Review Article

The Efficacy of Artificial Intelligence in Predicting the Postoperative Mortality Rate in Patients with Congenital Heart Disease: A Systematic Review and Meta-Analysis

Lies Dina Liastuti [1,2]*, Syafarrani Arma Bahirah [2], Averina Geffanie Suwana [2]

[1] Cipto Mangunkusumo National Central Public Hospital, Jakarta, Indonesia.
[2] Faculty of Medicine, University of Indonesia, Jakarta, Indonesia.

Citation: Liastuti LD, Bahirah SA, Suwana AG. The Efficacy of Artificial Intelligence in Predicting the Postoperative Mortality Rate in Patients with Congenital Heart Disease: A Systematic Review and Meta-Analysis. Res Heart Yield Transl Med 2025; 20(3):231-241.

https://doi.org/10.18502/jthc.v20i3.20115

Highlights

- Artificial intelligence (AI) models, especially Extreme Gradient Boosting (XGBoost) and Gradient Boosting Machine (GBM), achieved the highest predictive performance for postoperative mortality in congenital heart disease (CHD) patients, with pooled AUC values of 0.93 and 0.91, respectively.
- The meta-analysis demonstrated that AI-assisted prediction tools provide high specificity (0.96) and robust accuracy compared to conventional risk stratification systems.
- Integrating AI into clinical workflows can enhance perioperative decision-making and improve patient outcomes in CHD, though further multicenter validation is needed for broader implementation.

Article info:

Received: 9 Mar. 2025
Revised: 7 Apr. 2025
Accepted: 16 Apr. 2025

ABSTRACT

Background: Congenital heart disease (CHD) is a leading cause of morbidity and mortality in children requiring surgical intervention. Accurate prediction of postoperative mortality remains challenging because of the limitations of traditional risk stratification systems. Artificial intelligence (AI) has emerged as a promising tool for enhancing predictive accuracy in this field.

Objective: This systematic review and meta-analysis aimed to evaluate the efficacy of AI in predicting postoperative mortality in patients with CHD.

Methods: Following the PRISMA guidelines, we systematically searched four databases for relevant studies published up to July 16, 2024. Studies with retrospective, prospective, or cross-sectional designs that evaluated AI-based models for predicting mortality after CHD surgery were eligible for inclusion. Data were extracted, and study quality was assessed using the PROBAST tool. Pooled estimates for sensitivity, specificity, and the area under the curve (AUC) were calculated.

Results: Six studies involving 42,536 patients and evaluating 11 distinct AI models were included. The meta-analysis yielded a pooled AUC of 0.90 (95% CI, 0.88 to 0.93), with a pooled sensitivity of 0.43 (95% CI, 0.23 to 0.65) and a pooled specificity of 0.96 (95% CI, 0.92 to 0.98). Subgroup analysis revealed that the Extreme Gradient Boosting (AUC, 0.93) and Gradient Boosting Machine (AUC, 0.91) models had the highest predictive performance. All included studies were judged to have a low risk of bias.

Conclusion: The Extreme Gradient Boosting and Gradient Boosting Machine models demonstrate high specificity and promising accuracy for predicting postoperative mortality in patients with CHD, outperforming traditional scoring systems. Further multicenter, prospective studies are needed to enhance generalizability and support clinical implementation.

*** Corresponding Author:**
Lies Dina Liastuti
Cipto Mangunkusumo National Central Public Hospital- Faculty of Medicine, University of Indonesia 71 Pangeran Diponegoro Street, Central Jakarta, Special Capital Region of Jakarta 10430.
Tel: +62 813 811 97007
E-mail: dr.liesdina@gmail.com

Keywords: Congenital Heart Disease; Artificial Intelligence; Mortality Prediction; Postoperative Outcomes

Background

Congenital heart disease (CHD) is among the most common congenital anomalies, occurring in approximately 4 to 10 per 1000 live births.[1] Despite advancements in diagnostics, surgical techniques, and perioperative care, children with CHD who undergo anatomical corrective procedures continue to experience high rates of morbidity and mortality.[2] Surgery remains the cornerstone of CHD management, as untreated CHD is associated with substantial mortality. In developed countries, surgical interventions have markedly improved patient outcomes and reduced mortality rates.[3] Nonetheless, nearly 20% of pediatric patients with CHD are readmitted within 30 days after surgery, and 4.2% of those undergoing surgical procedures do not survive.[4] In addition, early mortality rates after neonatal cardiac surgery are estimated at approximately 10%.[5]

