Machine Learning-Based Prediction Models for Healthcare Outcomes in Patients Participating in Cardiac Rehabilitation

A SYSTEMATIC REVIEW

Xiarepati Tieliwaerdi, MD; Kathryn Manalo, DO; Abulikemu Abuduweili, MS; Sana Khan, MD; Edmund Appiah-kubi, MD; Brent A. Williams, PhD; Andrew C. Oehler, MD

Purpose: Cardiac rehabilitation (CR) has been proven to reduce mortality and morbidity in patients with cardiovascular disease. Machine learning (ML) techniques are increasingly used to predict healthcare outcomes in various fields of medicine including CR. This systemic review aims to perform critical appraisal of existing ML-based prognosis predictive model within CR and identify key research gaps in this area.

Review Methods: A systematic literature search was conducted in Scopus, PubMed, Web of Science, and Google Scholar from the inception of each database to January 28, 2024. The data extracted included clinical features, predicted outcomes, model development, and validation as well as model performance metrics. Included studies underwent quality assessments using the IJMEDI and Prediction Model Risk of Bias Assessment Tool checklist.

Summary: A total of 22 ML-based clinical models from 7 studies across multiple phases of CR were included. Most models were developed using smaller patient cohorts from 41 to 227, with one exception involving 2280 patients. The prediction objectives ranged from patient intention to initiate CR to graduate from outpatient CR along with interval physiological and psychological progression in CR. The best-performing ML models reported area under the receiver operating characteristics curve between 0.82 and 0.91, with sensitivity from 0.77 to 0.95, indicating good prediction capabilities. However, none of them underwent calibration or external validation. Most studies raised concerns about bias. Readiness of these models for implementation into practice is questionable. External validation of existing models and development of new models with robust methodology based on larger populations and targeting diverse clinical outcomes in CR are needed.

Key Words: cardiac rehabilitation • clinical prediction models • machine learning • systematic review

KEY PERSPECTIVE

What is novel?

- This systematic review is the first to critically appraise existing machine learning (ML)-based clinical prognostic models in cardiac rehabilitation (CR).
- It identifies key research gaps in ML application in CR and provides constructive suggestions for improving the quality of ML models specifically in CR.

What are the clinical and/or research implications?

- The review demonstrates various clinical challenges in the CR setting that ML models have been used to address and their corresponding outcomes.
- It reveals potential limitations in existing models and offers recommendations to enhance the readiness of current and future models for implementation in clinical settings as clinical decision-support tools.

Cardiovascular disease remains a leading cause of morbidity and mortality globally, directly causing over 19 million deaths worldwide in 2020 alone.[1] Cardiac rehabilitation (CR) is a comprehensive evidence-based intervention tailored to patients with cardiovascular conditions such as ischemic heart disease, heart failure, myocardial infarction, or those undergoing cardiovascular interventions such as coronary angioplasty or bypass grafting.[2-5] Participation in CR has been shown to significantly reduce morbidity and mortality, improve functional capacity, and enhance quality of life.[6-11] To optimize these benefits, extensive clinical research has been conducted to identify challenges within CR and develop strategies for improvement.[12-15] Within this body of research, conventional statistical methods (CSM), such as regression models, play a key role. These CSM serve two primary purposes: inference and prediction.[16] Inference focuses on identifying associations between patient characteristics and clinical outcomes, while prediction estimates the likelihood of future outcomes.[17] Although both tasks offer clinical value, CSM in CR have predominantly been used for inference, identifying risk factors linked to CR outcomes, with fewer studies emphasizing prediction.[18-20] For example, Kavanagh et al used regression models to identify peak oxygen intake as the strongest predictor of mortality. However, when predicting 1-year major adverse cardiovascular events, the model demonstrated only modest performance, with an area under the receiver operating characteristics (AUC-ROC) curve of 0.66, reflecting the limitations of CSM in prediction.[21] These CSM also emphasize parsimony and interpretability.[22] Parsimony refers to the preference for simpler

Author Affiliations: *Department of Medicine, Allegheny Health Network, Pittsburgh, Pennsylvania (Drs Tieliwaerdi, Manalo, Khan, and Appiah-kubi); Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania (Dr Abuduweili); and Allegheny Health Network, Allegheny Health Network Cardiovascular Institute, Pittsburgh, Pennsylvania (Drs Williams and Oehler)*

The authors declare no conflicts of interest

All authors have read and approved the manuscript.

All authors meet the 4 ICMJE criteria for authorship.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.jpnnjournal.com).

Correspondence: *Andrew C. Oehler, MD, Allegheny Health Network, Allegheny Health Network Cardiovascular Institute, 320 E. North Avenue, Pittsburgh, PA 15212 (Andrew.oehler@ahn.org).*

DOI: 10.1097/HCR.0000000000000943

models with fewer predictors to avoid overfitting, while interpretability ensures that results, like odds ratios, are intuitive in terms of quantifying the effects of patient characteristics on outcomes.[22] However, this simplicity can come at the cost of predictive accuracy, where more complex models, such as those in machine learning (ML), tend to excel.[23]

As a branch of artificial intelligence, ML holds great potential for achieving high predictive power in complex datasets, such as healthcare data, due to its ability to detect nonlinear relationships and recognize patterns and interactions among large combinations of variables and outcomes.[24] While this strength may come at the expense of interpretability, it offers significant clinical value in prediction tasks and complements CSM. For example, as wearable devices become more common in CR, particularly in home-based programs that generate millions of continuous data points such as vital signs, oxygen consumption, and cardiac monitoring, ML can efficiently process these large datasets, surpassing the capacity of CSM.[25] This enables the prediction of patient progression and helps providers guide personalized exercise therapy for subsequent sessions, ultimately optimizing patient outcomes.[26-29] Additionally, the capacity of ML to process multimodal data, a task technically unfeasible for CSM such as patient queries, provider handwritten notes, electrocardiographic tracings, and imaging like echocardiograms, opens possibilities for more comprehensive and personalized patient care through ML approaches.[30-32]

Recent literature has seen a proliferation of systematic reviews and meta-analyses aimed at evaluating the performance of ML-based clinical prediction models in different fields of medicine.[33-35] However, to the best of our knowledge, there has yet to be a systematic review of ML predictive models in CR. The primary aim of this review is to evaluate existing literature regarding ML predictive models in the context of CR and provide a systematic appraisal of these models from their development to validation. We aim to compare performance metrics, validation processes, and appropriateness of algorithms used in CR, as well as identify key research gaps in this area.

METHODS

LITERATURE SEARCH STRATEGY

The literature search and related statistical analyses were conducted in accordance with the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement.[36] We comprehensively searched publications from database inception to January 28, 2024 in Google Scholar, PubMed, Web of Science, and Scopus. Keywords used for the search included a combination of terms related to ML and CR. The full search strategy, including the specific combinations of search terms used in each database, was provided in Table 1, Supplementary Digital Content available at: http://links.lww.com/JCRP/A593.

