

Prediction models for postoperative delirium of cardiovascular surgery (PODOCVS): A protocol for a systematic review

Xuling Zhao, Yike Wang, Liju Li, Meijuan Lan, Xiaodi He

Submitted to: JMIR Research Protocols
on: April 02, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 23

0..... 23

0..... 23

Multimedia Appendixes 24

Multimedia Appendix 0..... 24

Prediction models for postoperative delirium of cardiovascular surgery (PODOCVS): A protocol for a systematic review

Xuling Zhao^{1*} MFA; Yike Wang^{2*} MFA; Liju Li¹ MFA; Meijuan Lan² MFA; Xiaodi He² MFA

¹ Zhejiang Taizhou Hospital Linhai CN

² The Second Affiliated Hospital of Zhejiang University School of Medicine Hangzhou CN

*these authors contributed equally

Corresponding Author:

Yike Wang MFA

The Second Affiliated Hospital of Zhejiang University School of Medicine
Jiefang Road 88
Hangzhou
CN

Abstract

Postoperative delirium of cardiovascular surgery (PODOCVS) is an acute brain dysfunction characterised by inattention, impaired consciousness, and cognitive disorders, and the severity and presence of these symptoms fluctuate over time. PODOCVS occurs during the early postoperative period and is associated with adverse outcomes, including prolonged mechanical ventilation, extended hospital stays, increased requirements for long-term care, exacerbation of preexisting cognitive impairments, new-onset dementia, heightened anxiety and depression, and premature mortality. Advances in its early diagnosis and treatment have mitigated some of the initial adverse effects of PODOCVS, but models for predicting risk in patients who have already developed PODOCVS remain inadequate for effective secondary prevention. Developing multivariable prediction models for stratifying PODOCVS risk would enable early, personalised interventions. To this end, this protocol aims to systematically review and critically evaluate the development, performance, and applicability of existing prediction models for PODOCVS. We will search various databases including Embase, PubMed, the Web of Science Core Collection, Wan Fang, Wei Pu, and the China National Knowledge Infrastructure (CNKI) for studies on multivariate prediction models for PODOCVS. A manual search of the included studies' reference lists will also be conducted to identify any additional relevant publications. No restrictions will be placed on publication status or language. The studies will undergo that meet the following criteria: studies with subject populations comprising adult cardiovascular surgery patients aged ≥ 18 years; studies involving the development and internal or external validation of predictive models for PODOCVS via multivariate analysis; studies with outcome measures focused on postoperative delirium. Two researchers will independently extract the data and assess the included studies' model quality using the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist and the Predictive Model Bias Risk Assessment Tool (PROBAST). This protocol ensures methodological transparency, stimulates discourse, and paves the way for informed interventions to improve overall health outcomes. PROSPERO registry: CRD42024578957.

(JMIR Preprints 02/04/2025:75368)

DOI: <https://doi.org/10.2196/preprints.75368>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/preprint/75368>

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://www.jmir.org/preprint/75368>



Original Manuscript

Prediction models for postoperative delirium of cardiovascular surgery (PODOCVS): A protocol for a systematic review

Zhao Xuling ^{a,†}, Wang Yike ^{b,†*}, Li Liju ^a, Lan Meijuan ^a, He Xiaodi ^a

^a Taizhou Hospital of Zhejiang Province, Taizhou, China

^b Department of Nursing, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China

[†] Contributed equally to this work and shared first authorship

* Corresponding author: Wang Yike, The Second Affiliated Hospital Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. E-mail: 2522080@zju.edu.cn

Abstract

Background

Postoperative delirium of cardiovascular surgery (PODOCVS) is an acute brain dysfunction characterised by inattention, impaired consciousness, and cognitive disorders, and the severity and presence of these symptoms fluctuate over time. PODOCVS occurs during the early postoperative period and is associated with adverse outcomes, including prolonged mechanical ventilation, extended hospital stays, increased requirements for long-term care, exacerbation of preexisting cognitive impairments, new-onset dementia, heightened anxiety and depression, and premature mortality. Advances in its early diagnosis and treatment have mitigated some of the initial adverse effects of PODOCVS, but models for predicting risk in patients who have already developed PODOCVS remain inadequate for effective secondary prevention. Developing multivariable prediction models for stratifying PODOCVS risk would enable early, personalised interventions.

Objective

The aims of this study are to systematically review and critically evaluate the development, performance, and applicability of existing prediction models for PODOCVS.