The risk of mortality in patients with CHD is closely linked to the complexity of surgical procedures. Accurate prediction of in-hospital mortality is critical for supporting clinical decision-making, optimizing procedural strategies, and improving patient outcomes. Several major risk stratification systems are currently used to predict mortality and morbidity in pediatric CHD surgery, including Risk Adjustment for Congenital Heart Surgery (RACHS-1), Aristotle Basic Complexity, Aristotle Comprehensive Complexity, and Society of Thoracic Surgeons–European Association for Cardio-Thoracic Surgery (STS-EACTS) Congenital Heart Surgery Mortality Categories. These frameworks are largely based on estimated procedural risks or complexities and rely heavily on expert opinion and consensus.[3,6] Nevertheless, these traditional tools primarily categorize surgical procedures and often fail to account for comprehensive individual patient risk factors, which may limit their predictive accuracy for specific cases. Accordingly, combining multiple clinical features to determine prognosis is essential. Developing a predictive model that integrates diverse and clinically relevant parameters holds substantial value.[3]

Machine learning (ML), an emerging technology, has gained momentum because of its ability to model complex nonlinear relationships, particularly when analyzing large datasets. Moreover, ML addresses common challenges in clinical research, including small sample sizes, risk of bias, insufficiently detailed descriptions of treatments and patient characteristics, missing data, and the lack of calibrated models.[1] Several studies have demonstrated that ML-assisted tools outperform traditional scoring systems. For example, Zeng et al[7] reported that an Extreme Gradient Boosting (XGBoost) model provided superior predictive performance for postoperative complications compared with conventional risk adjustment models in pediatric cardiac surgery.

The objective of this systematic review and meta-analysis is to synthesize recent evidence on the use of artificial intelligence (AI) in predicting mortality among patients with CHD following surgery.

Methods

Search Strategy and Selection Criteria

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A systematic search was conducted in four databases (PubMed, Cochrane, Embase, and Scopus) through July 16, 2024. Relevant studies identified by manual searching were also included. Two researchers independently reviewed titles and abstracts and selected pertinent citations for full-text review. Because this study involved only the retrieval and synthesis of data from published research, ethical approval was not required. The detailed search strategies for each database are presented in (Table 1).

Table 1. Search results from PubMed, Cochrane, Embase, Scopus, and ScienceDirect databases.

PubMed	
((("Artificial Intelligence"[Mesh]) AND "Mortality"[Mesh]) AND "Heart Defects, Congenital"[Mesh]) AND "surgery" [Subheading]	4
Cochrane	
Artificial Intelligence OR Machine Intelligence OR Machine Learning OR Computer Reasoning AND mortality OR death AND Congenital Heart Disease OR Congenital Heart Defect OR Malformation Of Heart AND surgery	51
Scopus	
artificial AND intelligence OR machine AND intelligence OR machine AND learning OR computer AND reasoning AND mortality OR death AND congenital AND heart AND disease OR congenital AND heart AND defect OR malformation AND of AND heart AND surgery OR operative AND procedures OR operations	50
Embase	
(('artificial intelligence':ab,ti OR 'machine learning':ab,ti OR 'information processing':ab,ti) AND 'mortality rate':ab,ti AND 'congenital heart disease':ab,ti OR 'congenital heart malformation':ab,ti) AND surgery:ab,ti	69
ScienceDirect	
((Artificial Intelligence OR Machine Learning OR Computer Reasoning) AND (mortality OR death) AND (Congenital Heart Disease OR Congenital Heart Defect OR Malformation of Heart) AND (surgery)	307

Study Eligibility

The Population, Intervention, Comparison, and Outcome (PICO) framework was used to guide screening and interpretation. The population of interest included postoperative patients with CHD. The intervention was the application of AI techniques, including but not limited to ML and deep learning. Comparators included traditional risk-scoring methods or no comparator. The primary outcome was mortality.

Studies were excluded if they were not retrospective, prospective, or cross-sectional in design; if they were published in languages other than English; or if they were unrelated to the research topic.

Data Extraction and Quality Assessment

A single reviewer independently extracted data on study characteristics and diagnostic outcomes using a standardized form. The extracted data included author names, publication year, study design, number of patients, population characteristics, AI method, and key results, including the area under the curve (AUC). The risk of bias was assessed for each study using the Prediction model Risk of Bias Assessment Tool (PROBAST), which is designed for reviews of AI diagnostic test accuracy. Four domains—participants, predictors, outcomes, and analysis—were evaluated for risk of bias and applicability concerns. A domain was categorized as having a low risk of bias if most signaling questions were answered “yes.” A study was considered to have an overall low risk of bias only if all domains were rated as low risk. Studies with a high risk of bias in one or more domains were considered to be at high risk of bias, and those with at least one unclear domain (while all others were low risk) were categorized as having an unclear risk of bias.

Statistical Analysis

We aggregated the numbers of true positives, false positives, true negatives, and false negatives to calculate pooled sensitivity and specificity for predicting postoperative mortality in patients with CHD. A meta-analysis was performed using MetaDTA, STATA version 17, and R Studio to generate forest plots and summary receiver operating characteristic curves. A subgroup analysis was conducted to compare the performance of different AI algorithms that were used in three or more studies.

