NARRATIVE REVIEW

Deep learning for electrocardiogram interpretation: Bench to bedside

Bas B. S. Schots[1] | **Camila S. Pizarro[1]** | **Bauke K. O. Arends[1]** | **Marish I. F. J. Oerlemans[1]** | **Dino Ahmetagić[1]** | **Pim van der Harst[1]** | **René van Es[1,2]**

[1]Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands

[2]Cordys Analytics B.V., Utrecht, The Netherlands

Correspondence
René van Es, Department of Cardiology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.
Email: r.vanes@umcutrecht.nl

Abstract

Background: Recent advancements in deep learning (DL), a subset of artificial intelligence, have shown the potential to automate and improve disease recognition, phenotyping and prediction of disease onset and outcomes by analysing various sources of medical data. The electrocardiogram (ECG) is a valuable tool for diagnosing and monitoring cardiovascular conditions.

Methods: The implementation of DL in ECG analysis has been used to detect and predict rhythm abnormalities and conduction abnormalities, ischemic and structural heart diseases, with performance comparable to physicians. However, despite promising development of DL algorithms for automatic ECG analysis, the integration of DL-based ECG analysis and deployment of medical devices incorporating these algorithms into routine clinical practice remains limited.

Results: This narrative review highlights the applications of DL in 12-lead ECG analysis. Furthermore, we review randomized controlled trials that assess the clinical effectiveness of these DL tools. Finally, it addresses different key barriers to widespread implementation in clinical practice, including regulatory hurdles, algorithm transparency and data privacy concerns.

Conclusions: By outlining both the progress and the obstacles in this field, this review aims to provide insights into how DL could shape the future of ECG analysis and enhance cardiovascular care in daily clinical practice.

KEYWORDS
artificial intelligence, cardiovascular diseases, deep learning, electrocardiogram, implementation

Bas B.S. Schots and Camila S. Pizarro contributed equally to this work.

1 | INTRODUCTION

Since its discovery by the Dutch scientist Willem Einthoven over a century ago, the electrocardiogram (ECG) has become an essential tool in clinical practice. Its operation is based on detecting the electrical signals that propagate through the body when cardiac cells depolarize and repolarize, using electrodes positioned on the body's surface. The ECG has been fundamental in diagnosing arrhythmias, myocardial ischemia and in identifying non-cardiac disorders, such as electrolyte imbalances.[1,2] Due to its low cost, noninvasive nature, and ease of acquisition, this instrument is widely available worldwide. However, detecting ECG changes associated with pathological processes requires a certain level of expertise, particularly when it comes to subtle changes that enable early diagnosis.[3] Moreover, great variability in ECG analysis and in accuracy of diagnoses have been observed among physicians when compared to experts.[4]

In recent years, the increase in availability of large-scale datasets combined with the increase in computing power has allowed for the use of artificial intelligence (AI) to become a prominent topic in many different areas of research, including medicine.[5] Furthermore, FDA regulatory bodies have approved AI-based products at an increasing rate, with a strong focus on machine learning (ML) and deep learning (DL) technologies.[6] Where ML methods require manual feature extraction, DL uses neural networks to automatically learn complex features and patterns from large datasets. This is especially promising in the clinical context, where much data is acquired, and expert opinion is not always readily available. Frequently used terms in the AI field are defined in Table 1.

DL methods have already made significant contributions in areas such as medical imaging, disease prediction and drug development.[5] For ECG analysis, it has already been shown that a variety of cardiac diseases can be accurately detected using DL, in some cases even outperforming medical experts.[7] Also, in situations where specialized expertise is not available, DL-based automatic ECG analysis could enhance the interpretability and accessibility of ECG data. Additionally, the automatic extraction of features from raw ECG data enables the identification of unknown (subtle) patterns associated with known diseases or patterns linked to pre-disease stages, that may be imperceptible to the human eye, offering predictive and prognostic value.

Deep learning for ECG analysis has been reviewed previously, highlighting its suitability for this task and presenting applications in different domains.[8–10] The current review provides an update on DL for ECG analysis and places additional emphasis on the implementation in clinical practice, including regulatory and practical barriers. Given that the 12-lead ECG is the most common method of ECG acquisition in clinical practice, this review focused exclusively on studies that included more than 10,000 12-lead ECGs. We first provide an overview of current DL clinical applications. Subsequently, we focus on the implementation of ECG-based DL models in daily clinical practice and highlight the ongoing challenges to implementation (Figure 1).

1.1 | Current clinical applications

To explore the applications of DL in ECG analysis, a comprehensive PubMed search was conducted. The articles identified were screened by two independent authors (BS and CP). To highlight the broad applicability of DL ECG analysis, cardiac disorders were grouped into three categories: rhythm and conduction disorders, structural and functional diseases and ischemic heart disease.

2 | RHYTHM AND CONDUCTION DISORDERS

Synchronized heart contraction and blood propulsion requires a stable propagation of electrical current through cardiomyocytes. Abnormalities in impulse generation or propagation can lead to rhythm and conduction disorders. Hemodynamic repercussions vary among different arrhythmias and identification of arrhythmias that require prompt treatment is important. However, ECG evaluation in the clinic is time consuming and observer dependent. Automation of this process using AI can therefore be of great importance. Many studies have focused on the automatic classification of rhythm disorders from ECGs using DL.[11] These models can be applied in an end-to-end fashion, directly identifying arrhythmias from the ECG input data. Most approaches adopt a convolutional neural network (CNN), a type of DL network in which convolutional operations are used to extract more complex features from the ECG waveforms. Placing these operations in series (layers) allows the algorithm to use complex features from the ECG waveform to predict specified outcomes. Other architectures have also been explored (Figure 2). Both single and multi-label classification approaches have shown promise.[11–25] For a deeper exploration of technical DL details in ECG analysis, readers are referred to previous reviews on DL ECG.[8–10]

Single-label, or binary classification targets just one specific abnormality at a time. DL algorithms have been shown to perform well at this task, accurately classifying conditions such as atrial fibrillation (AF), identifying premature

TABLE 1 Definitions of key artificial intelligence and deep learning terms used in ECG analysis.

