ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20230326000480
Subject review

A Review on Machine Learning Applications: CVI Risk Assessment

Ayşe Banu BİRLİK*, Hakan TOZAN, Kevser Banu KÖSE

Abstract: Comprehensive literature has been published on the development of digital health applications using machine learning methods in cardiovascular surgery. Many machine learning methods have been applied in clinical decision-making processes, particularly for risk estimation models. This review of the literature shares an update on machine learning applications for cardiovascular intervention (CVI) risk assessment. This study selected peer-reviewed scientific publications providing sufficient detail about machine learning methods and outcomes predicting short-term CVI risk in cardiac surgery. Thirteen articles fulfilling pre-set criteria were reviewed and tables were created presenting the relevant characteristics of the studies. The review demonstrates the usefulness of machine learning methods in high-risk CVI applications, identifies the need for improvement, and provides efficient support for future prediction models for the healthcare system.

Keywords: cardiovascular; decision-making; machine learning; prediction model; risk assessment

1 INTRODUCTION

In our technology-driven world, developments in digital transformation practices in the health field have rapidly accelerated. Artificial Intelligence (AI) is integral to this. AI refers to information technologies or models that can mimic human intelligence. Machine learning (ML) is a sophisticated subset of AI that employs a data-driven approach to extract profound insights from cumulative data, learning data behavior through an algorithmic framework. The presence of big data has accelerated data mining and expanded ML development into decision-making processes in research areas that include engineering, finance, management, and medicine [1-4]. AI and ML applications in health science can be applied in such sub-activity subjects as medical diagnosis and follow-up, cost estimation, resource planning, and emergency strategy management.

Health systems have expanding literature on machine learning-based algorithms and their potential clinical benefits in the diagnosis and treatment of diseases. Results analysis using clinical data efficiently is of great importance. ML methods for advancing fields such as outcome prediction, diagnosis, medical image interpretation, and treatment in medical science, utilizing large databases accumulated over time, are being developed [5-8]. Although ML use has begun in different segments of health systems to estimate by displaying significant information and estimations from data accumulation, few studies yet exist on its use in high-risk surgical cardiovascular interventions (CVI) [9].

Cardiovascular disease (CVD) affects the heart or blood vessels. Coronary artery disease (CAD), such as angina and myocardial infarction, is a type of CVD usually caused by atherosclerosis, an accumulation of plaque inside the artery walls. If the accumulation causes a blockage that narrows these vessels, a decrease in the blood flow which supplies the heart could result in heart attack or paralysis [10]. CAD is one of the most common causes of death in the world, and has been reported to significantly increase overall health care costs [11].

Coronary artery bypass grafting (CABG) is a surgical procedure to treat coronary artery disease that redirects blood around a section of a blocked or partially blocked artery in the heart. CABG is a high-cost, high-risk surgery with a death rate of approximately 3-5% [12] and high morbidity and loss of life (LoL) risks associated with intraoperative and postoperative complications. Still, CABG is the "gold standard" treatment for multi-core coronary artery disease, especially for three-vein or left main coronary artery disease, and remains the most commonly applied heart surgery globally [11, 12]. Therefore, the risk analysis and the careful selection of CABG candidates is essential for optimizing the procedure's results.

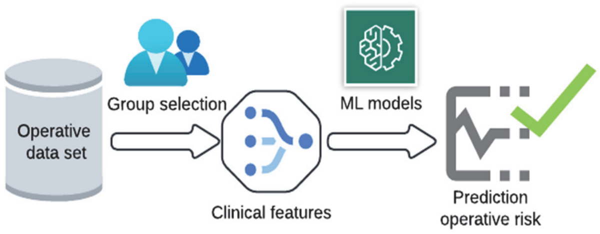

Figure 1 Basic workflow example of operative outcome prediction

Various risk classification models, such as the European Cardiac Operational Risk Assessment System (EuroSCORE) [15, 16] and Society of Thoracic Surgeons (STS) score [17], are outcome prediction models developed to support clinical decision-making following cardiac surgery. The Euro SCORE is used to assess operative risk, predicting in-hospital LoL following CVI. The risk scoring systems such as EuroSCORE tend to overestimate real risk, which may adversely affect clinical decision-making, leading to unsuccessful interventions and providing unrealistic information to patients and their families [18, 19]. Risk prediction models play an important role in the decision-making process and so demand the evaluation of new methods with different and more complex mathematical assumptions than previous approaches [9]. Understanding operational risk before CABG surgery allows both surgeons and patients to participate effectively in treatment decisions [20].

Careful risk-benefit analysis and patient selection optimize surgical results for adult patients with many comorbidities, and the evaluation of operative risk in CVI is increasingly emphasized. In addition to clinical assessments, surgeons will have the potential to decide on the best treatment plan for a particular patient, considering the risk rating generated by a prediction model. It has been reported that effective cost policies and interventions are urgently needed to achieve the world's sustainable development goals and achieve a 30% reduction in early deaths unrelated to infectious disease [21]. CVD is the leading cause of death in the World Health Organization's European Region, killing more than four million people a year and more than 1.9 million people in the European Union (EU); it will cost the EU economy an estimated 196 billion euros annually [22].

1.1 Related Work

Cardiovascular research throughout Europe is progressing at an ever-increasing rate. It must be supported with consistent, reliable, and regular funds to ensure that research and clinical studies can continue. The British Heart Foundation (BHF) provides about 100 million pounds for new research annually, making it the largest provider of cardiovascular research funds in Europe outside the EU [22]. Aware of the urgent need for an alternative risk estimation model, the Cardiothoracic Surgery Association (SCTS) has approved the "Risk Prediction in Adult Cardiac Surgery in the UK" project conducted by the Alan Turing Institute. Funded by the British Heart Foundation, this project uses machine learning for accurate risk estimation for patients who have undergone open-heart surgery and is [23]. Machine learning and artificial intelligence for risk modeling in heart surgery led researcher Arman Kilic at the Medical University of South Carolina's Surgery Innovation Center to develop predictive models that can be used to support individualized patient care and clinical decision-making [9, 24-26]. The Thoracic Surgery Foundation Research Award supported Kilic's work on the predictive benefit of the machine learning algorithm in estimating the risk of LoL in heart surgery [25]. These events emphasize the importance of developing risk prediction models in invasive cardiac surgery in scientific studies.

