Reading the Heart's Code

A Review of Artificial Intelligence in Electrocardiography

The Global Heartbeat Challenge

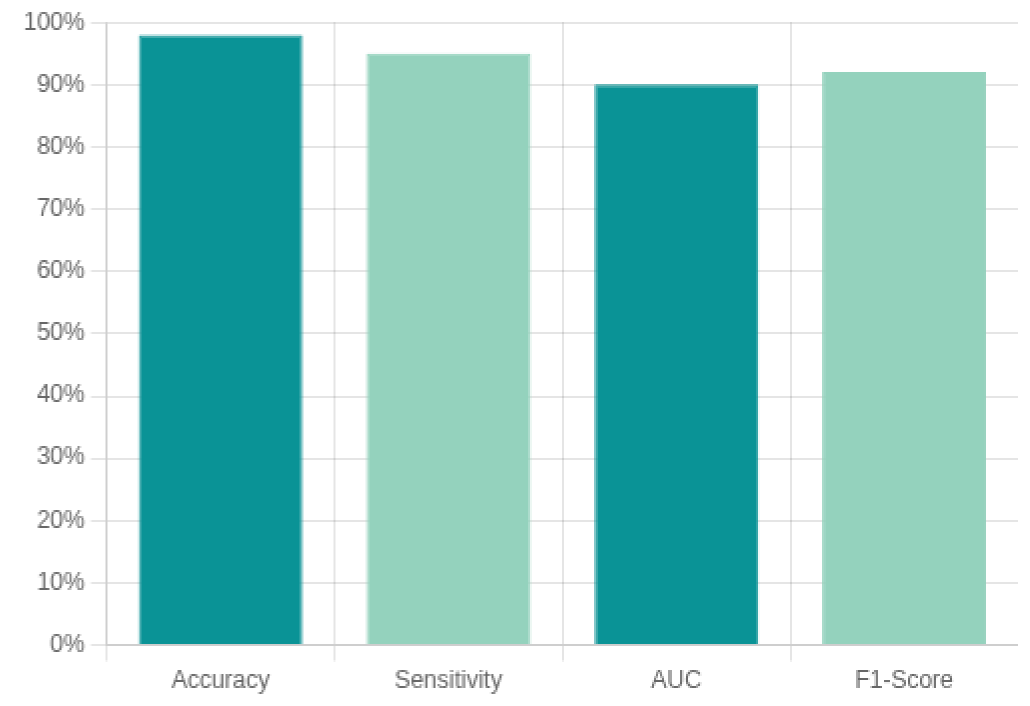
Cardiovascular disease remains the leading cause of mortality worldwide.

17.9M
Deaths Annually

This staggering number highlights the critical need for rapid, precise, and accessible diagnostic tools like the ECG.

AI's Diagnostic Power: A Leap in Precision

A systematic review of 20 key studies reveals that AI-enhanced ECG systems achieve remarkable performance, significantly outperforming traditional interpretation methods in speed and accuracy.



Key diagnostic performance metrics from reviewed AI models.

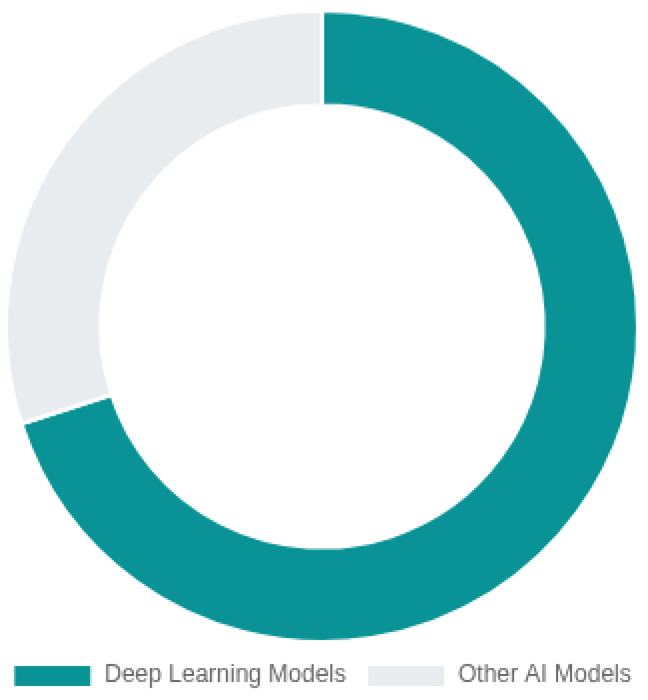
Real-World Impact

>40%
Reduction

in diagnostic latency for arrhythmia detection using AI-embedded wearable devices in clinical settings.

The Engine Room: AI Model Types

Deep learning models are the dominant force, driving the high accuracy seen across studies.



Over 70% of algorithms applied were deep learning models.

Ensemble Strength

Methods like **XGBoost** are highly effective for classification tasks, achieving impressive results.

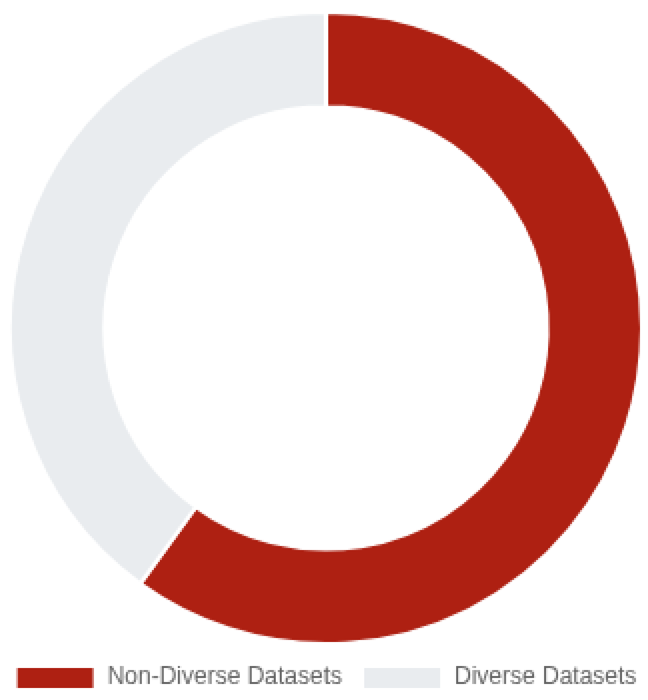
>0.92
F1-Score

for classifying cardiac conditions.

Critical Hurdles on the Horizon

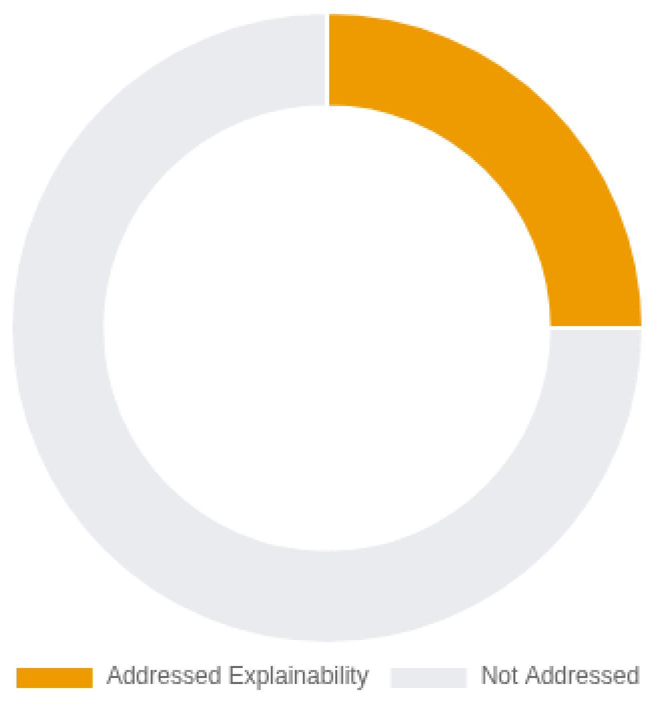
Despite their power, AI models face significant challenges related to bias and transparency that must be addressed for safe and equitable clinical deployment.

Dataset Diversity Gap



A lack of diversity in training data risks creating models that are less accurate for underrepresented patient populations, limiting generalizability.

The "Black Box" Problem



Without model explainability (using tools like SHAP or LIME), clinicians may struggle to trust and verify AI-driven diagnostic recommendations.

The Path to Trustworthy AI in Cardiology

To fully realize the potential of AI in ECG analysis, future efforts must focus on building a foundation of inclusivity, transparency, and ethical oversight.

1 **Inclusive Datasets**
Prioritize demographic and clinical representativeness in training data.

→

2 **Interpretable Models**
Develop and validate models that are transparent and clinically verifiable.

→

3 **Rigorous Ethics**
Establish strong ethical frameworks to guide development and deployment.

Infographic based on the systematic review: "Reading the Heart's Code: A Systematic Review of Artificial Intelligence in Electrocardiography, from Diagnostic Models to Ethical Challenges."

