Eksplorium *p*-ISSN 0854-1418
Volume 46 No. 2, June 2025: 469–482 *e*-ISSN 2503-426X

A Systematic Review of AI Techniques for ECG Analysis in Acute Coronary Syndrome Detection

Thirumurugan E [1], Dr. Nirmala Rajesh[2], Dr. Nagarathna Shenoy[3], Shashidhar Kotian M[4], Dr. Jeyakumar S[5]

[1]Research Scholar, Department of Allied Health Science, Srinivas University, Manglore, Karnataka, India.

E-mail: thiruahs1002@gmail.com

[2]Associate professor, Department of Physiology, School of Allied Health Science, Srinivas University, Manglore, Karnataka, India.

Email: nimmiraj.naidu@gmail.com

[3]Assistant professor, Department of Cardiology, Kasturba Medical College, MAHE, Manipal, Karnataka, India.

Email: nagarathna.nayak@manipal.edu

[4]Research director, Srinivas University, Manglore, Karnataka, India.

Email: mskot_70@yahoo.co.in

[5]Professor & Research supervisor, School of Health Sciences, Garden city university, Banglore, India.

Email: jeysubikshaa@gmail.com

Article Received: 20 Jan 2025, Revised: 15 June 2025, Accepted: 25 June 2025

Abstract: Electrocardiography (ECG) is an essential cardiovascular medicine tool, offering non-invasive information regarding cardiac electrical activity to diagnose many conditions. ECG interpretation, though, can be intricate, laborious, and variable across human interpreters. Artificial intelligence (AI), such as machine learning (ML) and deep learning (DL), has become a game-changer with the potential to enhance the precision, efficiency, and availability of ECG analysis. This systematic review synthesizes the evidence on the diagnostic performance of AI models in ECG interpretation for acute coronary syndrome (ACS), comparing AI's effectiveness with that of human experts and traditional methods. The methodology followed established systematic review guidelines, such as PRISMA, with comprehensive searches across databases like PubMed and Scopus. Included studies utilized high-level ML/DL architectures, e.g., CNNs, and were risk of bias assessed using tools such as QUADAS-2. Diagnostic accuracy measures, e.g., sensitivity, specificity, and Area Under the Curve (AUC), were estimated through bivariate random-effects models. Results show that AI models exhibit high diagnostic accuracy in ACS conditions. For Acute Coronary Syndromes (ACS), AI models recorded AUROCs of as much as 0.997, outperforming human experts in speed and accuracy. Even with these numbers, issues persist, such as AI-associated misdiagnosis, performance variability, and data-related issues like quality and imbalance. Patient privacy and algorithmic bias ethical issues also present challenges. The practical implementation of AI in clinical practice relies on continuous improvement, verification, and coordination among AI vendors, clinicians, and policymakers to overcome the technical and practical challenges.

Keywords: Artificial Intelligence, Electrocardiogram, Acute Coronary Syndrome, ECG Interpretation.

INTRODUCTION:

Electrocardiography (ECG) has been an integral stalwart in the diagnosis and management of cardiovascular disease for a long time. As a low-cost, non-invasive modality, ECG records the complex electrical activity of the heart, yielding critical information regarding its function and concomitant aberrations. Its ease of use and availability have cemented its place as a core tool in various medical settings, such as emergency rooms, intensive care units, and ambulatory clinics. More than 300 million ECGs are conducted every year worldwide, highlighting its ubiquitous presence in clinical practice [1, 2]. The historical importance of computerized ECG interpretation spans more than five decades; it has been a pivotal factor in forecasting cardiovascular morbidity and mortality and relieving the heavy burden on healthcare professionals. The timely and precise interpretation of ECG results is especially important in the early risk stratification of acute and time-dependent illnesses, including ST-elevation myocardial infarction (STEMI) and non-ST-elevation acute coronary syndromes (NSTE-ACS), which can profoundly affect cardiac performance and cause fatal complications [3,4]. The long-term applicability of ECG in today's clinical practice and its non-invasive nature make it a prime candidate for technological innovation to improve diagnostic accuracy and efficacy. The landscape of medical data analysis is being revolutionized with the advent and accelerated spread of artificial intelligence (AI), machine learning (ML), and deep learning (DL) technologies. These advanced computational methodologies demonstrate significant potential in improving the interpretation of complex medical data, particularly within the cardiology domain [5].

AI, a broad discipline encompassing ML and DL, is increasingly being utilized to augment diagnostic and prognostic capabilities by leveraging extensive datasets from diverse sources, including hospital records, electrocardiograms, and echocardiograms [6]. Recent advances in AI, specifically via advanced deep learning models and neural networks, have made it possible to analyze ECG signals with high accuracy and efficiency. Such models have the ability to learn and recognize subtle patterns and deviations in ECG recordings that are associated with different pathological conditions, and in many instances, achieve precision that can even surpass that of human practitioners in certain tasks [7]. The synergistic combination of increasing numbers of large ECG datasets and the ongoing development of increasingly sophisticated deep learning algorithms presents an unprecedented opportunity for highly accurate and automated ECG interpretation [1]. In addition, applications of AI in cardiovascular health go beyond diagnosis; they include prediction of adverse events, i.e., myocardial infarctions and tailoring of treatment regimens, thus significantly impacting the new paradigm of precision medicine [8]. Such transformative value is emphasizing the potential of AI in shaping cardiovascular care in the future.

Rationale for a Systematic Review on AI in ECG Analysis:

The area of artificial intelligence (AI) in electrocardiogram (ECG) analysis continues to grow quickly, generating heterogeneous research findings on a wide range of cardiovascular diseases. Although growth is welcomed, it has resulted in dispersed knowledge, limiting the ability to distinctly perceive AI's diagnostic performance, advantages, and issues for clinical use.

There is indeed a pressing demand for AI in analyzing ECGs because of the sheer quantity of data, with more than 300 million ECGs taken annually across the globe [1].

Manual interpretation, even crucial as it is, takes time and is prone to variability. This only highlights the need for automated AI-based solutions, which can handle big data efficiently, improve diagnostic consistency, and increase access to quality interpretation, particularly in underserved populations. AI also has promise in actively identifying diseases commonly missed by conventional means, such as myocardial infarction, and enabling early detection and treatment [4]. This functionality could change the paradigm to preventive care and alleviate health inequalities. In spite of the potential of AI, there are not many systematic reviews that evaluate its role in detecting specific conditions. Current studies frequently have inconsistencies in methods and performance measures [9].

Thus, this systematic review will attempt to present an up-to-date synthesis of the state of AI in ECG interpretation, filling fundamental gaps and presenting a better understanding of its applications and challenges.

Objectives of the Systematic Review:

- ✓ To assess the diagnostic performance of AI models in identifying certain conditions, like acute coronary syndromes (ACS), directly from ECG records, with a focus on important measures like sensitivity, specificity, and overall diagnostic performance.
- ✓ To comparatively analyze AI models' diagnostic performance and efficiency against human experts and conventional established methods of ECG interpretation.
- ✓ To recognize and discuss extensively the technical challenges, ethical implications (such as patient confidentiality, bias in algorithms, transparency, and accountability), and pragmatic obstacles (such as regulatory barriers, clinician acceptability, and cost) involved with the use of AI in ECG interpretation in the general field of cardiovascular healthcare.

METHODOLOGY:

2.1. Search Strategy and Databases:

A systematic and exhaustive search plan was devised and undertaken to locate all feasible studies relating to the use of artificial intelligence for electrocardiogram (ECG) interpretation.

The major electronic databases widely used were PubMed and Scopus, which were first used to seek out studies within the past ten years that relate to broader ECG interpretive research.

In the case of studies that focused particularly on acute myocardial infarction, the search space was expanded to include other databases, i.e., EBSCO and ProQuest.

The search terms used for these searches were "ECG", myocardial infarction, and "Artificial Intelligence," with language and publication date restrictions not placed to allow a complete review of the literature. The tactical use of multiple databases and more than one set of keywords was planned to maximize the retrieval of useful studies so that the scope and quality of the review were maximized.

2.2 Eligibility Criteria:

Study selection for this systematic review was directed by clearly stated eligibility criteria grouped by the PICOTS-SD framework, which includes Patients, Intervention/Exposure, Comparator, Outcomes, Time, Setting, and Study Design. This organization was used to maintain consistency and pertinence in the review process. The review only considered full-text, peer-reviewed articles written in English.

Included studies had to measure the efficacy of artificial intelligence (AI) in the interpretation of electrocardiograms (ECGs), with particular focus on those that reported quantitative results regarding misdiagnosis rates or comparative studies with regard to AI systems versus other methods.

In the setting of Acute Coronary Syndrome (ACS), only studies that specifically created or tested AI models for the detection of acute coronary syndrome using ECG recordings were included in this review. Some studies were systematically excluded in order to maintain the quality and relevance of this work. Such exclusions involved pilot or feasibility studies, studies that lacked available full-text articles, those articles not published in English. The strict adherence to such eligibility criteria avoided contamination of the synthesis with methodologically unsound and irrelevant studies.

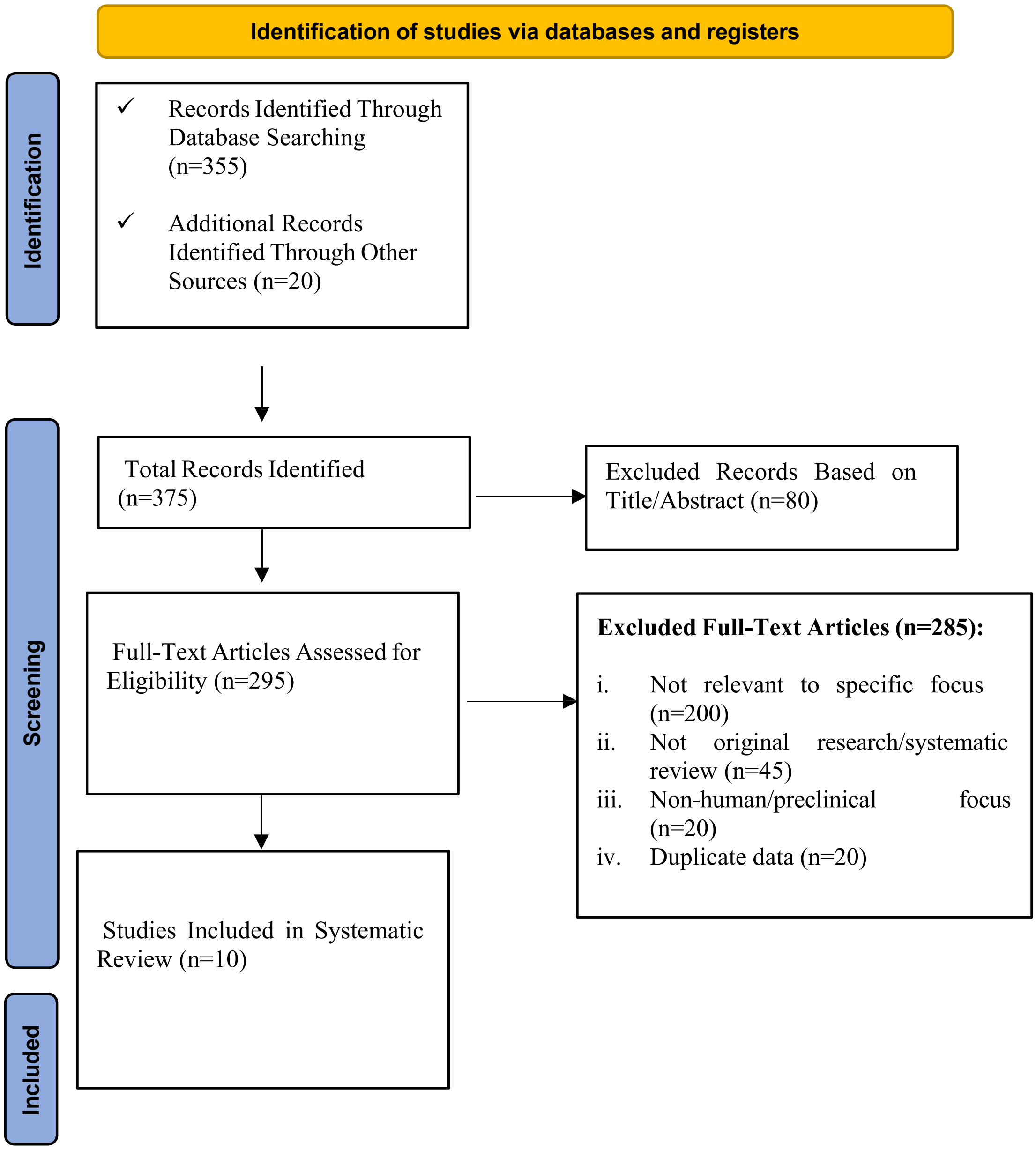
2.3. Process of Study Selection:

The process of study selection was systematic across several stages, strictly adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and recommendations within the Cochrane Handbook for Systematic Reviews of Intervention, version 6.5.6 **(Figure-1).** This multi-stage process facilitated transparency and effectively managed bias across the selection process. First, all the search results from different databases were all brought into reference management software, i.e., EndNote X9, for the easy detection and elimination of duplicate records.

After this deduplication process, titles and abstracts of the remaining possibly pertinent articles were then subjected to an initial screening by two or three independent reviewers.

This initial screening was intended to quickly filter out studies that failed to meet the preliminary inclusion criteria. Next, the complete texts of articles identified as potentially relevant following the initial review were retrieved for a systematic review. Independent reviewers made careful evaluations of each article against predetermined eligibility criteria in this full-text screening. Any reviewer disagreements about the inclusion or exclusion of a particular study were resolved by careful discussion and consensus, and commonly with the use of a third reviewer if needed. The sequential process of this undertaking, usually depicted by a PRISMA flow diagram, guaranteed that only studies that conformed to strict methodological and content standards were finally included in the systematic review.

Figure 1: PRISMA Flow Diagram for Study Selection



2.4. Data Extraction:

A standardized and rigorous data extraction procedure was followed for each study that passed the set eligibility criteria. This systematic method ensured a standardized way of gathering pertinent information, which is the base upon which to build further synthesis and analysis. For every study included, detailed data were extracted on each, including the first author's name and publication year, and the study's detailed characteristics such as the design (e.g., retrospective cohort, prospective observational), geographical region, particular intervention or exposure (e.g., form of AI model, ECG leads used), and the comparison process (e.g., echocardiography, human expert interpretation).

Furthermore, relevant subject demographics and patient populations' clinical characteristics were drawn out to offer background information for the studies.