Methods

An extensive systematic search will be performed across multiple databases, including Embase, PubMed, the Web of Science Core Collection, Wanfang Data, Wei Pu, and the China National Knowledge Infrastructure (CNKI), to identify studies related to multivariate predictive models for PODOCVS. A manual search of the included studies' reference lists will also be conducted to identify any additional relevant publications. No restrictions will be placed on publication status or language. This systematic review will include studies that meet the following criteria: studies with subject populations comprising adult cardiovascular surgery patients aged ≥ 18 years; studies involving the development and internal or external validation of predictive models for PODOCVS via multivariate analysis; studies with outcome measures focused on postoperative delirium. Two researchers will independently extract the data and assess the included studies' model quality using the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist and the Predictive Model Bias Risk Assessment Tool (PROBAST). Since this study will not involve patient data, ethics approval is not required. Our findings will be published in a peer-reviewed scientific

journal and the dataset will be made freely available.

Result

Literature searches were conducted in 20 May 2024 (updated up to 20 March, 2025) and data extraction and analysis are expected to be complete by the end of May 2025. We currently have a preliminary plan to publish the complete study results by August 2025, subject to any unforeseen delays or changes in the research timeline.

Conclusions

We present a protocol for the systematic review of prediction models for postoperative delirium in cardiac surgery (CS) patients. Aiming to identify, summarize, and critically appraise existing risk models globally, this review seeks to provide an up-to-date reference for stakeholders involved in CS patient care, policy-making, and research. Additionally, we aim to investigate whether machine learning (ML) models for PODOCVS offer more accurate predictions compared to traditional statistical models.

Trial Registration

PROSPERO CRD42024578957; <https://www.crd.york.ac.uk/PROSPERO/recorddashboard>

Keywords prediction models; postoperative delirium; cardiovascular surgery; acute brain dysfunction; machine learning; systematic review

Introduction

Background

Postoperative delirium of cardiovascular surgery (PODOCVS) is an acute brain dysfunction characterised by inattention, impaired consciousness, and cognitive impairment that fluctuates in presence and severity across time[1]. Studies show that PODOCVS affects approximately 26% to 52% of cardiovascular surgery (CS) patients[2]. As a common complication of cardiovascular surgery, PODOCVS worsens patient outcomes[3] and consumes enormous financial resources in excess of \$164 billion per year[4]. PODOCVS is independently associated with major surgical trauma, advanced age, cardiopulmonary bypass, compromised brain health, and transfer to the cardiovascular intensive care unit post-surgery[2, 5, 6]. PODOCVS occurs during the early postoperative period and is associated with adverse outcomes including prolonged mechanical ventilation times and hospital stays, long-term nursing needs, worsening of preexisting cognitive impairment, new dementia, increased anxiety and depression, and premature mortality[7]. In terms of pharmacological therapies, ramelteon and suvorexant can reportedly reduce delirium, but premedication to prevent delirium is not recommended for all patients[8]. However, due to the lack of specific treatments,

strategies for managing delirium currently focus only on its prevention and early detection[9, 10]. Therefore, identifying populations at high risk of developing delirium and providing targeted and effective interventions are crucial[9, 11]. Risk stratification and identification of vulnerable patients offer an extremely efficient method of protecting them from the initial adverse outcomes of PODOCVS. The pathogenesis of PODOCVS is multifactorial and, as such, multivariable prediction models for stratifying PODOCVS risk may enable early personalised treatment interventions[12-14]. This would significantly reduce the incidence of delirium, shorten hospital stays, cut medical costs, and reduce the incidence of postoperative complications and mortality risk[4].

At present, there are many PODOCVS prediction models used in clinical practice. These can be divided into traditional statistical models and machine learning (ML) models. Although statistical models (e.g., logistic regression) are favourable in terms of model interpretability, Seong et al.[15] indicate that machine learning (ML) is preferred for predictive accuracy. The predictive power of traditional statistics-based models (e.g., the E-PRE-DELIRIC model) varies greatly across regional cohorts with an area under the curve (AUC) of 0.54, which is below the receiver operating characteristic curve of 0.75[16, 17]. Hani[18] and Xue[19] have shown that the predictive accuracy of ML algorithms was superior to that of conventional statistical models in the area of PODOCVS[20]. ML models can handle complex, multidimensional data without the limitations of traditional statistical methods. Although there is an expanding body of published literature on the use of ML models in cardiovascular health care, whether their performance in predicting PODOCVS is superior to that of traditional statistical models remains to be verified [21].

The accuracy of predictive models for patients at risk of developing delirium has been reported to be insufficient when applied to CS patients[17]. Hence, predictive models for delirium risk should be constructed for CS patients specifically[22]. In recent years, an increasing number of studies have focused on developing or validating predictive models to estimate PODOCVS risk. However, healthcare professionals are uncertain as to which model to use for patients undergoing CS in specific patterns, settings, and populations. Hence, the quality of existing models and their applicability remain unclear. Therefore, it is critical to thoroughly evaluate the predictive performance, applicability, and quality of existing PODOCVS risk prediction models[23]. To this end, this study presents a protocol for the systematic review and critical assessment of the quality, performance, and applicability of current predictive models for PODOCVS risk.

Objectives

The systematic review aims to thoroughly evaluate the predictive performance, applicability, and quality of existing PODOCVS risk prediction models. To provide a robust scientific foundation for healthcare professionals to select appropriate delirium prediction models tailored to CS patients in specific clinical environments. The objectives of this project are as follows:

- The construction methodology and model validation approach of the existing PODOCVS risk prediction

model were thoroughly analyzed, and the predictors were ranked based on the prevalence and strength of the predictors.

- Comprehensively assess the predictive performance, applicability, and quality of existing PODOCVS risk prediction models.
- Inform stakeholders, such as policymakers and healthcare workers directly involved in the treatment of CS patients, on the available delirium prediction models and their setting-specific clinical utility, strengths and limitations.
- Determine whether PODOCVS ML models can provide more accurate predictions as compared to PODOCVS traditional statistical models.

Method

Study Design and Data Source

This protocol was developed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) guidelines (see Supplemental Material, Item 1)[24].

A systematic review will be conducted according to the recommended methods for systematic reviews and meta-analyses of prediction models[24] and will adhere to the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist[25]. The results of this review will be reported as per the updated PRISMA 2020 checklist[26]. The review will be registered in the International Prospective Register of Systematic Reviews (PROSPERO). Any modifications to the protocol will be amended accordingly.

Eligibility Criteria

This review follows the PICOTS (population, intervention model, comparator, outcomes, timing, and setting) approach to frame the review question and determine the inclusion criteria (see Table 1)[25, 26].

This systematic review will include studies with subject populations of CS patients aged 18 years and older, and studies using developed and internally or externally validated predictive models for PODOCVS risk based on multivariate analysis. The outcome metric is postoperative delirium. Studies will not be excluded based on their publication status and language. The exclusion criteria are as follows: studies focused solely on the external validation of existing models; studies that construct predictive models without conducting internal/external verification; studies for which data are unavailable through the accessible channels; studies lacking evidence of model performance in derivation or validation cohorts; studies that only analyse predictive risk factors without constructing predictive models; and various types of publications such as conference abstracts, reviews, comments, research protocols, animal model studies, and textbook materials.

Table 1: Eligibility criteria based on the PICOS

Population	The population of interest comprises CS patients aged 18 years and older.
Index model	This review will include studies involving the development and concurrent internal or external validation of predictive models for PODOCVS risk. Studies that focus exclusively on the external validation of existing models and studies that construct predictive models without conducting internal/external validation will be excluded from consideration.
Comparator	Not applicable .
Outcome	The outcome is defined as postoperative delirium.
Timing	Outcomes will be predicted using either preoperative, intraoperative, or postoperative conditions.
Setting	Predictive models are intended to perform risk stratification for the development of delirium, enabling the implementation of preventive measures and early intervention.

Search strategy

Once this protocol is published, the search will be performed in the following databases: including Embase, PubMed, the Web of Science Cor Collection, the China National Knowledge Infrastructure (CNKI), Wan Fang, and Wei Pu (VIP). The following search terms will be used: “cardiac surgical procedure”, “cardiac surgery”, “cardiovascular surgery”, “heart surgery”, “coronary artery bypass grafting,” “Coronary Artery Bypass Grafting (CABG)”, “extracorporeal circulation”, “type A aortic dissection”, “type B aortic dissection,” “subacute delirium”, “delirium”, “postoperative delirium”, “delirium of mixed origin”, “deliri*”, “psychosis”, “intensive care delirium”, “neurological complications”, “risk prediction”, “model”, “risk score”, and “risk assessment”. The retrieval time limit will range from each corresponding database’s inception until May 20, 2024 (updated until March 20, 2025). The search will be performed by two authors (W.Y.K. and Z.X.L.). We will perform both forward and backward citation searches for the included studies and relevant previous systematic reviews. The complete search strings are presented in Supplemental Material, Item 2.

Selection process

Endnote’s *Delete duplicates* function will be used, followed by the manual deletion of any remaining duplicates. Subsequently, the authors (Z.X.L. and W.Y.K.) will evaluate the acceptability of the titles and abstracts. In case of an eventual disagreement between the authors who analyze the eligibility of the documents, the opinion of a third author (L.L.J.) will be considered. Any inconsistencies will be resolved through discussion to arrive at a consensus. Thereafter, the selected full texts will finally be included in the review.

Data Extraction and Management

The following information will be identified from each selected study and summarized in tables for qualitative analysis: author name, country, aim of the study, model development sample size, external validation sample size, internal validation sample size, participants, follow-up, study design (prospective study or retrospective study), events per variable, main outcome, predictors in final model, modeling methodology, number of models,



model characteristics (i.e., modelling methods, model validation methods, and candidate predictors) and model performances (i.e., discrimination, calibration, and classification measures) (Table 2). Data will be extracted independently by two authors (Z.X.L. and W.Y.K.), who will enter the data into different Microsoft Excel spreadsheets. A third author will verify their agreement and resolve disagreements by re-analyzing the data (L.L.J.). A further limitation of the planned review is that we will not contact study authors to request unreported information, as we will explicitly document instances where information is missing.

Table 2 Information for data extraction and subsequent summary and appraisal

Domain	Key items
Source of data	Source of data (eg, cohort, case-control, randomised trial participants, registry data, etc)
Participants	Participant eligibility and recruitment method (eg, location, number of centres, setting, inclusion and exclusion criteria) Participant description (age, sex, Types of cardiovascular surgery, Postoperative follow-up time) Details of undergoing cardiovascular surgery How delirium diagnosis is defined (whether consistent for all participants, clinical history and physical signs, etc) Study dates
Outcome(s) to be predicted	Type of outcome (eg, single or combined endpoints) Definition and method for measurement of outcome (for example, is mortality disease-specific or all-cause, is cure/initial failure/relapse diagnosed based on clinical symptoms and/or diagnostic testing) Was the same outcome definition (and method for measurement) used in all patients? Time of outcome occurrence or summary of duration of follow-up Was the outcome assessed without knowledge of the candidate predictors (ie, blinded)?
Candidate predictors	Number and type of predictors (eg, demographics, patient history, physical examination, laboratory parameters, etc) Definition and method for measurement of candidate predictors (including whether defined and measured in a similar way for all participants)

	Timing of predictor measurement (eg, at patient presentation, at diagnosis, at treatment initiation or otherwise)
	Handling of predictors in the modelling (eg, continuous, linear, non-linear transformations or categorised)
Sample size	Number of participants and number of outcomes/events Events per candidate predictor
Missing data	Whether the authors describe a sample size calculation Number of participants with any missing value (including predictors and outcomes) Number of participants with missing data for each predictor
Model development	Handling of missing data (eg, complete-case analysis, imputation or other methods) Modelling method (eg, logistic, survival or other) Modelling assumptions satisfied Description of participants that were excluded from the analysis with justification Method for selection of predictors for inclusion in multivariable modelling (eg, all candidate predictors, preselection based on unadjusted association with the outcome) Method for selection of predictors during multivariable modelling (eg, full model approach, backward or forward selection) and criteria used (eg, p value, Akaike information criterion)
Model performance	Shrinkage of predictor weights or regression coefficients (eg, no shrinkage, uniform shrinkage, penalised estimation) Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test), discrimination (C-statistic, D-statistic, log-rank) and overall performance measures with confidence intervals Classification measures (eg, sensitivity, specificity, predictive values, net reclassification improvement) and whether a priori cut points were used
Model evaluation	Method used for testing model performance: development dataset only (apparent performance, random split of data, resampling methods, eg, bootstrap or cross-validation, none) or separate external validation For external validations; data source and participants to be described as per 'source of data' and 'participants' domains. Definitions and distributions (including missing data) of outcome and candidate predictors

Traditional prediction model	Traditional algorithm-based prediction models refer to those that employ classical statistical methods for forecasting future events or values, for example, linear regression, logistic regression.
Machine Learning Prediction model	Machine learning-based prediction models are capable of processing and analyzing extensive historical datasets, extracting underlying patterns and rules, and leveraging the acquired knowledge to forecast future scenarios, for example, extreme gradient boosting (XGBoost), support vector machine (SVM), adaptive boosting (AdaBoost), multilayer perceptron (MLP), neural network (NN), naive bayes (NB), gradient boosting machine (GBM) In case of poor external validation, whether model was updated or extended (eg, intercept recalibrated, predictor effects adjusted, or new predictors added)
Results	Final and other multivariable models presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with SEs or CIs) Any alternative presentation of the final prediction models, for example, sum score, nomogram, score chart, predictions for specific risk subgroups with performance Comparison of the definition and distribution of predictors (including missing data) for development and validation datasets
Interpretation and discussion	Study authors' interpretation of presented models (intended use, clinical utility, etc) Study authors' reported strengths and limitations

Adapted from CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) and Prediction model Risk Of Bias ASsessment Tool (PROBAST).
CI, confidence interval; SE, standard error; CS, cardiovascular surgery; PODOCVS: postoperative delirium of cardiovascular surgery.

Risk of Bias in Individual Studies

Two authors (Z.X.L. and W.Y.K.) will independently use the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist[25] to extract the relevant key information. Candidate studies with uncertain model characteristics (i.e., modelling methods, model validation methods, and candidate predictors) and model performances (i.e., discrimination, calibration, and classification measures) will be discussed

with a third author (L.L.J.) before inclusion.

We will utilise the Prediction Model Risk of Bias Assessment Tool (PROBAST) to evaluate the quality of the candidate studies, as evaluated independently by two authors (Z.X.L. and W.Y.K.). Similarly, uncertain cases will be discussed with a third author (L.L.J.) before inclusion. PROBAST[27] comprises four domains, namely, participants, predictors, outcome, and analysis. Each domain includes 20 signalling questions used to evaluate the risk of bias and applicability. These signalling questions rely on factual information and are categorised as either 'Yes' or 'Probably yes' (Y), 'No' or 'Probably no' (N), or 'No information' (NI). After evaluating each domain using signalling questions specific to that domain, they will be categorised as either 'High', 'Low', or 'Unclear'. By synthesising the results obtained from each dimension, an overall assessment will be made regarding the predictive models' risk of bias and applicability.

Assessing the models' predictive performance & accuracy

The evaluation of the predictive models' performance will be based on an assessment of their discrimination and calibration[28]. 'Discrimination' refers to a model's ability to distinguish between individuals with different outcomes or conditions. It assesses how well the model correctly ranks individuals based on their risk of experiencing a certain event or condition. Common methods for evaluating discrimination include the c-statistic (also referred to as the area under the receiver operating characteristic curve) and the concordance index[28]. 'Calibration' refers to the conformity between a model's predicted probabilities or risk evaluations and the observed actual outcomes. Calibration is often assessed by comparing the predicted and observed event rates among different risk groups, or by using calibration plots[20, 28].

Patient and public involvement

None.

Ethics and dissemination

Ethical approval is not required for this review, as it will exclusively use data from

published studies. The findings of this systematic review will be published in an open-access journal and disseminated at scientific conferences so that they are accessible to as many interested parties as possible.

Result

Literature searches were conducted in 20 May 2024 (updated up to 20 March, 2025) and data extraction and analysis are expected to be complete by the end of May 2025. We currently have a preliminary plan to publish the complete study results by August 2025, subject to any unforeseen delays or changes in the research timeline. The findings will be utilized to inform stakeholders, including policymakers and healthcare professionals directly involved in the treatment of CS patients, about the existing delirium prediction models, their setting-specific clinical utility, strengths, and limitations. Additionally, it serves as a reference for selecting between traditional statistical models and ML models in future research.

Discussion

Expected Results and Practical Implications

This protocol aims to provide a detailed description of the process for conducting a systematic review of existing global predictive models for PODOCVS, focusing on their predictive performance, applicability, and quality. Due to differences in the basic characteristics of study populations, protocols, surgical methods, and evaluation methods and frequency, the reported incidence of PODOCVS varies greatly[29-31]. PODOCVS can cause serious adverse outcomes[32], and prevention is the most effective strategy for minimising its occurrence and poor prognosis[2]. Accurate risk estimation for surgical patients using predictive models can aid clinical decision-making and inform policy, thereby guiding the optimal allocation of often-limited resources. By identifying, summarising, and evaluating published PODOCVS prediction models, this systematic review will serve as a comprehensive resource for PODOCVS stakeholders, including healthcare workers, policymakers, and researchers. Although numerous PODOCVS prediction models exist, most

of them have not been developed, validated, and assessed according to established guidelines for predictive research. This has given rise to significant biases in PODOCVS risk estimation, deficiencies in the statistical methods used, and a lack of internal and external validation.

While the clinical outcomes of PODOCVS are heterogeneous, many predictors of poor clinical outcomes have been identified. These include age (≥ 60 years old), sex (male), education level, history of living alone, frailty, body mass index (> 30 kg/m²), high EuroSCORE (≥ 14), disease severity, changes in lifestyle patterns before and after admission, personality traits (high irritability), and history of smoking and alcohol consumption[33-35]. These interrelated factors' relative influence on patient outcomes can be described through multivariable modelling and subsequently used to construct a predictive model to estimate CS patients' risk of developing postoperative delirium[36]. Therefore, to ensure the validity of the included predictive models, we will only select models that are internally or externally validated during development. Studies focused solely on the external validation of existing models or otherwise unvalidated predictive models will be excluded since several existing externally validated models have exhibited significant variability in their predictive accuracy across different populations. The reason for this may be that postoperative delirium in CS patients is closely related to the type of surgery that patients receive[2, 37]. The planned systematic review will concisely summarise key information presented across all identified predictive model studies. CS healthcare providers and policymakers can then use this information to assess a model's applicability to the CS patient population in different scenarios.

In recent years, ML-based predictive models have shown great success[38]. Machine learning models can process vast quantities of multi-dimensional and unstructured data to reveal the roles and relationships between multiple variables, providing an effective means for accurately predicting delirium risk in CS patients[9, 39, 40]. Although ML-based predictive models exhibit good predictive performance, research has shown that they lack explanatory power. The reason for this may be that patients' clinical data are not uniform and standardised, and key data are often omitted[41]. Which of the two model types (i.e., ML-based or statistical models) exhibits superior discrimination and calibration in the prediction

of PODOCVS risk has not been demonstrated. To some extent, this affects the selection and use of predictive models by clinical medical staff. Therefore, this review will also analyze whether ML models can provide more accurate predictions as compared to traditional statistical models.

Strengths of this study include its strict adherence to the updated PRISMA 2020 checklist and its planned adherence to the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist. The utilization of a consensus-based analytical approach fosters collective accountability in interpretative decisions, thereby strengthening the robustness and credibility of the findings.

Limitations

Some limitations of the study warrant consideration. We will only conduct qualitative description and not perform a meta-analysis. We anticipate that inconsistencies in the predictors included in the models, follow-up duration, clinical settings in which the models were developed, types of cardiovascular surgery, and geographic distribution of the study populations may pose challenges to conducting a meta-analysis. Even if each predictive model we included has undergone internal or external validation, its limited clinical application necessitates further evaluation of its generalizability to ensure both applicability and scientific rigor. Although the study employs a comprehensive methodology, one limitation is the omission of a search for grey literature, which could potentially result in the exclusion of studies reported in non-traditional publications. Another limitation of the planned review is that we will not contact study authors to request unreported data, although missing information will be explicitly documented.

Conclusion

We present a protocol for the systematic review of prediction models of postoperative delirium for CS patients. With the aim of identifying, summarising and appraising the available risk models, we hope to provide a current reference to stakeholders engaged in CS patient care, policy and research. In addition, we also hope to explore whether PODOCVS ML models can provide more accurate predictions as compared to PODOCVS traditional

statistical models.

Author Contributions

Conceptualization, Z.X.L., W.Y.K.; methodology, Z.X.L., W.Y.K. and L.L.J.; writing—original draft preparation Z.X.L., W.Y.K.; writing—review and editing, Z.X.L., W.Y.K.; supervision, L.M.J. and H.X.D. All authors have read and agreed to the published version of the manuscript.

Funding

None.

Competing interests

None declared.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication

Not applicable.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data Availability Statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Supplemental material

This content has been supplied by the author(s).

Supplementary Data

Supplementary Material 1

Supplementary Material 2

References

- [1] COHEN C L, ATKINS K J, EVERED L A, et al. Examining Subjective Psychological Experiences of Postoperative Delirium in Older Cardiac Surgery Patients [J]. *Anesth Analg*, 2023, 136(6): 1174-81.
- [2] CAI S, LI J, GAO J, et al. Prediction models for postoperative delirium after cardiac surgery: Systematic review and critical appraisal [J]. *Int J Nurs Stud*, 2022, 136: 104340.
- [3] SUGIMURA Y, SIPAHI N F, MEHDIANI A, et al. Risk and Consequences of Postoperative Delirium in Cardiac Surgery [J]. *Thorac Cardiovasc Surg*, 2020, 68(5): 417-24.
- [4] ZHANG Y, WAN D H, CHEN M, et al. Automated machine learning-based model for the prediction of delirium in patients after surgery for degenerative spinal disease [J]. *CNS Neurosci Ther*, 2023, 29(1): 282-95.
- [5] LIN J L, ZHENG G Z, CHEN L W, et al. A nomogram model for assessing predictors and prognosis of postoperative delirium in patients receiving acute type A aortic dissection surgery [J]. *BMC Cardiovasc Disord*, 2023, 23(1): 72.
- [6] EREMENKO A A, CHEMOVA E V. [Comparison of dexmedetomidine and propofol for short-term sedation in early postoperative period after cardiac surgery] [J]. *Anesteziol Reanimatol*, 2014, (2): 37-41.
- [7] YOKOYAMA C, YOSHITNAI K, OGATA S, et al. Effect of postoperative delirium after cardiovascular surgery on 5-year mortality [J]. *JA Clin Rep*, 2023, 9(1): 66.
- [8] NAGATA C, HATA M, MIYAZAKI Y, et al. Development of postoperative delirium prediction models in patients undergoing cardiovascular surgery using machine learning algorithms [J]. *Scientific reports*, 2023, 13(1): 21090.
- [9] HAN C, KIM H I, SOH S, et al. Machine learning with clinical and intraoperative biosignal data for predicting postoperative delirium after cardiac surgery [J]. *iScience*, 2024, 27(6): 109932.
- [10] JIN Z, HU J, MA D. Postoperative delirium: perioperative assessment, risk reduction, and management [J]. *Br J Anaesth*, 2020, 125(4): 492-504.
- [11] CURTIS M S, FORMAN N A, DONOVAN A L, et al. Postoperative delirium: why, what, and how to confront it at your institution [J]. *Curr Opin Anaesthesiol*, 2020, 33(5): 668-73.
- [12] YANG T, YANG H, LIU Y, et al. Postoperative delirium prediction after cardiac surgery using machine learning models [J]. *Comput Biol Med*, 2024, 169: 107818.
- [13] SADLONOVA M, HANSEN N, ESSELMANN H, et al. Preoperative Delirium Risk Screening in Patients Undergoing a Cardiac Surgery: Results from the Prospective Observational FINDERI Study [J]. *American Journal of Geriatric Psychiatry*, 2024.
- [14] ZHAO X, LI J, XIE X, et al. Online interpretable dynamic prediction models for postoperative delirium after cardiac surgery under cardiopulmonary bypass developed based on machine learning algorithms: A retrospective cohort study [J]. *J Psychosom Res*, 2024, 176: 111553.
- [15] CHOI S G, OH M, PARK D H, et al. Comparisons of the prediction models for undiagnosed diabetes between machine learning versus traditional statistical methods [J]. *Sci Rep*, 2023,

- 13(1): 13101.
- [16] GAO W, ZHANG Y, JIN J. Validation of E-PRE-DELIRIC in cardiac surgical ICU delirium: A retrospective cohort study [J]. *Nursing in critical care*, 2022, 27(2): 233-9.
 - [17] LEE A, MU J L, JOYNT G M, et al. Risk prediction models for delirium in the intensive care unit after cardiac surgery: a systematic review and independent external validation [J]. *Br J Anaesth*, 2017, 118(3): 391-9.
 - [18] MUFTI H N, HIRSCH G M, ABIDI S R, et al. Exploiting Machine Learning Algorithms and Methods for the Prediction of Agitated Delirium After Cardiac Surgery: Models Development and Validation Study [J]. *JMIR Med Inform*, 2019, 7(4): e14993.
 - [19] XUE X, CHEN W, CHEN X. A Novel Radiomics-Based Machine Learning Framework for Prediction of Acute Kidney Injury-Related Delirium in Patients Who Underwent Cardiovascular Surgery [J]. *Comput Math Methods Med*, 2022, 2022: 4242069.
 - [20] KANWAR M K, KILIC A, MEHRA M R. Machine learning, artificial intelligence and mechanical circulatory support: A primer for clinicians [J]. *J Heart Lung Transplant*, 2021, 40(6): 414-25.
 - [21] CHRISTODOULOU E, MA J, COLLINS G S, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models [J]. *J Clin Epidemiol*, 2019, 110: 12-22.
 - [22] LINDROTH H, BRATZKE L, PURVIS S, et al. Systematic review of prediction models for delirium in the older adult inpatient [J]. *BMJ Open*, 2018, 8(4): e019223.
 - [23] Jia Y Y, HU H T, HU J N, et al. A systematic review of prognostic prediction models for patients with chronic heart failure in China [J]. *Clinical Journal of Thoracic and Cardiovascular Surgery*: 1-10.
 - [24] MOHER D, SHAMSEER L, CLARKE M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement [J]. *Syst Rev*, 2015, 4(1): 1.
 - [25] MOONS K G, DE GROOT J A, BOUWMEESTER W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist [J]. *PLoS Med*, 2014, 11(10): e1001744.
 - [26] DEBRAY T P, DAMEN J A, SNELL K I, et al. A guide to systematic review and meta-analysis of prediction model performance [J]. *Bmj*, 2017, 356: i6460.
 - [27] WOLFF R F, MOONS K G M, RILEY R D, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies [J]. *Ann Intern Med*, 2019, 170(1): 51-8.
 - [28] DAMEN J A A, MOONS K G M, VAN SMEDEN M, et al. How to conduct a systematic review and meta-analysis of prognostic model studies [J]. *Clin Microbiol Infect*, 2023, 29(4): 434-40.
 - [29] PLASCHKE K, FICHTENKAMM P, SCHRAMM C, et al. Early postoperative delirium after open-heart cardiac surgery is associated with decreased bispectral EEG and increased cortisol and interleukin-6 [J]. *Intensive Care Med*, 2010, 36(12): 2081-9.
 - [30] SCHOEN J, MEYERROSE J, PAARMANN H, et al. Preoperative regional cerebral oxygen saturation is a predictor of postoperative delirium in on-pump cardiac surgery patients: a prospective observational trial [J]. *Crit Care*, 2011, 15(5): R218.
 - [31] BROWN C H. Delirium in the cardiac surgical ICU [J]. *Curr Opin Anaesthesiol*, 2014, 27(2): 117-22.
 - [32] Wang Y, Zhang C, WU Q C. Incidence and modifiable risk factors of postoperative delirium in

- patients undergoing cardiac surgery: a prospective observational study based on propensity score matching [J]. Chinese PLA Medical Journal: 1-12.
- [33] HUANG W, WU Q, ZHANG Y, et al. Development and validation of a nomogram to predict postoperative delirium in type B aortic dissection patients underwent thoracic endovascular aortic repair [J]. Front Surg, 2022, 9: 986185.
- [34] GREAVES D, PSALTIS P J, DAVIS D H J, et al. Risk Factors for Delirium and Cognitive Decline Following Coronary Artery Bypass Grafting Surgery: A Systematic Review and Meta-Analysis [J]. J Am Heart Assoc, 2020, 9(22): e017275.
- [35] TAN C, SAITO N, MIYAWAKI I, et al. Preoperative circadian physical activity rhythm and postoperative delirium in cardiovascular surgery patients [J]. Chronobiol Int, 2020, 37(7): 1059-66.
- [36] WILSON J, CHOWDHURY F, HASSAN S, et al. Prognostic prediction models for clinical outcomes in patients diagnosed with visceral leishmaniasis: protocol for a systematic review [J]. BMJ Open, 2023, 13(10): e075597.
- [37] Chinese International Exchange and Promotion Association for Medical and Health Care Branch of Cardiac Critical Care. Chinese expert consensus on the prevention and treatment of delirium after cardiac and macrovascular surgery [J]. Chin J Med, 2023, 103(45): 3635-44.
- [38] SULIMAN A, MASUD M, SERHANI M A, et al. Predictive performance of machine learning compared to statistical methods in time-to-event analysis of cardiovascular disease: a systematic review protocol [J]. BMJ Open, 2024, 14(4): e082654.
- [39] LUO W, PHUNG D, TRAN T, et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View [J]. J Med Internet Res, 2016, 18(12): e323.
- [40] ZHAO X, LI J, XIE X, et al. Online interpretable dynamic prediction models for postoperative delirium after cardiac surgery under cardiopulmonary bypass developed based on machine learning algorithms: A retrospective cohort study [J]. Journal of Psychosomatic Research, 2024, 176.
- [41] WANG S, ZHU X. Predictive Modeling of Hospital Readmission: Challenges and Solutions [J]. IEEE/ACM Trans Comput Biol Bioinform, 2022, 19(5): 2975-95.

Supplementary Files

Modification notes.

URL: <http://asset.jmir.pub/assets/d7f6e65b797ae1c806745c948cf205c1.docx>

protocol (revised version with tracked changes) - round 2.

URL: <http://asset.jmir.pub/assets/43b142359c910288b8447f07fce5ffae.docx>

Multimedia Appendixes

Untitled.

URL: <http://asset.jmir.pub/assets/00d203f5d3d1a6c5bf9dbe8946a62ace.zip>