Term	Description
Artificial intelligence (AI)	The concept of computer systems being able to carry out complex tasks mimicking human intelligence.
Machine learning (ML)	A subset of AI in which the model finds patterns in the data requiring manual feature extraction.
Deep learning (DL)	A subset of ML in which neural networks automatically learn features from large datasets without manual feature extraction
Feature	A property or characteristic of the data, for instance a waveform characteristic like the QRS width.
Label	The target variable that the model is predicting, for instance 'atrial fibrillation'
Deep neural network (DNN)	A type of neural networks that can automatically learn features from large datasets without manual feature extraction
Convolutional neural network (CNN)	A type of DNN in which convolutional layers are used that are especially well designed for pattern recognition
Residual neural network (Res-Net)	A special type of CNN that uses connections between layers to leverage information flow across deep networks
Recurrent neural network (RNN)	A type of neural network designed for sequential data analysis, where information from previous inputs is used to influence current processing. RNNs are particularly suited for tasks such as time-series analysis.
Long short-term memory (LSTM)	A type of RNN specifically designed to handle long-term dependencies in sequential data. LSTMs are effective for analysing ECG signals, capturing patterns across longer sequences for tasks like arrhythmia detection.
Auto Encoder	A neural network designed for reconstructing the input data from a compressed state, useful in ECG analysis for noise reduction and feature extraction.
Training	Network training is the process of feeding training data into the model to update its parameters and improve its predictions.
Secondary validation	Using a separate dataset often from another source or patient demographic group to validate the model's performance and test its generalization capabilities.
Area under the receiver operating curve (AUROC)	A performance metric that measures the classification performance at different thresholds. What is good and what is bad

ventricular complexes and other arrhythmias with high precision.[11–15] These models can enhance clinical care by improving diagnostic throughput and reducing observer variability, increasing the efficiency of clinical workflow. In a pivotal study, Attia et al. developed a ResNet-based CNN to detect paroxysmal AF from sinus rhythm ECGs acquired up to 31 days before diagnosis, achieving an AUROC of 0.87.[16] This showed how DL can identify intrinsic markers of AF onset from normal ECGs, beyond human capability. Building on this, Raghunath et al. trained a model to predict new-onset AF within 1 year from sinus rhythm ECGs, achieving an AUROC of 0.85.[17] This shift from detection to prediction highlights DL's potential to transition from diagnostics to preventive care.