INCLUSION AND EXCLUSION CRITERIA

In our review, we included studies that employed an ML model to predict clinical progression, health outcomes, or risk stratification in a cohort of adult patients participating in any phase of CR. We did not restrict it by the country of origin or publication source. For clarity, we defined ML as algorithms, such as random forest analysis, support vector machines, and neural networks, that are more complex than logistic regression models and capable of making decisions based on data patterns.

Our inclusion criteria were structured using the population, intervention, comparison, outcome, and time approach as follows: (1) Population: Adult patients aged 18 years or older engaging in CR programs; (2) Intervention: Studies used ML models for predicting outcomes in CR; (3) Comparison: Not applicable, due to the lack of a universally accepted prognostic model in CR; (4) Outcome: Studies reporting on clinical progression, health outcomes or risk stratification outcomes; (5) Time: Studies that harnessed features to predict outcomes after any given follow-up period were accepted.[37] In terms of study design, we did not restrict the types of studies included. Retrospective, prospective, cross-sectional, cohort, and case-control studies as well as randomized controlled trials were all considered for inclusion.

Exclusion criteria were: (1) studies not in English; (2) studies not involving ML-based predictive models: for example, studies reporting novel ML-based wearable devices used for monitoring biomarkers such as heart rate or blood pressure were excluded; (3) studies where predictive models are not developed from patients in CR; (4) studies focusing on identifying predictors associated with outcomes rather than developing a prognostic model. (5) Reviews, case reports, and studies not available in full text were excluded.

DATA EXTRACTION

Duplicate records were initially removed using auto-deduplication function in EndNote 21, followed by a manual check for complete deduplication. The screening of titles and abstracts was then carried out in EndNote 21, adhering to the inclusion and exclusion criteria previously outlined.

A team of 4 reviewers (X.T., K.M., S.K., and E.A.) assessed articles for eligibility first by screening titles and abstracts to ensure relevance in EndNote; each study was independently assessed by at least 2 reviewers. Both agreed-upon and conflicting articles were retained for second-round screening based on full-text review. The full-text evaluation was independently carried out by the reviewers (X.T. and A.A.). Disagreements were resolved through mutual consensus.

The data from the included articles were independently extracted into tables by 2 authors (S.K. and K.M.). They then reviewed and compared each other's work to identify and resolve any discrepancies. Any unresolved discrepancies were addressed through discussion or consultation with a third reviewer (X.T.). Variables were extracted and tabulated in Excel 2020, which includes (1) baseline characteristics of studies; (2) features and ML algorithms used for the development of the models; (3) objectives of the prediction models as well as evaluation index for their performance, including AUC-ROC or abbreviated as AUC, accuracy, precision, sensitivity, and specificity; (4) methods employed for model validation.

QUALITY ASSESSMENT

The quality of the included studies was evaluated independently by 2 reviewers (X.T. and A.A.) by using the IJMEDI checklist.[38] The IJMEDI checklist is developed specifically for use in medical research. The IJMEDI checklist distinguishes high-quality ML research, which involves rigorous model development, validation, and clinical relevance, from studies that merely apply ML methods to medical data without thorough consideration of algorithm suitability or real-world clinical applicability. It focuses on 6 key aspects: problem understanding, data understanding, data preparation, modeling, validation, and deployment, encompassing a total of 30 detailed questions. These questions were categorized into high- and low-priority based on their relative importance in assessing study quality. High-priority questions were scored as follows: 2 for adequately addressing the IJMEDI checklist

requirements (OK), 1 for moderately addressing with potential for improvement and requiring minor revisions (mR), and 0 for inadequately addressing, requiring major revisions (MR), in accordance with the IJMEDI checklist. Low-priority questions were scored as 0 for OK, 0.5 for mR, and 1 for MR. The point-scoring approach used in our study is not described in the original IJMEDI reference but was adopted by multiple publications to reflect the quality of studies being evaluated.[39-41] Studies were then classified into 3 quality categories: low (0-19.5 points), medium (20-34.5 points), or high quality (35-50 points).

We also employed the Prediction Model Risk of Bias Assessment Tool (PROBAST) checklist, which is designed for evaluating the risk of bias and applicability of diagnostic and prognostic prediction model studies.[42,43] Each study is evaluated in 4 key domains: participants, predictors, outcome, and analysis. Each domain encompasses signaling questions with responses guiding an overall judgment of bias risk as "Low", "High", or "Unclear". The signaling questions are answered as "Yes", "No", "Probably yes", "Probably no", or "No information", with "Yes" generally indicating lower bias. The overall risk of bias and applicability for each study is determined based on these domain judgments.

RESULTS

From the initial search that yielded 151 records, 84 were excluded following a review of the title and abstract, and 60 records were excluded after full-text review. In the end, 7 studies met the inclusion criteria and were included in our systemic review. The process of screening and study selection is shown in Figure 1.[44-50]

STUDY CHARACTERISTICS

The characteristics of studies using ML prediction models in CR are summarized in Table 1. Seven studies are included, one of which is a multicenter study conducted in Belgium and Ireland.[46] The rest are single-center studies from Australia, Belgium, Italy, Malaysia, and Chile.[44,45,47-50] Study designs include 5 retrospective studies, one cross-sectional study, and one that combines retrospective and prospective approaches.[44-50] Most studies target phase II of CR, with one addressing phases II and III and another focusing solely on phase III.[44,46] Participant ages were reported in 4 of the 7 studies, with mean ages ranging from 63 to 67.98 years.[44,47,49,50] The percentage of female participants is reported in 3 studies and varies from 19% to 27%, while the other 4 studies do not report sex distribution.[47,49,50] Patients included were typically referred to CR, with one study specifying the prerequisite of being employed prior to a cardiac event.[45] The studies generally excluded patients lacking post-rehabilitation data, those with contraindications to exercise, or those unwilling to participate. Functional sample sizes range significantly from 41 to 2280 across the studies.

MODEL CHARACTERISTICS

A summary of characteristics of ML models included in our review is presented in Table 2. Each study has different prediction goals, ranging from forecasting a patient's intention to begin CR to estimating their adherence to the program. Two studies aimed to predict patient prognosis in CR including 6-minute walk distance and various physiological and psychological outcomes such as cholesterol level and depression level changes after outpatient CR.[44,47] Additionally, 2 studies focused on predicting patient

Table 1

Characteristics of Studies Using a Machine Learning Prediction Model in Cardiac Rehabilitation

Study (yr)	Study Location	Rehab Phase	Study Design	Age (yr)	Female (%)	Cohort Inclusion Criteria	Cohort Exclusion Criteria	Functional Sample Size
Van et al (2010)[38]	Australia	II and III	Retrospective	Mean: 65.35	NR	Patients referred to CR	Patients did not have post-rehab data or where values in target attributes were missing	2280
Lofaro et al (2016)[44]	Italy	II	Retrospective	Mean: 67.98	19	Patients admitted for CR after CAD or MI	NR	129
De Cannière et al (2020)[41]	Belgium	II	Retrospective	Mean: 63	27	Patients eligible for CR	Patients who were unable to exercise or didn't complete the total study	89
Jahandideh et al (2021)[43]	Australia	II	Cross-sectional	Mean: 63	26	Patients referred to outpatient CR	Patients who were too ill or unable to speak English	217
Yuan et al (2023)[39]	Malaysia	II	Retrospective	≥18	NR	Patients referred to CR and employed prior to having a cardiac event	Patients who were pensioners or unemployed	184
Filos et al (2023)[40]	Belgium and Ireland	III	Retrospective	40-80	NR	Patients completed a phase II CR and enrolled in Internet-based home CR	Patients who dropped out or exercised very sparsely	41
Torres et al (2023)[42]	Chile	II	Retrospective and prospective	≥18	NR	Patients referred to CR	Patients who had any contraindication to physical exercise	227

Abbreviations: CAD, coronary artery disease; CR, cardiac rehabilitation; MI, myocardial infarction; NR, not reported.

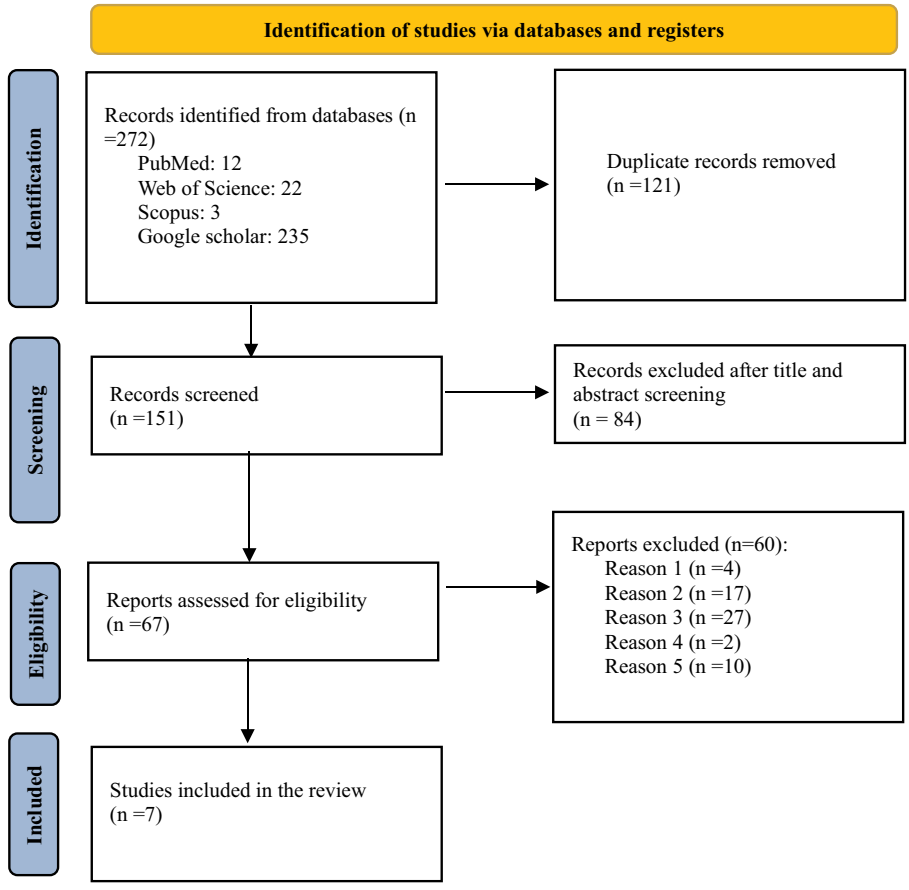

Figure 1. PRISMA flowchart of the review process for studies using a machine learning prediction model in cardiac rehabilitation.Adapted from PRISMA 2020 flow diagram for new systematic reviews, source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. This work is licensed under CC BY 4.0. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

disposition post-CR, such as likelihood of returning to work as well as the transition from one phase of CR to another.[45,48]

These studies utilized a variety of patient features to develop their models. These features encompassed anthropometric measurements such as body mass index and waist circumference, demographic information, and medical history especially cardiovascular health. Psycho-behavioral profiles were also considered in most of the studies including evaluation of anxiety and depression levels. Laboratory test results and physical fitness levels were commonly used. Two studies harnessed specific imaging like electrocardiogram and transthoracic echocardiogram.[47,48]

The number of features used across the studies varies, with one utilizing as few as 17 features and others up to 82.[49,50] Most studies used pre-processing techniques for feature selection to reduce the number of input variables to those that are most important to the predictive model to improve performance and reduce computational cost. Random forest and principal component analysis were most employed.[44,48,49]

Only 2 of the 7 studies reported methods to handle missing data.[44,50] In terms of ML algorithms, there was a wide array employed across these studies, from basic decision trees and support vector machines to advanced ensemble methods like AdaBoost and XGBoost. Three of 7 studies used a single algorithm, the other 4 compared multiple different algorithms, with decision trees, random forest, and support vector machines being the most adopted.[45,46,48,50]

PERFORMANCE AND VALIDATION

The best ML models, including their performance metrics and validation approaches, are outlined in Table 3. The performance of these models was evaluated using various metrics such as AUC-ROC curve, sensitivity and specificity, mean absolute error, and normalized mean squared error. The selection of best-performing ML algorithm varied based on each study's objectives. The best-performing models in their respective tasks reported AUC values between 0.82 and 0.91, sensitivity ranging from 0.77 to 0.95, demonstrating good predictive capabilities.[44-46] Six out of 7 studies employed internal validation techniques, with cross-validation being the most prevalent method.[44-48,50] However, none of the studies underwent external validation, meaning they were not validated on data from populations different from those used to develop the models, nor were their predictions calibrated against real-life observed outcomes.