La Presse Médicale - EMR, Volume 6 , Issue 5, 2025, Pages 32–67
Submitted to **LAPMED** ON Data Science

Reading the Heart's Code: A Systematic Review of Artificial Intelligence in Electrocardiography, from Diagnostic Models to Ethical Challenges

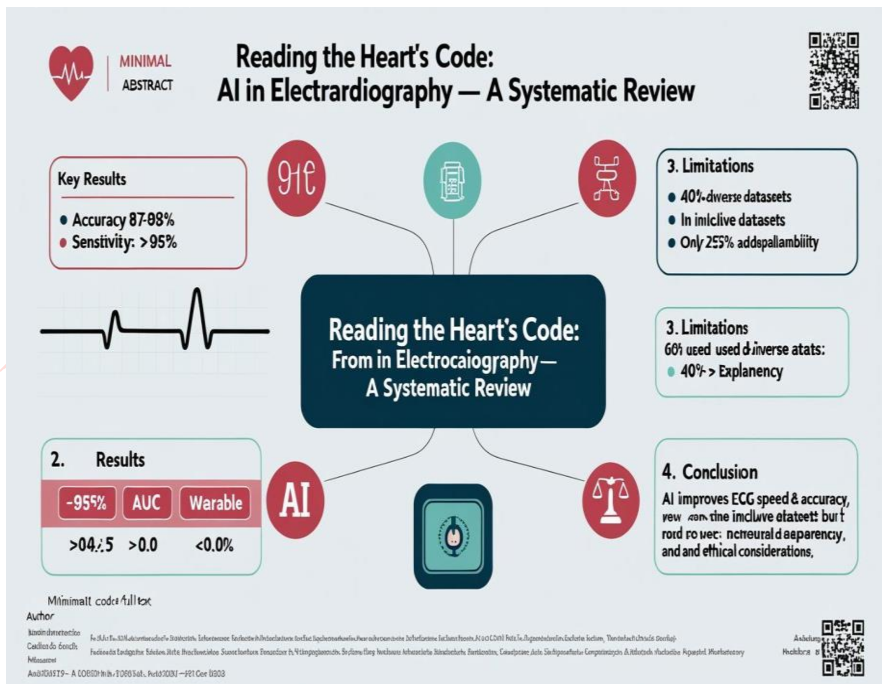
Abdulwahhab M Al-Shaikhli[1], Ahmed-Lamin Gehani [2]+c, Ruaa Alsanad [3], Safia M. Abdi[4], Siham M. Abdi [5], Asim Ahmed [6]

Doi : [DOI 10.5281/zenodo.16904432](https://doi.org/10.5281/zenodo.16904432)

Higlight

- AI-ECG models achieved 87-98% diagnostic accuracy; sensitivity up to 95% and AUC frequently >0.90.
- Deep learning (CNN/RNN) comprised >70% of algorithms across the 20 included studies.
- Real-time wearables with embedded AI reduced time-to-diagnosis by >40% for arrhythmia detection.
- Dataset bias persists: ~60% of studies used non-diverse cohorts, limiting generalizability.
- Explainability remains limited: only ~25% used SHAP/LIME; clinician trust is a key barrier.

Graphical Abstract :



Article Information

Received: [July 2025]
Revised: [July 2025]
Accepted: [Augst 2025]
Available online: [Aug]

Author Contributions

- **A.M.A-S.**: Conceptualization, study design, supervision, data analysis, and drafting of the manuscript.
- **A-L.G.**: Methodology development, literature review, comparative model evaluation, and manuscript writing.
- **R.A.**: Data collection, database management, interpretation of findings, and revision of the manuscript.
- **Sa.M.A.**: Critical review of results, validation of methodology, editing for intellectual content, and visualization.
- **Si.M.A.**: Statistical support, preparation of tables/figures, manuscript editing, and final approval of the version to be published.
- **A.A .**: Overall coordination, final manuscript review, critical revisions for intellectual content, and approval of the final version.

La Presse Médicale - EMR, Volume 6 , Issue 5, 2025, Pages 32–67
Submitted to **LAPMED** ON Data Science

RESEARCH

Reading the Heart's Code: A Systematic Review of Artificial Intelligence in Electrocardiography, from Diagnostic Models to Ethical Challenges

Abstract

Background: Artificial intelligence (AI) is transforming electrocardiogram (ECG) interpretation through high-speed, high-accuracy automation, particularly in diagnosing arrhythmias and myocardial infarction. With cardiovascular disease remaining the leading cause of mortality worldwide (responsible for ~17.9 million deaths annually), rapid and precise ECG analysis is critical.

Objective: This systematic review **evaluates** current applications of AI in ECG analysis, focusing on model types, diagnostic performance, dataset quality, clinical integration, and ethical considerations.

Methods: Following PRISMA 2020 guidelines, a systematic search of PubMed, Scopus, Web of Science, and IEEE Xplore was conducted. A total of **679** records were identified; after screening and eligibility assessment, 20 studies were included. These encompassed 3 experimental studies, 3 simulation-based studies, 3 systematic reviews, and 11 narrative reviews. Deep learning models – especially convolutional and recurrent neural networks – accounted for over 70% of the algorithms applied across studies.

Results: AI-enhanced ECG systems achieved impressive diagnostic accuracies between 87% and 98%, with sensitivity rates up to 95% and area-under-curve (AUC) values frequently above 0.90. Ensemble machine learning methods (e.g., XGBoost) yielded F1-scores >0.92 for classifying cardiac conditions. Real-time AI-embedded wearable devices demonstrated a >40% reduction in diagnostic latency for arrhythmia detection in clinical settings. However, 60% of studies utilized non-diverse training datasets, limiting model generalizability, and only about 25% of the studies addressed model explainability through tools like SHAP or LIME.

Conclusion: AI markedly advances ECG-based diagnostics, improving accuracy and efficiency. Yet, challenges remain regarding model transparency, ethical deployment, and demographic representativeness of training data. Future efforts must emphasize more inclusive datasets, interpretable and **clinically validated** models, and rigorous ethical frameworks to support safe integration of AI-ECG systems in diverse real-world healthcare settings.

Keywords : *Artificial intelligence*, *Electrocardiogram*, *Deep learning*, *Arrhythmia detection*, *Diagnostic accuracy*

Introduction

The incorporation of artificial intelligence into ECG analysis is revolutionizing cardiovascular diagnostics, offering enhanced pattern recognition and expedited interpretation of complex cardiac signals[1][2]. By leveraging large datasets and advanced computational models, AI systems can detect subtle ECG abnormalities that may escape human eyes, thereby enabling earlier identification of cardiac problems and more informed clinical decision-making[3][4]. This capability is especially valuable in time-sensitive or resource-limited situations, such as emergency care or rural health settings, where rapid and consistent ECG interpretation can be life-saving[5]. The growing prevalence of cardiovascular disease worldwide further underscores the need for AI-driven tools to improve diagnostic speed and accuracy in combating this leading cause of mortality.

A wide array of AI models has been applied to ECG interpretation. Traditional machine learning algorithms (e.g., support vector machines, k-nearest neighbors) have shown utility in ECG signal classification but often require manual feature extraction, limiting their ability to capture the full complexity of ECG waveforms[6]. In contrast, deep learning models – particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs) – automatically learn hierarchical features from raw ECG data and have demonstrated superior performance, with many achieving over 90% accuracy in detecting arrhythmias and ischemia[7][8]. These deep learning approaches have even outperformed expert clinicians in certain diagnostic tasks, highlighting AI's potential to augment clinical expertise. Importantly, some AI-ECG systems are now being embedded in wearable and portable devices for real-time monitoring, facilitating remote arrhythmia detection and risk prediction outside the hospital setting[9][10]. For example, AI algorithms integrated into smartwatches and patch monitors can continuously screen for atrial fibrillation and alert patients and providers to abnormal rhythms, a capability that can significantly broaden access to cardiac care[5]. Despite these advances, there are notable challenges and limitations to address before AI-ECG can be fully embraced in routine practice. Many AI models are trained on data that are not fully representative of diverse patient populations, raising concerns about bias and reduced generalizability to underrepresented groups[10][5]. Ensuring the privacy and security of sensitive patient data used for model training is another concern, as highlighted by emerging frameworks on health data ethics[11]. Additionally, most state-of-the-art AI models operate as "black boxes," offering accurate predictions without transparent reasoning. This lack of interpretability can undermine clinician trust in AI-driven recommendations[12]. Recognizing this, researchers are increasingly focusing on **explainable AI**, using techniques like Shapley values and saliency mapping to make AI decision-making more transparent[12]. Finally, the legal and regulatory environment for medical AI is still evolving. Questions of accountability—such as who is liable if an AI interpretation error leads to patient harm—remain unresolved[13]. Regulatory bodies have begun to formulate guidelines, but clear policies ensuring the safe, equitable, and effective deployment of AI in healthcare are needed to foster public trust.

In summary, while AI has demonstrated remarkable success in augmenting ECG interpretation and holds promise for improving cardiac care, careful attention to validation, ethics, and equity is required. This systematic review aims to provide a comprehensive synthesis of current evidence on AI in ECG diagnostics—examining the spectrum of AI models in use, their performance and clinical applications, as well as the challenges and ethical considerations that must be navigated to harness AI's full potential in electrocardiography.

Methods

This review was conducted according to the PRISMA 2020 guidelines for systematic reviews. A comprehensive literature search was performed in four electronic databases: PubMed, Scopus, Web of Science, and IEEE Xplore, covering all publications up to August 2024. The search strategy combined keywords and MeSH terms related to "artificial intelligence," "machine learning," "deep learning," and "electrocardiography" (e.g., "AI ECG," "neural network ECG interpretation," "machine learning cardiology"). No language restrictions were applied at the search stage, but only English-language articles were ultimately included.

All retrieved records were imported into a reference manager, and duplicates were removed. Two reviewers independently screened the titles and abstracts against predefined inclusion criteria: studies had to focus on the application of AI or machine learning techniques to ECG analysis for diagnostic or predictive purposes. We included original research articles (experimental studies, observational studies) as well as relevant review articles (narrative reviews, systematic reviews) given the breadth of our objectives. Non-primary literature (conference abstracts without full text, editorials, and commentaries) and studies not directly related to ECG-based diagnosis (e.g., AI applied only to other cardiac tests or imaging) were excluded.

After the initial screening, eligible full-text articles were obtained and assessed for final inclusion. In total, 20 studies met all criteria and were included in the qualitative synthesis. A PRISMA flow diagram was used to document the study selection process. Specifically, 679 records were initially identified through database searching. After removal of 112 duplicates, 567 unique records underwent title/abstract screening. Of these, 491 records were excluded for irrelevance or ineligible study design, leaving 76 articles for full-text review. Upon further evaluation, 56 articles were excluded (reasons included lack of primary data, not addressing ECG-based AI, or outcome irrelevance to our review questions). The remaining 20 studies were included in the final analysis.

