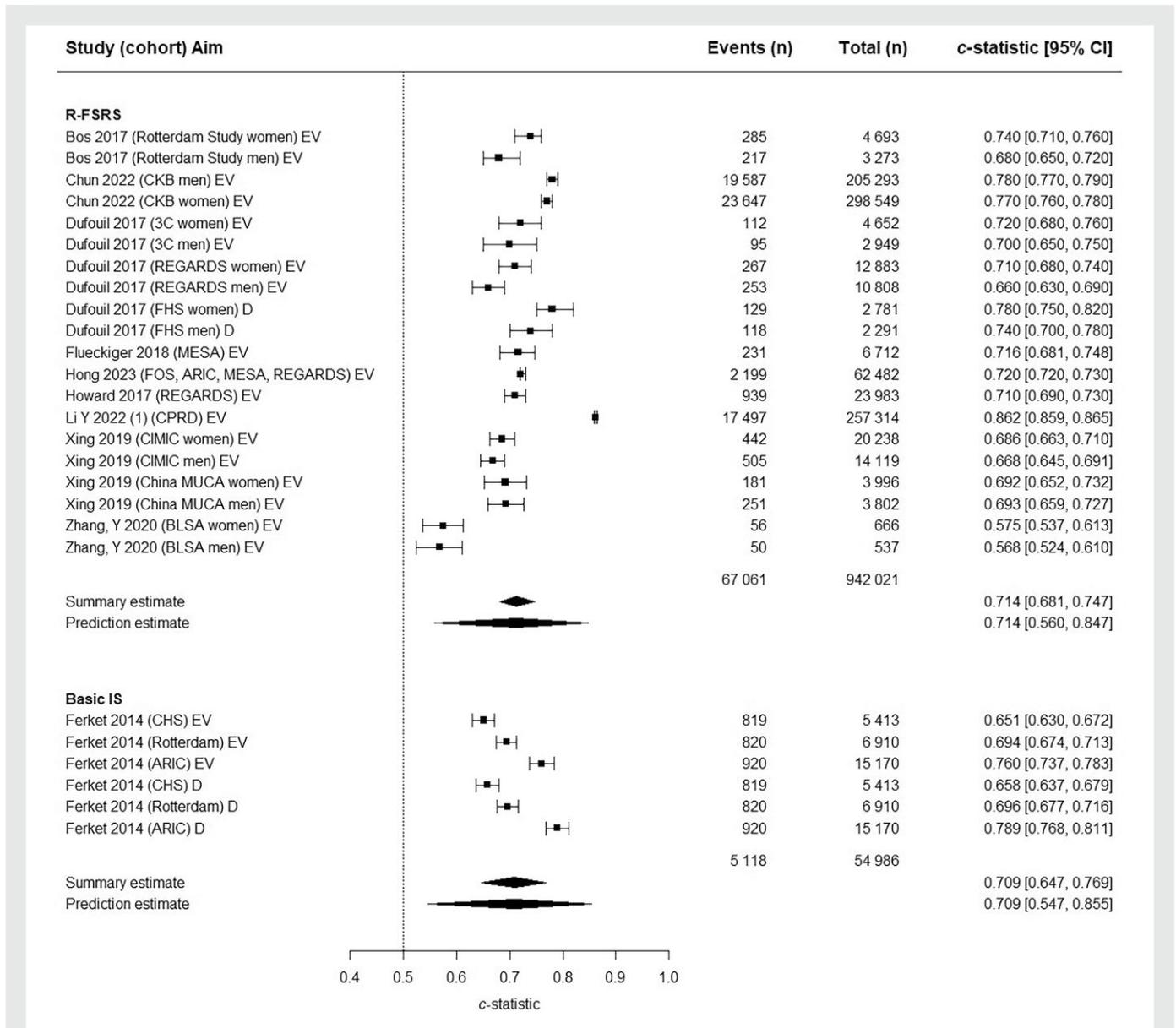


Figure 4 Forest plot of primary analysis of c-statistics. ARIC, Atherosclerosis Risk in Communities; BLSA, Beijing Longitudinal Study of Ageing; CHS, Cardiovascular Health Study; CIMIC, Community Intervention of Metabolic Syndrome in China; CKB, China Kadoorie Biobank; CPRD, Clinical Practice Research Datalink; D, Derivation; EV, External Validation; FHS, Framingham Heart Study; ICH, intracerebral haemorrhage; IS, ischaemic stroke; MESA, Multi-Ethnic Study of Atherosclerosis; REGARDS, Reasons for Geographic and Racial Differences in Stroke; R-FSRS, Revised Framingham Stroke Risk Score.

Clinical relevance

Given that the burden of stroke and related complications is rising,²⁻⁵ primary prevention is an increasingly important priority. Although there have been advances in reperfusion therapies in some geographies,⁷⁷ primary prevention remains the most effective method to reduce the overall burden of stroke.⁸ At present, stroke prevention is often encompassed within general cardiovascular disease prevention strategies.^{78,79} Yet the increasing incidence of stroke contrasts with the reductions seen in myocardial infarction (MI) and heart failure (HF),^{80,81} and whilst there are common risk factors, there are some notable differences between stroke and other conditions. For instance, evidence seems to suggest atrial fibrillation is a stronger risk factor for

stroke than MI,⁸² and there may be a more pronounced association between hypertension and stroke than other cardiovascular conditions.⁸³ Therefore, primary prevention of stroke may require a more tailored approach, and stroke-specific risk prediction models may prove useful to delineate those at particularly high risk in general medicine and primary care settings.

Guidelines for the primary prevention of stroke emphasize the use of risk prediction tools to aid in employing effective strategies to prevent incident stroke.¹⁵⁻¹⁷ They highlight the importance of simple and widely applicable tools and the need for further validation of current tools in different groups of the population.⁸ The models included in this review utilize variables that are readily available and easily obtainable in the community. Though many models were identified within this review,

there were significant shortcomings in the available evidence. First, the majority were only evaluated in the derivation cohort, and only a small number were externally validated. Second, there was a high level of potential bias and poor reporting of calibration metrics, in line with general trends of prediction model studies.⁸⁴ Third, we did not find any evidence of clinical utility analyses or clinical impact studies, limiting any conclusions regarding the effects of prediction models when used in a clinical pathway.