Notably, the major outcomes, particularly diagnostic test accuracy, were carefully recorded. This included extracting or computing key performance measures, such as the C-index, overall accuracy, sensitivity, specificity, Positive Likelihood Ratio (PLR), Negative Likelihood Ratio (NLR), and Diagnostic Odds Ratio (DOR), and their corresponding 95% Confidence Intervals (CI), wherever such data were present or could be obtained. For meta-analysis, data regarding performance based on test results, that is, true positives, true negatives, false positives, and false negatives, were gathered to enable the creation of 2x2 contingency tables. This extensive data extraction process was critical in ensuring that the findings from the review were strong and reliable.

2.5. Risk of Bias Assessment:

The methodological rigor and possible bias in each study included were carefully assessed using a standardized instrument. In particular, the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used for this strict evaluation. QUADAS-2 thoroughly reviews four key areas: selection of the patient, the index test (AI model), the reference standard (gold standard of diagnosis), and the flow and timing of the study.

For each included study, the risk of bias for these four domains was categorized into one of three groups: low, some concerns, or unclear, based on the specific criteria stated in the QUADAS-2 tool.

Inter-rater reliability was measured to guarantee objectivity and consistency of these assessments using Cohen's kappa statistic, which quantitates the level of agreement between independent reviewers. One of the key issues consistently found in the studies was a high risk of bias in patient selection. For instance, in ACS-detecting Acute Coronary Syndrome (ACS) studies using AI models, "some concerns" were mostly raised about patient selection. These issues were mostly attributed to possible selection bias due to retrospective participant recruitment or lack of sufficient detail about the recruitment processes. This persistent problem with patient selection bias is a major methodological weakness. When AI models are mostly trained and tested on datasets that do not capture the real-world clinical populations they are meant to treat, for example, highly controlled retrospective datasets or datasets from particular demographics, then the high performance reported by such models may not truly represent their generalizability or applicability in real-world clinical practice.

This limitation has obvious implications for the equality of the algorithms; biased datasets can cause disproportionate performance across different patient groups, risking amplifying current health disparities. To overcome this daunting challenge, more prospective, multi-centre studies using careful, transparent, and representative patient recruitment methods are needed to be undertaken. These measures are crucial to create AI models which are truly generalizable and fair in their application to diverse populations of patients.

2.6. Statistical Analysis:

The statistical Analysis performed in this systematic review, specifically in the context of the meta-analysis element, followed strict methodologies for synthesizing estimates of diagnostic accuracy and evaluating heterogeneity across studies. The accuracies of diagnostic tests were pooled on important metrics such as sensitivity, specificity, and the area under the curves (AUCs) of summary receiver operating characteristic (SROC) curves. Pooled estimates of specificity, sensitivity, and diagnostic odds ratio (DOR) were mainly calculated by employing a bivariate random-effects model. Such a model was chosen based on its ability to measure variability at the study level and across studies, providing a more stable overall estimate.

Heterogeneity, which is typical in AI-ECG studies, was thoroughly assessed using conventional statistical tests such as the chi-square test, Cochran Q-test, and the I^2 statistic. The I^2 statistic measures the proportion of variance between studies accounted for by heterogeneity instead of random error. Interpretation of the statistic was guided by conventional methods: an I^2 of 0–40% indicates low heterogeneity; 30–60% indicates moderate heterogeneity; 50–90% indicates high heterogeneity; and 75–100% indicates high heterogeneity. This repeated and widely documented heterogeneity is caused by variations in testing methods, sources of data, types of data, types of algorithm, types of study population, and ECG acquisition technique, impacting the generalizability and reproducibility of any aggregated outcome or general conclusion. This variability implies that comparative studies might be inherently complicated, and a "mean" performance will not necessarily correspond to any one clinical case. This broad methodological heterogeneity highlights an imperative shortfall in standardization among AI-ECG studies, that constitutes a serious obstacle to the process of translating research findings into trustworthy clinical practices. Subgroup analyses were performed wherever feasible to explore the determinants for this heterogeneity. These analyses investigated the impact of factors such as differences among AI model types (i.e., conventional machine learning versus deep learning), data sources (i.e., internal versus external validation datasets), and the number of ECG leads used (i.e., 12-lead versus single-lead configurations). Moreover, sensitivity analyses were undertaken, where required, to test the stability of the pooled estimates with regard to certain study characteristics or exclusions. Finally, publication bias, which may tend to skew the aggregate results of a systematic review, was examined by graphing the logarithms of Diagnostic Odds Ratios (DOR) against square roots of effective sample sizes. A P-value of below 0.05 for the slope coefficient would suggest the presence of significant publication bias.

3. RESULTS:

3.1 **Diagnostic Performance of AI in Acute Coronary Syndromes (ACS) including STEMI and NSTEMI:**

The timely and correct diagnosis of Acute Coronary Syndromes (ACS), involving ST-elevation myocardial infarction (STEMI) and non-ST-elevation acute coronary syndromes (NSTE-ACS), is essential for enabling timely intervention and improving patient outcomes [10]. AI-augmented electrocardiography (ECG) has become a valuable tool in this time-sensitive clinical situation.

A prospective multicenter study has shown that AI-ECG has comparable or even higher diagnostic accuracy and predictive value for Acute Myocardial Infarction (AMI) and 30-day Major Adverse Cardiac Events (MACE) than conventional risk assessment tools, including HEART score, GRACE 2.0 score, and high-sensitivity troponin values, as well as physician diagnoses in emergency departments. Particularly, the AI-ECG had an area under the receiver operating characteristic curve (AUROC) of 0.878 (95% CI, 0.868–0.888) for AMI and 0.866 (95% CI, 0.856–0.877) to predict 30-day MACE with its integration allowing for a large net reclassification improvement of 19.6% [11].

3.1.1 STEMI Detection Performance:

AI models have consistently outperformed human experts and traditional ECG algorithms in STEMI detection. For instance, a CNN-LSTM model by Chen et al. (2022) had an AUROC and accuracy of 99% in STEMI diagnosis from 12-lead mini-ECG devices with a significantly quicker response time of 37.2 ± 11.3 seconds than online doctors, who had a response time of 113.2 ± 369.4 seconds [12]. Zhao et al. (2020) had described an AI program for STEMI diagnosis, based on more than 36 million ECGs, with an AUROC of 0.997 and accuracy of 94%, significantly better than the 80.3% accuracy of 15 cardiologists in a comparison study [13]. In addition, Choi et al. (2022) reported that an AI program (CNN) evaluated through a mobile app outperformed doctors in detecting STEMI with a higher area under the curve (AUC) (0.919 compared to 0.856, $p = 0.004$) as well as improved sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) [14]. Liu et al. (2021) also showed that a deep learning model (DLM) performed with high diagnostic accuracy for STEMI using ECG data alone, with AUC being 0.997, sensitivity of 98.4%, and specificity of 96.9%, which surpassed physician performance [15].

3.1.2 Detection Performance in NSTE-ACS:

AI models have also shown strong effectiveness in detecting NSTE-ACS. An Artificial Neural Network (ANN) developed by Al Zaiti et al. (2020) accurately identified persistent myocardial ischemia from pre-hospital 12-lead ECGs with an AUROC of 0.82, a high negative predictive value of 94% and an acceptable specificity of 76%. This model achieved a 37% improvement in sensitivity against expert clinicians and a 52% improvement compared to commercial ECG algorithms [16]. Wu et al. (2019) also presented an ANN model that attained high AUROC (0.984), sensitivity (90.91%), specificity (93.33%), and accuracy (92.86%) for NSTEMI identification from clinical, laboratory, and ECG data [17]. Moreover, a comparative study on six different machine learning algorithms for NSTEMI diagnosis by Qin et al. (2023) showed the eXtreme Gradient Boosting (XGBoost) model performed best, with an accuracy of 95% and precision of 94% using multimodal data [18].

3.1.3 Detection of Occlusion Myocardial Infarction (OMI) in NSTEMI Patients: A specialized AI model designed by Herman et al. (2024) for detecting OMI on standard 12-lead ECGs in NSTEMI patients exhibited superior accuracy (AUROC of 0.938) and reduced diagnostic time compared to cardiologists applying conventional ECG STEMI criteria [19].

3.1.4 Integrated Data Approaches:

Research integrating AI-ECG with other clinical data sources, such as cardiac troponin levels or comprehensive clinical features, has consistently enhanced diagnostic performance for ACS. For example, the combination of a DLM with traditional cardiac troponin I for NSTEMI identification generated an increased AUC of 0.978. In addition, an SVM system that considered clinical, laboratory, electrocardiographic, and echocardiographic information had an excellent accuracy rate of almost 100% (99.13%) for diagnosing ACS. These results reflect AI's high potential to optimize the diagnostic pathways for ACS by processing and combining varied medical data [20,21].

3.2 **AI vs. Human Experts and Standard Diagnostic Scores:**

An important test of AI in ECG interpretation is its comparative diagnostic performance with human clinicians and standard diagnostic scores. In multiple cardiovascular conditions, AI algorithms have regularly displayed diagnostic performance that is frequently on par with, or even better than, that of human professionals (cardiologists, emergency physicians) and conventional risk stratification strategies.

In the diagnosis of Acute Myocardial Infarction (AMI), AI-ECG has shown diagnostic performance and prediction for AMI as well as 30-day Major Adverse Cardiac Events (MACE) which were similar to or superior to conventional risk stratification models, including HEART score, GRACE 2.0 score, and high-sensitivity troponin, and evaluations by emergency department physicians [3].

In the case of STEMI detection, AI models have consistently outperformed human specialists and commercial ECG algorithms.

For example, AI algorithms have demonstrated significantly quicker response times (e.g., 37.2 ± 11.3 seconds for AI vs. 113.2 ± 369.4 seconds for online physicians).

In addition, AI models have shown better AUROC values (up to 0.997) and accuracies (up to 99%) than cardiologists, whose comparative test accuracy was found to be approximately 80.3% [20,21]. A computer algorithm evaluated through a smartphone application even outperformed doctors in the detection of STEMI, demonstrating a greater AUC (0.919 vs. 0.856, $p = 0.004$) and better sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) [14]. For non-ST elevation acute coronary syndrome (NSTE-ACS), an Artificial Neural Network (ANN) had a 37% increase in sensitivity compared to expert clinicians and a 52% increase compared to commercial ECG algorithms in detecting myocardial ischemia from pre-hospital 12-lead ECGs [16]. Likewise, an AI model for Occlusion Myocardial Infarction (OMI) detection in NSTEMI patients showed better accuracy and shorter diagnostic time than cardiologists following routine STEMI criteria [17].

In addition, the integration of multimodal data and AI models—such as coupling deep learning with cardiac troponin or using Support Vector Machine (SVM) systems that measure clinical, laboratory, ECG, and echocardiographic data—tended to produce diagnostic performance greater than conventional pathways or evaluation from single data modalities (Table-1) [19].

This depicts the capability of AI to excel in interpreting ECGs as well as its capacity to combine information from different sources towards making overall and precise diagnoses.

Cardiovascular Condition	Predominant AI Model Types	Pooled Sensitivity (95% CI)	Pooled Specificity (95% CI)	AUROC/SROC-AUC (95% CI)	Diagnostic Odds Ratio (DOR) (95% CI)	Key Specific Findings
Acute Coronary Syndromes (ACS)	CNN-LSTM, ANN, XGBoost, DLM, SVM	Varies widely by specific ACS type, often >0.90	Varies widely by specific ACS type, often >0.90	Up to 0.997	N/A	Consistently high accuracy; often outperforms human experts and traditional scores; faster response times.

Table-1: Diagnostic Efficacy of Artificial Intelligence Models in the Analysis of Electrocardiograms for Acute Coronary Syndrome.

3.3. Detection of Key Predictors and Feature Importance in AI Models:

AI models, especially deep learning-based models, have the capability to detect and utilize a select set of ECG features or incorporate patient information to render extremely accurate diagnostic predictions. This ability stems from their power to acquire complicated patterns that are either delicate or below the threshold of human detection, thus greatly improving their overall performance. Convolutional neural networks (CNNs), which are widely used in ECG classification, have resulted in strong feature importance rankings. These models always pick out major predictors for cardiac ailments, including ventricular rate, QRS duration, and P-R interval [22,23]. This implies that AI algorithms, similar to human specialists, develop the ability to focus on particular, clinically significant ECG parameters, possibly with more accuracy and reliability, particularly when interpreting large datasets. Along with these established parameters, AI models, particularly those based on deep learning algorithms and neural networks, are skilled at identifying small nuances and anomalies in ECG signals characteristic of many conditions, including electrolyte disturbances. This ability to extract subtle features from raw ECG waveforms, frequently without needing manual feature engineering, is a major strength, partially explaining their high diagnostic accuracy. Further, AI algorithms can be trained on identifying distinctive patterns of coronary artery disease by processing massive image databases, depicting their strength in extracting detailed, high-level features from raw clinical data for early detection and risk stratification [8]. This sophisticated pattern recognition enables AI to detect diagnostic information that traditional interpretation techniques might otherwise miss.

4. DISCUSSION:

The systematic review of the use of artificial intelligence (AI) in electrocardiogram (ECG) interpretation shows an evolutionary shift in cardiovascular diagnosis, marked by great promise and several challenges. The repeated attainment of high diagnostic performance across different cardiovascular conditions, such as acute coronary syndromes, highlights the strong potential of AI. Artificial intelligence algorithms often surpass the ability of human experts and conventional scoring algorithms in terms of speed and diagnostic accuracy, especially in acute conditions like acute coronary syndromes, where prompt interpretation is crucial to achieve desirable outcomes among patients. Such superior performance can be traced back to the ability of AI to analyze vast amounts of data and recognize patterns in ECG signals that might go unnoticed to human operators, improving diagnostic reliability and broadening access to high-quality interpretation. With the enormous amount of ECG data produced on a yearly basis, combined with variation and time-consuming human interpretation, there exists a strong argument to utilize AI-based solutions. This makes AI both an evolutionary and incremental step in cardiovascular diagnosis, ready to handle a rising diagnostic load [1,3]. However, cost-effectiveness as a barrier does not hold back the implementation of AI in clinical ECG interpretation. Although remarkable performance records, the potential for AI-driven misdiagnoses is still a cause of concern. The wide range of AI model performance across different studies and clinical settings with implications for limited generalizability and stability [5] demonstrates some of the limitations. The range is usually compounded by data quality concerns, class imbalance in datasets, and potential for models overfitting training data, thus limiting their performance with new data [23].

One of the salient challenges in the literature is methodological heterogeneity that exists in many studies. Differences in study populations, ECG acquisition protocols, AI model training methods, and assessment processes are some of the factors responsible for the inconsistencies seen in reported results [6]. The absence of standardization makes it difficult to compare different studies directly and hinders the successful translation of research evidence into valid clinical practices. Overcoming this challenge requires constructing and following focused methodological guidelines and reporting requirements for AI in medical imaging and signal processing. Standardizing these would allow for greater comparability and replicability, enabling solid meta-analyses to enhance the credibility of clinical uptake. In addition, a key issue is the reported "bias blind spot" in most studies, most notably concerning patient selection. The risk of patient selection bias—commonly resulting from a lack of details in sampling or the use of retrospective case-control designs—means AI models are likely to be trained and validated largely on unrepresentative datasets [8]. If these models are trained from data within certain demographics or clinical environments, their seemingly high performance is unlikely to translate to practical usefulness when deployed with a more heterogeneous, real-world population. This scenario poses ethical questions over algorithmic bias and equity because unrepresentative datasets have the potential to lead to uneven performance between different patient populations, actually aggravating health disparities rather than helping to eradicate them. Mitigating these issues needs a combined effort at more prospective, multi-centre studies using stringent, transparent, and representative recruitment methods to create truly generalizable and equitable AI models. Ethical implications go beyond algorithmic bias to include such critical patient confidentiality, data protection, transparency, and accountability concerns. The use of large amounts of sensitive patient data by AI systems raises the risk of unauthorized access and data breaches. Further, the "black box" quality of most deep learning algorithms makes it difficult to interpret and explain their suggestions, diminishing clinician confidence and patient trust [4]. Establishing responsibility for mistakes created through AI creates challenging legal and ethical issues. Such issues highlight that AI is used to support human expertise as opposed to automating it, allowing capable professionals to utilize the strength of AI while ensuring that good ethics dictate the incorporation of these innovative technologies.

Conclusion: The systematic review makes visible the extensive potential of artificial intelligence (AI) to revolutionize electrocardiogram interpretation in cardiovascular medicine. AI demonstrates excellent diagnostic accuracy for acute coronary syndrome and tends to outperform conventional methods and human professionals. Its capacity to process large ECG data enables it to identify fine patterns and early detection of diseases, thereby enabling proactive cardiovascular care. Nevertheless, the implementation of AI in clinical practice is not without challenges, such as fears of misdiagnosis, variability in performance, and data quality. There is a requirement of standardization in research and emphasis on algorithmic fairness to avoid health inequalities. Ethical concerns such as patient privacy and transparency of algorithms are important, which calls for the creation of Explainable AI. AI needs to be viewed as a supplement to human expertise and not a substitute. Working together among AI developers, clinicians, and policymakers is necessary to resolve technical issues, create regulatory guidelines, and provide successful training. Future research must focus on enhancing AI's performance while ensuring fairness, transparency, and ethical governance for safe integration into cardiovascular healthcare.

Reference:

[1] A Meta-Transfer Learning Approach to ECG Arrhythmia Detection, accessed June 4, 2025, https://par.nsf.gov/servlets/purl/10379679

[2] www.medrxiv.org, accessed June 4, 2025, https://www.medrxiv.org/content/10.1101/2025.01.09.25320300v1.full.pdf

[3] Artificial intelligence applied to electrocardiogram to rule out acute ..., accessed June 4, 2025, https://academic.oup.com/eurheartj/article/46/20/1917/8037874

[4] AI in ECG-Based Electrolyte Imbalance Prediction: A Systematic Review and Meta Analysis, accessed June 4, 2025, https://www.medrxiv.org/content/10.1101/2025.01.09.25320300v1

[5] ARTIFICIAL INTELLIGENCE MISDIAGNOSES IN ECG ... - JACC, accessed June 4, 2025, https://www.jacc.org/doi/full/10.1016/S0735-1097%2825%2900606-0

[6] Artificial Intelligence and Its Role in the Diagnosis and Prediction of ..., accessed June 4, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC12029043/

[7] AI in ECG-Based Electrolyte Imbalance Prediction: A Systematic Review and Meta Analysis, accessed June 4, 2025, https://www.medrxiv.org/content/10.1101/2025.01.09.25320300v1.full-text

[8] Ethical and clinical implications of AI integration in ... - ResearchGate, accessed June 4, 2025, https://www.researchgate.net/publication/384462224_Ethical_and_clinical_implications_of_AI_integration_in_cardiovascular_healthcare/fulltext/66fadd7b869f1104c6c1568e/Ethical-and-clinical-implications-of-AI-integration-in-cardiovascular-healthcare.pdf?origin=scientificContributions

[9] Efficacy of AI Models in Detecting Heart Failure Using ECG Data: A ..., accessed June 4, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11891813/