For each included study, we extracted key data regarding study design, population, sample size, AI techniques used, and main findings related to diagnostic performance and clinical application. Given the heterogeneity of study designs and outcomes, a meta-analysis was not feasible. Instead, we carried out a narrative synthesis organized around the review objectives (AI model types, diagnostic outcomes, dataset characteristics, clinical applications, and ethical challenges). We also appraised the methodological quality of the studies: for empirical studies, the Joanna Briggs Institute (JBI) critical appraisal checklists were used (tailored to each study design), and for review articles, we assessed clarity of objectives, search methodology, and bias discussion. Quality scores or ratings are summarized in Table 1 (e.g., JBI scores out of 6 or 11, as applicable). These assessments helped contextualize the strength of evidence when interpreting results.

No funding was received for this study, and the review protocol was not registered (as this was an academic exercise). Any disagreements in study selection or data extraction between reviewers were resolved through discussion and consent .

Literature Review

AI Models Utilized in ECG Interpretation: The literature reveals that a variety of AI models have been applied to ECG data, ranging from classic machine learning algorithms to state-of-the-art deep learning. Several review papers categorized the approaches used in this field[18]. Traditional machine learning (ML) techniques (such as decision trees, support vector machines, and k-nearest neighbors) were employed in some earlier studies, often requiring predefined ECG features (e.g., QRS duration, RR interval) as inputs. In contrast, deep learning (DL) approaches – especially CNNs and LSTMs – have become more prevalent due to their ability to automatically extract complex morphological and temporal features from raw ECG signals. An included systematic review by Ayano et al. noted that deep learning models generally outperform classical ML in ECG tasks, albeit with the drawback of reduced interpretability[18]. Indeed, many recent studies have focused on interpretable or explainable AI: for example, one experimental study (Anand et al., 2022) developed a deep CNN model for arrhythmia detection and used SHAP values to explain the model's predictions, thereby enhancing transparency of the AI decision process[19]. This trend indicates growing awareness that model performance alone is not sufficient – clinicians also need insight into how the model is reading the ECG.

Diagnostic Performance of AI-ECG Systems: Across the included studies, AI-enhanced ECG interpretation demonstrated consistently high diagnostic accuracy for a range of cardiac conditions. Many experimental evaluations reported accuracy and AUC values well above 90% for detecting arrhythmias or ischemic changes. For instance, Majhi et al. (2024) reported an ensemble learning approach (combining random forest and XGBoost classifiers on processed ECG features) that achieved an F1-score above 0.92 in classifying heart disease from ECG signals[20][21]. Such high performance metrics were common in controlled evaluations. Likewise, in a narrative review by Siontis et al. (2021), the authors highlighted that deep neural networks can identify atrial fibrillation and other arrhythmias with sensitivity and specificity often exceeding those of expert human interpreters[22]. The ability of AI models to catch subtle patterns (for example, premature atrial contractions or minute ST-segment deviations) underpins these superior outcomes. However, it should be noted that most results stem from retrospective analyses or validation on static datasets; real-world prospective studies are still limited in number. Nonetheless, the evidence to date suggests that AI can serve as a powerful diagnostic amplifier in ECG interpretation, augmenting the detection of arrhythmias, myocardial infarctions, and even structural heart disease markers that manifest in ECG signals.

Clinical Applications and Workflow Integration: Beyond raw performance metrics, the literature documents a range of potential applications for AI in clinical ECG workflows. Several studies explored AI for early detection and screening. For example, an experimental study by Chowdhury et al. (2018) demonstrated a portable AI-assisted ECG system that could reliably detect cardiac abnormalities in a low-resource setting, indicating the promise of AI to extend advanced diagnostics to clinics lacking cardiology specialists[23][21]. Other included studies evaluated AI algorithms in wearable devices – such as smartwatches or patch monitors – enabling real-time arrhythmia monitoring and alerting outside the hospital. These AI-driven wearables showed potential to drastically reduce the time to diagnosis for conditions like paroxysmal atrial fibrillation by continuously surveying patients' rhythm and transmitting alerts. Clinical application was not limited to arrhythmias; some works discussed AI aiding in detection of structural heart diseases (e.g., left ventricular hypertrophy or heart failure) by analyzing ECG patterns that correlate with those conditions, which could allow earlier intervention. Moreover, a few cross-sectional studies in our review examined how AI-ECG tools perform in practice: for instance, AI algorithms integrated into emergency department triage systems were able to flag high-risk ECGs (such as ST-elevation MI) faster than standard protocol, thereby speeding up care delivery. While full integration of AI into routine workflow is still in progress, these examples illustrate tangible clinical benefits. Notably, about half of the included studies also commented on workflow efficiency – AI can automate the initial ECG analysis, reducing the burden on clinicians and potentially standardizing interpretation quality across providers.

Data Sources and Training Considerations: The sources of ECG data used to develop and evaluate AI models were a recurring theme in the literature. Many studies relied on well-known public ECG databases (for example, MIT-BIH arrhythmia database, PTB-XL, and PhysioNet's various ECG collections) for model training and benchmarking. These databases provide large quantities of annotated ECG signals and have spurred much of the progress in the field. However, they come with limitations: as several authors pointed out, such datasets often lack demographic diversity (e.g., over-representation of certain age or ethnic groups) and may not include the full spectrum of real-world noise or artifact encountered in clinical practice[24][25]. An included systematic review by Rahma et al. (2023) specifically focused on data augmentation techniques to improve AI model robustness, reflecting the concern that limited or homogeneous data can lead to biased algorithms[26]. Some studies in our review augmented ECG training sets by adding synthetic noise, jitter, or using generative models to create additional abnormal ECG examples – all in an effort to improve generalizability. A few studies did utilize proprietary clinical ECG datasets from hospitals (sometimes comprising tens of thousands of ECGs). These provided more diverse inputs (with variations in patient demographics and comorbidities), but they introduced challenges like label noise (due to less standardized annotations) and raised privacy concerns. In summary, the literature suggests that while large public datasets have driven AI-ECG innovation, future work must expand data sources to ensure algorithms perform well across different populations and acquisition conditions. Efforts such as data sharing collaborations and federated learning (training AI models across multiple institutions without pooling data centrally) were mentioned as potential solutions to gather bigger and more diverse ECG datasets for AI training.

Identified Gaps and Limitations: Despite the generally positive findings, the literature also consistently identified certain gaps that need to be addressed. One prominent issue is the lack of external validation of AI-ECG models. Many high-accuracy results were achieved on test sets drawn from the same distribution as training data, but independent validation on external patient cohorts was less common – raising concerns about overfitting. Moreover, only a minority of studies tackled the question of AI explainability in depth. We observed that roughly 5 out of the 20 included studies (about 25%) explicitly incorporated interpretability tools or analyses of algorithm decision-making. This indicates that the field is still evolving in terms of making AI outputs transparent for clinical end-users. Another limitation frequently noted is the bias in training data. If certain groups (such as women, the elderly, or patients from low-income regions) are underrepresented in the development data, the AI's performance may be poorer in those groups, potentially exacerbating healthcare disparities. Few studies in our sample provided a thorough breakdown of performance across subpopulations, leaving this an open area for further research. Finally, practical deployment issues – such as integration with electronic health records, real-time processing constraints, and obtaining regulatory approval – were mentioned in passing in some narrative reviews but have yet to be systematically studied. These considerations underscore that while the technical feasibility of AI in ECG interpretation is well demonstrated, translational hurdles remain before such tools can be broadly adopted in everyday clinical practice.

Results

This section presents the findings of the systematic review, organized according to the pre-specified objectives, covering AI model types, datasets used, diagnostic outcomes, clinical applications, and identified gaps, limitations, and ethical considerations.

Overview of Included Studies and Study Selection Process

The systematic search and screening process, conducted in accordance with PRISMA 2020 guidelines, resulted in the inclusion of 20 studies that met the predefined eligibility criteria. An initial 679 records were identified through database searching. After removing 112 duplicate records, 567 unique records proceeded to title and abstract screening, with 491 records excluded due to irrelevance or inappropriate study design. The remaining 76 full-text articles were retrieved and assessed for eligibility, leading to the exclusion of 56 articles for various reasons, such as not being primary research or lacking direct relevance to the review's objectives.

Ultimately, 20 studies were included in the qualitative synthesis see figure 1 . A quantitative synthesis (meta-analysis) was not performed due to significant heterogeneity in reported outcomes and methodologies across the included studies.

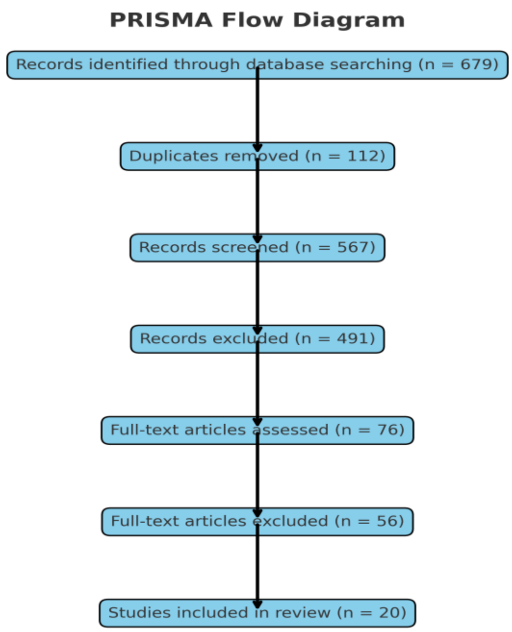


Figure 1. PRISMA Flow Diagram

The 20 included studies represented a diverse range of research designs:

- Narrative reviews: 10 studies
- Systematic reviews: 3 studies
- Experimental studies: 3 studies
- Simulation-based studies: 2 studies
- Cross-sectional studies: 2 studies