This review underscores that no current model is ready for clinical use in primary prevention pathways for stroke. Whilst there is no lack of interest in developing models, we would recommend a greater emphasis on validating, updating and prospectively testing existing models across diverse populations. Incorporating routinely collected data and assessing calibration early in model design and embedding evaluation of clinical utility in studies are essential to determine which models should be considered for prospective testing. Furthermore, the effect of the implementation of models within pathways on at least the control of risk factors for stroke is essential before one can consider translating stroke prediction models to clinical practice.

Strengths and limitations

We employed a comprehensive search strategy to identify relevant articles and models. We aimed to maximize applicability by only including models from the general population that incorporated variables readily available in such settings.

We also acknowledge the limitations of our study. Meta-analysis of calibration performance was not possible due to a lack of calibration reporting. We did not present meta-regression or subgroup meta-analysis to investigate heterogeneity between studies based on study-level characteristics or subgroups in the absence of available individual patient data, given that such analyses would be prone to ecological bias,⁸⁵ and are inferior to subgroup results derived with patient-level data.²⁶ The funnel plot demonstrated between-study heterogeneity, which may reflect differences in the characteristics, quality or population of studies, as populations varied in age, sex, comorbidities and stroke incidence. To address between-study heterogeneity, including variation in baseline stroke incidence, we used a Bayesian random-effects meta-analysis for all pooled estimates. Nonetheless, heterogeneity was considerable, and the 95% PI was large, which limits the robustness of the assessment for how the included prediction models may perform in a new dataset. Poor reporting of race and ethnicity across studies limited our ability to assess the performance of prediction models by these important sociodemographic factors, especially given the reported racial differences in stroke and other cardiovascular risk factors.⁸⁶ Missing data was common and is a frequently observed shortfall in prediction modelling research,⁸⁷ even in models recommended for use in healthcare.⁸⁸ In the absence of patient-level data and model coefficients across all cohorts to quantify the relative importance of predictor variables, we were unable to assess which predictors contributed the most to prediction across all models.

Conclusion

Many models have been derived for stroke prediction, however, they are rarely externally validated, and studies are limited by a high risk of bias, poor reporting of calibration and a lack of clinical utility analysis or prospective validation. Thus, the evidence base is insufficient to translate these models to clinical practice.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Authorship

M.H. organized and contributed to all aspects of the work. E.R. and T.Y. assisted with searches and data extraction. J.W. and H.R. assisted with the statistical analyses. C.W., A.C., G.F.R. and G.Y.H.L. reviewed and revised the manuscript. R.N. and C.P.G. reviewed and revised the manuscript and oversaw all aspects of the work. All authors gave final approval and agreed to be accountable for all aspects of work, ensuring integrity and accuracy.

Author contributions

Mohammad Haris (Conceptualization, Methodology, Data collection, Data analysis, Manuscript writing and editing), Elizabeth Romer (Data collection, Manuscript review), Tanina Younsi (Data collection, Manuscript review), Jianhua Wu (Data analysis), Harriet Larvin (Data analysis), Chris Wilkinson (Manuscript review and editing), Alan Cameron (Manuscript review and editing), Giulio F. Romiti (Manuscript review and editing), Gregory Y. H. Lip (Manuscript review and editing), Ramesh Nadarajah (Conceptualization, Methodology, Manuscript review and editing), and Chris P. Gale (Conceptualization, Methodology, Manuscript review and editing)

Funding

None

Conflict of interest: C.P.G. received honoraria for consulting and/or speeches from AstraZeneca, Amgen, Bayer, Boehringer Ingelheim, Daiichi Sankyo, Vifor, Pharma, Menarini, Wondr Medical, Raisio Group and Oxford University Press. A.C. has received research grants from Pfizer and honoraria from BMS, Pfizer, AstraZeneca and Boehringer Ingelheim. G.F.R. served as a consultant for Boehringer Ingelheim and received an educational grant from Anthos, although no fees were received personally. G.Y.H.L. has served as a consultant for BMS/Pfizer, Medtronic, Boehringer Ingelheim, Anthos and Daiichi-Sankyo, although no fees were received personally. All the other authors declared no competing interests.

The authors had full responsibility for data collection, data interpretation, and writing of the report. All data can be found in the manuscript or [supplementary material](#).

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

- Pu L, Wang L, Zhang R, Zhao T, Jiang Y, Han L. Projected global trends in ischemic stroke incidence, deaths and disability-adjusted life years from 2020 to 2030. *Stroke* 2023;**54**:1330–1339.
- Hilkens NA, Casolla B, Leung TW, de Leeuw FE. Stroke. *The Lancet* 2024;**403**:2820–2836.
- Feigin VL, Stark BA, Johnson CO, Roth GA, Bisignano C, Abady GG, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Neurol* 2021;**20**:795–820.
- Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. *Circulation* 2019;**139**:e56–e528.
- Feigin VL, Owolabi MO, Feigin VL, Abd-Allah F, Akinyemi RO, Bhattacharjee NV, et al. Pragmatic solutions to reduce the global burden of stroke: a world stroke organization–lancet neurology commission. *Lancet Neurol* 2023;**22**:1160–1206.
- Katan M, Luft A. Global burden of stroke. *Semin Neurol* 2018;**38**:208–211.
- Rochmah TN, Rahmawati IT, Dahlui M, Budiarto W, Bilqis N. Economic burden of stroke disease: a systematic review. *Int J Environ Res Public Health* 2021;**18**:7552.
- Meschia JF, Bushnell C, Boden-Albala B, Braun LT, Bravata DM, Chaturvedi S, et al. Guidelines for the primary prevention of stroke: a statement for healthcare

- professionals from the American Heart Association/American stroke association. *Stroke* 2014;**45**:3754–3832.
9. Rodgers H, Price C. Stroke unit care, inpatient rehabilitation and early supported discharge. *Clin Med (Lond)* 2017;**17**:173–177.
 10. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *Circulation* 2014;**129**:S49–S73.
 11. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;**357**:j2099.
 12. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham study. *Stroke* 1991;**22**:312–318.
 13. D'Agostino RB, Wolf PA, Belanger AJ, Kannel WB. Stroke risk profile: adjustment for antihypertensive medication. The Framingham study. *Stroke* 1994;**25**:40–43.
 14. Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. *BMJ* 2013;**346**:f2573.
 15. Owolabi MO, Thrift AG, Mahal A, Ishida M, Martins S, Johnson WD, et al. Primary stroke prevention worldwide: translating evidence into action. *Lancet Public Health* 2022;**7**:e74–e85.
 16. Strokes: progressive stroke pathway [Internet]. [cited 2024 Jul 16]. Available from: <https://www.gov.scot/publications/progressive-stroke-pathway/>
 17. Steiger N, Cifu AS. Primary prevention of stroke. *JAMA* 2016;**316**:658–659.
 18. Asowata OJ, Okeunle AP, Olaiya MT, Akinyemi J, Owolabi M, Akpa OM. Stroke risk prediction models: a systematic review and meta-analysis. *J Neurol Sci* 2024;**460**:122997.
 19. Richards A, Cheng EM. Stroke risk calculators in the era of electronic health records linked to administrative databases. *Stroke* 2013;**44**:564–569.
 20. Chahine Y, Magoon MJ, Maidu B, del Álamo JC, Boyle PM, Akoum N. Machine learning and the conundrum of stroke risk prediction. *Arrhythm Electrophysiol Rev* 2023;**12**:e07.
 21. Karmali KN, Persell SD, Perel P, Lloyd-Jones DM, Berendsen MA, Huffman MD. Risk scoring for the primary prevention of cardiovascular disease - karmali, KN - 2017 | cochrane library. *Cochrane Database Syst Rev* 2017;**3**:CD006887.
 22. Krittanawong C, Virk HUH, Bangalore S, Wang Z, Johnson KW, Pinotti R, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep* 2020;**10**:16057.
 23. Sheikh A, Nurmatov U, Al-Katheeri HA, Huneiti RAA. Risk prediction models for atherosclerotic cardiovascular disease: a systematic assessment with particular reference to Qatar. *Qatar Med J* 2021;**2021**:42.
 24. de Jong Y, Ramspek CL, van der Endt VHW, Rookmaaker MB, Blankstijn PJ, Vernooij RWM, et al. A systematic review and external validation of stroke prediction models demonstrates poor performance in dialysis patients. *J Clin Epidemiol* 2020;**123**:69–79.
 25. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016;**5**:210.
 26. Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;**356**:i6460.
 27. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;**170**:W1–33.
 28. Bridge J, Blakey JD, Bonnett LJ. A systematic review of methodology used in the development of prediction models for future asthma exacerbation. *BMC Med Res Methodol* 2020;**20**:22.
 29. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res* 2019;**28**:2768–2786.
 30. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation* 2010;**121**:1768–1777.
 31. Khan SS, Ning H, Shah SJ, Yancy CW, Carnethon M, Berry JD, et al. 10-Year risk equations for incident heart failure in the general population. *J Am Coll Cardiol* 2019;**73**:2388–2397.
 32. Snell KI, Ensor J, Debray TP, Moons KG, Riley RD. Meta-analysis of prediction model performance across multiple studies: which scale helps ensure between-study normality for the C-statistic and calibration measures? *Stat Methods Med Res* 2018;**27**:3505–3522.
 33. Riley RD, Tierney JF, Stewart LA. Eds. *Individual Participant Data Meta-Analysis: a Handbook for Healthcare Research*. Hoboken, NJ: Wiley; 2021. p1. (Wiley series in statistics in practice).
 34. R Core Team. *R: A Language and Environment for Statistical Computing [Internet]*. Vienna, Austria: R Foundation for Statistical Computing; 2018. Available from: <https://www.R-project.org/>.
 35. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;**36**:1–48.
 36. Arafa A, Kokubo Y, Sheerah HA, Sakai Y, Watanabe E, Li J, et al. Developing a stroke risk prediction model using cardiovascular risk factors: the suita study. *Cerebrovasc Dis* 2022;**51**:323–330.
 37. Assmann G, Schulte H, Cullen P, Seedorf U. Assessing risk of myocardial infarction and stroke: new data from the prospective cardiovascular münster (PROCAM) study. *Eur J Clin Invest* 2007;**37**:925–932.
 38. Borglykke A, Andreassen AH, Kuulasmaa K, Sans S, Ducimetière P, Vanuzzo D, et al. Stroke risk estimation across nine European countries in the MORGAM project. *Heart* 2010;**96**:1997–2004.
 39. Camen S, Palosaari T, Reinikainen J, Sprünker NA, Niiranen T, Gianfagna F, et al. Cardiac troponin I and incident stroke in European cohorts. *Stroke* 2020;**51**:2770–2777.
 40. Chambless LE, Heiss G, Shahar E, Earp MJ, Toole J. Prediction of ischemic stroke risk in the atherosclerosis risk in communities study. *Am J Epidemiol* 2004;**160**:259–269.
 41. Chien KL, Su TC, Hsu HC, Chang WT, Chen PC, Sung FC, et al. Constructing the prediction model for the risk of stroke in a Chinese population: report from a cohort study in Taiwan. *Stroke* 2010;**41**:1858–1864.
 42. Chun M, Clarke R, Zhu T, Clifton D, Bennett DA, Chen Y, et al. Development, validation and comparison of multivariable risk scores for prediction of total stroke and stroke types in Chinese adults: a prospective study of 0.5 million adults. *Stroke Vasc Neurol* 2022;**7**:328–336.
 43. Di Castelnuovo A, Veronesi G, Costanzo S, Zeller T, Schnabel RB, de Curtis A, et al. NT-proBNP (N-Terminal pro-B-type natriuretic peptide) and the risk of stroke. *Stroke* 2019;**50**:610–617.
 44. Ferket BS, van Kempen BJH, Wieberdink RG, Steyerberg EW, Koudstaal PJ, Hofman A, et al. Separate prediction of intracerebral hemorrhage and ischemic stroke. *Neurology* 2014;**82**:1804–1812.
 45. Flueckiger P, Longstreth W, Herrington D, Yeboah J. Revised framingham stroke risk score, nontraditional risk markers, and incident stroke in a multiethnic cohort. *Stroke* 2018;**49**:363–369.
 46. Foraker RE, Greiner M, Sims M, Tucker KL, Towfighi A, Bidulescu A, et al. Comparison of risk scores for the prediction of stroke in African Americans: findings from the Jackson heart study. *Am Heart J* 2016;**177**:25–32.
 47. Harada A, Ueshima H, Kinoshita Y, Miura K, Ohkubo T, Asayama K, et al. Absolute risk score for stroke, myocardial infarction, and all cardiovascular disease: japan arteriosclerosis longitudinal study. *Hypertens Res* 2019;**42**:567–579.
 48. Hilvo M, Dhar I, Lääperi M, Lysne V, Sulo G, Tell GS, et al. Primary cardiovascular risk prediction by LDL-cholesterol in Caucasian middle-aged and older adults: a joint analysis of three cohorts. *Eur J Prev Cardiol* 2022;**29**:e128–e137.
 49. Hong C, Pencina MJ, Wojdyla DM, Hall JL, Judd SE, Cary M, et al. Predictive accuracy of stroke risk prediction models across black and white race, sex, and age groups. *JAMA* 2023;**329**:306–317.
 50. Howard G, McClure LA, Moy CS, Howard VJ, Judd SE, Yuan Y, et al. Self-reported stroke risk stratification: the REasons for geographic and racial differences in stroke (REGARDS) study. *Stroke* 2017;**48**:1737–1743.
 51. Hung CY, Lin CH, Lan TH, Peng GS, Lee CC. Development of an intelligent decision support system for ischemic stroke risk assessment in a population-based electronic health record database. *PLoS One* 2019;**14**:e0213007.
 52. Hung CY, Lin CH, Lee CC. Improving young stroke prediction by learning with active data augment in a large-scale electronic medical claims database. *Annu Int Conf IEEE Eng Med Biol Soc* 2018;**2018**:5362–5365.
 53. Hunter E, Kelleher JD. Age specific models to capture the change in risk factor contribution by age to short term primary ischemic stroke risk. *Front Neurol* 2022;**13**:803749.
 54. Jung KJ, Hwang S, Lee S, Kim HC, Jee SH. Traditional and genetic risk score and stroke risk prediction in Korea. *Korean Circ J* 2018;**48**:731–740.
 55. Lee S, Lee H, Kim HS, Koh SB. Incidence, risk factors, and prediction of myocardial infarction and stroke in farmers: a Korean nationwide population-based study. *J Prev Med Public Health* 2020;**53**:313–322.
 56. Li X, Wang Y, Xu J. Development of a machine learning-based risk prediction model for cerebral infarction and comparison with nomogram model. *J Affect Disord* 2022;**314**:341–348.
 57. Li Y, Mamouei M, Salimi-Khorshidi G, Rao S, Hassaine A, Canoy D, et al. Hi-BEHT: hierarchical transformer-based model for accurate prediction of clinical events using multimodal longitudinal electronic health records. *IEEE J Biomed Health Inform* 2023;**27**:1106–1117.
 58. Lolak S, Attia J, McKay GJ, Thakkinstant A. Comparing explainable machine learning approaches with traditional statistical methods for evaluating stroke risk models: retrospective cohort study. *JMIR Cardio* 2023;**7**:e47736.
 59. Majed B, Tafflet M, Kee F, Haas B, Ferrieres J, Montaye M, et al. External validation of the 2008 framingham cardiovascular risk equation for CHD and stroke events in a European population of middle-aged men. *The PRIME study. Prev Med* 2013;**57**:49–54.
 60. Marrugat J, Subirana I, Ramos R, Vila J, Marín-Ibañez A, Guembe MJ, et al. Derivation and validation of a set of 10-year cardiovascular risk predictive functions in Spain: the FRESKO study. *Prev Med* 2014;**61**:66–74.
 61. Teoh D. Towards stroke prediction using electronic health records. *BMC Med Inform Decis Mak* 2018;**18**:127.
 62. Wannamethee SG, Shaper AG, Lennon L, Morris RW. Metabolic syndrome vs framingham risk score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. *Arch Intern Med* 2005;**165**:2644–2650.

63. Wu Y, Fang Y. Stroke prediction with machine learning methods among older Chinese. *Int J Environ Res Public Health* 2020;**17**:1828.
64. Yang K, Chen M, Wang Y, Jiang G, Hou N, Wang L, et al. Development of a predictive risk stratification tool to identify the population over age 45 at risk for new-onset stroke within 7 years. *Front Aging Neurosci* 2023;**15**:1101867.
65. Yatsuya H, Iso H, Li Y, Yamagishi K, Kokubo Y, Saito I, et al. Development of a risk equation for the incidence of coronary artery disease and ischemic stroke for middle-aged Japanese - Japan public health center-based prospective study. *Circ J* 2016;**80**:1386–1395.
66. Yatsuya H, Iso H, Yamagishi K, Kokubo Y, Saito I, Suzuki K, et al. Development of a point-based prediction model for the incidence of total stroke: japan public health center study. *Stroke* 2013;**44**:1295–1302.
67. Zhang XF, Attia J, D'Este C, Yu XH, Wu XG. A risk score predicted coronary heart disease and stroke in a Chinese cohort. *J Clin Epidemiol* 2005;**58**:951–958.
68. Zhang Y, Fang X, Guan S, Wu X, Liu H, Wang C, et al. Validation of 10-year stroke prediction scores in a community-based cohort of Chinese older adults. *Front Neurol* 2020;**11**:986.
69. Dufouil C, Beiser A, McLure LA, Wolf PA, Tzourio C, Howard VJ, et al. Revised framingham stroke risk profile to reflect temporal trends. *Circulation* 2017;**135**:1145–1159.
70. Hung C-Y, Chen W-C, Lai P-T, Lin C-H, Lee C-C. Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. *Annu Int Conf IEEE Eng Med Biol Soc* 2017;**2017**:3110–3113.
71. Sun L, Pennells L, Kaptoge S, Nelson CP, Ritchie SC, Abraham G, et al. Polygenic risk scores in cardiovascular risk prediction: a cohort study and modelling analyses. *PLoS Med* 2021;**18**:e1003498.
72. Xing X, Yang X, Liu F, Li J, Chen J, Liu X, et al. Predicting 10-year and lifetime stroke risk in Chinese population. *Stroke* 2019;**50**:2371–2378.
73. Bos D, Ikram MA, Leening MJG, Ikram MK. The revised framingham stroke risk profile in a primary prevention population: the rotterdam study. *Circulation* 2017;**135**:2207–2209.
74. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the framingham heart study. *Circulation* 2008;**117**:743–753.
75. Li Y, Salimi-Khorshidi G, Rao S, Canoy D, Hassaine A, Lukasiewicz T, et al. Validation of risk prediction models applied to longitudinal electronic health record data for the prediction of major cardiovascular events in the presence of data shifts. *Eur Heart J Digit Health* 2022;**3**:535–547.
76. Vu T, Kokubo Y, Inoue M, Yamamoto M, Mohsen A, Martin-Morales A, et al. Machine learning approaches for stroke risk prediction: findings from the suita study. *J Cardiovasc Dev Dis* 2024;**11**:207.
77. Widimsky P, Snyder K, Sulzenko J, Hopkins LN, Stetkarova I. Acute ischaemic stroke: recent advances in reperfusion treatment. *Eur Heart J* 2023;**44**:1205–1215.
78. Samarasekera EJ, Clark CE, Kaur S, Patel RS, Mills J; Guideline Committee. Cardiovascular disease risk assessment and reduction: summary of updated NICE guidance. *BMJ* 2023;**381**:1028.
79. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. *Circulation* 2019;**140**:e596–e646.
80. Renoux C, Patenaude V, Suissa S. Incidence, mortality, and sex differences of non-valvular atrial fibrillation: a population-based study. *J Am Heart Assoc* 2014;**3**:e001402.
81. Ruff CT, Giugliano RP, Braunwald E, Hoffman EB, Deenadayalu N, Ezekowitz MD, et al. Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: a meta-analysis of randomised trials. *Lancet* 2014;**383**:955–962.
82. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the framingham study. *Stroke* 1991;**22**:983–988.
83. Kjeldsen SE. Hypertension and cardiovascular risk: general aspects. *Pharmacol Res* 2018;**129**:95–99.
84. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: the achilles heel of predictive analytics. *BMC Med* 2019;**17**:230.
85. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Anti-lymphocyte antibody induction therapy study group. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;**21**:371–387.
86. Kang DS, Yang PS, Kim D, Jang E, Yu HT, Kim TH, et al. Racial differences in ischemic and hemorrhagic stroke: an ecological epidemiological study. *Thromb Haemost* 2024;**124**:883–892.
87. Nijman S, Leeuwenberg AM, Beekers I, Verkouter I, Jacobs J, Bots ML, et al. Missing data is poorly handled and reported in prediction model studies using machine learning: a literature review. *J Clin Epidemiol* 2022;**142**:218–229.
88. Tsvetanova A, Sperrin M, Peek N, Buchan I, Hyland S, Martin GP. Missing data was handled inconsistently in UK prediction models: a review of method used. *J Clin Epidemiol* 2021;**140**:149–158.