[10] (PDF) Artificial Intelligence and Its Role in the Diagnosis and Prediction of Adverse Events in Acute Coronary Syndrome: A Narrative Review of the Literature - ResearchGate, accessed June 4, 2025, https://www.researchgate.net/publication/390064140_Artificial_Intelligence_and_Its_Role_in_the_Diagnosis_and_Prediction_of_Adverse_Events_in_Acute_Coronary_Syndrome_A_Narrative_Review_of_the_Literature

[11] The diagnostic value of electrocardiogram-based machine learning in long QT syndrome: a systematic review and meta-analysis - Frontiers, accessed June 4, 2025, https://www.frontiersin.org/journals/cardiovascularmedicine/articles/10.3389/fcvm.2023.1172451/full

[12] Chen K.W., Wang Y.C., Liu M.H., Tsai B.Y., Wu M.Y., Hsieh P.H., Wei J.T., Shih E.S.C., Shiao Y.T., Hwang M.J., et al. Artificial intelligence-assisted remote detection of ST-elevation myocardial infarction using a mini-12-lead electrocardiogram device in prehospital ambulance care. Front. Cardiovasc. Med. 2022;9:1001982. doi: 10.3389/fcvm.2022.1001982.

[13] Zhao Y., Xiong J., Hou Y., Zhu M., Lu Y., Xu Y., Teliewubai J., Liu W., Xu X., Li X., et al. Early detection of ST-segment elevated myocardial infarction by artificial intelligence with 12-lead electrocardiogram. Int. J. Cardiol. 2020;317:223–230. doi: 10.1016/j.ijcard.2020.04.089

[14] Choi Y.J., Park M.J., Ko Y., Soh M.S., Kim H.M., Kim C.H., Lee E., Kim J. Artificial intelligence versus physicians on interpretation of printed ECG images: Diagnostic performance of ST-elevation myocardial infarction on electrocardiography. Int. J. Cardiol. 2022;363:6–10. doi: 10.1016/J.IJCARD.2022.06.012.

[15] Liu W.C., Lin C.S., Tsai C.S., Tsao T.P., Cheng C.C., Liou J.T., Lin W.S., Cheng S.M., Lou Y.S., Lee C.C., et al. A deep learning algorithm for detecting acute myocardial infarction. EuroIntervention. 2021;17:765–773. doi: 10.4244/EIJ-D-20-01155.

[16] Al-Zaiti S., Besomi L., Bouzid Z., Faramand Z., Frisch S., Martin-Gill C., Gregg R., Saba S., Callaway C., Sejdić E. Machine learning-based prediction of acute coronary syndrome using only the pre-hospital 12-lead electrocardiogram. Nat. Commun. 2020;11:3966. doi: 10.1038/s41467-020-17804-2

[17] Wu C.C., Hsu W.D., Islam M.M., Poly T.N., Yang H.C., Nguyen P.A., Wang Y.C., Li Y.J. An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. Comput. Methods Programs Biomed. 2019;173:109–117. doi: 10.1016/J.CMPB.2019.01.013.

[18] Qin L., Qi Q., Aikeliyaer A., Hou W.Q., Zuo C.X., Ma X. Machine learning algorithm can provide assistance for the diagnosis of non-ST-segment elevation myocardial infarction. Postgrad. Med. J. 2022;99:442–454. doi: 10.1136/POSTGRADMEDJ-2021-141329.

[19] Herman R., Meyers H.P., Smith S.W., Bertolone D.T., Leone A., Bermpeis K., Viscusi M.M., Belmonte M., Demolder A., Boza V., et al. International evaluation of an artificial intelligence–powered electrocardiogram model detecting acute coronary occlusion myocardial infarction. Eur. Heart J. Digit. Health. 2024;5:123. doi: 10.1093/EHJDH/ZTAD074.

[20] Doudesis D., Lee K.K., Boeddinghaus J., Bularga A., Ferry A.V., Tuck C., Lowry M.T.H., Lopez-Ayala P., Nestelberger T., Koechlin L., et al. Machine learning for diagnosis of myocardial infarction using cardiac troponin concentrations. Nat. Med. 2023;29:1201–1210. doi: 10.1038/s41591-023-02325-

[21] Berikol G.B., Yildiz O., Özcan T. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. J. Med. Syst. 2016;40:84. doi: 10.1007/s10916-016-0432-6.

[22] Yoo Jin Choi, Min Ji Park, Yura Ko, Moon-Seung Soh, Hyue Mee Kim, Chee Hae Kim, Eunkyoung Lee, Joonghee Kim, Artificial intelligence versus physicians on interpretation of printed ECG images: Diagnostic performance of ST-elevation myocardial infarction on electrocardiography, International Journal of Cardiology, Volume 363, 2022, Pages 6-10, https://doi.org/10.1016/j.ijcard.2022.06.012.

[23] A systematic review on the impact of artificial intelligence on ..., accessed June 4, 2025,https://www.researchgate.net/publication/386587425_A_systematic_review_on_the_impact_of_artificial_intelligence_on_electrocardiograms_in_cardiology