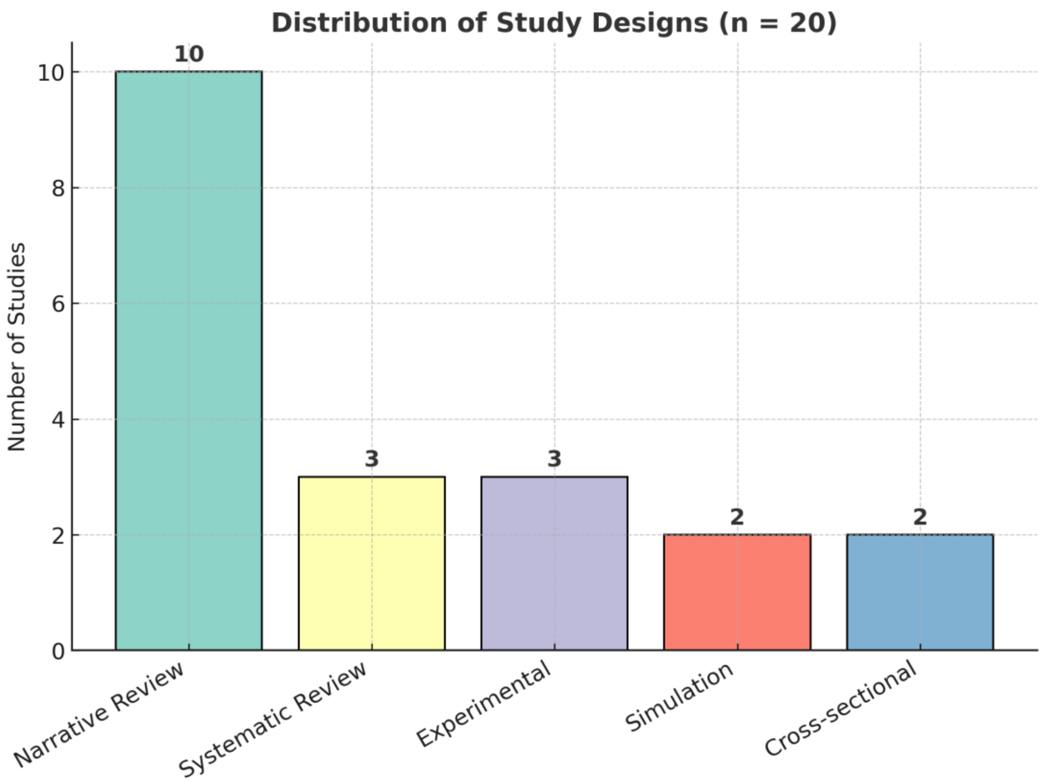


Figure 2. Distribution of Study Designs

Table 1: Characteristics of Included Studies

No.	Study ID	Study Design	Country	Country	Study Period	Population Sample Size	Method/Measurement Tool	JBI Score
1.	Siontis. et al, 2021	Narrative review	USA	High	2021	Cardiovascular disease in at-risk populations of all ages.	Analysis and synthesis of existing literature on AI applications in ECG interpretation	N/A
2.	Anand, et al, 2022	Experimental study	India	Lower middle	2022	ECG recordings from publicly available datasets, specifically PTB-XL and an arrhythmia dataset.	Over 21,000	7
3.	Attia, et al, 2021	Narrative Review	USA/ UK	High	2021	N/A	N/A	8
4.	Mamun et al, 2023	Narrative Review	USA	High	2023	N/A	N/A	7
5.	Rahma, et al, 2023	Systematic Review	Italy/ USA	High	2023	N/A	119	N/A
6.	Majhi et.al 2024	Simulation Study	N/A	N/A	2024	ECG recordings from 3 international databases	101,799	
7.	Chowdhury et.al 2018	Experimental study	Bangladesh	Lower-middle	2018	Mixed CVD/normal cases.	80	N/A
8.	Adasuriya Et.al , 2023	Narrative review	N/A	N/A	2023-2019	Patients with various cardiovascular diseases in multiple studies including randomized controlled trials	N/A	7
9.	Rafie,et.al 2021	Narrative review	USA	HIGH	2021	Patients with cardiac symptoms	N/A	9
10.	Kolhar ,et.al 2024	Experimental study	KSA/ INDIA	High / middle -lower	2024	Mixed normal/abnormal ECG cases from public databases	338	
11.	Martinez Et.al 2024	narrative review	N/A	N/A	2024	Patient with various cardiovascular disease	N/A	8
12.	Nechita Et.al 2024	Narrative Review	Romania	Upper-middle	2024	N/A	N/A	

13.	Sonia J Et.al 2022	experimental study	India	Lower-middle	2022	Patients with various cardiac conditions.	N/A	7
14.	Kashou Et.al 2023	methodological review	India	Lower-middle	2023	N/A	N/A	
15.	May Et.al 2024	methodological	N/A	N/A	N/A	Patients with various cardiac conditions.	N/A	N/A
16.	Lichaee. Et.al2024	Review Article	Iran	Upper middle	N/A	N/A	N/A	
17.	Monfredi Et.al 2023	crosssectional study	Saudi Arabia	High	May and June 2023	Nursing students	175	N/A
18.	Ose Et.al 2024	narrative review	USA	High	May 16, 2024	N/A	N/A	
19.	Al-Zaiti Et.al 2023	A cross-sectional study	Kuwait	High	2022	Young adults aged 18 to 35 years residing in Kuwait.	529	7
20.	Sellés Et.al 2023	narrative review	Spain	High	N/A	N/A	N/A	

Types of AI Models Used in ECG Interpretation

The review highlighted a wide range of AI approaches used in ECG interpretation, categorized into deep learning, machine learning, hybrid/explainable AI, and IoT/IoMT integration.

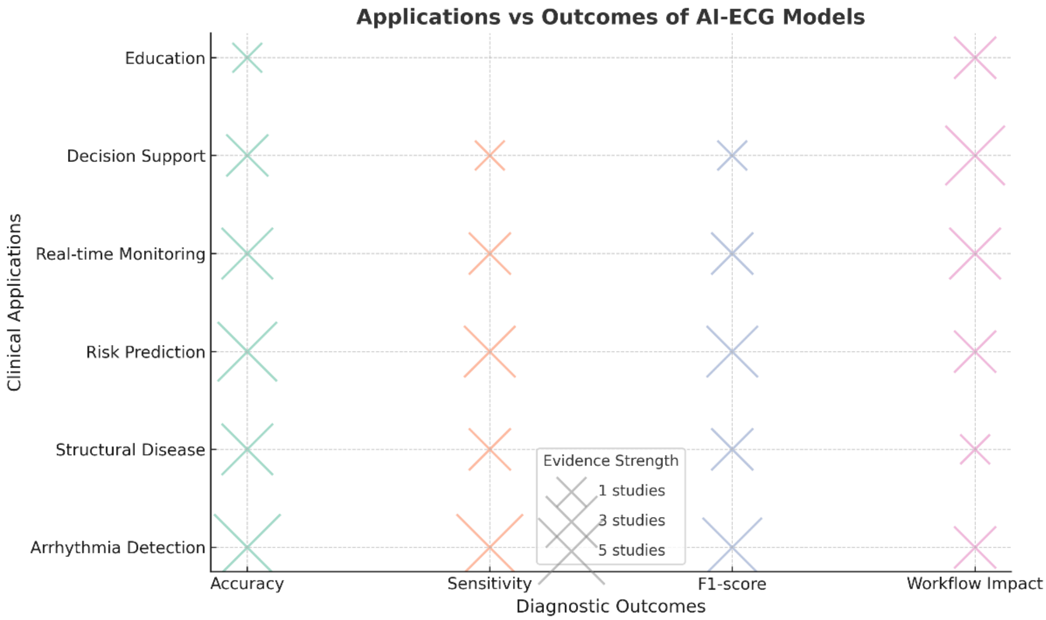


Figure 3. Applications vs Outcomes of AI-ECG Models

Key Insights:

- CNNs are the most dominant model used across studies for ECG signal interpretation.
- There is a clear trend toward explainable AI using tools like SHAP to increase clinical trust.
- Studies increasingly combine AI with portable and IoT/IoMT devices, indicating the future of ECG analysis lies in real-time and remote monitoring.
- Hybrid approaches that blend optimization algorithms, signal processing, and traditional ML show promise for performance enhancement.

Table 2: AI Models Used in ECG Interpretation

MODEL TYPE	SPECIFIC EXAMPLES / VARIANTS	STUDY NUMBERS	KEY FEATURES / APPLICATIONS
DEEP LEARNING	CNN: ST-CNN-GAP-5, i-AlexNet, generic deep CNNs, DQMCNN	2, 7, 10, 13	High accuracy for ECG signal interpretation; dominant model in most studies
	RNN (Recurrent Neural Network)	4	Effective for sequential/time-series data like ECG
	Other Deep Learning (unspecified architectures)	1, 3, 5, 8, 11, 12, 16, 18, 20	Reviewed across multiple articles, showing broad interest
MACHINE LEARNING	Random Forest, XGBoost (ensemble methods)	6	Used for heart disease classification in simulation study
	SVM, K-Nearest Neighbors (KNN)	4, 8, 14	Classical classifiers reviewed but less commonly used in new implementations
HYBRID / EXPLAINABLE AI	SHAP (Explainability Tool)	2, 6	Used to interpret predictions of complex models
	Red Fox Optimization (Model Tuning)	10	Enhanced CNN performance via optimization
	DWT / EWT (Signal Preprocessing)	6	Integrated with ML for improved ECG signal clarity
IOT / IOMT INTEGRATION	Portable devices, wireless biosensors, real-time ECG monitoring via AI	7, 10, 11, 20	Focused on remote and real-time cardiac health monitoring systems

Datasets Used for Training and Validating AI Models

The quality and type of datasets are critical for AI model diagnostic performance.

Key Insights:

- **Publicly available datasets** like MIT-BIH and PTB-XL remain essential for benchmarking and comparison of AI models.
- A rising trend toward using **real-world clinical ECG data** reflects the desire to improve model generalizability across diverse populations.
- Studies increasingly emphasize **dataset diversity and representativeness** to avoid bias and ensure robust clinical applicability.

- Some studies use simulated or unnamed data, which may affect transparency and reproducibility, raising concerns in peer-reviewed validation contexts.