Multi-label classification refers to models that can predict several rhythm abnormalities from a single ECG that may occur simultaneously. Studies have shown that DL approaches are suitable for this task.[11,18–21] For example, Kashou et al. developed a CNN trained on over two million ECGs to classify 66 distinct labels, including 23 rhythm disorders, achieving an AUROC >0.98 across all labels.[22] Similarly, Ribeiro et al. utilized a ResNet-based CNN to classify six types of arrhythmias and conduction disorders, performing comparably to medical professionals.[23] These studies underscore the capacity of DL systems to handle multiple abnormalities simultaneously, assisting clinicians and reducing their workload. Other approaches aim to optimize performance by moving toward lead-specific prediction, where DL networks are designed with separate CNN branches for each of the 12 leads or using only a subset of the 12 leads as input.[24,25]
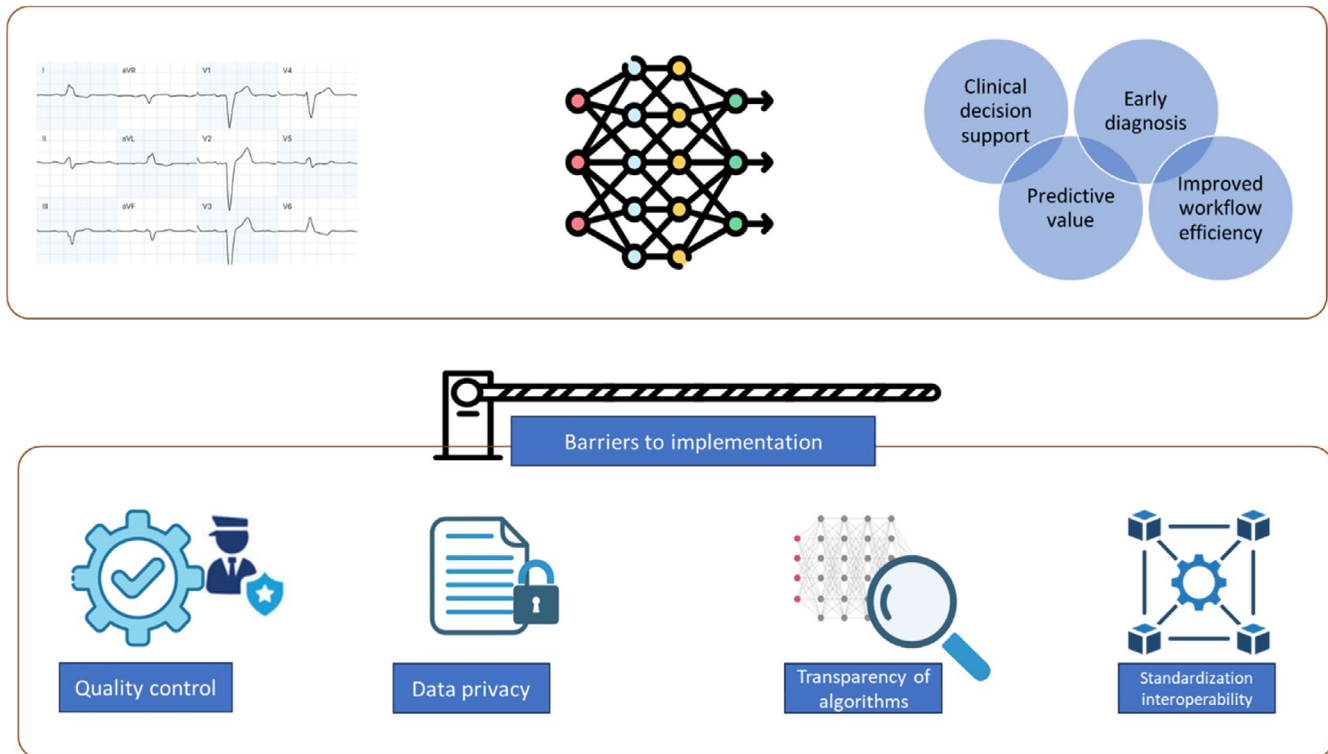
FIGURE 1 Schematic overview of applications of DL-enabled ECG analysis and the barriers to implementation of DL tools to clinical practice.
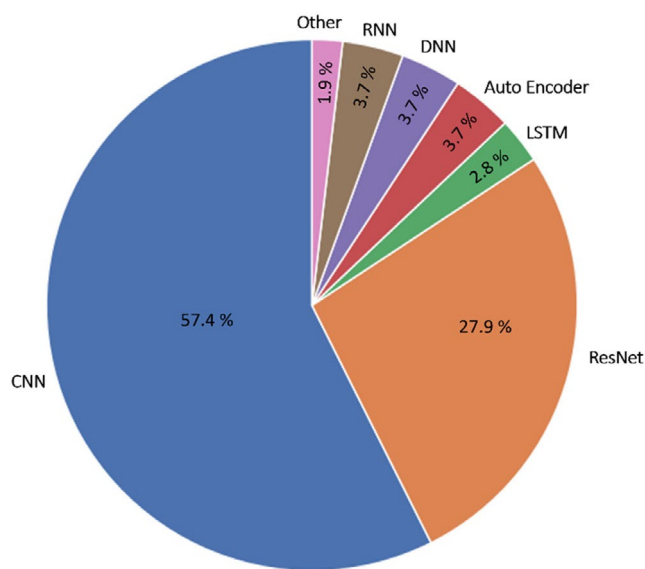

FIGURE 2 DL architectures used in DL ECG analysis. Overview of model architectures used in DL ECG analysis. CNN, convolutional neural network; DNN, deep neural network; LSTM, long short-term memory; ResNet, residual network; RNN, recurrent neural network.

Often only a subset of these rhythm abnormalities can be consistently predicted, making the model noncomprehensive. Others have investigated more comprehensive triage systems for rhythms disorders, with the aim of classifying the patients in categories of urgency, rather than predicting the specific pathology.[26–28] Such models may be used to reduce clinical workload by improving the prioritization of ECGs and have also been tested in a prospective implementation setting, proving to be more accurate than the currently employed non-DL based algorithm.[29]

3 | STRUCTURAL AND FUNCTIONAL DISEASES

3.1 | Valvular diseases

The cost-effectiveness of a screening program with echocardiography for diagnosis of valvular heart disease (VHD) in the general population has not been demonstrated.[30] DL applications may enable ECG-based diagnosis of VHD. Several strategies have been developed to detect VHD, with aortic stenosis (AS) and mitral regurgitation (MR) being the most studied diseases.[31–35] Kwon et al. were among the pioneers in using DL for the detection of significant AS in ECGs, achieving an AUROC of 0.884 in an external validation set.[36] Comparable results were obtained when ECG-echocardiogram pairs were used in CNN models for detecting moderate–severe AS, highlighting the potential of ECG as a screening tool for identifying

patients at high risk of valvular disease and helping to optimize the use of diagnostic resources in settings where access to echocardiography is limited[35,37]

3.2 | Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy (HCM) affects 1 in 200–500 people, with potential outcomes including heart failure and sudden cardiac death.[38] Therefore, early detection might help reduce the risk of adverse outcomes. Ko et al. developed a CNN model to detect hypertrophic cardiomyopathy (HCM) from the ECG, later validated in an independent cohort. The model achieved a high performance to detect HCM, even in patients with a normal ECG.[39] Recently, this model was further validated using data from an international cohort of HCM patients and non-HCM controls, with comparable results.[40] Interestingly, ECGs were analysed without selection for tracing quality or any pre-processing, making the results comparable with real-world ECG acquisition.[40] Automatic identification of HCM patients from electronic health record data has also been reported.[41] These models could provide an improvement in screening and early diagnosis for HCM, allowing earlier intervention and better disease management. With increasing options for HCM-related therapies, such as cardiac myosin inhibitors, earlier diagnosis is increasingly relevant for these patients.[42,43]

3.3 | Left ventricular dysfunction

Left ventricular dysfunction is an independent predictor of mortality in heart failure patients with reduced (HFrEF) and preserved (HFpEF) ejection fraction.[44,45] It is currently assessed using imaging techniques, but ECG DL allows detecting both ventricular systolic (LVSD) and ventricular diastolic (LVDD) dysfunction, characteristic markers of HFrEF and HFpEF, respectively.[46–51] Attia et al. demonstrated that an AI-enabled ECG model could accurately predict LVSD (ejection fraction ≤35%).[52] Validation using routine ECGs from a broad patient population showed its potential as an effective screening tool for asymptomatic individuals.[53] Yagi et al. investigated the generalizability of a CNN model designed to detect LVSD from 12-lead ECGs across multiple institutions. The model performed well on internal data, but showed varying performance on external data, underscoring the need for models to be tested across diverse populations before clinical deployment.[54] An external validation in a European population confirmed the model's potential in detecting LVSD.[55] It achieved a high negative predictive value (96%) but a lower positive predictive value (39.9%),

indicating that further testing is required in broader, non-hospitalized populations. Importantly, they also explored the model's ability to predict future LVSD in patients with preserved ejection fraction (≥50%), demonstrating moderate success.[55] This predictive capability highlights the potential of AI-ECG models not only in current diagnosis but also in identifying patients at risk for future LVSD.

4 | ISCHEMIC HEART DISEASE

Prompt diagnosis of myocardial infarction (MI) is vital in reducing mortality and morbidity.[56,57] Using DL to complement ECG assessment of patients with suspected acute coronary syndrome could improve diagnostic accuracy and potentially reduce door-to-balloon time.[58] Chang et al. proposed an algorithm to detect acute ST-elevation MI (STEMI) and achieved an AUROC of 0.98 in an external test set, outperforming board-certified physicians and commercial ECG algorithms.[59] Their model was also effective in prehospital settings, reducing response times for paramedics by quickly identifying STEMI.[58] However, ST-segment elevation remains an imperfect surrogate marker for coronary occlusion, with 25%–30% of patients presenting with occlusion on angiography, despite not having ST-segment elevation.[60] A DL model by Herman et al. has shown the ability to detect occlusive MI based on more complex ECG patterns.[61] The model's performance was superior to conventional STEMI criteria and comparable to specialized ECG experts, highlighting the potential for such models to improve patient outcomes by capturing those with silent or atypical presentations. Others have extended this creating DL models to detect MI without ST-segment elevation (NSTEMI), which showed that AI could identify MI cases even in patients without classical ST-segment elevation.[62,63] These studies highlight the broader applicability of DL models, not just in STEMI detection, but in providing diagnostic insights for NSTEMI patients, who often present with ambiguous ECG findings.

5 | RANDOMIZED CONTROLLED TRIALS

Randomized controlled trials (RCTs) are considered essential in demonstrating the effectiveness of an intervention and are a crucial tool to evaluate implementation of DL tools in daily clinical practice. To date, few RCTs investigating the use of DL algorithms for ECG analysis have been conducted.

Yao et al. aimed to assess the potential of an ECG-based, DL-powered clinical decision support tool to

enable early diagnosis of LVSD (EF ≤50% within 90 days of ECG acquisition),[64] using the network developed by Attia.[52] This RCT was conducted in 120 care teams from 45 clinics and hospitals in the USA. ECGs from 22,641 patients were acquired, with 11,573 in the intervention group and 11,068 in the control group. The intervention group received model-assisted alerts indicating a high risk of LVSD, while the control group followed routine clinical care. LVSD diagnosis rate in the intervention arm was significantly higher, indicating the added benefit of screening individuals using DL ECG analysis.

Adedinsewo et al. conducted an RCT to evaluate DL-based cardiomyopathy screening in pregnant and post-partum women in Nigeria, a country with the highest global incidence of peripartum cardiomyopathy.[65,66] The primary endpoint was the detection of LSVD, defined as ejection fraction <50%. A total of 587 women were assigned to the intervention arm and 616 to the control arm. Patients in the intervention arm received DL-based cardiomyopathy screening, while the control group received usual care with echocardiography at the physician's discretion. However, no significant difference between the two arms was found.

Lin et al. looked at the ability of their model to predict overall mortality in a multisite RCT involving 39 physicians and 15,965 patients from the emergency and inpatient departments. It was shown that implementation of their algorithm significantly reduced the 90-day mortality rate from 4.3% to 3.6%. Furthermore, physicians were more likely to escalate clinical care in the case where the DL model flagged the patient as high-risk for mortality.[67]

In the ARISE study, an ECG-based supporting tool for reducing door-to-balloon time for STEMI patients was tested.[68] Among the 43,234 patients, 145 were diagnosed with STEMI, including 77 in the intervention group and 68 in the control group; the intervention group demonstrated a significantly shorter door-to-balloon time of 82 min compared to 96 min in the control group ($p = 0.002$). They concluded that DL-assisted ECG triage decreased the ECG-to-balloon time for patients presenting to the ED as well as for inpatients.[68] These results underscore the potential of ECG-AI triage to enhance STEMI patient treatment.

More RCTs have been proposed and are currently under consideration in different application areas for cardiovascular research, for example, for AF risk assessment and coronary artery disease diagnosis.[69,70]

5.1 | Barriers to the implementation of DL in clinical practice and ongoing efforts to overcome them

Integration of DL into daily clinical practice remains an ongoing challenge and is yet to become a widespread reality. Similar to the development of other medical devices, development of DL algorithms needs to follow a structured approach[71] (Figure 3).