A meta-analysis comparing the accuracy of discriminating ML models vs LR in predicting operational LoL following heart surgery revealed that ML models were more accurate in predicting operative LoL [27]. Scoping research was conducted to evaluate the extent and potential limitations of the use of ML in perioperative anesthetic therapy, particularly in heart surgery patients. ML-based prediction models behaved similarly to classic statistical approaches [28]. In addition to their research to evaluate the predictive efficacy of ML methods, Penny-Dimri et al. [29] reviewed studies investigating adverse events as the primary outcome, such as acute kidney injury, prolonged mechanical ventilation, and LoL. They presented a systematic analysis distinguishing between in-hospital and 30-day LoL.

There will be ongoing exploration of machine learning approaches to data modeling that will feature larger analyses and methods consisting of different algorithms. Themethodological development of models must inform future work in the development of risk estimation tools following CVI. Consequently, this has motivated a comprehensive examination of the methodologies behind operative risk prediction models to facilitate the widespread adoption of risk assessment tools in clinical practice. This study explores the benefits reported from the improved models and the areas needing improvement. For this reason, this review of the literature determines the scope and characteristics of the existing machine learning practices used to predict operative risk after CABG. Unlike related studies, this review focused on the kind of operation, and papers assessing the operational LoL of CABG patients were reviewed.

Our study concludes it is necessary to conduct additional research to find the optimal algorithms for a variety of prediction tasks. To minimize inefficiencies in research requires careful planning and detailed records of the methodologies used in model development and validation.

The second section remaining in this document describes our methodology, search strategy, and study selection. The third section presents a comprehensive examination of the models and the findings. The fourth section summarizes study conclusions.

2 METHODOLOGIES

2.1 Literature Analysis

The literature research was compiled until 01 November 2022 using PubMed, Web of Science, ProQuest, and IEEE Xplore databases. The search was made in English without date and geography restrictions. Search terms for keywords include "machine learning", "risk estimation", "risk modeling", "mortality", "operative mortality", "coronary artery bypass grafting", "cardiovascular", "cardiovascular intervention", and "heart surgery" or any combination of these. To estimate mortality after CABG surgery, articles that do not focus on machine learning-based studies have been excluded. Of the full-text articles evaluated for eligibility, 13 publications have been included in the review.

The extraction items list was based on the CHARMS checklist and was revised following lengthy debate among the authors [30-33]. Relevant features were extracted for each article: type of cardiac surgery; age; sex (%); source of data; the size of datasets; operative loss of life (%); type of ML methods; several input variables; the processing of missing data; validation model; the split ratio of training to testing sample; type of statistical analysis; the area under the receiver operating characteristic curve (AUROC). In this study, AUROC, a generalization known as the C-index, was utilized to give insight into the reciprocal interactions of multiple models. In this review study, it is considered that $\geq 0.7 - < 0.8 = acceptable$ value is satisfactory for the interpretation of AUROC [28].

3 RESULTS AND DISCUSSION

Most risk-scoring systems were developed using a biostatistical approach based on linear relationship assumptions and a generalized linear model [34], and risk classification scores were developed from coefficients from the Logistic Regression (LR) analysis of a dataset [35]. Because of its flexible interpretation structure and low computational complexity, LR modeling is an accepted approach to risk prediction models [36]. For logistic regression, one of the generalized linear models, the dependent variable is a linear combination of independent variables applied to a nonlinear link function [37].

Table 1 Study features

Study	Surgery	Age	Source of data	Postoperative Survival Data / %
Lippmann & Shahian [43]	CABG	not reported	*STS	96.6
Tu et al. [44]	CABG	not reported	Retrospective	97.0
Chong et al. [46]	On-pump CABG	64.0 (10.4) training set 63.3 (9.7) testing set	*Medical Health Record	92.5
Ghavidel et al. [49]	On-pump CABG	45-60 (58.1%); 61-75 (39.5%) > 76 (2.4%); 63.4	*Medical Health Record	96.2
Nouei et al. [50]	CABG combined with other surgical procedure	58.62 (10.18) survival group 61.82 (10.72) death group	*Medical Health Record	96.7
Mendes et al. [37]	On-pump or off-pump CABG	60.4 (9.6) training set 61.1 (9.8) testing set	Prospective	91.4
Jamaati et al. [51]	CABG	57 all patients	Prospective	87.8
Nouei et al. [52]	combined with other surgical procedures were excluded CABG	58.24 (9.74) survival group 62.07 (9.47) death group	*Medical Health Record	96.5
Geltser et al. [41]	CABG	63 all patients	Retrospective	95.0
Huang et al. [42]	CABG	71.27 (4.78) survival group 74.30 (5.60) death group	* NHIRD	39.1
Mori et al. [53]	CABG	65.3 (10.2) all patients	*STS Adult Cardiac Surgery Database	98.1
Khalaji et al. [54]	CABG	67.34 (9.67) all patients	Retrospective	97.2
Zea-Vera et al. [55]	CABG	65 all patients	*Local STS Adult Cardiac Surgery Database	98.0

CABG, Coronary Artery Bypass Grafting; NHIRD, National Health Insurance Research Database; STS, Society of Thoracic Surgeons; ACSD, Adult Cardiac Surgery Database; * Retrospective