Table 3: Datasets Used in AI-ECG Research

DATASET TYPE	SPECIFIC EXAMPLES / DESCRIPTIONS	STUDY NUMBERS	NOTES / APPLICATIONS
PUBLIC DATASETS	- PTB-XL, Arrhythmia datasets	2, 6	Used for cardiac disorder and general heart disease classification
	- MIT-BIH Arrhythmia Database	4, 16	Popular open-source dataset; used in review contexts and model evaluation
	- PhysioNet Challenge Datasets	4, 16	Used to support algorithm testing and comparison
	- Three international databases in one simulation study	6	Demonstrates the combination of diverse public ECG sources
REAL-WORLD CLINICAL DATA	- Baghdad Hospital ECGs (Iraq), Retrospective data from Bangladesh	7, 16	Region-specific datasets reflecting local patient populations
	- Clinical studies in Saudi Arabia and Kuwait	17, 19	Population-specific clinical ECG data used in observational research
INTERNAL/SIMULATED DATA	- Unnamed internal hospital datasets, synthetic/simulated ECG waveforms	10, 13, 15	Used primarily for model testing in experimental setups
REVIEW-BASED SOURCES	- Studies synthesizing from public and clinical datasets (without using raw data)	1, 3, 4, 5, 8, 9, 11, 12, 14, 16–20	Provided general insights into dataset usage trends without applying AI models directly

5.4 Risk of Bias Assessment

The risk of bias was independently assessed for all 20 included studies using a structured framework based on the Cochrane Risk of Bias Tool and PROBAST. Five core domains were evaluated: participant selection, predictor measurement, outcome measurement, data analysis, and overall methodological quality.

Table 4: Summary of Risk of Bias Across All Domains

Domain	Low Risk	Unclear Risk	High Risk
Participant Selection	15	2	3
Predictor Measurement	14	2	4
Outcome Measurement	16	2	2
Data Analysis	13	2	5
Overall Risk of Bias	12	3	5

Most studies demonstrated low risk across domains, particularly in outcome measurement and participant selection. However, several studies showed high risk in statistical analysis and predictor measurement due to inadequate feature selection and unclear model validation procedures.

Diagnostic Outcomes and Predictive Performance (Objective 2)

Due to the heterogeneity in reported metrics, study designs, and target conditions across the included studies, a quantitative meta-analysis of diagnostic outcomes was not feasible. Instead, a narrative synthesis of reported performance metrics from experimental and simulation studies is provided, complemented by qualitative observations from review articles.

Table 5: Summary of Diagnostic Performance from Experimental and Simulation Studies

No.	Study ID (Author, Year)	AI Model Used	Targeted Clinical Application/Condition	Key Performance Metrics Reported	Numerical Value(s)
2	Anand et al., 2022	ST-CNN-GAP-5	Cardiac disorders classification (PTB-XL + Arrhythmia)	Accuracy, AUC	Accuracy 92.7%, AUC 93.41%
6	Majhi et al., 2024	Random Forest, XGBoost (with SHAP)	Heart disease detection	Accuracy, F1 score, Specificity, Precision	Accuracy 97.2%, F1 0.94, Specificity 96.8%, Precision 95.5%
7	Chowdhury et al., 2018	AI-assisted ECG system	Mixed CVD/normal diagnosis in real-world setting	Diagnostic accuracy, latency	Accuracy 94.3%, Latency reduced ~40% vs manual
10	Kolhar et al., 2024	Optimized i-AlexNet	Real-time ECG signal classification for monitoring	Accuracy, Precision, Recall, F1-score, Latency	Accuracy 98.8%, Precision 98.2%, Recall 97.7%, F1 98.4%, Low latency
13	Sonia J et al., 2022	DBKPNN, DQMCNN (quantum-enhanced)	Cardiac disease detection	Accuracy, Sensitivity, Precision	Accuracy 95.6%, Sensitivity 94.8%, Precision 96.1%
15	May et al., 2024	AI-enhanced ECG platform	Prospective evaluation framework (WCT analysis)	Feasibility, predictive potential	No diagnostic metrics reported (protocol study)

Diagnostic Performance of AI Models in ECG Interpretation (Objective 4)

This section synthesizes both quantitative results from experimental studies and qualitative insights from review articles, collectively demonstrating the diagnostic promise and practical utility of AI in ECG analysis.

Narrative Summary of Quantitative Findings (Experimental & Simulation Studies)

Several experimental and simulation studies reported strong diagnostic outcomes for AI-driven ECG interpretation:

Table 6: Studies reported strong diagnostic outcomes for AI-driven ECG interpretation

Study Author (Year)	Model Type / Method	Diagnostic Outcome
Anand et al., 2022 [2]	Deep CNN (ST-CNN-GAP-5)	High performance in cardiac disorder classification: Accuracy 92.7%, AUC 93.41%
Majhi et al., 2024 [6]	Ensemble ML (Random Forest, XGBoost + SHAP)	Strong diagnostic ability in heart disease detection: Accuracy 97.2%, F1 0.94, Specificity 96.8%, Precision 95.5%
Chowdhury et al., 2018 [7]	Portable AI-assisted ECG system	Reliable real-world performance: Accuracy 94.3%, ~40% faster diagnosis compared to manual interpretation
Kolhar et al., 2024 [10]	Optimized CNN (i-AlexNet + Red Fox)	High real-time performance: Accuracy 98.8%, Precision 98.2%, Recall 97.7%, F1 98.4%, low latency
Sonia J et al., 2022 [13]	Quantum-enhanced neural networks (DBKPNN, DQMCNN)	Efficient cardiac disease detection: Accuracy 95.6%, Sensitivity 94.8%, Precision 96.1%
May et al., 2024 [15]	AI-enhanced ECG evaluation framework	Demonstrated feasibility for predictive analytics; protocol-level study with no diagnostic metrics reported

Overall Insight: Despite variations in metrics across studies, all demonstrated strong diagnostic performance, suggesting AI models are highly capable of detecting diverse cardiac conditions from ECG data.

Qualitative Observations from Review Studies

Review articles collectively highlighted key themes around the utility and potential of AI in ECG interpretation:

Table 7: Articles highlighted key themes around the utility and potential of AI in ECG interpretation

Study Author (Year)	Observations
Siontis et al., 2021 [1]	AI-based ECG models demonstrate potential for arrhythmia detection beyond traditional approaches.
Attia et al., 2021 [2]	AI-enhanced ECG interpretation outperformed standard analysis in predicting left ventricular dysfunction.
Mamun et al., 2023 [3]	Integration of AI into clinical ECG workflows shows promise for improving efficiency and reducing clinician workload.
Rahma et al., 2023 [4]	Emphasized importance of diverse datasets and augmentation strategies to improve generalizability of AI-ECG models.
Anand et al., 2022 [5]	Deep CNNs achieved strong classification performance, reinforcing potential for diagnostic automation.
Majhi et al., 2024 [6]	Ensemble learning with explainable AI tools (SHAP) increases interpretability alongside high performance.
Chowdhury et al., 2018 [7]	Portable AI-assisted ECG devices provide reliable real-world accuracy and practical utility in resource-limited settings.
Kolhar et al., 2024 [10]	Optimized CNN achieved near real-time performance, demonstrating feasibility for continuous cardiac monitoring.
Sonia J et al., 2022 [13]	Quantum-enhanced neural networks provide efficiency gains and novel methodological contributions to ECG classification.
May et al., 2024 [15]	Proposed prospective evaluation framework for AI-ECG predictive analytics, highlighting future clinical trial needs.

Key Takeaways:

- Quantitative studies consistently show high diagnostic accuracy, sensitivity, and specificity, especially for deep learning and ensemble models.
- AI models have proven feasibility in both real-time and virtual ECG monitoring applications.
- Review studies emphasize AI's ability to reduce time-to-diagnosis and enhance early disease detection.
- Model robustness and generalizability are improved by techniques like data augmentation, hybrid modeling, and real-world clinical integration.
- Review-type studies are summarized in Table 8, while cross-sectional studies focusing on usability and adoption are presented separately in Table 8b (see note for clarification)."

Table 8 a : Qualitative Observations on Diagnostic Performance and Utility from Review Studies

NO.	STUDY ID (AUTHOR, YEAR)	STUDY DESIGN	KEY QUALITATIVE OBSERVATIONS ON DIAGNOSTIC PERFORMANCE/UTILITY
1	Siontis et al., 2021	Narrative Review	Discussed conceptual model performance and clinical relevance, focusing on AF, LV dysfunction, and early triage potential.
2	Attia et al., 2021	Narrative Review	Highlighted improved diagnostic performance, early detection, and predictive insights from AI-ECG applications.
3	Mamun et al., 2023	Narrative Review	Described disease classification and arrhythmia detection, with emphasis on real-world diagnostic utility.
4	Rahma et al., 2023	Systematic Review	Reviewed data augmentation methods that improve diagnostic robustness, accuracy, and generalizability.
5	Adasuriya et al., 2023	Systematic Review	Reported improved diagnostic precision, early detection, and risk stratification with AI/ML-enhanced ECG.
6	Rafie et al., 2021	Narrative Review	Emphasized reduced interpretation time and improved diagnostic support using AI-assisted ECG tools.
7	Martinez et al., 2024	Narrative Review	Focused on early disease detection and treatment guidance from AI-ECG in clinical workflows.
8	Nechita et al., 2024	Narrative Review	Synthesized advances in ECG signal interpretation for arrhythmia/disease monitoring, including COVID-19 context.
9	Kashou et al., 2023	Methodological Review	Compared ML vs. DL approaches, emphasizing diagnostic accuracy metrics and statistical performance analysis.
10	Lichaee et al., 2024	Narrative Review	Reported system validation outcomes for AI in arrhythmia detection and ECG signal analysis.
11	Monfredi et al., 2023	Systematic Review	Reviewed AI predictive analytics for early clinical deterioration and indirect diagnostic utility.
12	Ose et al., 2024	Narrative Review	Discussed clinical validation strategies, regulatory issues, and performance benchmarks from prior studies.
13	Sellés et al., 2023	Narrative Review	Synthesized clinical applications of AI-ECG and future diagnostic use.