QUALITY ASSESSMENT

The scores for each dimension and the total score of each study according to the IMED checklist are summarized in Table 4. The included studies had an average score of 30.8, with a range from 26 to 35. Most of the studies fell into the medium-quality category, and one study stood out as being of high quality.[44] Most studies demonstrated a discernible bias in the "Data Preparation," "Validation," and "Deployment" dimensions, with lower scores that suggest

Table 2

Characteristics of Machine Learning Prediction Models Used in Cardiac Rehabilitation

Study (yr)	Object of Prediction	Features Used for Prediction		Missing Data Handling	Feature Selection	ML Algorithms
		Type	Number			
Van et al (2010)[38]	Physiological and psychosocial outcomes of CR	Anthropometric measurements, sociodemographic, psycho-behavioral profile, medical history, laboratorial tests, physical fitness level	49	Cases with >25% of missing data were discarded, otherwise were replaced by mean or mode	RF, PCA	DT
Lofaro et al (2016)[44]	Patient-specific CR exercise program	Anthropometric measurements, socio-demographics, medical history, laboratory tests and images (PFTs), physical fitness level	17	Variables with >50% of missing data were discarded	None	Lasso regression, SVM, RF, Bagged FDA, Boost C 5.0, Bagged CART
De Cannière et al (2020)[41]	6-min walk distance	Anthropometric measurements, psycho-behavioral profile, medical history, laboratory tests and images (EKG, TTE), physical fitness level	NR	NR	None	SVM, linear regression
Jahandideh et al (2021)[43]	Individual's intention to engage in CR	Anthropometric measurements, socio-demographics, medical history, perceived need for CR	82	NR	RF	RF
Yuan et al (2023)[39]	Ability to return to work after CR.	Anthropometric measurements, socio-demographic and psychosocial profile, medical history, laboratory tests, physical fitness level	30	NR	Recursive feature elimination	DT, RF, AdaBoost, XGBoost, CatBoost, SVM, Complement Naive Baye
Filos et al (2023)[40]	6-mo adherence to home-based CR	Anthropometric measurements, psycho-behavioral profile, laboratory tests, physical fitness level	52	NR	None	DT and RF
Torres et al (2023)[42]	Probability of progression from CR phase II to III	Anthropometric measurements, psycho-behavioral profile, laboratory tests and images (EKG and TTE), physical fitness level	44	NR	PCA, correlation analysis	XGboost, gradient boosting, SVM, RF, KNN

Abbreviations: CR, cardiac rehabilitation; DT, decision trees; EKG, electrocardiogram; FDA, flexible discriminant analysis; KNN, k-nearest neighbors; LASSO, least absolute shrinkage and selection operator; ML, machine learning; NR, not reported; PCA, principal component analysis; PFT, pulmonary function test; RF, random forest; SVM, support vector machine; TTE, transthoracic echocardiogram; XGBoost, Extreme Gradient Boosting.

these areas may have impacted the overall quality and reliability of the outcomes.

The percentage of studies rated by the level of concern regarding risk of bias and applicability for each domain according to the PROBAST checklist is presented in Figure 2. It reveals a predominantly low risk of bias in "Participants" and "Predictors" at 100% and 86% of included studies, respectively, with similarly low applicability concerns. However, a significant 71% of studies were considered to have a high overall risk of bias, which may be attributed to the "Analysis" domain, where 57% of studies show high risk of bias. A detailed evaluation of each study is provided in Table 2, Supplemental Digital Content available at: http://links.lww.com/JCRP/A593.

DISCUSSION

This review has evaluated the performance of 22 ML-based clinical models in 7 studies aiming to predict healthcare outcomes for patients participating in CR. The prediction objectives ranged from patient intention to initiate CR to graduate from outpatient CR as well as interval physiological and psychological changes during the program. The best-performing ML models in their respective tasks reported AUC between 0.82 and 0.91 and sensitivity from 0.77 to 0.95, demonstrating good prediction capabilities in general.[51] The majority of the included studies were rated as medium quality according to the IJMEDI checklist and there were high concerns for bias per PROBAST assessment. Meta-analysis was not conducted as the included ML models were highly heterogenous in terms of targeted population, prediction objectives, outcome measurement, and validation.

An ideal clinical prediction model should correctly distinguish between patients who will develop certain events and those who will not without misclassification in any case.[42,52] Its quality is associated with 2 properties of the model: discrimination and calibration. Discrimination is the model's capacity to correctly separate individuals at higher risk of an event from those at lower risk. Calibration refers to the model's ability to estimate absolute risks accurately.[53] Discrimination is typically measured by the AUC values. It can also be assessed by sensitivity and specificity.[54] However, sensitivity and specificity vary as the cut point used to determine "positive" and "negative" test results changes. The ROC curve is a graph of the sensitivity of a test versus its false-positive rate (1-specicificy) for all potential cut points. The AUC value represents average prediction accuracy after balancing the inherent tradeoffs that exist between sensitivity and specificity across the spectrum of varying cut points.[55] A higher AUC indicates better discrimination ability. An AUC of 0.5 suggests no discrimination, equivalent to random guessing, while an AUC of

Table 3

Performance and Validation of Machine Learning Prediction Models in Cardiac Rehabilitation

Study (yr)	Best ML Algorithm	Performance Metrics	Validation Methods	External Validation	Practice Implementation
Van et al (2010)[38]	DT	AUC: 0.815; accuracy: 0.769; precision: 0.734; sensitivity: 0.769	Internal validation: 10-fold cross-validation	None	None
Lofaro et al (2016)[44]	Lasso Regression	Accuracy: 0.935; precision: 0.941; sensitivity: 0.9	5-fold cross-validation	None	None
De Cannière et al (2020)[41]	SVM	MAE: 42.8 m	20-fold-validation;	None	None
Jahandideh et al (2021)[43]	RF	Accuracy: 0.715	NR	None	None
Yuan et al (2023)[39]	AdaBoost	AUC: 0.924; accuracy: 0.864; sensitivity: 0.928; specificity: 0.733	10-fold cross-validation	None	None
Filos et al (2023)[40]	DT	Sensitivity: 0.945; precision: 0.80	10-fold cross-validation	None	None
Torres et al (2023)[42]	XGboost	NMSE: 0.008; R^2: 92%	10-fold cross-validation	None	None

Abbreviations: AUC, area under curve; DT, decision tree; LASSO, least absolute shrinkage and selection operator; MAE, mean absolute error; ML, machine learning; NMSE, normalized mean squared error; RF, random forest; SVM, support vector machine; XGBoost, Extreme Gradient Boosting.

1.0 indicates perfect discrimination.[56] Van et al reported an AUC of 0.815, which suggests good discrimination in predicting post-rehab deterioration.[44] Yuan et al reported an AUC of 0.923, reflecting the model's excellent effectiveness in predicting patient likelihood of return-to-work.[45] Two studies reported sensitivity and precision rather than AUC as performance metrics, reflecting the model's performance for a particular cut point instead of all possible thresholds.[46,50] One study reported accuracy only, measuring the proportion of true results, including true positives and true negatives. High accuracy can sometimes be misleading if the outcomes that the ML model aims to predict occur frequently in the study population. Jahandideh et al claimed an accuracy of 0.715 in differentiating highly motivated patients for CR initiation but did not report the distribution of motivation levels in the study population.[49] This could lead to overestimation of accuracy in a predominantly motivated group or underestimation in a less motivated one. Overall, the reviewed studies used various evaluation metrics, the appropriateness of which is generally acceptable but could benefit from a more comprehensive measurement approach.