Table 2 Study methodological features

Study	No. input variables	Processing of missing data	Validation model	Split rate train:test	Statistical analysis
Lippmann & Shahian [43]	33 categorical variables 3 continuous variables	mean substitution or statistical mode	Data splitting k-fold cross-validation	50:50	Chi square test
Tu et al. [44]	17 categorical variables	Not reported	Data splitting and k-fold cross-validation	65:35	not reported
Chong et al. [46]	14 categorical variables 4 continuous variables	Filling in the data (categorical; coded as missing continuous; mean substitution)	Data splitting and k-fold cross-validation	75:25	Root mean square error
Ghavidel et al. [49]	40 categorical variables 4 continuous variables	Missing excluded from the model	Data splitting and k-fold cross-validation	70:30	Chi square test Student t test
Nouei et al. [50]	40	Missing excluded from the model	Data splitting	70:30	Pearson Chi-square test Student t test
Mendes et al. [37]	4 categorical variables 8 continuous variables	Not reported	Data splitting	80:20	not reported
Jamaati et al. [51]	17	Not reported	Not reported	Not reported	Hosmer-Lemeshow goodness-of-fit statistic
Nouei et al. [52]	40	Missing excluded from the model	Data splitting	70:30	Pearson Chi-square test Student t test
Geltser et al. [41]	5 categorical variables 2 continuous variables	Not reported	Data splitting	75:25	Mann-Whitney test Chi-square analysis Fisher's exact test
Huang et al. [42]	72	Not reported	Cross-validation	Not reported	Pearson Chi-square analysis
Mori et al. [53]	23 categorical variables 3 continuous variables	Filling in the data (categorical; lowest risk value, continuous; average assignment)	Data splitting	70:30	The sum of mean squared error
Khalaji et al. [54]	11	Missing excluded from the model	Data splitting and k-fold cross-validation	70:30	Pearson Chi-square test Student t test Fisher's exact test Independent sample t-test
Zea-Vera et al. [55]	80	Not reported	Validation Split with Bootstrapping	80:20	DeLong's test Bootstrapping

Table 3 Model performance features

Study	Model type	*AUROC
Lippmann & Shahian [43]	ANN Naïve Bayes LR	0.76 0.75 0.76
Tu et al. [44]	ANN LR	0.78 0.77
Chong et al. [46]	ANN LR	0.89 0.81
Ghavidel et al.[49]	EEFDT EECDT LR EuroSCORE	0.90 0.86 0.78 0.77
Nouei et al. [50]	LGFAS MLP(ANN) LR	0.91 0.70 0.72
Mendes et al. [37]	ANN LR	0.85 0.86
Jamaati et al. [51]	SVM LR EuroSCORE	0.98 0.84 0.72
Nouei et al. [52]	ANFIS LR	0.82 0.62
Geltser et al. [41]	ANN RF LR	0.85 0.71 0.83
Huang et al. [42]	RF MARS CART XGBoost LR	0.71 0.71 0.69 0.72 0.72
Mori et al. [53]	XGBoost LR	0.79 0.79
Khalaji et al. [54]	RF Naïve Bayes SVM XGBoost KNN LR	0.78 0.78 0.74 0.79 0.72 0.81
Zea-Vera et al. [55]	XGBoost	0.77

ANFIS; Adaptive Neuro-Fuzzy Inference System; ANN, Artificial Neural Networks; AUROC, Area Under the Receiver Operating Characteristic Curve; CART, Classification And Regression Trees; EECDT, Entropy Error Crisp Decision Tree; EEFDT, Entropy Error Fuzzy Decision Tree; EuroSCORE, European System for Cardiac Operative Risk Evaluation; KNN, K-Nearest Neighbors; LGFAS, Lookup Genetic Fuzzy Annealing System; LR, Logistic Regression; MARS, Multivariate Adaptive Regression Splines; MLP, Multilayer Perceptron; NB, Naïve Bayesian; RF, Random Forest; SVM, Support Vector Machine; XGBoost, eXtreme Gradient Boosting. *belong to the mortality risk classification obtained using preoperative variables

Such models can oversimplify complex relationships with nonlinear interactions that involve numerous risk factors, leading to missed non-linear relationships and misidentification of the model during the development of points [38]. Advanced statistical classifiers have been suggested to overcome the constraints of LR analysis and increase the effectiveness of risk prediction models [39]. Pre-surgical risk models are designed to make use of more data while making fewer suppositions about clinical data status [35]. Previous studies in cardiothoracic surgery have developed ML algorithms that estimate in-hospital LoL after cardiac operations better than standard operating risk scores [40-42]. However, some researchers have reported that the probabilities figured by a trained neural network are no worse than the regression model's predictions [43, 44].

When properly trained, artificial neural networks (ANN) can be used as predictive models of classification. ANNs are data processing systems inspired by the biological nervous system, consisting of nonlinear computational elements [45]. They work in a nonlinear way that can better define the interaction between health risk factors and are used in many research subjects, such as speech recognition, prediction of protein structure, converting text to speech, and forecasting stock prices. There is also potential use for them in the classification of cancer disease, diagnostic estimation, electrocardiogram interpretation, and use of various health systems such as CVI in medical decision-making processes [46-48].

Differences in variable types, handling of missing data fields, operational procedure selection criteria, and geographic area differences can all have an impact on the accuracy of estimation of different risk models [34]. The applications of ANNs are particularly prominent in classification models due to their advantages, such as not requiring prior assumptions about underlying frequency distributions, their ability to model non-linear relationships, and their robustness to missing data and feature errors [41, 46]. The study features corresponding to each article are given in Tab. 1 and Tab. 2. In addition, the AUROC values of the models are presented in Tab. 3.

The AUROC values for ML models range from 0.69 to 0.98 and 0.62 to 0.86 for LR, and the number of predictors enclosed in the models ranges from 7 to 72. For the verification method, both sample-splitting and k-fold cross were performed [46, 49]. Statistical analysis methods have been reported in 8 studies. In the data pre-processing, the missing data are excluded or the data is filled with the mean imputation method. Fig. 2 presents the number of patients included in the studies and the number of in patients. The sample size was 563 to 378,572 patients, and the operative LoL ranged from 2.0% to 60.9% depending on the sample size.