Investments in AI development have surged in recent decades. In 2023, private investment in AI reached €9 billion across Europe. Additionally, the EU, through its Digital Europe Program, is set to invest €7.6 billion to accelerate digital transformation over the next 7 years. This growing interest in AI is further demonstrated by the sharp rise in machine learning publications, including DL, which increased sevenfold between 2015 and 2022.[72] Similarly, the number of AI-related medical devices in the cardiovascular field approved by the FDA has risen, from just 5 in 2016 to 60 by 2022.[72]

Moreover, openness from clinicians and patients to integrate AI in daily clinical practice has also been studied. Recently, Arends et al., investigating barriers and facilitators to AI usage in clinical practice, showed that potential
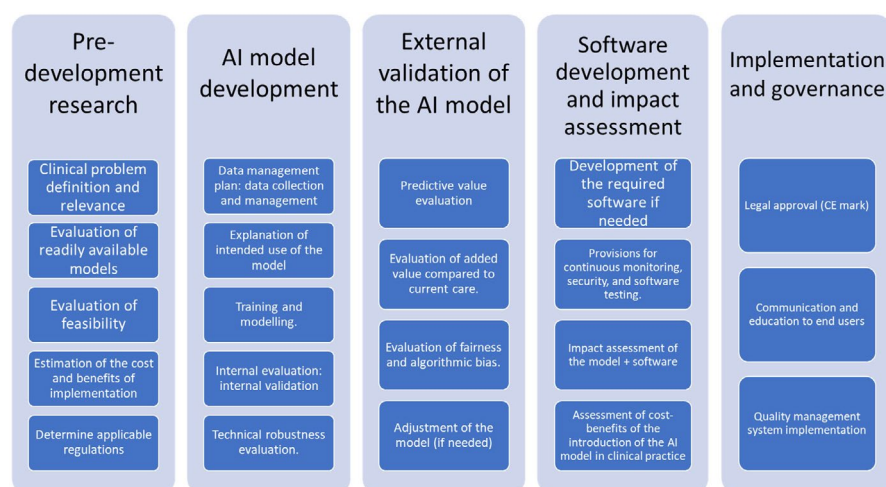

FIGURE 3 Overview of the steps needed for AI development and implementation.

end-users of DL ECG analysis had a positive perception of automatization for ECG analysis and were open to integrating AI in their work [Arends, 2024]. Despite this surge in investment and innovation, implementation of AI developments in healthcare continues to face prominent issues.

6 | CONCERN FOR PATIENT SAFETY: QUALITY CONTROL OF ALGORITHMS

One barrier to AI implementation is related to the concern of physicians for the accuracy of AI algorithms and how this may affect patient safety. In the study of Arends et al., poor model performance was identified as the most important barrier [Arends, 2024]. Critical evaluation of developed algorithms is crucial for their implementation. Often, external validation is considered an appropriate method to determine models' reliability and generalizability. However, external validation does not guarantee model usefulness.[73] AI algorithms are considered medical devices; therefore, development and certification of an AI algorithm are regulated and standardized processes. Industry standards are created to provide developers with guidelines to produce compliant and high-quality products. New standards with basic safety and performance requirements are being developed, applying to all ECG devices, including computerized analysis.[74] The 80601-2-86 standard includes ECG algorithm testing requirements, testing methods and data sets based on the intended used of the algorithm and not the type of ECG device. Furthermore, the use of medical devices in clinical practice is regulated by law. In Europe, CE certification indicated that a medical device complies with the Medical Device Regulation and can be commercialized in the European Economic Union, while in the USA, the FDA oversees regulation. In other regions, regulation follows local laws and standards.[75,76]

In august 2024 the new regulation focused on ensuring safety, transparency and accountability has taken effect in Europe, the European Artificial Intelligence act.[77] This aims to contribute to the clarification of requirements and obligations for developers and deployers, and it adopts a risk-based approach. High risk AI systems, including AI-driven diagnostic tools, medical devices integrated with AI, patient monitoring and risk assessment, must meet strict requirements. These include implementation of risk management frameworks throughout the AI system lifecycle; representative, and bias-free data for model development; transparency; human oversight; and post marketing monitoring detect any potential harm or safety

issues after deployment.[77] This act marks considerable progress toward enhancing patient safety. However, while it includes provisions to reduce barriers—such as allowing companies to test high-risk AI applications under regulatory supervision without the full regulatory burden—it may still complicate implementation. Additionally, given the rapidly evolving nature of AI, the current regulations will require continuous revision to keep pace with technological advancements and ensure ongoing relevance and effectiveness.

7 | DATA SHARING AND PRIVACY

One of the key factors in developing and training AI algorithms is the availability and quality of data. The amount of data needed for training varies based on factors such as the type of learning and the complexity of the model. The amount of data used to develop models ranged from 10,000 to 2.5 million ECGs among studies included in this review. As a result, effective strategies for data sharing, including international collaboration, are crucial to ensure the development and deployment of high-quality AI tools. However, this raises concerns about patients' rights to data protection and privacy. Since 2018, data sharing within the European Union has been governed by the General Data Protection Regulation (GDPR), which sets strict standards to safeguard individuals' rights and privacy, including limiting data sharing with non-EU countries.[78] Despite these protections, some GDPR principles, such as data minimization and the right to erasure, may conflict with AI's reliance on large-scale datasets, posing challenges to balancing privacy with technological advancement.

Anonymization of data ensures compliance with current EU regulations. However, due to the detailed and granular nature of many datasets, particularly in fields like healthcare, achieving effective anonymization on a large scale is challenging, leaving a significant risk of re-identification. This risk arises from the possibility that even anonymized data, when combined with other data sources, can inadvertently reveal an individual's identity.[79] To address these data-sharing challenges, initiatives such as anonymized benchmarking databases have emerged. These databases provide access to de-identified datasets while minimizing privacy risks. For example, several international datasets are available for research purposes, including anonymized ECG data (Table 2). Another promising approach to overcoming data-sharing hurdles is the use of federated learning methods, which allows AI models to be trained across multiple decentralized datasets without directly sharing the data between institutions.[80] Instead, models are trained locally, and only the learned

TABLE 2 Electrocardiography databases.

Database Name	Subjects	No of recordings	No of leads	Duration	Sampling (Hz)	Diagnostic class
MIT-BIH Arrhythmia Database[88]	47	48	2	30 min	360	Beat annotations (normal beat, LBBB beat, RBBB beat, atrial premature beat, etc) and rhythm annotations (atrial flutter, nodal rhythm, paced rhythm, etc)
PTB Diagnostic ECG Database[92]	294	549	15	10 sec	1000	MI (148), cardiomyopathy/HF (18), BBB (15), dysrhythmia (14), myocardial hypertrophy (7), valvular heart disease (6), myocarditis (4), miscellaneous (4) and healthy controls (52)2)
PTB-XL[93]	18,869	21,799	12	10 sec	500	71 classes grouped in five main categories: Normal ECG (9514), MI (5469), ST/T Change (5235), conduction disturbance (4898) and hypertrophy (2649)
Long-Term ST Database[94]	80	86	2	21–24 h	250	ST segment changes, including ischemic ST episodes, axis-related non-ischemic ST episodes, episodes of slow ST level drift and episodes containing mixtures of these phenomena
European ST-T Database[95]	79	90	2	2 h	250	Changes in ST segment and T-wave morphology
MIMIC-III Waveform Database[87]	30,000	67,830	NA	NA	500	Waveforms are unannotated
Lobachevsky University Electrocardiography Database (LUDB)[96]	200	200	12	10 s	500 Hz	Specified heart rate types, position of the electrical axis, specific type of conduction abnormalities, types of extrasystole, types of hypertrophies and ischemia location.
The ECG-ID Database[97]	90	310	1	20 s	500 Hz	10 annotated beats (unaudited R- and T-wave peaks annotations from an automated detector)
PhysioNet computing in cardiology challenge 2017	NA	8528 (train), 3658 (test)	1	9–60 s	300 Hz	Normal (5154), Atrial fibrillation (771), Other rhythm (2557), Noisy (46)
St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database	32	74	12	30 min	257 Hz	Acute MI, transient ischemic attack, prior MI, CAD, with hypertension, SV ectopy, AF, SVTA, WPW, AV block, BBB
Shandong Provincial Hospital (SPH) Database[98]	24,666	25,770	12	10–60 s	500 Hz	44 primary classes and 15 modifier, including: Normal ECG, AF, VPCs

Abbreviations: AF, atrial fibrillation; AV, atrioventricular; CAD, coronary artery disease; LBBB, left bundle branch block; MI, myocardial infarction; RBBB, right bundle branch block; SV, supra ventricular; SVTA, supraventricular tachyarrhythmia; VPCs, ventricular premature complexes; WPW, wolf Parkinson white.

parameters are shared. This method preserves privacy while enabling the collaborative development of robust AI tools.

8 | TRANSPARENCY AND MODEL INTERPRETABILITY

Constant development of increasingly complex models raises concerns about their black-box nature. Transparency in AI has ethical, legal and societal implications and encompasses concepts like explainability, disclosure of the data used for training; and sufficient interpretability to be oversighted by humans.[81] Several articles included in the GDPR support the concept of 'right to explanation', granting individuals the right to meaningful information about the logic involved in automated decision-making; transparency in processing of personal data; and right to human intervention to challenge or review automatic decisions.[78] The new European AI act has further specified the obligation that must be followed to comply with existing regulations, particularly for high-risk systems.[77] Nevertheless, gaps such as the level of explainability that an AI system must reach or how to ensure human accountability for decisions made by AI remain and continuous reevaluation and specification of regulations may help in the acceptance of integration of AI systems in clinical practice.

Explainable AI provides mechanisms that make the decision-making process of AI models understandable.[82] In the context of AI-enabled interpretation of the ECG, different techniques can be used to identify ECG segments or leads that are the most influential in the model's decision, helping to make the results more transparent and interpretable for clinicians. Post hoc techniques, such as class activation mapping (CAM) and saliency maps can be applied to CNNs to interpret models' prediction. These methods highlight the temporal locations in the ECG that are most relevant to the model's output.[83,84] However, a key limitation of these techniques is that they do not reveal specific features (e.g. QRS morphology) that influence the prediction. Therefore, their use may be limited, or even lead to various kinds of bias. Furthermore, the validity of such techniques has been questioned.

Recently, Van der Leur et al., introduced a novel approach using a variational auto-encoder (VAE), to improve explainability of ECG analysis. Unlike post hoc methods, this technique incorporates explainability directly into the model's design.[49] The VAE was trained to compress the median-beat 12-lead ECG into a set of explanatory factors. These factors can subsequently be used in common, interpretable prediction models, such as Cox regression, allowing for the identification of factors associated with the predicted outcome. Using the decoder part of the VAE,

the factors associated with specific outcomes can be visualized in a reconstructed ECG, providing insights into what patterns the model has learned. This approach offers explainability both at the individual prediction level and at the model level.[49]

9 | STANDARDIZATION AND INTEROPERABILITY

Interoperability, particularly regarding Electronic Health Records (EHRs), is essential for deploying AI models in clinical settings. Addressing this issue through consistent ECG data standards and interoperable systems is crucial for ensuring the accuracy, reliability and generalizability of AI-driven ECG analysis.

Currently, ECG data can be obtained from different devices, including wearables, and in diverse settings, such as ambulatory ECGs, resting ECGs and stress tests. Data standardization enables consistent interpretation across studies. Key techniques for ECG data standardization include signal processing, amplitude and time normalization, and data annotation using standardized formats like MIT-BIH annotation guidelines.[85] Format standardization can be achieved using common formats like the European Data Format (EDF), DICOM or Waveform Database format (WFDB). WFDB, an open-source set of file standards for reading and storing signal data, is widely adopted by both research and clinical applications, notably by PhysioNet databases.[86–88]

Another important aspect requiring standardization is the reporting system in AI studies. The lack of uniformity among studies makes comparison and pooled analysis challenging. Adopting standard guidelines for AI data reporting with transparency and completeness is essential for critical assessment and bias identification. Furthermore, it may facilitate development of robust meta-analyses, which could aggregate data and provide stronger evidence supporting the use of AI tools in clinical practice. In recent years various guidelines to standardize reporting of studies have been developed, from algorithm development and validation (TRIPOD-AI and STARD-AI)[89,90] to large-scale randomized trials (CONSORT-AI).[91]

10 | LIMITATIONS AND CONCLUSIONS

This is not a comprehensive review; the detection and prediction of various other diseases have been investigated using DL but are not included in this work, nor have all papers on specific disease been included due to the shear abundance of available literature.

We can conclude that DL-enabled ECG analysis is a developing field, with various potential applications in diagnosis and prediction of cardiac and noncardiac diseases. Furthermore, implementation of DL tools in clinical practice is being investigated prospectively in different randomized studies. The results of those studies will help to evaluate the feasibility of implementing these algorithms in real-world clinical practice and their utility compared to current systems. However, implementation of AI systems into the workflow of healthcare systems remains to face several barriers, including concerns related to accuracy and safety, data protection, transparency of the algorithm and interoperability. Despite advances in regulations further investigation and specific guidelines are needed to ensure an increase in deployment of novel AI developments into clinical practice.

CONFLICT OF INTEREST STATEMENT

The authors declare the following potential conflicts of interest: The Department of Cardiology at UMC Utrecht may receive royalties in the future from sales of deep learning ECG algorithms developed by Cordys Analytics, a spin-off company. Additionally, René van Es is the Chief Scientific Officer (CSO) and a shareholder of Cordys Analytics. These affiliations and potential financial interests have been disclosed and are being managed in accordance with institutional policies.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID

Bas B. S. Schots https://orcid.org/0009-0009-0291-7940
Bauke K. O. Arends https://orcid.org/0009-0009-7494-4542
René van Es https://orcid.org/0000-0001-9950-4388

REFERENCES

1. Wagner GS, Macfarlane P, Wellens H, et al. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part VI: acute ischemia/infarction: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol*. 2009;53(11):1003-1011. doi:10.1016/j.jacc.2008.12.016

2. Surawicz B, Childers R, Deal BJ, et al. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part III: intraventricular conduction disturbances: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical ardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol*. 2009;53(11):976-981. doi:10.1016/j.jacc.2008.12.013

3. Breen CJ, Kelly GP, Kernohan WG. ECG interpretation skill acquisition: a review of learning, teaching and assessment. *J Electrocardiol*. 2022;73:125-128. doi:10.1016/j.jelectrocard.2019.03.010

4. Salerno SM, Alguire PC, Waxman HS. Competency in interpretation of 12-lead electrocardiograms: a summary and appraisal of published evidence. *Ann Intern Med*. 2003;138(9):751-760. doi:10.7326/0003-4819-138-9-200305060-00013

5. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28(1):31-38. doi:10.1038/s41591-021-01614-0

6. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med*. 2020;3(1):118. doi:10.1038/s41746-020-00324-0

7. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. *Eur Heart J*. 2021;42(46):4717-4730. doi:10.1093/eurheartj/ehab649

8. Somani S, Russak AJ, Richter F, et al. Deep learning and the electrocardiogram: review of the current state-of-the-art. *Europace*. 2021;23(8):1179-1191. doi:10.1093/europace/euaa377

9. Musa N, Gital AY, Aljojo N, et al. A systematic review and meta-data analysis on the applications of deep learning in electrocardiogram. *J Ambient Intell Humaniz Comput*. 2023;14(7):9677-9750. doi:10.1007/s12652-022-03868-z

10. Petmezas G, Stefanopoulos L, Kilintzis V, et al. State-of-the-art deep learning methods on electrocardiogram data: systematic review. *JMIR Med Inform*. 2022;10(8):e38454. doi:10.2196/38454

11. Jo YY, Kwon JM, Jeon KH, et al. Detection and classification of arrhythmia using an explainable deep learning model. *J Electrocardiol*. 2021;67:124-132. doi:10.1016/j.jelectrocard.2021.06.006

12. Bender T, Seidler T, Bengel P, Sax U, Krefting D. Application of pre-trained deep learning models for clinical ECGs. *Stud Health Technol Inform*. 2021;283:39-45. doi:10.3233/shti210539