Results

The detailed search strategies for each database are presented in (Table 1) of the supplementary material. The database search yielded 481 records. After screening titles and abstracts, 25 articles were selected for full-text evaluation. Of these, six studies met the eligibility criteria and were included in the final analysis. The PRISMA flow diagram is shown in (Figure 1).

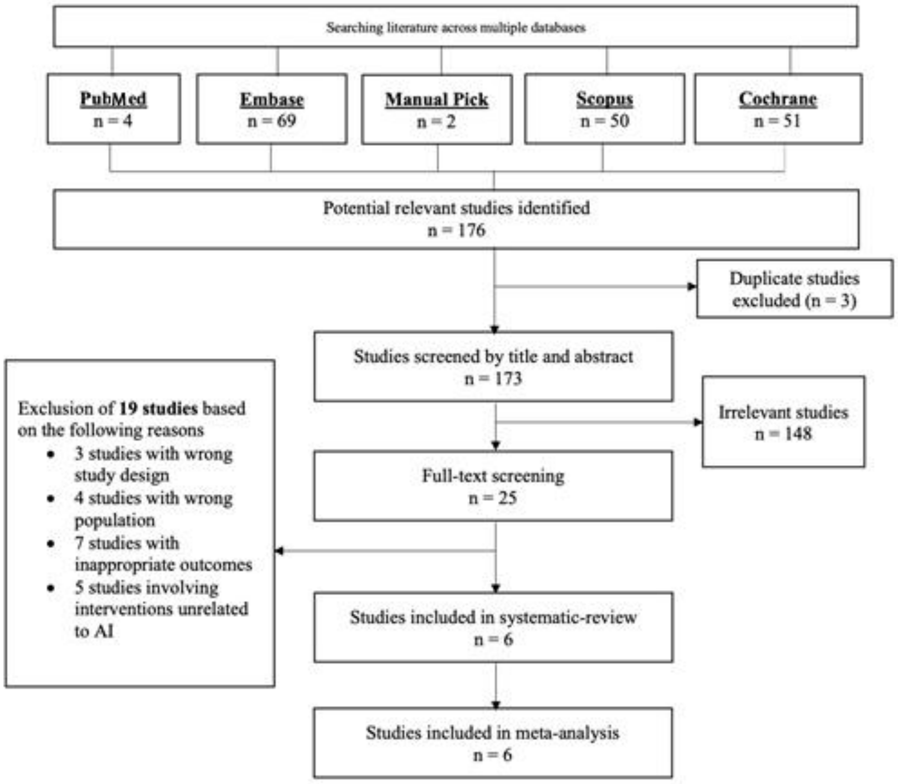

Figure 1. The image depicts the flow diagram of the study selection.

The Prediction Model Risk of Bias Assessment Tool (PROBAST) was utilized to evaluate risk of bias in the included studies. As depicted in (Table 2), all studies were judged to have a low risk of bias.

Table 2. Summary of the risk of bias and applicability concerns

Studies	Domain 1	Domain 2	Domain 3	Domain 4	Overall
Zurn C, 2023[8]	+	+	+	+	+
Du X, 2022[3]	+	+	-	+	-
Weiss AJ, 2023[9]	+	+	+	+	+
Zhou Y, 2021[10]	+	+	+	+	+
Jalali A, 2020[11]	+	-	+	-	-
Tong C, 2024[1]	+	+	+	-	-

Judgement

X High risk

\+ Low risk

\- Some concerns

Domains:

D1: Bias due to participants

D2: Bias due to predictors

D3: Bias due to outcomes

D4: Bias due to analyses

The meta-analysis included five retrospective studies and one prospective study. Of these, three were conducted in China, two in the United States, and one in Germany. All studies involved patients with postoperative CHD. Each study reported AUC as an outcome measure for the AI models assessed. (Table 3) summarizes the study and patient characteristics, with additional details provided in the supplementary material.