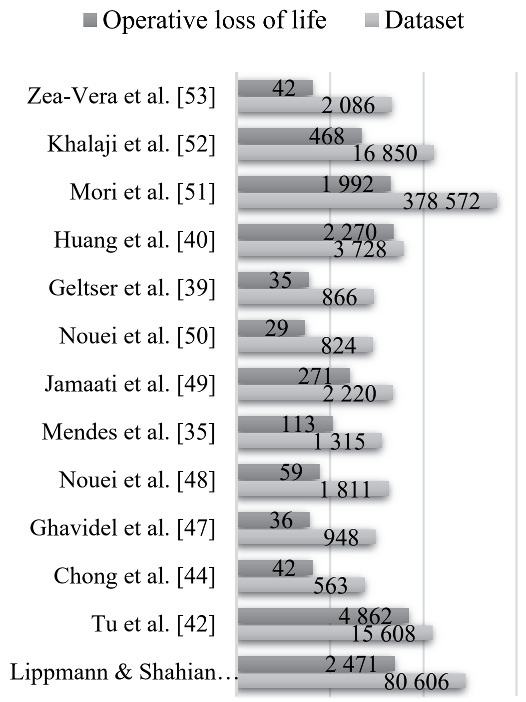

Figure 2 Sample size and loss of life numbers used in the modeling

Mori et al. [53] used the largest dataset with the number of patients meeting the criteria through the STS ACSD database. In Fig. 3 the gender ratios among the number of patients included in the studies are given.

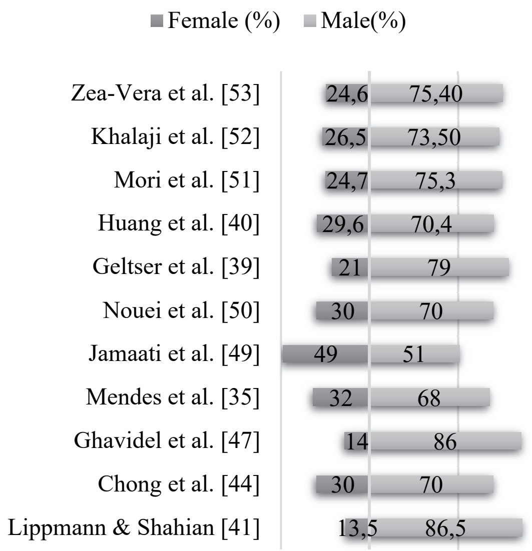

Figure 3 Percentage values of the sample size by gender

Tu et al. [44] improved ANN and LR models to predict the risk of in-hospital LoL following CABG surgery, reporting that ANNs offer a new perspective on statistical conduct and that the relationships between patient attributes and LoL are similar in ANNs and LR models. Different predictive variables were included in this study in addition to those used in prior studies.

Jamaati et al. [51] reported that the best result was the SVM model with an AUROC value of 0.98, with 17 preoperative variables that they used to predict LoL risk in their study. They reported that the SVM and LR models showed good calibration results. No validation method was included in their study.

Chong et al. [46] developed an ANN model using 18 preoperative variables to predict 7 major adverse outcomes, including LoL, in 563 patients with CABG. Although they determined 21 variables for the ML and LR models in their study, they excluded 3 variables from the model because of an insufficient number of patients with liver disease and episodes of prior endotracheal intubation or prior cardiopulmonary resuscitation (CPR). The ANN model performed better than the LR model. It has been reported that ANN's tend to perform better and can be considered a useful instrument for assigning patients to low-risk groups.

Huang et al. [42] reported that the XGBoost model had the best predictive ability without variable selection. Using different feature selection methods, they investigated which variables would affect the LoL risk and evaluated the results after feature selection for each classification model. They concluded that to be innovative and more practical in clinical trials, developing a survival model to predict LoL in adult patients who have had their first CABG yielded better prediction when limited to10 variables.

Mendes et al. [37] compared the ANN and LR models for prognostic assessment of the need for reintubation, prolonged mechanical ventilation, and LoL after CABG surgery. They used operative variables such as cardiopulmonary bypass and total graft count in addition to preoperative variables in the models.

Ghavidel et al. [49] showed that the EEFDT (19 variables) and EECDT (15 variables) models they developed to estimate the risk of LoL in CABG surgery outperformed the LR analysis (10 variables) and EuroSCORE (17 variables). They reported that there was no statistically significant difference in the discriminative ability of the EEFDT and EECDT (P = .066) models. Even though the database they used belonged to a referral institution for central cardiac surgery, they recommended using an independent test dataset to analyze the model's performance. The AUROC values that resulted from the LoL in Tab. 3 showed minor differences between models. Nouei et al. [50, 52] compared the results of the multilayer perceptron (MLP) neural network, a class of feedforward neural networks, LR analysis and LGFAS model. Considering the area under the curve, it was reported that the LGFAS model was better in estimating the risk of death compared to LR and ANNS. The superior accuracy of their model can be explained by fuzzy set theory and the many fuzzy concepts it contains which, by nature, are uncertain.They obtained promising results in their studies but stipulated that other aspects of the findings should be considered more.Mori et al. [53] researched the effect of the model on performance by including intraoperative features in the present preoperative model for 7 different outcomes after CABG surgery. To test the LR and XGBoost models, they used only preoperative, only intraoperative, or pre+intraoperative features as variables. They reported that models utilizing pre+intraoperative variables had better AUROC values in all outcomes than models with only intraoperative or only preoperative variables.

Geltser et al. [41] identified 7 risk factors with great potential for in-hospital LoL prediction outcomes, including LR, RF, and ANN models. Model validations were performed using the AUROC, sensitivity, and specificity parameters. They reported that the ANN model including RWT (relative wall thickness) and LVRMI (LV relative mass index) predictors showed maximum prognostic accuracy.

Khalaji et al. [54] evaluated the ability of ML models to predict LoL after CABG. They tested six different ML models after identifying the important variables using the RF technique and claimed that the LR model had the best one-year LoL prediction ability (AUC = 0.81), with perfect discrimination. Furthermore, all ML models, except for LR, performed well in predicting one-year LoL, according to the AUC interpretation (0.7 < AUC < 0.8). Other investigations utilizing LR, RF, and XGBoost models have shown the impact of mechanical ventilation duration in CABG LoL.