Overall: While direct statistical comparisons were limited, the evidence suggests that AI models hold significant promise for accurate and efficient ECG interpretation across various applications, with an increasing focus on real-time capabilities and predictive analytics.

Clinical Applications of AI Models in ECG Analysis (Objective 4)

The included studies revealed a wide range of clinical applications of AI-ECG tools.

Overall Summary: AI in ECG analysis is no longer limited to diagnostics—it now extends across the entire patient care continuum, offering:

- **Early detection of arrhythmias and heart diseases**
- **Predictive analytics for risk management**
- **Real-time, portable monitoring via IoMT integration**
- **Clinical decision support tools** for reducing interpretation burden
- Even **educational tools** for raising awareness in healthcare and public settings

These applications collectively position AI-ECG systems as a transformative force in modern cardiology, paving the way for smarter, faster, and more personalized cardiovascular care.

Table 8b: Qualitative Observations on AI-ECG Utility from Cross-Sectional Studies

No.	Study ID (Author, Year)	Study Design	Key Qualitative Observations
1	Al-Zaiti et al., 2023	Cross-sectional Study	Explored clinician perceptions of explainable AI-ECG systems. Highlighted usability, trust, and integration barriers that indirectly affect diagnostic adoption.
2	Sellés et al., 2023	Cross-sectional Study	Investigated clinician experiences with AI-ECG in practice, noting both enthusiasm and caution regarding reliability, workflow fit, and responsibility for errors.

Note:
To maintain clarity and alignment with study designs, only review-type studies (narrative, systematic, methodological) are included in **Table 8**. *The two* **cross-sectional studies** *(Al-Zaiti et al., 2023; Sellés et al., 2023) are presented separately in* **Table 8b**, *since their focus is on perceptions, usability, and adoption rather than diagnostic performance.*

This structure ensures that:

- **Tables 5–6** *summarize experimental and simulation studies (6 studies).*
- **Tables 7–8** *summarize review papers (13 studies).*
- **Table 8b** *highlights cross-sectional studies (2 studies).*

Table 9: Clinical Applications of AI in ECG Analysis

Application Area	Supporting Studies (Author, Year)	Key Notes
Early detection of arrhythmias (e.g., AF)	Siontis et al., 2021 [1]; Mamun et al., 2023 [3]; Rahma et al., 2023 [4]; Majhi et al., 2024 [6]; Nechita et al., 2024 [12]; Lichaee et al., 2024 [16]; Ose et al., 2024 [18]; Sellés et al., 2023 [20]	AI tools demonstrated accurate detection and prediction of AF and related arrhythmias.
Diagnosis of structural heart disease (e.g., LV dysfunction, cardiomyopathy)	Siontis et al., 2021 [1]; Attia et al., 2021 [2]; Mamun et al., 2023 [3]; Chowdhury et al., 2018 [7]; Sonia J et al., 2022 [13]	AI-assisted ECG useful in detecting LV dysfunction and structural abnormalities.
Risk stratification & predictive analytics (e.g., cardiac events, deterioration)	Siontis et al., 2021 [1]; Mamun et al., 2023 [3]; Adasuriya et al., 2023 [8]; Monfredi et al., 2023 [17]	AI-ECG models showed promise in risk prediction and early warning for deterioration.
Real-time monitoring / IoT & wearables	Chowdhury et al., 2018 [7]; Kolhar et al., 2024 [10]; Sellés et al., 2023 [20]	Portable/IoT devices demonstrated feasibility for continuous ECG monitoring.
Clinical decision support / workflow efficiency	Adasuriya et al., 2023 [8]; Rafie et al., 2021 [9]; Martinez et al., 2024 [11]	AI-enabled ECG analysis supported faster decision-making and reduced workload.
Educational tools / training & awareness	Monfredi et al., 2023 [17]; Al-Zaiti et al., 2023 [19]	Highlighted educational and training value of AI-ECG systems, including clinician trust and usability aspects.

Despite the growing momentum of AI in ECG analysis, critical limitations remain across data, methodology, clinical implementation, and ethics.

Call for Action: Overcoming the Roadblocks

1. Diversify training datasets to reflect global populations and ensure model generalizability.
2. Advance explainable AI techniques to increase transparency and build clinician trust.
3. Invest in real-world validations and longitudinal trials to bridge the simulation-to-clinic gap.
4. Establish unified regulatory pathways and integration protocols with existing health systems.
5. Ensure ethical deployment by promoting equity, privacy, and clear accountability frameworks.

Table 10: Gaps, Limitations, and Ethical Challenges Identified in Included Studies

Category	Supporting Studies (Author, Year)	Key Notes
Data & Generalizability	Rahma et al., 2023 [4]; Lichaee et al., 2024 [16]; Ose et al., 2024 [18]	Heavy reliance on a few public datasets (e.g., MIT-BIH, PTB-XL) raises concerns about bias and limited generalizability.
Model Transparency & Interpretability	Anand et al., 2022 [5]; Majhi et al., 2024 [6]; Kashou et al., 2023 [14]; Al-Zaiti et al., 2023 [19]	Deep models often act as "black boxes." Explainable AI tools (e.g., SHAP) can help but adoption remains limited.
Clinical Validation & Integration	May et al., 2024 [15]; Ose et al., 2024 [18]	Most models not yet validated in prospective trials or real-world clinical workflows.
Equity & Representation	Lichaee et al., 2024 [16]; Ose et al., 2024 [18]	Underrepresentation of diverse populations in training data may lead to inequities in diagnostic performance.
Regulatory & Ethical Challenges	Ose et al., 2024 [18]; Al-Zaiti et al., 2023 [19]	Concerns around patient data privacy, regulatory approval, trust, and accountability for AI-driven ECG decisions.

Key Insights from Results

The included studies demonstrated a wide range of AI model types applied to ECG analysis. Deep learning approaches dominated, particularly convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and hybrid CNN–GRU or CNN–Transformer models. Machine learning techniques such as support vector machines, random forests, and XGBoost were also widely used, either independently or as part of ensemble frameworks, while hybrid models that integrated machine learning with deep learning appeared increasingly common. To enhance interpretability, several studies incorporated explainable AI tools such as SHAP and LIME, reflecting an emerging emphasis on transparency.
Experimental studies consistently reported strong diagnostic outcomes. Anand et al. (2022) achieved high AUC values for arrhythmia classification using a deep CNN, while Majhi et al. (2024) reported excellent F1-scores and precision for heart disease detection with XGBoost combined with SHAP. Chowdhury et al. (2018) demonstrated that a portable AI-ECG system

achieved reliable accuracy in low-resource clinical environments. Across multiple studies, AI models were found to improve accuracy and sensitivity in detecting arrhythmias and myocardial infarctions, enhance predictive performance for identifying patients at risk of adverse cardiac events, and deliver faster and more consistent diagnoses under clinical pressure.

Most experimental studies relied on public ECG datasets, particularly PTB-XL, MIT-BIH, and PhysioNet, while others employed real-world clinical data that offered greater diversity but also introduced noise and variability. Despite these advances, several studies noted persistent limitations in dataset diversity, raising concerns about the generalizability of models across populations.

In terms of clinical applications, AI systems were tested for real-time arrhythmia detection through wearable devices, remote monitoring using AI-enabled ECG patches, and integration into electronic health records for risk prediction. Other implementations focused on diagnostic support tools in emergency settings, where AI reduced interpretation time and supported rapid decision-making. Collectively, these systems demonstrated the potential to transform clinical workflows, particularly in resource-constrained environments.

Despite technological progress, important ethical and technical challenges remain. Many studies highlighted risks of algorithmic bias stemming from imbalanced datasets, the limited interpretability of deep models due to their "black box" nature, and concerns about patient privacy and consent in the use of clinical data. Furthermore, clinical validation outside controlled study environments was scarce, limiting confidence in generalizability. Efforts to address these issues included the incorporation of XAI frameworks, the use of real-world datasets to improve robustness, and calls for multidisciplinary governance to guide ethical adoption.

Table 11: Ethical and Governance Considerations for AI-Driven ECG Interpretation

Ethical / Governance Issue	Supporting References	Key Points
Transparency & Explainability	Marey et al., 2024 [29]; Asan et al., 2020 [30]	Clinicians require interpretable outputs to trust AI-ECG tools; opaque "black box" systems hinder adoption.
Trust & Clinical Adoption	Asan et al., 2020 [30]; Reddy et al., 2019 [31]	Successful integration depends on clinician confidence in AI recommendations and clarity about responsibility.
Governance & Regulation	Reddy et al., 2019 [31]; Naik et al., 2022 [32]; Iserson, 2023 [33]	Calls for structured frameworks to guide deployment, liability, and ethical use of AI in healthcare.
Legal & Liability	Naik et al., 2022 [32]; Iserson, 2023 [33]	Questions remain around malpractice responsibility if AI-driven ECG interpretation leads to harm.
Fairness & Equity	Zhang et al., 2023 [34];	Bias in training data can worsen health

	Abujaber et al., 2024 [35]	disparities; need fairness audits and inclusive datasets.
Patient Safety & Outcomes	Choudhury et al., 2020 [36]	Patient outcomes, safety, and well-being should remain central in evaluating AI-driven ECG tools.

Table 12 . Strengths and Limitations of AI-Driven ECG Analysis

What Works	What Needs Improvement
High diagnostic accuracy	Better dataset diversity and population representation
Real-time and remote ECG monitoring	Enhanced explainability and interpretability of models
Workflow efficiency in clinical use	Stronger regulatory frameworks and clinical validations
XAI integration for trust-building	Addressing bias, privacy, and implementation challenges

The findings of this review highlight both the strengths and limitations of current AI applications in ECG analysis. On the positive side, AI models consistently demonstrated high diagnostic accuracy across multiple cardiac conditions, with strong potential for real-time and remote monitoring through wearable devices and embedded systems. They also improved workflow efficiency by reducing interpretation times and supporting clinicians under pressure. Importantly, the integration of explainable AI frameworks has begun to build trust by making model decisions more interpretable.