Table 3. Summary of the studies included in the meta-analysis

No.	Study	Sample Size	Population	AI model	Outcome
1	Zurn C, 2023[8]	1456 patients	Patients from the departments of pediatric cardiology and cardiac surgery who underwent congenital heart surgery	. Logistic Regression . Random Forest	. AUC: 0.9486 . Sensitivity: 0.85 . Specificity: 0.8948
2	Du X, 2022[3]	24685 patients	Patients aged 0–18 years who were diagnosed with CHD and underwent CHD surgery	. Extreme Gradient Boosting (XGBoost)	. AUC: 0.874 (0.848-0.901) . Sensitivity: 0.751 . Specificity: 0.879
3	Weiss AJ, 2023[9]	6392 patients	Patients who underwent cardiac surgeries, including CHD surgery	. Extreme Gradient Boosting (XGBoost)	. AUC: 0.978 (0.964-0.989) . Recall: 0.795 (0.686-0.900) . Precision: 0.756 (0.646-0.861)
4	Zhou Y, 2021[10]	381 patients	Patients of orthotopic heart transplantation, including patients with CHD	. Random Forest . AdaBoost . Logistic Regression . Support Vector Machine . Extreme Gradient Boosting (XGBoost) Gradient Boost Machine . Artificial Neural Network . Naive	. Random Forest: AUC: 0.801 (0.697-0.891); Sensitivity: 0.268 (0.059-0.529); Specificity: 0.927 (0.866-0.976) . AdaBoost: AUC: 0.641 (0.476-0.788); Sensitivity: 0.260 (0.059-0.500); Specificity: 0.894 (0.821-0.954) . Logistic Regression: . AUC: 0.688 (0.549-0.816); Sensitivity: 0.201 (0.000-0.429); Specificity: 0.917 (0.851-0.966) . Support Vector Machine: AUC: 0.714 (0.574-0.834); Sensitivity: 0.000 (0.000-0.000); Specificity: 1.000 (1.000-1.000) . Extreme Gradient Boosting: AUC: 0.769 (0.662-0.869); Sensitivity: 0.138 (0.000-0.353); Specificity: 0.953 (0.902-0.989) . Gradient Boosting Machine: AUC: 0.819 (0.737-0.889); Sensitivity: 0.271 (0.077-0.533); Specificity: 0.916 (0.845-0.966) . Artificial Neural Network: AUC: 0.755 (0.639-0.851); Sensitivity: 0.066 (0.000-0.214); Specificity: 0.988 (0.962-1.000) . Naive: AUC: 0.500 (0.500-0.500); Sensitivity: 0.000 (0.000-0.000); Specificity: 1.000 (1.000-1.000)

5	Jalali A, 2020[11]	549 patients	Newborns with single ventricle physiology who underwent either an MBTS or an RV-to-PA shunt during the Norwood procedure	. Deep Neural Network . Gradient Boosting Machine . Random Forest . Decision Tree	. Deep Neural Network: AUC: 0.95; Precision: 0.94; Recall: 0.86 . Gradient Boosting Machine: AUC: 0.90; Precision: 0.87; Recall: 0.78 . Random Forest: AUC: 0.84; Precision: 0.71; Recall: 0.27 . Decision Tree: AUC: 0.55; Precision: 0.43; Recall: 0.10
6	Tong C, 2024[1]	9073 patients	Pediatric patients who underwent congenital heart surgery	. Light Gradient Boosting Machine . Logistic Regression . Support Vector Machine . Random Forest . CatBoost	. Light Gradient Boosting Machine: AUC: 0.893 (0.884-0.895); Sensitivity: 0.763; Specificity: 0.871 . Logistic Regression: AUC: 0.887 (0.879-0.901); Sensitivity of 0.773; Specificity of 0.855 . Support Vector Machine: AUC: 0.883 (0.876-0.886); Sensitivity: 0.878; Specificity: 0.708 . Random Forest: AUC: 0.890 (0.874-0.906); Sensitivity: 0.758; Specificity: 0.866 . CatBoost: AUC: 0.892 (0.876-0.893); Sensitivity: 0.761; Specificity: 0.866

CHD: congenital heart disease; AUC: area under the curve

The AUC for postsurgical mortality among patients with CHD was reported across six studies, involving 11 AI models, including Random Forest, Logistic Regression, Extreme Gradient Boosting (XGBoost), Gradient Boosting Machine, Support Vector Machine, AdaBoost, Artificial Neural Network, Naive Bayes, Deep Neural Network, Decision Tree, and CatBoost. The pooled AUC was 0.90 (95% CI, 0.88 to 0.93) (Figure 2).

The pooled sensitivity of ML models for predicting postoperative mortality among patients with CHD was 0.43 (95% CI, 0.23 to 0.65), and the pooled specificity was 0.96 (95% CI, 0.92 to 0.98) (Figure 3).

A subgroup analysis was conducted to compare the performance of different ML models in predicting mortality among patients with CHD after surgery. Subgroup analyses were performed only for AI models included in more than one study. Extreme Gradient Boosting and Gradient Boosting Machine showed the best performance, with area under the AUC values of 0.93 (95% CI, 0.45 to 0.98) and 0.91 (95% CI, 0.55 to 0.98), respectively (Table 4).

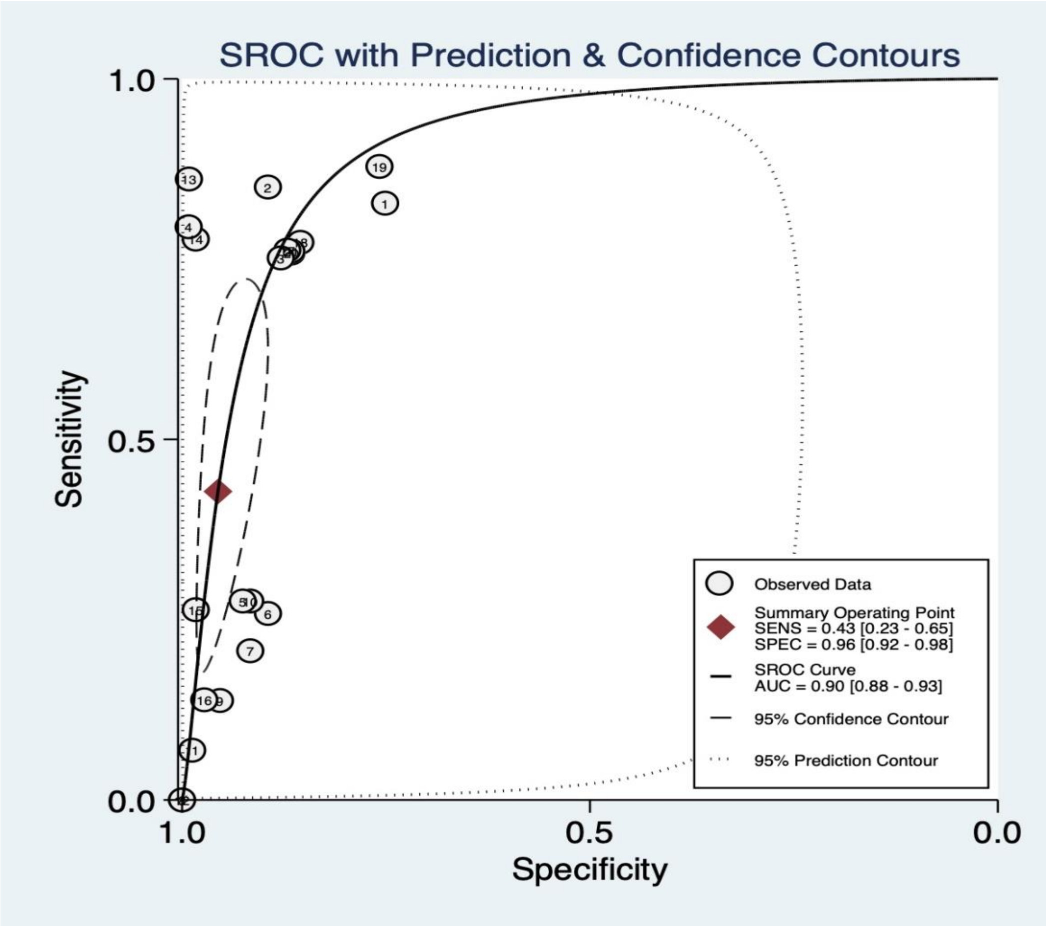

Figure 2. The summary receiver operating characteristic (SROC) curve shows the performance of artificial intelligence (AI) in predicting postsurgical mortality among patients with congenital heart disease (CHD).

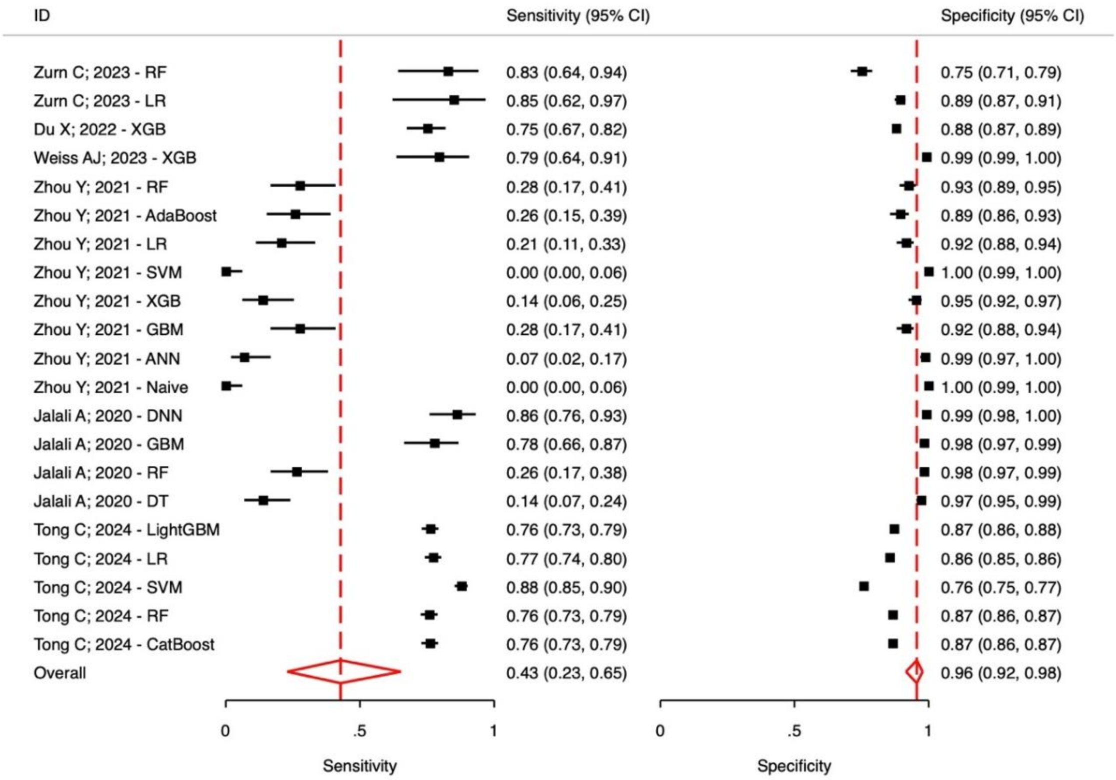

Figure 3. The image illustrates the forest plots of the pooled sensitivity and specificity for the performance of artificial intelligence (AI) in predicting mortality among CHD patients postoperatively.

Table 4. Subgroup analysis of each artificial intelligence (AI) type in predicting postoperative mortality rate among congenital heart disease (CHD) patients

Subgroup	Study	Sensitivity	Specificity	PLR	NLR	DOR	AUC
All combined	6	0.43 (0.23-0.65)	0.96 (0.92-0.98)	9.7 (5.8-16.3)	0.60 (0.41-0.87)	16 (8.0-33.0)	0.90 (0.88-0.93)
Random Forest	4	0.55 (0.27-0.80)	0.91 (0.79-0.97)	6.5 (3.8-11.1)	0.49 (0.27-0.88)	13 (7.0-24.0)	0.87 (0.83-0.89)
Logistic Regression	3	0.61 (0.21-0.91)	0.889 (0.85-0.92)	5.22 (2.3-676)	0.46 (0.11-0.87)	16.90 (2.66-57.80)	0.89 (0.78-0.92)
Extreme Gradient Boosting (XGBoost)	3	0.56 (0.15-0.90)	0.96 (0.84-0.99)	21.6 (1.72-88.0)	0.481 (0.47-0.93)	89.80 (1.94-533.0)	0.93 (0.45-0.98)
Gradient Boosting Machine	3	0.62 (0.29-0.87)	0.94 (0.82-0.98)	13.0 (2.26-41.1)	0.42 (0.13-0.8)	47.6 (3.0-226.0)	0.91 (0.55-0.98)
Support Vector Machine	2	0.21 (0.00-0.99)	0.98 (0.18-1.00)	6.06 (1.2-10.1)	0.65 (0.03-1.00)	14.70 (2.8-45.9)	0.885 (0.21-0.98)

Discussion

The present meta-analysis identified 6 studies that examined the role of AI in predicting postoperative mortality in patients with CHD. These studies included a total of 42,536 patients who underwent surgery and evaluated 11 distinct AI models. The pooled AUC was 0.90 (95% CI, 0.88 to 0.93), with a pooled sensitivity of 0.43 (95% CI, 0.23 to 0.65) and a pooled specificity of 0.96 (95% CI, 0.92 to 0.98). Subgroup analyses revealed that the XGBoost and Gradient Boosting Machine models demonstrated the best overall performance for predicting postoperative mortality.

The emergence of ML in health care has transformed approaches to predicting and managing outcomes in CHD, including postoperative mortality.[12] Studies have shown that ML can outperform existing risk-scoring systems.[3] Current consensus-based evaluation methods, such as the Society of Thoracic Surgeons–European Association for Cardio-Thoracic Surgery (STAT) score, are static and do not account for perioperative or postoperative changes in a patient's condition. In contrast, ML models can integrate dynamic perioperative data, which enables more accurate risk prediction and supports clinical decision-making. The early identification of high-risk patients by ML may facilitate targeted preventive measures or more intensive postoperative monitoring. For instance, research by Zürn et al[8] found that combining the STAT score with postoperative markers significantly improved survival prediction, particularly in the first 24 hours after surgery. Similarly, Du et al[3] reported that the XGBoost model generated more accurate predictions than conventional models, such as the STS-EACTS or the RACHS-1, both of which rely primarily on procedural complexity rather than individual patient characteristics. Weiss et al[9] further demonstrated that STS scoring may not adequately account for certain or combined procedures, whereas ML approaches can incorporate these complexities into mortality risk prediction.