Zea-Vera et al. [55] created and evaluated a decision tree-based machine learning system to predict mortality, significant morbidity, high total hospitalization costs, and 30-day readmission outcomes following isolated CABG. Preoperative ML models outperformed the standard STS model in predicting LoL or significant morbidity in patients undergoing isolated CABG, mostly by incorporating intraoperative data like as cross-clamp and bypass times as additive predictive variables.

The calibration of risk estimates, that is, whether they are reliable or not, was rarely evaluated in studies. Of the 13 studies, 3 considered the evaluation of model performance in terms of calibration for the accuracy of risk estimates [43, 51, 53]. More calibration procedures are required to identify situations in which modern methods outperform traditional approaches. Model validation processes are frequently weak or poorly described, preventing fair model comparison in clinical evaluations [37, 46].

We have encountered instances that deviate from our assumptions prior to the study. The consideration of large amounts of data sets that have not yet been structured has been raised as another research topic. While diversity is expected in prediction variables, perioperative variables are focused on except [53] and [55] studies. In addition, the new models developed in the studies were not analyzed using an independent database in different centers. The evaluation of variables in terms of their clinical impact for upcoming studies will be beneficial for improving model performance.

4 CONCLUSIONS

ML methods are a new trend in healthcare. Predictive analytical solutions build models using data mining techniques. Increasingly, risk estimation modeling studies for multidisciplinary decision-making processes used in the clinical field have attracted a lot of attention. It has been proposed that comparing new studies with the performances of existing models will improve estimation results through the formation of large datasets. Estimating operative risk in adult CVI is critical for decision-making processes, in the clinic and in institutional planning and health system management based on outcome assessments. This literature review highlighted research results of predictive models for CABG and suggested the creation of local databases and increasing data accumulation to improve it. Comparing the existing estimation models provides up-to-date information about research on ML-based methods to predict operative risk after CABG.

The application of ML models to forecast outcomes following CABG surgery represents an innovative and pragmatic approach in clinical research. Even though current findings show that ML models produce better results, optimal methods to improve prediction performance are still being researched. Different sets of variables used in ML algorithms show important or statistically significant patient characteristics under investigation. The prediction performance of the models changes when the number of variables available before surgery is reduced or when intraoperative variables are added. This modification emphasizes the potential value of patient characteristics in model improvement.

The application of more complex mathematical models capable of capturing nonlinear relationships among variables predicting operative risk has the potential to enhance the accuracy of risk models. Additionally, databases must include information on patients who have been rejected for surgery. Developing descriptive characteristics that include a CABG surgical patient at a higher risk of LoL will be a guide for future studies.

Quality issues in clinical registry data limit model performance improvement. A risk model that updates the predicted risk for CVI may be possible with appropriate data integration. Risk reclassification may improve the model's predictive performance, particularly for patients at extremely high risk. Comparison results indicate there are still research points in areas with dynamic data flows in the current estimation methods. This points to the fact that ML applications need to be accelerated to improve not only risk models in adult CVI but also sufficient data extraction. Even though ML use is not yet widespread in clinical practice, data mining and ML techniques promise to monitor and improve the quality of CVI outcomes. Unlike traditional statistics, ML focuses on developing automated clinical decision systems that improve estimate accuracy, replacing conventional systems used. More successful predictive algorithms may be developed by collecting more data for machine learning-based clinical decision process analysis.

The development of ML algorithms in high-risk subjects of the health system will facilitate digital health breakthroughs with a revolutionary impact on health policy management. Algorithms that make estimations and reveal hidden and significant information in data will catalyze developments in health informatics.

Several limitations of this review should be noted. Despite conducting a systematic search, we may have overlooked articles due to the wide range of topics and abbreviations in the literature. This may have resulted in the incorrect exclusion of studies from the initial selection. Furthermore, because we only summarized the data descriptively without conducting meta-analysis, no definitive conclusions about the AUROC values between the different methodologies can be drawn. Still, this article provides an overview of current ML applications for risk estimation of LoL following CABG surgery, pointing to areas where additional research is required.

We plan future studies to investigate the influence of predictive variables on the performance of models and to compare the accuracy of prediction models developed with different methodologies. In addition, we will analyze the results of this literature review with our development of a machine learning-based hybrid model; the model is adapted to regional reality for short-term survival risk assessment based on preoperative clinical and laboratory data in patients undergoing CABG.

5 REFERENCES

[1] Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & De Felice, F. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability (Switzerland)*, *12*(2). https://doi.org/10.3390/su12020492

[2] Tozan, H. & Donmez, S. (2015). A Genetic Algorithm Based Approach to Provide Solutions for Emergency Aid Stations Location Problem and a Case Study for Pendik/Istanbul. *Journal of Homeland Security and Emergency Management*, *12*(4), 915-940. https://doi.org/10.1515/jhsem-2015-0025

[3] Öztürk, N. & Tozan, H. (2015). A fuzzy-based decision support model for selecting the best dialyser flux in haemodialysis. *Journal of Healthcare Engineering*, *6*(3), 303-324. https://doi.org/10.1260/2040-2295.6.3.303

[4] Öztürk, N., Tozan, H., & Vayvay, Ö. (2020). A new decision model approach for health technology assessment and a case study for dialysis alternatives in Turkey. *International Journal of Environmental Research and Public Health*, *17*(10). https://doi.org/10.3390/ijerph17103608

[5] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, *380*(14), 1347-1358. https://doi.org/10.1056/nejmra1814259

[6] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230-243. https://doi.org/10.1136/svn-2017-000101

[7] Çelik Ertuğrul, D. & Abdullah, S. A. H. M. E. D. (2022). A Decision-Making Tool for Early Detection of Breast Cancer on Mammographic Images. *Tehnicki Vjesnik*, *29*(5), 1528-1536. https://doi.org/10.17559/TV-20211221131838