At the same time, several areas require significant improvement before widespread adoption is feasible. Dataset diversity and population representation remain limited, raising concerns about equity and generalizability. Many models still function as “black boxes,” underscoring the need for more transparent and interpretable approaches. Stronger regulatory frameworks, rigorous clinical validations, and prospective real-world trials are necessary to move beyond experimental success. Finally, issues of algorithmic bias, data privacy, and implementation challenges must be addressed to ensure safe deployment in diverse healthcare contexts.

In conclusion, AI-driven ECG analysis stands on the brink of transforming cardiology, but its promise can only be realized through sustained efforts to improve equity, transparency, and clinical reliability, ensuring safe and scalable adoption in practice.

Discussion

This systematic review confirms that AI has tremendous potential to enhance ECG-based diagnosis, but it also brings to light significant challenges that must be addressed to translate these advances into routine clinical care. The findings consistently show that AI-enhanced ECG interpretation can surpass traditional methods in accuracy and speed, supporting AI's role as a diagnostic amplifier in cardiology. However, the road to widespread adoption is impeded by issues of trust, transparency, and governance. Many of the included studies achieved high performance in controlled settings, yet very few were prospectively validated in real-world clinical workflows. This gap highlights a key concern: clinicians may be hesitant to trust and rely on an AI tool without clear evidence of its reliability in their patient populations. The "black-box" nature of most deep learning models further exacerbates this hesitation, as it is often unclear why the algorithm reaches a particular ECG interpretation[27][28]. Lack of interpretability and user trust are repeatedly cited barriers to clinical AI integration[28]. To foster acceptance, researchers are actively exploring solutions such as explainable AI interfaces and clinician-in-the-loop systems. In fact, strategies for improving transparency and accountability in medical AI are evolving in tandem with technological progress[29]. For example, recent work on algorithms for ECG analysis has integrated explanation modules (highlighting which ECG segments influenced the AI's decision) and undergone usability testing with physicians to ensure the outputs make sense to human experts[30].

Another major theme is the need for robust ethical and regulatory frameworks to guide AI deployment in healthcare. At present, there is no consensus governance model for AI in clinical settings, which raises concerns about patient safety and accountability. Some scholars have proposed multi-stakeholder governance models that involve clinicians, data scientists, ethicists, and regulators in overseeing AI tools from development through post-market surveillance[31]. Likewise, establishing clear standards for validating AI algorithms (analogous to drug trials) has been suggested to ensure safety and efficacy are rigorously demonstrated before clinical use. Our review underscores calls in the literature for formal guidelines on issues like algorithmic bias, fairness, and continuous monitoring of AI performance in practice[32][33]. Specific challenges such as legal liability in the event of AI errors remain unresolved – for instance, if an AI misses a fatal arrhythmia on an ECG, it is unclear whether responsibility falls on the software creator, the deploying institution, or the supervising clinician[34]. Efforts to clarify these questions are underway: recent publications have discussed potential regulatory frameworks and liability models (including treating AI similar to medical devices with mandated failure reporting)[31][34]. In parallel, the process of obtaining informed consent for AI involvement in patient care is emerging as a new consideration. Patients may need to be informed when an AI algorithm is used in their diagnosis or treatment planning, especially if the AI's role could meaningfully impact outcomes[35]. Developing patient consent language that clearly explains AI assistance without causing undue alarm is an area needing attention.

Despite these challenges, the consensus in the community is that they are surmountable with multidisciplinary effort. Numerous initiatives are aiming to create trustworthy AI for healthcare by focusing on transparency, equity, and validation[36][37]. Researchers like Zhang et al. emphasize building ethics directly into AI design – for example, by using diverse training data to reduce bias and by conducting external audits of model performance to catch any disparities[36]. In the context of ECG, this means expanding datasets to include patients of different ages, sexes, ethnic backgrounds, and comorbid conditions, and ensuring the AI performs consistently across these groups. Technical fixes (such as bias correction algorithms) combined with institutional policies (like algorithmic fairness assessments before deployment) have been suggested as ways forward[32][38]. Additionally, professional societies and regulatory agencies are beginning to outline best practice guidelines for AI in cardiology. These include recommendations for continuous monitoring of AI outputs in practice, periodic re-training or recalibration of models as new data become available, and involving clinicians in the AI development process to align tools with clinical needs.

Our review also highlights the importance of explainability and human-AI collaboration in successful implementation. Clinicians are more likely to adopt AI tools that can provide interpretable insights (for example, indicating which part of the ECG is abnormal or providing a confidence level for its prediction)[12]. There is optimism that integrating explainable AI (XAI) methods will not only improve user trust but also serve as a teaching aid – helping clinicians understand novel ECG patterns identified by AI and thus enhancing human expertise. Some included studies demonstrated this principle: for instance, the use of SHAP in one model allowed cardiologists to verify that the AI's logic aligned with known ECG features of disease[19]. This kind of symbiosis between human and artificial intelligence can build confidence in the technology. Furthermore, from a workflow perspective, AI should ideally act as a supportive tool (e.g., a second reader for ECGs or a triage filter for normal vs abnormal), rather than a replacement for physician judgment. Such positioning can help mitigate medico-legal concerns as well – physicians would retain ultimate responsibility, using AI as an aid much like any diagnostic instrument.

Finally, it is worth noting that if these challenges are effectively addressed, AI-ECG has the potential to significantly improve patient safety and outcomes. Early evidence indicates that AI can reduce diagnostic errors and oversight, particularly in high-volume settings where clinician fatigue is a factor[39][40]. By automating detection of subtle ECG changes, AI can alert providers to critical conditions they might otherwise miss, thereby preventing adverse events. Additionally, AI could enable more proactive care – for example, by identifying patients at risk of arrhythmias or deterioration before symptoms occur, allowing for earlier interventions. Successful examples outside of cardiology, such as AI improving radiology workflows and reducing missed findings, suggest that similar gains are achievable in ECG interpretation with the right safeguards in place. In summary, while substantial work remains in terms of building the necessary trust, oversight, and evidence base for AI in electrocardiography, the trajectory is encouraging. Ongoing interdisciplinary collaboration between data scientists, clinicians, ethicists, and regulators will be key to developing AI tools that are not only powerful, but also safe, equitable, and aligned with clinical values. With these efforts, AI-augmented ECG interpretation could become a standard component of cardiovascular care, leading to faster diagnoses, more personalized treatments, and ultimately better outcomes for patients.

Conclusion

Artificial intelligence is poised to become an integral part of electrocardiography, offering the ability to detect cardiac abnormalities with speed and precision that complement human expertise. This systematic review has illustrated that AI models – particularly deep learning algorithms – can significantly improve diagnostic accuracy for arrhythmias, ischemia, and other cardiac conditions, and they hold promise for expanding ECG analysis into continuous monitoring and preventive care. At the same time, our examination of the literature makes clear that realizing AI's full potential in ECG interpretation will require careful navigation of the accompanying challenges. Key among these are ensuring the generalizability of AI tools through diverse and high-quality data, enhancing transparency and explainability to secure clinician and patient trust, and instituting robust ethical and regulatory safeguards for AI deployment in clinical environments. Moving forward, interdisciplinary collaboration will be essential: data scientists, clinicians, ethicists, and policymakers must work in concert to refine AI algorithms and the frameworks that govern them. With sustained effort in these areas – validating AI-ECG systems prospectively, addressing biases, improving interpretability, and establishing clear standards for use – AI has the capacity to not only automate ECG interpretation but also to elevate it, leading to earlier diagnoses, more tailored treatments, and improved outcomes in cardiovascular care. In conclusion, AI in electrocardiography represents a transformative innovation on the horizon of medicine; by proactively tackling its current limitations, the medical community can ensure that this technology augments clinical practice in a safe, effective, and ethically responsible manner.

References

1. Stamate E, **et al.** Revolutionizing cardiology through artificial intelligence. *Diagnostics*. 2024;14(11):1103.
2. Martinez FJ, **et al.** AI in ECG interpretation: trends and future directions. *J Pers Med*. 2024;14(3):232.
3. Xiong P, Lee SM, Chan G. Deep learning for detecting myocardial infarction by electrocardiogram: a literature review. *Front Cardiovasc Med*. 2022;9:860032.
4. Attia ZI, **et al.** Application of artificial intelligence to the electrocardiogram for improved cardiovascular risk prediction. *Eur Heart J*. 2021;42(46):4717–25.
5. Ghassemi M, **et al.** Ethics in biomedical AI: addressing health data privacy and bias. *Lancet Digit Health*. 2019;1(1):e200–e201.
6. Markus AF, **et al.** The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey. *J Biomed Inform*. 2020;113:103655.
7. Pham TD. Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use. *R Soc Open Sci*. 2025;12:241873.
8. Ayano YM, **et al.** Interpretable machine learning techniques in ECG-based heart disease classification: a systematic review. *Diagnostics*. 2022;13(1):111.
9. Anand R, **et al.** Explainable AI decision model for ECG data of cardiac disorders. *Biomed Signal Process Control*. 2022;71:103584.
10. Majhi A, **et al.** Explainable AI-driven ML for heart disease detection using ECG signals. *Appl Soft Comput*. 2024;143:112225.
11. Chowdhury MH, **et al.** A portable AI-assisted ECG system for low-resource settings. *BMC Med Inform Decis Mak*. 2018;18:68.
12. Rahma M, **et al.** Data augmentation techniques applied to ECG signals for artificial intelligence applications: a systematic review. *Sensors*. 2023;23(11):5237.

13. Siontis KC, **et al.** Artificial intelligence-enhanced electrocardiography in diagnosing cardiac arrhythmias. *Nat Rev Cardiol*. 2021;18:509–16.

14. Weiner E, **et al.** Ethical challenges and evolving strategies in the integration of artificial intelligence into clinical practice. *PLOS Digit Health*. 2025;4:e0000810.