Model calibration is typically assessed after its discrimination is deemed acceptable. Calibration refers to how well a model's predicted absolute probability aligns with the true likelihood of the outcomes.[53,57,58] None of the included ML models underwent calibration, so their ability to predict absolute risk remains uncertain. Although they have demonstrated overall good discrimination, calibration is essential to prove their capability in clinical decision support. This limitation contributed to lower scores in the deployment section of the IJMEDI checklist and a high risk of bias per the PROBAST checklist as shown in Figure 2 and Table 4. Another major deficiency is the lack of external validation, which tests model efficacy in a different population than it was initially derived from. Without robust external validation, a model's generalizability is questionable.[57] One notable example is that the National Institute for Health and Care Excellence recommended an independent external validation of QRISK2 and the Framingham risk score, which were performed subsequently and demonstrated systematic miscalibration of the Framingham risk scores and led to the need for different treatment thresholds in the UK cohort.[59] Although some cases may not require immediate external validation if the sample size is large and representative of predictors and target outcomes on top of appropriate internal validation, most studies in our review, with populations ranging from 41 to 227 (except one with 2280 patients), would benefit from external validation as a key step toward implementation into clinical practice.[60]

In addition to methodological robustness in model development, a useful clinical prediction model should address clinically significant issues. Low enrollment and adherence remain major challenges for CR.[15] Khatanga et al used logistic regression to identify factors such as surgical diagnosis, non or former tobacco use, and intensity of physician recommendation as independent predictors for CR participation, whereas factors including anxiety, depression, or executive function had no significant impact.[61] Other studies have suggested that age, low socioeconomic status, gender, CR center location, and psycho-behavioral factors including lack of motivation and reduced self-efficacy are barriers to participating in CR based on CSM.[13,62,63] The study by Jahandideh et al developed a model predicting individual motivation to participate in CR based on factors such as demographics, medical history, perceived need, outcome expectations, self-efficacy, and barriers.[49] The findings revealed that the significance of these variables varied, influencing motivation differently for each individual. Based on the model's predictions, providers could identify the most significant barriers or motivators, enabling more targeted, individualized interventions. The Filos et al study emphasized that adherence during the first 6 weeks of a home-based CR program is crucial for predicting long-term adherence.[46] Patients adhering during this phase had a 92% chance of continued participation, while those struggling had a non-adherence risk of 56%. This provides decision-makers with insights to allocate resources and supervision more intensively during the early phase of CR, enabling early identification of patients at risk for low adherence. Future research could leverage the capacity of ML for remote monitoring and real-time big data analysis to predict optimal CR modalities, including single or mixed models that combine home-based, facility-based, and community-based programs for better outcomes. Studies could also

Table 4

Results of the Quality Assessment According to the IJMEDI Checklist of Machine Learning Prediction Models in Cardiac Rehabilitation[a,b]

Study (yr)	Van et al (2010)[38]	Lofaro et al (2016)[44]	De Cannière et al (2020)[41]	Jahandideh et al (2021)[43]	Yuan et al (2023)[39]	Filos et al (2023)[40]	Torres et al (2023)[42]
Problem understanding (10)	8.5	8.5	9.5	9	9.5	9.5	9
1	2	2	2	2	2	2	2
2	2	2	2	2	2	2	2
3	0.5	0.5	1	1	1	1	1
4	2	2	2	2	2	2	2
5	2	2	2	2	2	2	2
6	0	0	0.5	0	0	0	0
Data understanding (6)	4	3	4	4	3	3	4
7	1	2	2	2	1	1	0
8	1	1	1	1	1	1	2
9	2	0	1	1	1	1	2
Data preparation (8)	6	1	1	3	4	2	3
10	0	0	0	0	0	0	0
11	2	0	0	1	1	0	1
12	2	1	1	2	2	2	1
13	2	0	0	0	1	0	1
Modeling (6)	6	6	6	6	6	6	6
14	2	2	2	2	2	2	2
15	2	2	2	2	2	2	2
16	2	2	2	2	2	2	2
Validation (12)	8	6	6.5	4	7	6	8.5
17	2	1	1	1	2	2	2
18	2	2	2	1	1	1	2
19	0	0	0	0	0	0	0
20	2	1	1	1	2	2	2
21	0	0	0	0	0	0	0
22	2	2	2	1	2	1	2
23	0	0	0.5	0	0	0	0.5
Deployment (8)	2.5	1.5	3	3	3.5	2.5	3
24	1	1	1	1	1	1	1
25	0.5	0.5	0.5	1	1	0.5	0.5
26	0.5	0	1	0.5	1	0.5	1
27	0.5	0	0.5	0.5	0.5	0.5	0.5
28	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0
Total (50)	35	26	30	30	33	29	33.5

[a]Items in bold indicate high priority.

[b]High-priority questions were scored as 2 for adequately addressing the IJMEDI checklist requirements (OK), 1 for moderately addressing with potential for improvement and requiring minor revisions (mR), and 0 for inadequately addressing, requiring major revisions (MR). Low-priority questions were scored as 0 for OK, 0.5 for mR, and 1 for MR.

explore the ideal balance of current interventions for each individual, such as exercise training, health behavior modification, patient education, and nutritional and psychological counseling, providing valuable insights for decision-makers to promote adherence and optimize outcomes.

STRENGTHS AND LIMITATIONS OF THIS STUDY

The major strengths of this review include a comprehensive search and meticulous selection of studies, focusing on ML-based clinical prediction models in CR. Key characteristics such as predictors, data processing, ML algorithms, and

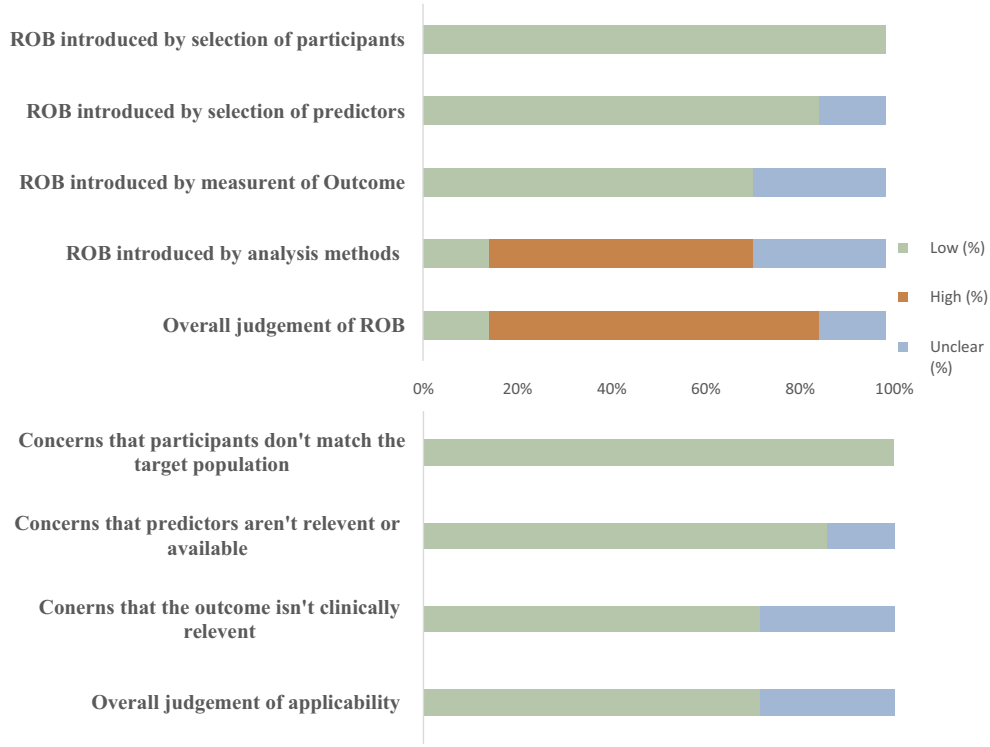

Figure 2. Results of the risk of bias assessment according to PROBAST checklist for studies using a machine learning prediction model in cardiac rehabilitation. Abbreviation: ROB, risk of bias.

their performance were extensively extracted and analyzed. A significant highlight is the use of specialized ML-targeted quality assessment tools, namely the PROBAST and IJMEDI checklists, ensuring a thorough and rigorous evaluation of model quality and bias. However, the review has several limitations. The scope was confined to only 7 studies, featuring 22 ML models, possibly not representing the full spectrum of research in this area. Geographical limitations are evident, with no studies included from North America or Asia, potentially affecting the applicability of the findings. The high heterogeneity among the models precluded a comparative analysis and meta-analysis, limiting our ability to determine the effective prediction model in specific tasks related to CR. Notably, the lack of calibration and external validation of these models raises concerns regarding their potential for clinical practice implementation.

RECOMMENDATIONS FOR FUTURE STUDIES

First, future research should focus on developing new prediction models utilizing larger patient samples that encompass greater diversity in sociodemographic characteristics, risk factors, and healthcare outcomes. In contrast to the abundance of ML prediction models in other fields of cardiovascular diseases, there is a paucity of ML-based models related to CR.[64-67] It is likely due to a shortage in large datasets in the CR population compared to the general population with cardiovascular diseases, where large datasets like the Cleveland, University of California repository and Massachusetts Institute of Technology-Beth Israel Hospital arrhythmia database are available.[64,68,69] The development of large datasets in CR is challenging due to low utilization of CR and limited communication and collaboration among CR programs.[15,70] The American Association of Cardiovascular and Pulmonary Rehabilitation CR registry represents a meaningful step forward in this area and may serve as a foothold for future ML-based research in CR. Machine learning is often considered to be "data hungry," where larger sample sizes with more cases of relevant predictors and target outcomes generally lead to better performance.[71] The lack of a large sample pool has posed a significant challenge for developing effective ML prediction models. To address this, collaboration among multiple centers is essential, as modeled by the American Association of Cardiovascular and Pulmonary Rehabilitation data registry. Additionally, expanding research to include not just facility-based CR populations but also those community-based or home-based CR programs could be a viable solution, as novel methods for the delivery of CR via virtual and remote options are becoming more common.[72] Second, exploring more advanced ML algorithms is recommended, with a strong emphasis on strict adherence to guidelines from data preparation through to model validation.[42,73] Algorithms such as artificial neural networks should be considered due to their potential superior accuracy and lesser dependence on feature selection.[74] Third, conducting calibration and external validation of existing datasets is highly valuable. Additionally, performing head-to-head comparisons between different ML-based prediction models and CSM, targeting similar tasks in comparable populations is also beneficial. These steps will significantly increase the potential for ML models to be integrated into clinical practice as effective decision-support tools.

CONCLUSION

There's a scarcity of ML-based clinical prognostic models for predicting healthcare outcomes in CR participants.

While current models show good predictive capacities, the lack of standard practices in methodologic reporting and the absence of external validation limit the ability to compare models and diminish the clinical relevance of any individual model. Future research should focus on developing new prediction models aimed at various outcomes in more diverse populations using robust methodological approaches. Additionally, enhancing the generalizability of existing models through external validation is necessary. There is still a long journey ahead before these models can be fully embraced in clinical settings.

REFERENCES

1. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart Disease and Stroke Statistics—2023 update: a report from the American Heart Association. *Circulation*. 2023;147(8):e93–e621. doi:10.1161/CIR.0000000000001123.
2. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol*. 2022;79(17):e263–e421. doi:10.1016/j.jacc.2021.12.012.
3. Thomas RJ, Balady G, Banka G, et al. 2018 ACC/AHA clinical performance and quality measures for cardiac rehabilitation: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures. *Circ Cardiovasc Qual Outcomes*. 2018. 11(4). doi:10.1161/HCQ.0000000000000037.
4. Lawton JS, Tamis-Holland JE, Bangalore S, et al. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2022. 145(3). doi:10.1161/CIR.0000000000001038.
5. Virani SS, Newby LK, Arnold SV, et al. 2023 AHA/ACC/ACCP/ASPC/NLA/PCNA Guideline for the Management of Patients With Chronic Coronary Disease: a Report of the American Heart Association/American College of Cardiology Joint Committee on Clinical Practice Guidelines. *Circulation*. 2023. 148(9). doi:10.1161/CIR.0000000000001168.
6. Hamm LF, Sanderson BK, Ades PA, et al. Core competencies for cardiac rehabilitation/secondary prevention professionals. *J Cardiopulm Rehabil Prev*. 2011;31(1):2-10. doi:10.1097/HCR.0b013e318203999d.
7. Anderson L, Thompson DR, Oldridge N, et al. Exercise-based cardiac rehabilitation for coronary heart disease. *Cochrane Database Syst Rev*. 2016. Published online January 5. doi:10.1002/14651858.CD001800.pub3.
8. Long L, Mordi IR, Bridges C, et al. Exercise-based cardiac rehabilitation for adults with heart failure. *Cochrane Database Syst Rev*. 2019. 2019(1). doi:10.1002/14651858.CD003331.pub5.
9. Uithoven KE, Smith JR, Medina-Inojosa JR, Squires RW, Olson TP. The role of cardiac rehabilitation in reducing major adverse cardiac events in heart transplant patients. *J Card Fail*. 2020;26(8):645-651. doi: 10.1016/j.cardfail.2020.01.011.
10. Francis T, Kabboul N, Rac V, et al. The effect of cardiac rehabilitation on health-related quality of life in patients with coronary artery disease: a Meta-analysis. *Cana J Cardiol*. 2019;35(3):352-364. doi:10.1016/j.cjca.2018.11.013.
11. Grochulska A, Glowinski S, Bryndal A. Cardiac rehabilitation and physical performance in patients after myocardial infarction: preliminary research. *J Clin Med*. 2021;10(11):2253. doi:10.3390/jcm10112253.
12. Thomas RJ, Sapir O, Gomes PF, Iftikhar U, Smith JR, Squires RW. Advances, challenges, and progress in cardiac rehabilitation in chronic CVD management. *Curr Atheroscler Rep*. 2023;25(6):247-256. doi: 10.1007/s11883-023-01100-7.
13. Chindhy S, Taub PR, Lavie CJ, Shen J. Current challenges in cardiac rehabilitation: strategies to overcome social factors and attendance barriers. *Expert Rev Cardiovasc Ther*. 2020;18(11):777-789. doi:10.1080/14779072.2020.1816464.
14. Ozemek C, Squires RW. Enrollment and adherence to early outpatient and maintenance cardiac rehabilitation programs. *J Cardiopulm Rehabil Prev*. 2021;41(6):367-374. doi:10.1097/HCR.0000000000000645.
15. Ritchey MD, Maresh S, McNeely J, et al. Tracking cardiac rehabilitation participation and completion among medicare beneficiaries to inform the efforts of a national initiative. *Circ Cardiovasc Qual Outcomes*. 2020. 13(1). doi:10.1161/CIRCOUTCOMES.119.005902.
16. Rosenberg MA. Deep learning and biased prediction: more questions than answers? *Circ Heart Fail*. 2024;17(1):E011368. doi:10.1161/CIRCHEARTFAILURE.123.011368/ASSET/92677776-AB0A-40A0-81B2-7743F9C89C9F/ASSETS/CIRCHEARTFAILURE.123.011368.FP.PNG.
17. Ij H. Statistics versus machine learning. *Nat Methods*. 2018;15(4):233. doi: 10.1038/nmeth.4642.
18. Peters AE, Keeley EC. Trends and predictors of participation in cardiac rehabilitation following acute myocardial infarction: data from the behavioral risk factor surveillance system. *J Am Heart Assoc*. 2018;7(1). doi:10.1161/jaha.117.007664/asset/dc54c664-f95e-432c-8011-dd25a15359db/assets/images/large/jah32840-fig-0001.jpg.
19. Gaalema DE, Savage PD, Rengo JL, et al. Patient characteristics predictive of cardiac rehabilitation adherence. *J Cardiopulm Rehabil Prev*. 2017;37(2):103-110. doi:10.1097/HCR.0000000000000225.
20. Park LG, Schopfer DW, Zhang N, Shen H, Whooley MA. Participation in cardiac rehabilitation among patients with heart failure. *J Card Fail*. 2017;23(5):427-431. doi:10.1016/J.CARDFAIL.2017.02.003.
21. Kavanagh T, Mertens DJ, Hamm LF, et al. Prediction of long-term prognosis in 12 169 men referred for cardiac rehabilitation. *Circulation*. 2002;106(6):666-671. doi:10.1161/01.CIR.0000024413.15949.ED/ASSET/90F74775-0D83-4834-B085-87F5E82F3E17/ASSETS/GRAPHIC/G9FF4.JPEG.
22. Coelho MTP, Diniz-Filho JA, Rangel TF. A parsimonious view of the parsimony principle in ecology and evolution. *Ecography*. 2019;42(5):968-976. doi:10.1111/ECOG.04228.
23. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. *Electronic Markets*. 2021;31(3):685-695. doi:10.1007/s12525-021-00475-2.
24. Bakshi K, Bakshi K Considerations for artificial intelligence and machine learning: approaches and use cases. IEEE Aerospace Conference Proceedings, 2018 March; 2018: 1-9. doi:10.1109/AERO.2018.8396488.
25. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593.
26. Raja MA, Loughran R, Caffery FM. A review of applications of artificial intelligence in cardiorespiratory rehabilitation. *Inform Med Unlocked*. 2023;41:101327. doi:10.1016/J.IMU.2023.101327.
27. Karlsson P, Strand R, Kullberg J, et al. A detailed analysis of body composition in relation to cardiopulmonary exercise test indices. *Scientific Reports 2024*. 2024;14(1):1-11. doi:10.1038/s41598-024-72973-0.
28. Fu Z, Hong S, Zhang R, Sensors SD. 2021 undefined. Artificial-intelligence-enhanced mobile system for cardiovascular health management. *mdpi.comZ Fu, S Hong, R Zhang, S DuSensors, 2021•mdpi.com*; 2021. Published online 2021. doi:10.3390/s21030773.
29. Hijazi S, Page A, Kantarci B, Soyata T. Machine learning in cardiac health monitoring and decision support. *Computer (Long Beach Calif)*. 2016;49(11):38-48. doi:10.1109/MC.2016.339.
30. Barry T, Farina JM, Chao CJ, et al. The role of artificial intelligence in echocardiography. *J Imaging*. 2023;9(2):50. doi:10.3390/JIMAGING9020050.
31. Chang A, Cadaret LM, Liu K. Machine learning in electrocardiography and echocardiography: technological advances in clinical cardiology. *Curr Cardiol Rep*. 2020;22(12):1-7. doi:10.1007/S11886-020-01416-9/METRICS.
32. Manz CR, Parikh RB, Small DS, et al. Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: a stepped-wedge cluster randomized clinical trial. *JAMA Oncol*. 2020;6(12):e204759–e204759. doi:10.1001/JAMAONCOL.2020.4759.
33. Smith LA, Oakden-Rayner L, Bird A, et al. Machine learning and deep learning predictive models for long-term prognosis in patients

with chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Lancet Digit Health*. 2023;5(12):e872–e881. doi:10.1016/S2589-7500(23)00177-2.
34. Zhang Z, Yang L, Han W, et al. Machine learning prediction models for gestational diabetes mellitus: meta-analysis. *J Med Internet Res*. 2022;24(3):e26634. doi:10.2196/26634.
35. Chowdhury MZI, Naeem I, Quan H, et al. Prediction of hypertension using traditional regression and machine learning models: a systematic review and meta-analysis. *PLoS One*. 2022;17(4): e0266334. doi:10.1371/journal.pone.0266334.
36. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021;372:n160. doi:10.1136/bmj.n160.
37. Riva JJ, Malik KMP, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. *J Can Chiropr Assoc*. 2012;56(2):167–171.
38. Cabitza F, Campagner A. The need to separate the wheat from the chaff in medical informatics: introducing a comprehensive checklist for the (self)-assessment of medical AI studies. *Int J Med Inform*. 2021;153:104510. doi:10.1016/j.ijmedinf.2021.104510.
39. Zhou Y, Tong GY, lei SX, et al. Machine learning predictive models for acute pancreatitis: a systematic review. *Int J Med Inform*. 2022;157:104641. doi:10.1016/j.ijmedinf.2021.104641.
40. Moulaei K, Sharifi H, Bahaadinbeigy K, Haghdoost AA, Nasiri N. Machine learning for prediction of viral hepatitis: a systematic review and meta-analysis. *Int J Med Inform*. 2023;179:105243. doi:10.1016/j.ijmedinf.2023.105243.
41. Wang M, Francis F, Kunz H, et al. Artificial intelligence models for predicting cardiovascular diseases in people with type 2 diabetes: a systematic review. *Intell Based Med*. 2022;6:100072. doi:10.1016/j.ibmed.2022.100072.
42. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170(1):W1–W33. doi:10.7326/M18-1377.
43. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51-58. doi:10.7326/M18-1376.
44. Van A, Gay VC, Kennedy PJ, Barin E, Leijdekkers P. Understanding risk factors in cardiac rehabilitation patients with random forests and decision trees. In: Conferences in Research and Practice in Information Technology Series; 2010.
45. Choo JIA, Varathan KD, Suhaimi A, Lee WAN. Predicting return to work after cardiac rehabilitation using machine learning models. *J Rehabil Med*. 2023:jrm00348.doi:10.2340/jrm.v54.2432.
46. Filos D, Claes J, Cornelissen V, Kouidi E, Chouvarda I. Predicting adherence to home-based cardiac rehabilitation with data-driven methods. *Appl Sci*. 2023;13(10):6120. doi: 10.3390/app13106120.
47. Cannière H D, Corradi F, Smeets CJP, et al. Wearable monitoring and interpretable machine learning can objectively track progression in patients during cardiac rehabilitation. *Sensors*. 2020;20 (12):3601. doi:10.3390/s20123601.
48. Torres R, Zurita C, Mellado D, et al. Predicting cardiovascular rehabilitation of patients with coronary artery disease using transfer feature learning. *Diagnostics*. 2023;13(3):508. doi:10.3390/diagnostics13030508.
49. Jahandideh S, Jahandideh M, Barzegari E. Individuals' intention to engage in outpatient cardiac rehabilitation programs: prediction based on an enhanced model. *J Clin Psychol Med Settings*. Published online 2021 1-10.
50. Lofaro D, Groccia MC, Guido R, Conforti D, Caroleo S, Fragomeni G. Machine learning approaches for supporting patient-specific cardiac rehabilitation programs. In: 2016 Computing in Cardiology Conference (CinC). IEEE; 2016:149-152.
51. Li F, He H. Assessing the accuracy of diagnostic tests. *Shanghai Arch Psychiatry*. 2018;30(3):207-212. doi:10.11919/j.issn.1002-0829.218052.
52. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63. doi:10.7326/M14-0697.
53. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models. *JAMA*. 2017;318(14):1377. doi:10.1001/jama.2017.12126.
54. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*. 2007;115(5):654-657. doi:10.1161/CIRCULATIONAHA.105.594929.
55. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. *Radiology*. 2003;229(1):3-8. doi: 10.1148/radiol.2291010898.
56. Corbacioglu Ş, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. *Turk J Emerg Med*. 2023;23(4):195. doi:10.4103/tjem.tjem_182_23.
57. Futoma J, Simons M, Panch T, Doshi-Velez F, Celi LA. The myth of generalisability in clinical research and machine learning in health care. *Lancet Digit Health*. 2020;2(9):e489–e492. doi:10.1016/S2589-7500(20)30186-2.
58. Huang Y, Li W, Macheret F, Gabriel RA, Ohno-Machado L. A tutorial on calibration measurements and calibration models for clinical prediction models. *Journal of the American Medical Informatics Association*. 2020;27(4):621-633. doi:10.1093/jamia/ocz228.
59. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. *BMJ*. 2012;344: e4181.
60. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. Published online January 8. 2024;e074819. doi:10.1136/bmj-2023-074819.
61. Khadanga S, Savage PD, Gaalema DE, Ades PA. Predictors of cardiac rehabilitation participation. *J Cardiopulm Rehabil Prev*. 2021;41(5):322-327. doi:10.1097/HCR.0000000000000573.
62. Oehler AC, Holmstrand EC, Zhou L, et al. Cost analysis of remote cardiac rehabilitation compared with facility-based cardiac rehabilitation for coronary artery disease. *Am J Cardiol*. 2024;210:266-272. doi:10.1016/j.amjcard.2023.08.061.
63. Graversen CB, Eichhorst R, Ravn L, Christiansen SSR, Johansen MB, Larsen ML. Social inequality and barriers to cardiac rehabilitation in the rehab-North register. *Scand Cardiovasc J*. 2017;51(6):316-322. doi:10.1080/14017431.2017.1385838.
64. Dinesh KG, Arumugaraj K, Santhosh KD, Mareeswari V. Prediction of cardiovascular disease using machine learning algorithms. In: 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), IEEE; 2018:1-7.
65. Kavitha M, Gnaneswar G, Dinesh R, Sai YR, Suraj RS. Heart disease prediction using hybrid machine learning model. In: 2021 6th International Conference on Inventive Computation Technologies (ICICT), IEEE; 2021:1329-1333.
66. Naushad SM, Hussain T, Indumathi B, Samreen K, Alrokayan SA, Kutala VK. Machine learning algorithm-based risk prediction model of coronary artery disease. *Mol Biol Rep*. 2018;45(5):901-910. doi: 10.1007/s11033-018-4236-2.
67. Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*. 2019;7:81542-81554. doi:10.1109/ACCESS.2019.2923707.
68. Srinivasan S, Gunasekaran S, Mathivanan SK, Bam MB, Jayagopal P, Dalu GT. An active learning machine technique based prediction of cardiovascular heart disease from UCI-repository database. *Sci Rep*. 2023;13(1):13588. doi: 10.1038/s41598-023-40717-1.
69. Rath A, Mishra D, Panda G, Satapathy SC. Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomed Signal Process Control*. 2021;68:102820. doi:10.1016/j.bspc.2021.102820.
70. Turk-Adawi K, Sarrafzadegan N, Grace SL. Global availability of cardiac rehabilitation. *Nat Rev Cardiol*. 2014;11(10):586-596. doi:10.1038/nrcardio.2014.98.
71. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. Published online March 18. 2020;m441. doi:10.1136/bmj.m441.
72. Thomas RJ, Beatty AL, Beckie TM, et al. Home-based cardiac rehabilitation: a scientific statement from the American Association of Cardiovascular and Pulmonary Rehabilitation, the American Heart Association, and the American College of Cardiology. *J Cardiopulm Rehabil Prev*. 2019;39:208-225.

73. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63. doi: 10.7326/M14-0697.

74. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. *PLoS One*. 2019;14(2):e0212356. doi:10.1371/journal.pone.0212356.