[8] Sun, Y. L. & Zhang, D. L. (2019). Machine learning techniques for screening and diagnosis of diabetes: A survey. *Tehnicki Vjesnik*, *26*(3), 872-880. https://doi.org/10.17559/TV-20190421122826

[9] Kilic, A., Habib, R. H., Miller, J. K., Shahian, D. M., Dearani, J. A., & Dubrawski, A. W. (2021). Supplementing Existing Societal Risk Models for Surgical Aortic Valve Replacement With Machine Learning for Improved Prediction. *Journal of the American Heart Association*, *10*(22). https://doi.org/10.1161/jaha.120.019697

[10] Thomas, H., Diamond, J., Vieco, A., Chaudhuri, S., Shinnar, E., Cromer, S., Perel, P., Mensah, G. A., Narula, J., Johnson, C. O., Roth, G. A., & Moran, A. E. (2018). Global Atlas of Cardiovascular Disease 2000-2016: The Path to Prevention and Control. *Global Heart*, *13*(3), 143-163. https://doi.org/10.1016/j.gheart.2018.09.511

[11] Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., Aboyans, V., Adetokunboh, O., Ärnlöv, J., Afshin, A., Agrawal, A., Kiadaliri, A. A., Ahmadi, A., Ahmed, M. B., Aichour, A. N., Aichour, I., Aichour, M. T. E., Aiyar, S., Al-Eyadhy, A., … Murray, C. J. L. (2017). Global, regional, and national age-sex specifc mortality for 264 causes of death, 1980-2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1151-1210. https://doi.org/10.1016/S0140-6736(17)32152-9

[12] Lee, T. S., Li, S. J., Jiang, Y., Shia, B. C., & Chen, M. (2020). Cost analysis of coronary artery bypass grafting surgery under single-payer reimbursement in Taiwan. *International Journal of Applied Science and Engineering*, *17*(4), 419-428. https://doi.org/10.6703/IJASE.202012_17(4).419

[13] Matkovic, M., Tutus, V., Bilbija, I., Milin Lazovic, J., Savic, M., Cubrilo, M., Aleksic, N., Atanasijevic, I., Andrijasevic, V., & Putnik, S. (2019). Long Term Outcomes of The Off-Pump and On-Pump Coronary Artery Bypass Grafting in a High-Volume Center. *Scientific Reports*, *9*(1), 1-6. https://doi.org/10.1038/s41598-019-45093-3

[14] Melly, L., Torregrossa, G., Lee, T., Jansens, J. L., & Puskas, J. D. (2018). Fifty years of coronary artery bypass grafting. *Journal of Thoracic Disease*, *10*(3), 1960-1967. https://doi.org/10.21037/jtd.2018.02.43

[15] Nashef, S. A. M., Roques, F., Sharples, L. D., Nilsson, J., Smith, C., Goldstone, A. R., & Lockowandt, U. (2012). Euroscore II. *European Journal of Cardio-Thoracic Surgery*, *41*(4), 734-745. https://doi.org/10.1093/ejcts/ezs043

[16] Nashef, S. A. M., Roques, F., Michel, P., Gauducheau, E., Lemeshow, S., & Salamon, R. (1999). European system for cardiac operative risk evaluation (EuroSCORE). *European Journal of Cardio-Thoracic Surgery*, *16*(1), 9-13. https://doi.org/10.1016/S1010-7940(99)00134-7

[17] Edwards, F. H., Clark, R. E., & Schwartz, M. (1994). Coronary artery bypass grafting: The Society of Thoracic Surgeons National Database experience. *The Annals of Thoracic Surgery*, *57*(1), 12-19. https://doi.org/10.1016/0003-4975(94)90358-1

[18] Ad, N., Holmes, S. D., Patel, J., Pritchard, G., Shuman, D. J., & Halpin, L. (2016). Comparison of EuroSCORE II, Original EuroSCORE, and The Society of Thoracic Surgeons Risk Score in Cardiac Surgery Patients. *Annals of Thoracic Surgery*, *102*(2), 573-579. https://doi.org/10.1016/j.athoracsur.2016.01.105

[19] Gummert, J. F., Funkat, A., Osswald, B., Beckmann, A., Schiller, W., Krian, A., Beyersdorf, F., Haverich, A., & Cremer, J. (2009). EuroSCORE overestimates the risk of cardiac surgery: Results from the national registry of the german society of thoracic and cardiovascular surgery. *Clinical Research in Cardiology*, *98*(6), 363-369. https://doi.org/10.1007/s00392-009-0010-8

[20] Karim, M. N., Reid, C. M., Cochrane, A., Tran, L., Alramadan, M., Hossain, M. N., & Billah, B. (2017). Mortality risk prediction models for coronary artery bypass graft surgery: Current scenario and future direction. *Journal of Cardiovascular Surgery*, *58*(6), 931-942. https://doi.org/10.23736/S0021-9509.17.09965-7

[21] Roth, G. A., Mensah, G. A., & Fuster, V. (2020). The global burden of cardiovascular diseases and risks: A compass for global action. *Journal of the American College of Cardiology*, *76*(25), 2980-2981. https://doi.org/10.1016/j.jacc.2020.11.021

[22] Nicholls, M. (2019). Funding cardiovascular research in Europe. *European Heart Journal*, *40*(2), 80-82. https://doi.org/10.1093/eurheartj/ehy817

[23] *The Alan Turing Institue*. (2022). https://www.turing.ac.uk/research/research-projects/risk-prediction-adult-cardiac-surgery

[24] Kilic, A. (2020). Artificial Intelligence and Machine Learning in Cardiovascular Health Care. *Annals of Thoracic Surgery*, *109*(5), 1323-1329. https://doi.org/10.1016/j.athoracsur.2019.09.042

[25] Kilic, A., Goyal, A., Miller, J. K., Gjekmarkaj, E., Tam, W. L., Gleason, T. G., Sultan, I., & Dubrawksi, A. (2020).