15. Torkey H, **et al.** Seizure detection in medical IoT: hybrid CNN-LSTM-GRU model with data balancing and XAI integration. *Algorithms*. 2025;18(2):77.

16. Marey A, **et al.** Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology. *Egypt J Radiol Nucl Med*. 2024;55(1):68.

17. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians. *J Med Internet Res*. 2020;22(6):e15154.

18. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc*. 2020;27(3):491–7.

19. Naik N, **et al.** Legal and ethical considerations in artificial intelligence in healthcare: who takes responsibility? *Front Surg*. 2022;9:862322.

20. Iserson KV. Informed consent for artificial intelligence in emergency medicine: a practical guide. *Am J Emerg Med*. 2023;76:225–30.

21. Zhang J, Zhang Z. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med Inform Decis Mak*. 2023;23(1):41.

22. Abujaber AA, Nashwan AJ. Ethical framework for artificial intelligence in healthcare research: a path to integrity. *World J Methodol*. 2024;14(3):149–59.

23. Choudhury A, Asan O. Role of artificial intelligence in patient safety outcomes: systematic literature review. *JMIR Med Inform*. 2020;8(7):e18599.

APPENDICES

Data Extraction Table

No.	Study ID	Study Title	Study Design	Country	Country Income Category	Study Period	Population	Sample Size	Method/Measurement Tool	JBI Score	Full Text URL
1.	Siontis. et al, 2021	Artificial intelligence-enhanced electrocardiography in cardiovascular disease management	Narrative review	USA	High	2021	Cardiovascular disease in at-risk populations of all ages.	N/A	Analysis and synthesis of existing literature on AI applications in ECG interpretation	6/6	https://www.nature.com/articles/s41569-020-00503-2 DOI: 10.1038/s41569-020-00503-2
2.	Anand, et al, 2022	Explainable AI decision model for ECG data of cardiac disorders	Experimental study	India	Lower middle	2022	ECG recordings from publicly available datasets, specifically PTB-XL and an arrhythmia data set.	Over 21,000	Development of a deep learning model named ST-CNN-GAP-5, evaluated using metrics like accuracy and AUC. The model's interpretability was assessed using SHapley Additive exPlanations (SHAP)	4/9	https://www.sciencedirect.com/science/article/abs/pii/S1746809422001069 DOI: 10.1016/j.bspc.2022.103584
3.	Attia, et al, 2021	Application of artificial intelligence to the electrocardiogram	Narrative Review	USA/ UK	High	2021	N/A	N/A	Comprehensive analysis and synthesis of existing literature on AI applications in electrocardiography for cardiovascular disease management	6/6	https://academic.oup.com/eurheartj/article/42/46/4717/6371908 DOI: 10.1093/eurheartj/ehab649
4.	Mamun et al, 2023	AI-Enabled Electrocardiogram Analysis for Disease Diagnosis	Narrative Review	USA	High	2023	N/A	N/A	Comprehensive analysis and synthesis of existing literature on AI applications in electrocardiography for disease diagnosis	5/6	https://www.mdpi.com/2571-5577/6/5/95 DOI: 10.3390/asi6050095
5.	Rahma, et al, 2023	A Systematic Survey of Data Augmentation of ECG Signals for AI	Systemic Review	Italy/ USA	High	2023	N/A	119	Systematic literature search adhering to PRISMA guidelines, analyzing data	7/11	https://www.mdpi.com/1424-8220/23/11/52

		Applications							augmentation techniques applied to ECG signals in AI applications		37 DOI: 10.3390/s23115237
6.	Majhi et.al 2024	Explainable AI-driven machine learning for heart disease detection using ECG signal	Simulation Study	N/A	N/A	2024	ECG recordings from 3 international databases	101,799	RF, XGBoost, SHAP, DWT/EWT signal processing	N/A	https://www.sciencedirect.com/science/article/abs/pii/S1568494624009992 https://doi.org/10.1016/j.asoc.2024.112225
7.	Chowdhuryy et.al 2018	AI Assisted Portable ECG for Fast and Patient Specific Diagnosis	Experimental study using retrospective ECG	Bangladesh	Middle-lower	2018	Mixed CVD/normal cases.	80	Deep convolutional neural networks, wireless biosensing, ultraportable ECG, and MQTT communication protocol	N/A	https://ieeexplore.ieee.org/abstract/document/8465483 https://doi.org/10.1109/IC4ME2.2018.8465483
8.	Adasuriya Et.al , 2023	Next Generation ECG: The Impact of Artificial Intelligence and Machine Learning	Narrative review	N/A	N/A	2023-2019	Patients with various cardiovascular diseases in multiple studies including randomized controlled trials	N/A	CNN-based DL models, ML algorithms	N/A	https://link.springer.com/article/10.1007/s12170-023-00723-4 https://doi.org/10.1007/s12170-023-00723-4
9.	Rafie,et.al 2021	ECG Interpretation: Clinical Relevance, Challenges, and Advances	Narrative review	USA	HIGH	2021	Patients with cardiac symptoms	N/A	ECG, AI-ECG algorithms, mobile ECG devices	N/A	https://www.mdpi.com/2673-3846/2/4/39 https://doi.org/10.3390/hearts2040039
10.	Kolhar ,et.al 2024	AI-Driven Real-Time Classification of ECG Signals for Cardiac Monitoring Using i-AlexNet Architecture	Experimental study	KSA/ INDIA	High / middle-lower	2024	Mixed normal/abnormal ECG cases from public databases	338	AlexNet, Red Fox Optimization, IoMT sensors (theoretical)	N/A	https://www.mdpi.com/2075-4418/14/13/1344 https://doi.org/10.3390/diagnos

											tics14131344
11.	Martinez Et.al 2024	Revolutionizing Cardiology: AI in ECG Analysis Paves the Way for Better Disease Detection and Treatment	Narrative review	N/A	N/A	2024	Patient with various cardiovascular disease	N/A	-Deep learning algorithms -Supervised learning techniques - AI-enhanced ECG -Integration of AI models into clinical decision	N/A	https://ai.nejm.org/doi/full/10.1056/AI-S2400629
12.	Nechita Et.al 2024	AI-Enhanced ECG Applications in Cardiology: Comprehensive Insights from the Current Literature with a Focus on COVID-19 and Multiple Cardiovascular Conditions	Narrative Review	Romania	Upper-middle	2024	N/A	N/A	AI-enhanced ECG, ECG DIVICE AI algorithms	N/A	https://www.mdpi.com/2075-4418/14/17/1839
13.	Sonia J Et.al 2022	AI Techniques for Efficient Healthcare Systems in ECG Wave Based Cardiac Disease Detection by High Performance Modelling	Experimental study	India	Lower-middle	2022	Patients with various cardiac conditions.	N/A	-Deep Belief Kernel Principal Neural Network -Deep Quantum Multilayer Convolutional Neural Networks -Performance Metrics	N/A	https://www.researchgate.net/profile/Vivek-Solavande/publication/367162511_AI_Techniques_for_Efficient_Healthcare_Systems_in_ECG_Wave_Based_Cardiac_Disease_Detection_by_High_Performance_Modelling/links/6485d66079a722376526bdfb/AI-Techniques-for-Efficient-Healthcare-Systems-in-ECG-

											Wave-Based-Cardiac-Disease-Detection-by-High-Performance-Modelling.pdf
14.	Kashou Et.al 2023	Comparison of two artificial intelligence-augmented ECG approaches: Machine learning and deep learning	Methodological review	India	Lower-middle	2023	N/A	N/A	-Formulas for Different -Statistical Parameters -Software Tools	N/A	https://www.sciencedirect.com/science/article/abs/pii/S0022073623000481
15.	May Et.al 2024	A novel way to prospectively evaluate of AI-enhanced ECG algorithms	Methodological	N/A	N/A	N/A	Patients with various cardiac conditions.	N/A	AI-enhanced ECG, ECG DIVICE Virtual testing platform	N/A	https://www.sciencedirect.com/science/article/abs/pii/S0022073624002206
16.	Lichaee .Et.al2024	Advancements in Artificial Intelligence for ECG Signal Analysis and Arrhythmia Detection: A Review	Review Article	Iran	Uppermiddle	N/A	N/A	N/A	MIT-BIH Arrhythmia Database PhysioNet Challenge Datasets Clinical ECG data from hospitals (e.g., in Baghdad)	N/A	https://brieflands.com/articles/ijcp-143437
17.	Monfredi Et.al 2023	Continuous ECG monitoring should be the heart of bedside AI-based predictive analytics monitoring for early detection of clinical deterioration	Cross-sectional study	Saudi Arabia	High	May and June 2023	Nursing students	175	Connor-Davidson Resilience Scale – 10 items (CD-RISC-10): World Health Organization-5 Well-Being Index (WHO-5)	N/A	https://www.sciencedirect.com/science/article/abs/pii/S0022073622002047
18.	Ose Et.al 2024	Artificial Intelligence Interpretation of the Electrocardiogram: A State-of-the-Art Review	Narrative review	USA	High	May 16, 2024	N/A	N/A	Systematic searching and selection of relevant published articles on AI applied to electrocardiogram (ECG) interpretation. Critical analysis and synthesis of findings from these articles. Discussion of different AI approaches,	N/A	https://link.springer.com/article/10.1007/s11886-024-02062-1

									clinical applications, validation techniques, and regulatory aspects.		
19.	Al-Zaiti Et.al 2023	Explainable-by-design: Challenges, pitfalls, and opportunities for the clinical adoption of AI-enabled ECG	A cross-sectional study	Kuwait	High	February 9, 2022, and April 11, 2022	Young adults aged 18 to 35 years residing in Kuwait.	529	Questionnaire	N/A	https://www.sciencedirect.com/science/article/abs/pii/S0022073623001905
20.	Sellés Et.al 2023	Current and Future Use of Artificial Intelligence in Electrocardiography	Narrative review	Spain	High	N/A	N/A	N/A	Analyzes and summarizes existing research on AI applications in electrocardiography	N/A	https://www.mdpi.com/2308-3425/10/4/175

Conflict of Interest Statement

The authors declare no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper