

PROTOCOL

Open Access



Prognosis prediction models for post-stroke depression: a protocol for systematic review, meta-analysis, and critical appraisal

Lu Zhou¹ , Lei Wang¹, Gao Liu¹ and EnLi Cai^{1*} 

Abstract

Introduction Post-stroke depression (PSD) is a prevalent complication that has been shown to have a negative impact on rehabilitation outcomes and quality of life and poses a significant risk for suicidal intention. However, models for discriminating and predicting PSD in stroke survivors for effective secondary prevention strategies are inadequate as the pathogenesis of PSD remains unknown. Prognostic prediction models that exhibit greater rule-in capacity have the potential to mitigate the issue of underdiagnosis and undertreatment of PSD. Thus, the planned study aims to systematically review and critically evaluate published studies on prognostic prediction models for PSD.

Methods and analysis A systematic literature search will be conducted in PubMed and Embase through Ovid. Two reviewers will complete study screening, data extraction, and quality assessment utilizing appropriate tools. Qualitative data on the characteristics of the included studies, methodological quality, and the appraisal of the clinical applicability of models will be summarized in the form of narrative comments and tables or figures. The predictive performance of the same model involving multiple studies will be synthesized with a random effects meta-analysis model or meta-regression, taking into account heterogeneity.

Ethics and dissemination Ethical approval is considered not applicable for this systematic review. Findings will be shared through dissemination at academic conferences and/or publication in peer-reviewed academic journals.

Systematic review registration PROSPERO CRD42023388548.

Keywords Stroke, Depression, Models, Protocols, Guidelines

Introduction

Stroke is a serious public health concern worldwide, with elevated rates of mortality, disability, and recurrence [1]. Post-stroke depression (PSD) refers to any depressive state that occurs after a stroke, which is the most common neuropsychiatric disorder [2]. The prevalence of PSD ranges from 11 to 41%, with a cumulative incidence

of 65%, of which roughly 14% are diagnosed with major depressive disorder (MDD) [1, 3, 4]. PSD is increasingly becoming a research hotspot due to its severe negative effects and economic burden [5].

While recovery from depression after a stroke within a year improves functional outcomes and quality of life [6, 7], PSD is linked to higher mortality, poorer recovery, more pronounced cognitive impairments, heavier financial burden, and lower quality of life than stroke without depression [8], indicating that depression hinders functional recovery after a stroke [9]. PSD can manifest at any point following a cerebrovascular event. It affects roughly one-third of stroke survivors and is notably associated with compromised functional recovery and

*Correspondence:

EnLi Cai
caienliycm@sina.cn

¹ School of Nursing, Yunnan University of Chinese Medicine, Kunming, China



heightened mortality rates. Thus, early screening and risk stratification interventions for stroke survivors at risk for depression are essential to adequately understand the mechanisms and development of symptomatology and even to change the prognosis. However, PSD arises from the complex interplay of neurobiological and psychosocial factors [1], exhibiting differential effects across various time frames post-stroke. The intricate interaction mechanisms and dynamic evolution of these factors throughout the development of PSD have posed enduring challenges within academic discourse. Consequently, this complexity contributes to suboptimal predictive dynamics and precision in PSD assessment.

Specifically, the diagnosis of PSD is primarily reliant on the Diagnostic and Statistical Manual of Mental Disorders (DSM) guidelines, in conjunction with a range of instruments measuring depression [10], but instruments have limitations in screening for PSD, such as insufficient clinical applicability or poor specificity [11]. In addition, PSD has been frequently under-diagnosed and under-treated due to the pathophysiological mechanisms of PSD not being fully understood [12], causing a sub-optimal prognosis for stroke survivors [13]. Nevertheless, the heightened administration of pharmacological treatment involving antidepressants, specifically escitalopram and fluoxetine, has demonstrated effectiveness in individuals who exhibit a high risk of PSD. But it is essential to note that such treatment may pose an excessive risk of harm in those who exhibit a lower risk of PSD. Notably, fluoxetine cannot improve depressive symptoms in PSD patients [13, 14], and these therapies lack risk stratification. Thus, identifying prediction variables (e.g., biomarkers or psychosocial factors, as well as demographic and clinical characteristics of patients) associated with an increased risk of PSD occurrence and then developing multivariable prediction models is one of the promising PSD prevention strategies [15].

Currently, the construction, validation, and updating of predictive models are gaining attention in clinical research [16]. Prediction models are formal combinations of multiple predictors that estimate the probability of an individual currently having a certain disease (diagnostic model) or having a certain outcome in the future (prognostic model) through a mathematical formula [17], from which risks for specific endpoints can be calculated for individual patients to facilitate the dissemination of preventive interventions, provide patient counseling, and establish clinical guidelines and policies [18, 19]. This study will focus on prognostic models.

Previous work has shown prediction models provide more accurate and less variable estimates of risk compared to more subjectively made predictions [20], but the methodology of model development is key to ensuring

predictive performance. Although an increasing number of prognostic prediction models for PSD have been published [21, 22], there has been limited advancement in the development of prognostic models for the stratification of PSD and MDD in stroke survivors [9], which are mainly based on clinical characteristics and biological markers ignoring psychosocial data support [23] causing the limited clinical predictive value. In addition, most of the existing prediction models are opportunistic and have been rarely used or even mentioned in clinical guidelines [24]. Only a small proportion of these models have been evaluated for their performance in data from other participants. Further, research design flaws, insufficient statistical methods, and incomplete reporting hinder the clinical application of these models. According to the PROGRESS group, significant heterogeneity exists among studies, the inclusion and exclusion criteria are too narrow, stroke type (ischemic or hemorrhagic) is not reported, blinding is rarely reported, preset cutoff values are not reported, multiple predictive models are rarely compared in the same population, and the appraisal of models across different languages, races/ethnicities, and cultures is lacking [15]. These factors point to significant waste in research, including both financial and scientific resources [25].

As the research on PSD prognostic prediction models continues to grow annually, there are varying emphases on the content, format, performance, and modeling approaches. The abundance of available clinical research data poses challenges for clinicians in extracting evidence, making it difficult to discern the most targeted predictive prognostic models to assist clinical decision-making and determine best practices from independently published literature. Furthermore, after preliminary searches in the PROSPERO database, Cochrane systematic review database, and JBI evidence synthesis, no completed or ongoing systematic reviews or scoping reviews were identified.

Thus, a comprehensive review and overview of existing PSD models is necessary to clarify their predictive performance, advantages, disadvantages, usage characteristics, and methodology. This will provide evidence-based support for practitioners in selecting models, while also promoting the development, validation, and updating of prognosis prediction models for PSD.

Research aims

The planned study aims to conduct a systematic review of all available evidence regarding the current prognostic models for PSD and to identify which prognostic prediction models have been developed, establishing the most effective and best performance model to predict PSD,

while informing clinical decision-making. The specific aims of this systematic review are:

1. To ascertain the existing prognostic prediction models for PSD.
2. To qualitatively characterize the qualitative properties of the included prognostic prediction models.
3. To summarize and compare the current prognostic models and their predictive performance.
4. To critically appraise the studies identified for inclusion, particularly the research methodology and reporting methods.

Methods and analysis

The present protocol was formulated in adherence to the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) guidelines and was duly registered with PROSPERO, the international prospective register of systematic reviews (Supplementary Material 1 contains the PRISMA-P checklist for reference) [26, 27].

A systematic review of prognostic prediction modeling studies for PSD will be conducted and will be in accordance with the guidelines established by the Cochrane Prognosis Methods Group (PMG) and PROgnosis REsearch Strategy (PROGRESS) throughout all stages of the process [28–30]. Certain specific steps and models, for instance, framing, critical appraisal, and the assessment of the risk of bias, will be conducted by employing the CHARMS checklist (critical appraisal and data extraction for systematic reviews of prediction modeling studies) [31] and the PROBAST (Prediction model Risk

Of Bias ASsessment Tool) with four domains (i.e., participants, predictors, outcome, and analysis) [32, 33]. Moreover, for predictive modeling studies applying machine learning techniques, study selection and evidence appraisal will be based on the metrics and statements highlighted and extended in the Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis—Artificial Intelligence (TRIPOD-AI) and PROBAST-AI being developed [34].

Eligibility criteria

The outline of the review data and study selection were defined according to the CHARMS checklist (key items to guide the framing of the review aim, search strategy, and study inclusion and exclusion criteria) [31] detailed in Table 1.

The inclusion criteria are as follows: (1) studies that develop or validate prognostic models (e.g., machine learning and Cox models), whether or not they include external validation; (2) study populations that involve research on ischemic or hemorrhagic stroke; (3) primary outcome measures indicating whether PSD occurred; (4) secondary outcome measures related to PSD, such as functional status, health status, quality of life, or mortality; (5) studies with a sample size of adequate power to detect small effects, with a goodness-of-fit statistic of over 0.99 for both closed and non-closed models; (6) cross-sectional and longitudinal primary research or literature research; (7) studies that report statistical models or instruments and their prediction indicators for predicting an individual's risk of a future outcome (i.e., prognostic prediction model); (8) other names for prediction

Table 1 The CHARMS checklist

Item	Comments
1. Prognostic versus diagnostic prediction model	The aim is to predict future events (prognostic prediction model)
2. Intended scope of the review	Models to inform healthcare professionals' clinical decision-making
3. Type of prediction modeling studies	All study types, i.e., prediction model development studies with and without external validation and external model validation studies with or without model updating
4. Target population to whom the prediction model applies	Survivors with stroke diagnosed according to each included study
5. Outcome to be predicted	<ul style="list-style-type: none"> • Assess the presence of post-stroke depression (morbidity) Mortality (primary) • PSD-related outcomes (secondary) PSD-related functional status PSD-related health status PSD-related quality of life
6. Time span of prediction	Depending on the duration of treatment and follow-up, the duration of prediction may differ across studies. However, it is anticipated that the majority of studies will assess endpoints within a range of 1 to 12 months
7. Intended moment of using the model	The proposed model is intended for utilization within clinical or hospital environments during the period of stroke diagnosis. Its purpose is to assist in the process of screening, providing targeted treatment, and offering programmatic support to stroke survivors who are at the highest risk of developing PSD

models include prognostic models, prognostic or prediction indices or rules, risk or clinical prediction models, and predictive models; and (9) prediction models are used to estimate the probability of a specific outcome occurring and can be reported using either absolute probability or relative risk score terms [35].

The exclusion criteria are as follows: (1) diagnostic prediction models; (2) evaluation of the predictive value of more than one variable, but without reporting subgroups or evaluation outcomes; (3) study population not related to stroke or studying a combined population with missing grouping results, or the study population only includes patients with individual or multiple complications of vascular damage (infarction, WMH, atrophy), such as cognitive impairment or dementia; (4) targeting depression occurring before stroke onset; and (5) presented as literature reviews, meta-reviews, protocols, theses, quality improvement activities, editorial comments, or letters, or not available in full text.

There will be no restrictions on year or language. In instances where multiple studies reported results from the identical cohort concerning a specific outcome measure, the data from the study that encompassed the largest patient population will be selected for analysis. Alternatively, if the studies involved an equal number of patients, the data from the earliest published study will be utilized.

Information sources

A search will be conducted in the following electronic databases: Ovid MEDLINE® Epub Ahead of Print, In-Process, Other Non-Indexed Citations, Daily Update; Embase Classic—Ovid®; Coverage: 1946 to present. The reference list of the included studies will undergo a meticulous manual search to identify any additional potentially relevant citations and a manual search will be conducted with the Google Scholar web search engine.

Search strategy

The search strategy will be devised for MEDLINE using the OvidSP platform, incorporating Medical Subject Headings (MeSH) and relevant keywords to enhance the efficacy of the search process (MeSH terms are available in Supplementary Material 2). Specifically, subject indexing terms will include a combination of the following five aspects of the PICOS system search construct [35, 36]: #1 Population search AND #2 Index search AND #3 Comparator search AND #4 Outcomes search NOT #5 Study design-exclusion filter.

All model development studies will be back-citation-searched to identify potentially relevant external validation studies. Subsequently, a comprehensive review of all retrieved studies will be performed to ascertain their suitability for inclusion in the analysis. References identified

by the search strategy will be entered into Endnote bibliographic software to screen the selected articles.

Study records

Data management

Upon exportation from electronic databases, all search results will be subsequently imported into Covidence, a systematic review management platform, to facilitate efficient and organized review and analysis [37], available at <https://www.covidence.org>, and duplicates will be removed.

Study selection

Based on the established eligibility criteria for article selection, one author (L.Z.) will test the retrieve strategy across all the databases while two authors (G.L. and L.W.) will independently screen the titles and abstracts. The search results will be then screened a second time, in duplicate. Potential disagreements regarding the inclusion of an article will be resolved through a discussion but, in case of differences, a third researcher (EL.C.) decides whether to include an article. If there is no sufficient data to determine eligibility, additional information will be obtained from the study authors; if missing data cannot be obtained, studies will be excluded from the analysis. But the report with the highest risk of bias will also be removed if data from the identical samples are related to the same model testing.

Data collection process

The data will be extracted independently across the included studies by two reviewers (L.Z. and G.L.) using a standardized electronic form developed with reference to the CHARMS checklist (relevant items to extract from individual studies in a systematic review of prediction models for purposes of description or assessment of risk of bias or applicability) that is available in Supplementary Material 3 [31]. Moreover, the data items in the checklist will adapt to the specific clinical question, for instance, aims; data source; participants; stakeholders; algorithms; predicted outcomes; potential predictors; sample size; missing data; model development; model performance, including properties of discrimination with confidence intervals, calibration, classification, and overall performance; final multivariable models; interpretation of presented models; and model evaluation. Through discussions between the co-investigator (EL.C.) and two reviewers (L.Z. and L.W.), the two data collection sheets will be reconciled into one data set. Any disagreement or uncertainty will be resolved by discussion among reviewers to reach a consensus, if required, by consulting another author of the review team (EL.C.).

Critical appraisal

PROBAST will be used to analyze the methodological quality and relevance of participants, predictors, and outcomes from each included study to the review topic in a systematic assessment [16]. With a total of 20 signaling questions, this instrument comprises four domains: participants, predictors, results, and analysis. Domains were scored as “high,” “low,” or “unclear” risk of bias. Two reviewers (L.Z. and G.L.) will independently apply the tool to rate the risk of bias and applicability of each included study of the 10 studies. The kappa coefficient for inter-rater reliability should be over 0.8 [38]. Any disagreement will be resolved by discussion. Graphical representations will be utilized to present the findings of each study.

Data synthesis

Evidence synthesis

The initial methodology will involve utilizing a narrative synthesis approach to systematically detail the characteristics and quantitative data obtained from the studies that have been included. Specifically, the qualitative/heterogeneous outcomes of studies, including predictors, performance measures, classification measures, measures of uncertainty, and a descriptive analysis of key items [30], will be summarized qualitatively. Results will be presented in tabular form with each study to facilitate comparison.

Meta-analysis

The homogeneous outcomes of the same prediction model which meet the following criteria will be statistically analyzed in meta-analysis: (1) across ≥ 2 studies; (2) the identical category of prediction modeling study, specifically either development or validation; and (3) the follow-up periods for the primary outcome(s) are considered similar. While conducting the meta-analysis, it is possible to combine re-scaled measures of model performance which have similar outcomes. It will be typically accomplished via a random-effects meta-analysis approach, using restricted maximum likelihood estimation. Additionally, the Hartung-Knapp-Siddik-Jonkman method will be used to derive confidence intervals. Where feasible, 95% prediction intervals will be estimated. The performance of the prognostic prediction model will be based on the following measures [30, 39], detailed in Table 2. Additionally, where possible, we will employ multivariate meta-analysis for jointly synthesizing calibration and discrimination performance, while accounting for their correlation.

Sensitivity analysis and investigation of heterogeneity

To ascertain the robustness of the findings, sensitivity analyses will be conducted, wherein studies deemed to have a significant or uncertain risk of bias will be excluded. The I^2 statistic for univariate meta-analysis models and sub-group analyses will be employed to explore heterogeneity between studies. Between-effects heterogeneity will be estimated via restricted maximum-likelihood I^2 and tau² statistics. Potential sources of considerable between-effects heterogeneity will be investigated by conducting a meta-regression analysis ($p < 0.05$). If possible, the sub-group analysis will be based on:

1. Stroke types—ischemic or hemorrhagic.
2. Risk factors—biomarkers or psychosocial factors.
3. Depression types—PSD or MDD.
4. Modeling techniques—machine learning or non-machine learning.
5. Follow-up duration.
6. Region—based on the Organisation for Economic Co-operation and Development classification, that is, low/middle-income and high-income countries.

The meta-analysis process will be conducted in the metareg module in Stata 13.0 regarding the Meta-analysis of Observational Studies in Epidemiology (MOOSE) group guidelines [40].

Reporting findings

The findings of this systematic review will be reported in adherence to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) guideline [41, 42] and the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [43].

Discussion

The planned study will be the first systematic review to evaluate existing evidence regarding prognostic prediction models (including machine learning algorithms, statistical models, and clinical risk scales) aimed at post-stroke depression for secondary prevention. The occurrence mechanism of PSD is complex and diverse. Currently, there is a lack of a gold standard for diagnosing PSD, and screening instruments have certain limitations, resulting in a relatively high rate of missed diagnoses. Although numerous PSD prediction models have been developed at this stage, most of the prediction models are not developed, validated, and assessed based on guidelines for predictive research [34]. This has led to significant biases in risk estimation and serious deficiencies in statistical methods, as well as a lack of internal

Table 2 Summary of measuring performance of prognostic prediction models

Items	Performance measures/statistics	Values for better performance	Visualization
Discrimination	AUC-ROC	Higher	ROC curve
	ACC	Higher	
	BER	Lower	
	D statistics	Lower	
	MCC	Higher	
	F1 score	Higher	
	Log-rank	Lower, $p > .05$	
	AUC-PRC	Higher	
Calibration	Calibration curve, slope, and intercept	Slope closer to 1 and intercept closer to 0	Calibration plot
	Reliability-deviation	Lower	
	Reliability-within-bin variation	Lower	
	Reliability-within-bin covariance	Higher	
	Resolution	Higher	
	Predictive range	Higher	
	E _{max}	Lower	
	E _{avg}	Lower	
	Hosmer-Lemeshow test	Higher	
	Total O:E ratios	Higher	
	Classification	RMSE via neighborhood estimate	
R ² statistic		Higher	
TPR		Higher	
TNR		Higher	
NRI		Higher	
IDI		Higher	
Overall performance	Brier score	Lower	DCA plot
	Brier skill score	Higher	
	Prediction squared error	Lower	
	Decision Curve Analysis	/	

AUC-ROC Area under receiver operating characteristic curve, TPR Sensitivity, ACC Accuracy, BER Balanced error rate, MCC Matthews correlation coefficient, TNRSpecificity, AUC-PRC Area under precision-recall curve, RMSE Root-mean-squared error, NRI Net reclassification improvement, IDI Integrated discrimination improvement

and external validation [11], affecting the performance and applicability of the models and resulting in less-than-ideal accuracy and precision in clinical PSD prediction. Additionally, at present, there is a lack of systematic reviews and evaluations of PSD prediction models, which hinders relevant practitioners in selecting, promoting, and applying these models. This systematic review refers to details of the foundation and evidence for further studies, which aimed at developing, verifying, implementing, and assessing prognostic prediction models for PSD within the four domains of the PROGRESS prognosis research framework [44]. Regarding the TRIPOD-AI and PROBAST-AI tool, incorporating insights from these forthcoming extensions could enhance the review’s comprehensiveness and relevance, especially concerning

machine learning-based prognostic models, ultimately contributing to more robust and applicable prognostic models for PSD in secondary prevention.

The findings will facilitate the early identification of people at high risk for PSD, the identification of the most effective current prognostic prediction models based on the shown predictive accuracy, and the stratification of PSD severity to estimate the risk of MDD after stroke. This will be a significant step towards informing the clinical management of patients with an established stroke diagnosis. It is essential for accurate identification of PSD, translation of clinical research of high-quality evidence, and savings in healthcare resources. Additionally, it will promote the consideration of the broad continuum of risk related to this condition in routine clinical practice.

At a health service level, prediction models with good performance and high clinical applicability would support a personalized risk-stratified model of care, which would ultimately better direct finite health resources to stroke survivors at high risk of PSD and most likely to benefit from intervention.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13643-024-02544-x>.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements

None declared.

Patient and public involvement

The study was conducted without the involvement of patients and the general public in its design and conception.

Authors' contributions

EL.C. and L.Z. conceived the study idea, and L.Z. was responsible for developing and writing the first draft of the systematic review protocol and manuscript. G.L. and L.W. contributed to the data curation. EL.C., L.Z., G.L., and L.W. provided critical insights at all stages. All authors approved and contributed to the final manuscript.

Funding

The study was supported by the Yunnan University of Chinese Medicine Research Program (202101AZ070001-221), and a grant was provided to the corresponding author (Enli Cai) to support stroke-related research.

Availability of data and materials

This content has been supplied by the authors.

Declarations

Ethics approval and consent to participate

Ethical approval is not applicable to this study. The findings will be disseminated by publishing in peer-reviewed publications and/or academic conferences.

Consent for publication

All authors approved and contributed to the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 1 September 2023 Accepted: 23 April 2024

Published online: 22 May 2024

References

- Guo J, Wang J, Sun W, Liu X. The advances of post-stroke depression: 2021 update. *J Neurol*. 2022;(269):1–14.
- Villa RF, Ferrari F, Moretti A. Post-stroke depression: mechanisms and pharmacological treatment. *Pharmacol Ther*. 2018;184:131–44.
- Mortensen JK, Andersen G. Pharmacological management of post-stroke depression: an update of the evidence and clinical guidance. *Expert Opin Pharmacother*. 2021;22(9):1157–66.
- Hackett ML, Pickles K. Part I: frequency of depression after stroke: an updated systematic review and meta-analysis of observational studies. *Int J Stroke*. 2014;9(8):1017–25.
- James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbas-tabar H, Abd-Allah F, Abdela J, Abdelalim A. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1789–858.
- Medeiros GC, Roy D, Kontos N, Beach SR. Post-stroke depression: a 2020 updated review. *Gen Hosp Psychiatry*. 2020;66:70–80.
- Nys G, Van Zandvoort M, Van Der Worp H, De Haan E, De Kort P, Jansen B, Kappelle L. Early cognitive impairment predicts long-term depressive symptoms and quality of life after stroke. *J Neurol Sci*. 2006;247(2):149–56.
- Shi YZ, Xiang YT, Yang Y, Zhang N, Wang S, Ungvari GS, Chiu HF, Tang WK, Wang YL, Zhao XQ. Depression after minor stroke: the association with disability and quality of life—a 1-year follow-up study. *Int J Geriatr Psychiatry*. 2016;31(4):421–7.
- Hirt J, van Meijeren LC, Saal S, Hafsteinsdóttir TB, Hofmeijer J, Kraft A, Meyer G, Janneke M. Predictive accuracy of the Post-Stroke Depression Prediction Scale: a prospective binational observational study. *J Affect Disord*. 2020;265:39–44.
- Guo J, Wang J, Sun W, Liu X. The advances of post-stroke depression: 2021 update. *J Neurol*. 2022;269(3):1236–49.
- Pollett S, Johansson MA, Reich NG, Brett-Major D, Del Valle SY, Venkatraman S, Lowe R, Porco T, Berry IM, Deshpande A. Recommended reporting items for epidemic forecasting and prediction research: the EPiFORGE 2020 guidelines. *PLoS Med*. 2021;18(10):e1003793.
- Ladwig S, Ziegler M, Südmeyer M, Werheid K. The Post-Stroke Depression Risk Scale (PoStDeRiS): development of an acute-phase prediction model for depression 6 months after stroke. *J Acad Consult Liaison Psychiatry*. 2022;63(2):144–52.
- Choi-Kwon S, Han SW, Kwon SU, Kang D-W, Choi JM, Kim JS. Fluoxetine treatment in poststroke depression, emotional incontinence, and anger proneness: a double-blind, placebo-controlled study. *Stroke*. 2006;37(1):156–61.
- Yi Z, Liu F, Zhai S. Fluoxetine for the prophylaxis of poststroke depression in patients with stroke: a meta-analysis. *Int J Clin Pract*. 2010;64(9):1310–7.
- Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG, Group P. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med*. 2013;10(2):e1001381.
- Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. Group P. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51–8.
- Riley RD, Moons KG, Snell KI, Ensor J, Hoof L, Altman DG, Hayden J, Collins GS, Debray TP. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ*. 2019;364:k4597.
- Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, Voysey M, Wharton R, Yu L-M, Moons KG. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14(1):1–11.
- Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, Malats N, Briggs A, Schroter S, Altman DG. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med*. 2013;10(2):e1001380.
- Kattan MW, Yu C, Stephenson AJ, Sartor O, Tombal B. Clinicians versus nomogram: predicting future technetium-99m bone scan positivity in patients with rising prostate-specific antigen after radical prostatectomy for prostate cancer. *Urology*. 2013;81(5):956–61.
- Liu R, Yue Y, Jiang H, Lu J, Yuan Y, Wang Q. A risk prediction model of PSD in stroke survivors. *Eur Psychiatry*. 2015;30(S1):1–1.
- Ryu YH, Kim SY, Kim TU, Lee SJ, Park SJ, Jung H-Y, Hyun JK. Prediction of poststroke depression based on the outcomes of machine learning algorithms. *J Clin Med*. 2022;11(8):2264.
- Liu R, Yue Y, Jiang H, Lu J, Wu A, Geng D, Wang J, Lu J, Li S, Tang H. A risk prediction model for post-stroke depression in Chinese stroke survivors based on clinical and socio-psychological features. *Oncotarget*. 2017;8(38):62891.

24. Vickers AJ, Cronin AM. Everything you always wanted to know about evaluating prediction models (but were too afraid to ask). *Urology*. 2010;76(6):1298–301.
25. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *The Lancet*. 2009;374(9683):86–9.
26. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, Stewart L. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev*. 2012;1(1):1–9.
27. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4(1):1–9.
28. Riley RD, Ridley G, Williams K, Altman DG, Hayden J, De Vet H. Prognosis research: toward evidence-based results and a Cochrane methods group. *J Clin Epidemiol*. 2007;60(8):863–5.
29. Moons KG, Hooft L, Williams K, Hayden JA, Damen JA, Riley RD. Implementing systematic reviews of prognosis studies in Cochrane. *Cochrane Database Syst Rev*. 2018;10:ED000129.
30. Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, Riley RD, Moons KG. A guide to systematic review and meta-analysis of prediction model performance. *BMJ*. 2017;356:i6460.
31. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11(10):e1001744.
32. Moons KG, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170(1):W1–33.
33. Fernandez-Felix BM, López-Alcalde J, Roqué M, Muriel A, Zamora J. CHARMS and PROBAST at your fingertips: a template for data extraction and risk of bias assessment in systematic reviews of predictive models. *BMC Med Res Methodol*. 2023;23(1):1–8.
34. Collins GS, Dhiman P, Navarro CLA, Ma J, Hooft L, Reitsma JB, Logullo P, Beam AL, Peng L, Van Calster B. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open*. 2021;11(7):e048008.
35. Cooray SD, Boyle JA, Soldatos G, Wijeyaratne LA, Teede HJ. Prognostic prediction models for pregnancy complications in women with gestational diabetes: a protocol for systematic review, critical appraisal and meta-analysis. *Syst Rev*. 2019;8(1):1–10.
36. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Serv Res*. 2014;14(1):1–10.
37. Babineau J. Product review: Covidence (systematic review software). *J Can Health Libr Assoc*. 2014;35(2):68–71.
38. Kvalseth TO. A coefficient of agreement for nominal scales: an asymmetric version of Kappa. *Educ Psychol Meas*. 1991;51(1):95–101.
39. Huang C, Li S-X, Caraballo C, Masoudi FA, Rumsfeld JS, Spertus JA, Normand SLT, Mortazavi BJ, Krumholz HM. Performance metrics for the comparative analysis of clinical risk prediction models employing machine learning. *Circ Cardiovasc Qual Outcomes*. 2021;14(10):e007526.
40. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*. 2000;283(15):2008–12.
41. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *J British Surg*. 2015;102(3):148–58.
42. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–73.
43. Moher D, Liberati A, Tetzlaff J, Altman DG. Group* P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9.
44. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KG, Steyerberg EW. Prognosis Research Strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ*. 2013;346:e5595.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.