Predictive Utility of a Machine Learning Algorithm in Estimating Mortality Risk in Cardiac Surgery. *Annals of Thoracic Surgery*, *109*(6), 1811-1819. https://doi.org/10.1016/j.athoracsur.2019.09.049

[26] Kilic, A., Goyal, A., Miller, J. K., Gleason, T. G., & Dubrawksi, A. (2021). Performance of a Machine Learning Algorithm in Predicting Outcomes of Aortic Valve Replacement. *Annals of Thoracic Surgery*, *111*(2), 503-510. https://doi.org/10.1016/j.athoracsur.2020.05.107

[27] Benedetto, U., Dimagli, A., Sinha, S., Cocomello, L., Gibbison, B., Caputo, M., Gaunt, T., Lyon, M., Holmes, C., & Angelini, G. D. (2020). Machine learning improves mortality risk prediction after cardiac surgery: Systematic review and meta-analysis. *Journal of Thoracic and Cardiovascular Surgery*. https://doi.org/10.1016/j.jtcvs.2020.07.105

[28] Rellum, S. R., Schuurmans, J., van der Ven, W. H., Eberl, S., Driessen, A. H. G., Vlaar, A. P. J., & Veelo, D. P. (2021). Machine learning methods for perioperative anesthetic management in cardiac surgery patients: a scoping review. *Journal of Thoracic Disease*, *13*(12), 6976-6993. https://doi.org/10.21037/jtd-21-765

[29] Penny-Dimri, J. C., Bergmeir, C., Perry, L., Hayes, L., Bellomo, R., & Smith, J. A. (2022). Machine learning to predict adverse outcomes after cardiac surgery: A systematic review and meta-analysis. *Journal of Cardiac Surgery*, *37*(11), 3838-3845. https://doi.org/10.1111/jocs.16842

[30] Di Tanna, G. L., Wirtz, H., Burrows, K. L., & Globe, G. (2020). Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PLoS ONE*, *15*(7 July), 1-23. https://doi.org/10.1371/journal.pone.0235970

[31] Bazoukis, G., Stavrakis, S., Zhou, J., Bollepalli, S. C., Tse, G., Zhang, Q., Singh, J. P., & Armoundas, A. A. (2021). Machine learning versus conventional clinical methods in guiding management of heart failure patients - a systematic review. *Heart Failure Reviews*, *26*(1), 23-34. https://doi.org/10.1007/s10741-020-10007-3

[32] Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, *110*, 12-22. https://doi.org/10.1016/j.jclinepi.2019.02.004

[33] Moons, K. G. M., de Groot, J. A. H., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., Reitsma, J. B., & Collins, G. S. (2014). Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. *PLoS Medicine*, *11*(10). https://doi.org/10.1371/journal.pmed.1001744

[34] Nilsson, J., Ohlsson, M., Thulin, L., Höglund, P., Nashef, S. A. M., & Brandt, J. (2006). Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks. *Journal of Thoracic and Cardiovascular Surgery*, *132*(1). https://doi.org/10.1016/j.jtcvs.2005.12.055

[35] Penny-Dimri, J. C., Bergmeir, C., Reid, C. M., Williams-Spence, J., Cochrane, A. D., & Smith, J. A. (2021). Machine Learning Algorithms for Predicting and Risk Profiling of Cardiac Surgery-Associated Acute Kidney Injury. *Seminars in Thoracic and Cardiovascular Surgery*, *33*(3), 735-745. https://doi.org/10.1053/j.semtcvs.2020.09.028

[36] Sajid, M. R., Almehmadi, B. A., Sami, W., Alzahrani, M. K., Muhammad, N., Chesneau, C., Hanif, A., Khan, A. A., & Shahbaz, A. (2021). Development of nonlaboratory-based risk prediction models for cardiovascular diseases using conventional and machine learning approaches. *International Journal of Environmental Research and Public Health*, *18*(23). https://doi.org/10.3390/ijerph182312586

[37] Mendes, R. G., De Souza, C. R., Machado, M. N., Correa, P. R., Di Thommazo-Luporini, L., Arena, R., Myers, J., Pizzolato, E. B., & Borghi-Silva, A. (2015). Predicting reintubation, prolonged mechanical ventilation and death in post-coronary artery bypass graft surgery: A comparison between artificial neural networks and logistic regression models. *Archives of Medical Science*, *11*(4), 756-763. https://doi.org/10.5114/aoms.2015.48145

[38] Suzuki, S., Yamashita, T., Sakama, T., Arita, T., Yagi, N., Otsuka, T., Semba, H., Kano, H., Matsuno, S., Kato, Y., Uejima, T., Oikawa, Y., Matsuhama, M., & Yajima, J. (2019). Comparison of risk models for mortality and cardiovascular events between machine learning and conventional logistic regression analysis. *PLoS ONE*, *14*(9), 1-14. https://doi.org/10.1371/journal.pone.0221911

[39] Damen, J. A. A. G., Hooft, L., Schuit, E., Debray, T. P. A., Collins, G. S., Tzoulaki, I., Lassale, C. M., Siontis, G. C. M., Chiocchia, V., Roberts, C., Schlüssel, M. M., Gerry, S., Black, J. A., Heus, P., Van Der Schouw, Y. T., Peelen, L. M., & Moons, K. G. M. (2016). Prediction models for cardiovascular disease risk in the general population: Systematic review. *BMJ (Online)*, *353*. https://doi.org/10.1136/bmj.i2416

[40] Allyn, J., Allou, N., Augustin, P., Philip, I., Martinet, O., Belghiti, M., Provenchere, S., Montravers, P., & Ferdynus, C. (2017). A comparison of a machine learning model with EuroSCORE II in predicting mortality after elective cardiac surgery: A decision curve analysis. *PLoS ONE*, *12*(1), 1-12. https://doi.org/10.1371/journal.pone.0169772

[41] Geltser, B. I., Shahgeldyan, K. J., Rublev, V. Y., Kotelnikov, V. N., Kriger, A. B., & Shirobokov, V. G. (2020). Machine learning methods for prediction of hospital mortality in patients with coronary heart disease after coronary artery bypass grafting. *Kardiologiya*, *60*(10), 38-46. https://doi.org/10.18087/CARDIO.2020.10.N1170

[42] Huang, Y. C., Li, S. J., Chen, M., Lee, T. S., & Chien, Y. N. (2021). Machine-Learning Techniques for Feature Selection and Prediction of Mortality in Elderly CABG Patients. *Healthcare (Switzerland)*, *9*(5), 1-11. https://doi.org/10.3390/healthcare9050547

[43] Lippmann, R. P. & Shahian, D. M. (1997). Coronary artery bypass risk prediction using neural networks. *Annals of Thoracic Surgery*, *63*(6), 1635-1643. https://doi.org/10.1016/S0003-4975(97)00225-7

[44] Tu, J. V., Weinstein, M. C., McNeil, B. J., & Naylor, C. D. (1998). Predicting mortality after coronary artery bypass surgery: What do artificial neural networks learn? *Medical Decision Making*, *18*(2), 229-235. https://doi.org/10.1177/0272989X9801800212

[45] Zou, J., Han, Y., & So, S. S. (2008). Overview of artificial neural networks. *Methods in Molecular Biology*, *458*, 15-23. https://doi.org/10.1007/978-1-60327-101-1_2

[46] Chong, C. F., Li, Y. C., Wang, T. L., & Chang, H. (2003). Stratification of adverse outcomes by preoperative risk factors in coronary artery bypass graft patients: an artificial neural network prediction model. *AMIA Annual Symposium Proceedings 2003*, 160-164.

[47] Hart, G. R., Roffman, D. A., Decker, R., & Deng, J. (2018). A multi-parameterized artificial neural network for lung cancer risk prediction. *PLoS ONE*, *13*(10), 1-13. https://doi.org/10.1371/journal.pone.0205264

[48] Tripathy, R. K., Bhattacharyya, A., & Pachori, R. B. (2019). A Novel Approach for Detection of Myocardial Infarction from ECG Signals of Multiple Electrodes. *IEEE Sensors Journal*, *19*(12), 4509-4517. https://doi.org/10.1109/JSEN.2019.2896308

[49] Ghavidel, A. A., Javadikasgari, H., Maleki, M., Karbassi, A., Omrani, G., & Noohi, F. (2014). Two new mathematical models for prediction of early mortality risk in coronary artery bypass graft surgery. *Journal of Thoracic and Cardiovascular Surgery*, *148*(4), 1291-1298. https://doi.org/10.1016/j.jtcvs.2014.02.028

[50] Nouei, M. T., Kamyad, A. V., Sarzaeem, M. R., &

Ghazalbash, S. (2014). Developing a Genetic Fuzzy System for Risk Assessment of Mortality After Cardiac Surgery. *Journal of Medical Systems*, *38*(10). https://doi.org/10.1007/s10916-014-0102-5

[51] Jamaati, H., Najafi, A., Kahe, F., Karimi, Z., Ahmadi, Z., Bolursaz, M., Masjedi, M., Velayati, A., & Hashemian, S. M. (2015). Assessment of the EuroSCORE risk scoring system for patients undergoing coronary artery bypass graft surgery in a group of Iranian patients. *Indian Journal of Critical Care Medicine*, *19*(10), 576-579. https://doi.org/10.4103/0972-5229.167033

[52] Nouei, M. T., Kamyad, A. V., Sarzaeem, M. R., & Ghazalbash, S. (2016). Fuzzy risk assessment of mortality after coronary surgery using combination of adaptive neuro-fuzzy inference system and K-means clustering. *Expert Systems*, *33*(3), 230-238. https://doi.org/10.1111/exsy.12145

[53] Mori, M., Durant, T. J. S., Huang, C., Mortazavi, B. J., Coppi, A., Jean, R. A., Geirsson, A., Schulz, W. L., & Krumholz, H. M. (2021). Toward Dynamic Risk Prediction of Outcomes after Coronary Artery Bypass Graft: Improving Risk Prediction with Intraoperative Events Using Gradient Boosting. *Circulation: Cardiovascular Quality and Outcomes*, *June*, 734-745. https://doi.org/10.1161/CIRCOUTCOMES.120.007363

[54] Khalaji, A., Behnoush, A. H., Jameie, M., Sharifi, A., Sheikhy, A., Fallahzadeh, A., Sadeghian, S., Pashang, M., Bagheri, J., Ahmadi Tafti, S. H., & Hosseini, K. (2022). Machine learning algorithms for predicting mortality after coronary artery bypass grafting. *Frontiers in Cardiovascular Medicine*, *9*(August), 1-9. https://doi.org/10.3389/fcvm.2022.977747

[55] Zea-Vera, R., Ryan, C. T., Havelka, J., Corr, S. J., Nguyen, T. C., Chatterjee, S., Wall, M. J., Coselli, J. S., Rosengart, T. K., & Ghanta, R. K. (2022). Machine Learning to Predict Outcomes and Cost by Phase of Care After Coronary Artery Bypass Grafting. *Annals of Thoracic Surgery*, *114*(3), 711-719. https://doi.org/10.1016/j.athoracsur.2021.08.040

Contact information:

Ayşe Banu BİRLİK, Lecturer
(Corresponding author)
1) Istanbul Medipol University,
Graduate School of Engineering and Natural Sciences,
Department of Healthcare System Engineering,
34810, Istanbul, Turkey
2) Istinye University,
Department of Medical Services and Techniques,
34810, Istanbul, Turkey
E-mail: banu.birlik@istinye.edu.tr

Hakan TOZAN, Professor
American University of the Middle East,
College of Engineering and Technology,
Egaila, Kuwait
E-mail: hakan.tozan@aum.edu.kw

Kevser Banu KÖSE, Assistant Professor
Istanbul Medipol University,
School of Engineering and Natural Sciences,
Department of Biomedical Engineering,
34810, Istanbul, Turkey
E-mail: kbkose@medipol.edu.tr