Another challenge in CHD management is patient heterogeneity, as individuals with the same diagnosis can have highly variable clinical manifestations. ML algorithms address this by integrating multidimensional data, including demographic information, echocardiographic findings, laboratory results, and postoperative markers, into predictive models. Further, ML can assign variable importance to different predictors, which helps clinicians prioritize modifiable risk factors. By way of example, Du et al[3] found that low oxygen saturation, the need for mechanical ventilation, and unplanned reintervention were key predictors of mortality. In another study, Zürn et al[8] identified elevated postoperative serum lactate level as a strong prognostic factor, reporting that each 1-

mmol/L increase was associated with a 2.16-fold higher mortality risk (OR, 2.16). Age also remains a critical determinant. As highlighted by Yeh et al,[14] mortality risk is highest within the first 5 years of life and declines thereafter. These findings underscore the clinical utility of ML for refining perioperative risk assessment.

Another strength of ML is its reproducibility, providing consistent predictions independent of a physician's expertise or training. This capability reduces interobserver variability, which is an inherent limitation of subjective clinical judgment. It is important to note that ML is not intended to replace clinical assessment but to complement it by offering an additional layer of decision support. Moreover, ML models can be integrated with continuously updated electronic health record (EHR) data, allowing for real-time adjustment of risk predictions as a patient's condition evolves.[8,9]

ML methods can automatically identify and select key feature combinations, enabling efficient and accurate assessment and prediction of disease progression. This capability provides improved scientific insight for clinical practice.[12] For instance, Du et al[3] found that preoperative oxygen saturation had the greatest impact on ML model performance for predicting mortality. The same study identified preoperative mechanical ventilation and unplanned cardiac reintervention in neonates with CHD as additional strong predictors of postoperative mortality. In another study, Zürn et al[8] identified serum lactate level as a strong predictor of postoperative mortality in patients with CHD. The authors reported that for each 1-mmol/L increase in the average serum lactate level during the first 24 hours after surgery, the mortality risk increased by 2.16-fold (OR, 2.16).

Age is another significant risk factor for mortality in this population. Research by Yeh et al[13] demonstrated that most deaths among patients with CHD occur within the first 5 years of life. This finding is supported by their subsequent 2015 study, which reported a declining trend in mortality risk with increasing age.[14] In this meta-analysis, subgroup analysis demonstrated that the XGBoost and Gradient Boosting Machine models had the best performance, achieving AUC values of 0.93 (95% CI, 0.45 to 0.98) and 0.91 (95% CI, 0.55 to 0.98), respectively. The XGBoost algorithm is an ensemble method based on classification and regression trees that is recognized for its computational efficiency and scalability. It also effectively handles missing data, which makes it suitable for identifying complex relationships between predictor variables and clinical outcomes. For instance, Li et al[15] reported that the XGBoost model successfully ranked 20 critical predictors from a set of 44 variables. These characteristics make XGBoost well-suited for clinical applications compared with other ML methods, although it remains susceptible to overfitting when applied to datasets with a limited number of features.

The Gradient Boosting Machine model benefits from its sequential learning structure, which improves model calibration and discrimination compared with other algorithms, such as Random Forest. Kong et al[16] emphasized that such ensemble methods generally outperform parametric approaches like logistic regression when applied to large, complex datasets. However, the accuracy of an ML model is highly dependent on its development process. For instance, Allyn et al[9] developed a model using data from a single institution and restricted their feature set to those in the EuroSCORE, an approach that risks omitting important multimodal and institution-specific predictors. Ideally, ML development should incorporate a comprehensive, multimodal feature set. Furthermore, while using hospital- and patient-specific data enables more accurate, personalized risk assessment and can identify unique center-level patterns, models trained on such specific data may lack generalizability and not be replicable across different institutions.[9]

The current systematic review and meta-analysis provides further insight into the role of AI in predicting postoperative mortality in patients with CHD. To our knowledge, it is the first to perform a subgroup analysis comparing the performance of different AI algorithms. Several limitations must be acknowledged. First, there was variability in the outcome measures across the included studies, with reported mortality outcomes ranging from 30-day to 1-year mortality. Second, the study population was restricted to patients with CHD who had undergone surgery, and we did not apply criteria based on the specific type of surgery performed, which resulted in a heterogeneous surgical population. These limitations were a direct consequence of the small number of relevant studies available for inclusion in this meta-analysis.

Conclusion

This systematic review and meta-analysis highlight the promising role of AI, particularly Extreme Gradient Boosting and Gradient Boosting Machine models, in predicting postoperative mortality among patients with CHD. With a pooled AUC of 0.90 and a high specificity of 0.96, these models demonstrate potential as reliable adjuncts to existing clinical risk stratification methods. Unlike traditional scoring systems, AI models provide individualized predictions by integrating a broader range of clinical variables.

Several limitations should be noted. The included studies differed in outcome measures (in-hospital, 30-day, and 1-year mortality), introducing methodological heterogeneity. The relatively small number of studies and the lack of multicenter prospective validation underscore the need for further research. Future investigations should focus on integrating AI into clinical workflows to enhance perioperative decision-making and patient outcomes, particularly in high-risk CHD populations.

Declarations:

Ethical Approval

This study was a systematic review and meta-analysis based solely on previously published data; no human participants or patient records were directly involved.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Acknowledgment

The authors would like to thank the Faculty of Medicine, Universitas Indonesia, and Cipto Mangunkusumo National Central Public Hospital (RSCM) for institutional support during manuscript preparation. Appreciation is also extended to all researchers whose studies were included in this meta-analysis.

References

1. Tong C, Du X, Chen Y, Zhang K, Shan M, Shen Z, Zhang H, Zheng J. Machine learning prediction model of major adverse outcomes after pediatric congenital heart surgery: a retrospective cohort study. Int J Surg. 2024 Apr 1;110(4):2207-2216. doi: 10.1097/JS9.0000000000001112.
2. Jacobs JP, Mayer JE, Mavroudis C, et al. The Society of Thoracic Surgeons Congenital Heart Surgery Database: 2016 Update on Outcomes and Quality. Ann Thorac Surg. 2016;101(3):850-862. doi:10.1016/J.ATHORACSUR.2016.01.057
3. Du X, Wang H, Wang S, He Y, Zheng J, Zhang H, et al. Machine Learning Model for Predicting Risk of In-Hospital Mortality after Surgery in Congenital Heart Disease Patients. Rev Cardiovasc Med. 2022;23(11):376. doi: 10.31083/j.rcm2311376.
4. Feudtner C, Pati S, Goodman DM, Kahn MG, Sharma V, Hutto JH, et al, Hall M, Shah SS. State-level child health system performance and the likelihood of readmission to children's hospitals. J Pediatr. 2010 Jul;157(1):98-102.e1.
 doi: 10.1016/j.jpeds.2010.01.049.
5. Triedman JK, Newburger JW. Trends in Congenital Heart Disease: The Next Decade. Circulation. 2016;133(25):2716-2733. doi:10.1161/CIRCULATIONAHA.116.023544
6. Hörer J, Kasnar-Samprec J, Cleuziou J, et al. Mortality Following Congenital Heart Surgery in Adults Can Be Predicted Accurately by Combining Expert-Based and Evidence-Based Pediatric Risk Scores. World J Pediatr Congenit Heart Surg. 2016;7(4):425-435. doi:10.1177/2150135116656001
7. Zeng X, An J, Lin R, et al. Prediction of complications after paediatric cardiac surgery. Eur J Cardiothorac Surg. 2020;57(2):350-358. doi:10.1093/EJCTS/EZZ198
8. Zürn C, Hübner D, Ziesenitz VC, et al. Model-driven survival prediction after congenital heart surgery. Interdiscip Cardiovasc Thorac Surg. 2023;37(3):89. doi:10.1093/ICVTS/IVAD089
9. Weiss AJ, Yadaw AS, Meretzky DL, et al. Machine learning using institution-specific multi-modal electronic health records improves mortality risk prediction for cardiac surgery patients. JTCVS Open. 2023;14:214-251. doi:10.1016/j.xjon.2023.03.010
10. Zhou Y, Chen S, Rao Z, et al. Prediction of 1-year mortality after heart transplantation using machine learning approaches: A single-center study from China. Int J Cardiol. 2021;339:21-27. doi:10.1016/j.ijcard.2021.07.024

11. Jalali A, Lonsdale H, Do N, Peck J, Gupta M, Kutty S, et al. Deep Learning for Improved Risk Prediction in Surgical Outcomes. Sci Rep. 2020 Jun 9;10(1):9289. doi: 10.1038/s41598-020-62971-3.
12. Zhu X, Xie B, Chen Y, Zeng H, Hu J. Machine learning in the prediction of in-hospital mortality in patients with first acute myocardial infarction. Clinica Chimica Acta. 2024;554:117776. doi:10.1016/J.CCA.2024.117776
13. Yeh SJ, Chen HC, Lu CW, et al. Prevalence, mortality, and the disease burden of pediatric congenital heart disease in Taiwan. Pediatr Neonatol. 2013;54(2):113-118. doi:10.1016/J.PEDNEO.2012.11.010
14. Yeh SJ, Chen HC, Lu CW, et al. National database study of survival of pediatric congenital heart disease patients in Taiwan. J Formos Med Assoc. 2015;114(2):159-163. doi:10.1016/J.JFMA.2012.10.006
15. Li R, Shen L, Ma W, Yan B, Chen W, Zhu J, et al. Use of machine learning models to predict in-hospital mortality in patients with acute coronary syndrome. Clin Cardiol. 2023 Feb;46(2):184-194. doi: 10.1002/clc.23957.
16. Kong G, Lin K, Hu Y. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU. BMC Med Inform Decis Mak. 2020;20(1):251. doi:10.1186/s12911-020-01271-2