13. Yang T, Yu L, Jin Q, Wu L, He B. Localization of origins of premature ventricular contraction by means of convolutional neural network from 12-Lead ECG. *IEEE Trans Biomed Eng*. 2018;65(7):1662-1671. doi:10.1109/tbme.2017.2756869

14. Aufiero S, Bleijendaal H, Robyns T, et al. A deep learning approach identifies new ECG features in congenital long QT syndrome. *BMC Med.* 2022;20(1):162. doi:10.1186/s12916-022-02350-z

15. Bajaj S, Bennett MT, Rabkin SW. Identifying premature ventricular complexes from outflow tracts based on PVC configuration: a machine learning approach. *J Clin Med.* 2023;12(17):5558. doi:10.3390/jcm12175558

16. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet.* 2019;394(10201):861-867. doi:10.1016/S0140-6736(19)31721-0

17. Raghunath S, Pfeifer JM, Ulloa-Cerna AE, et al. Deep neural networks can predict new-onset atrial fibrillation from the 12-Lead ECG and help identify those at risk of atrial fibrillation-related stroke. *Circulation.* 2021;143(13):1287-1298. doi:10.1161/CIRCULATIONAHA.120.047829

18. Bortolan G. 3D ECG display with deep learning approach for identification of cardiac abnormalities from a variable number of leads. *Physiol Meas.* 2023;44(2):024002. doi:10.1088/1361-6579/acb4dc

19. Gliner V, Keidar N, Makarov V, Avetisyan AI, Schuster A, Yaniv Y. Automatic classification of healthy and disease conditions from images or digital standard 12-lead electrocardiograms. *Sci Rep.* 2020;10(1):16331. doi:10.1038/s41598-020-73060-w

20. Yildirim O, Talo M, Ciaccio EJ, Tan RS, Acharya UR. Accurate deep neural network model to detect cardiac arrhythmia on more than 10,000 individual subject ECG records. *Comput Methods Prog Biomed.* 2020;197:105740. doi:10.1016/j.cmpb.2020.105740

21. Yoon T, Kang D. Bimodal CNN for cardiovascular disease classification by co-training ECG grayscale images and scalograms. *Sci Rep.* 2023;13(1):2937. doi:10.1038/s41598-023-30208-8

22. Kashou AH, Ko WY, Attia ZI, Cohen MS, Friedman PA, Noseworthy PA. A comprehensive artificial intelligence-enabled electrocardiogram interpretation program. *Cardiovasc Digit Health J.* 2020;1(2):62-70. doi:10.1016/j.cvdhj.2020.08.005

23. Ribeiro AH, Ribeiro MH, Paixão GMM, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun.* 2020;11(1):1760. doi:10.1038/s41467-020-15432-4

24. Zhang J, Liang D, Liu A, et al. MLBF-net: a multi-lead-branch fusion network for multi-class arrhythmia classification using 12-lead ECG. *IEEE J Transl EngHealth Med.* 2021;2021(9):1900211. doi:10.1109/JTEHM.2021.3064675

25. Zhou F, Chen L. Leadwise clustering multi-branch network for multi-label ECG classification. *Med Eng Phys.* 2024;130:104196. doi:10.1016/j.medengphy.2024.104196

26. van de Leur RR, Blom LJ, Gavves E, et al. Automatic triage of 12-lead ECGs using deep convolutional neural networks. *J Am Heart Assoc.* 2020;9(10):e015138. doi:10.1161/jaha.119.015138

27. Smith SW, Walsh B, Grauer K, et al. A deep neural network learning algorithm outperforms a conventional algorithm for emergency department electrocardiogram interpretation. *J Electrocardiol.* 2019;52:88-95. doi:10.1016/j.jelectrocard.2018.11.013

28. Chen CY, Lin YT, Lee SJ, et al. Automated ECG classification based on 1D deep learning network. *Methods.* 2022;202:127-135. doi:10.1016/j.ymeth.2021.04.021

29. van de Leur RR, van Sleuwen M, Zwetsloot PM, et al. Automatic triage of twelve-lead electrocardiograms using deep convolutional neural networks: a first implementation study. *Eur Heart J Digit Health.* 2024;5(1):89-96. doi:10.1093/ehjdh/ztad070

30. Motazedian P, Prosperi-Porta G, Hibbert B, et al. Cost-effectiveness of population screening for aortic stenosis. *Eur Heart J Qual Care Clin Outcomes.* 2024;qcae043. doi:10.1093/ehjqcco/qcae043

31. Kalmady SV, Salimi A, Sun W, et al. Development and validation of machine learning algorithms based on electrocardiograms for cardiovascular diagnoses at the population level. *NPJ Digit Med.* 2024;7(1):133. doi:10.1038/s41746-024-01130-8

32. Kwon JM, Kim KH, Akkus Z, Jeon KH, Park J, Oh BH. Artificial intelligence for detecting mitral regurgitation using electrocardiography. *J Electrocardiol.* 2020;59:151-157. doi:10.1016/j.jelectrocard.2020.02.008

33. Lin YT, Lin CS, Tsai CS, et al. Comprehensive clinical application analysis of artificial intelligence-enabled electrocardiograms for screening multiple valvular heart diseases. *Aging (Albany NY).* 2024;16(10):8717-8731. doi:10.18632/aging.205835

34. Vaid A, Argulian E, Lerakis S, et al. Multi-center retrospective cohort study applying deep learning to electrocardiograms to identify left heart valvular dysfunction. *Commun Med.* 2023;3(1):24. doi:10.1038/s43856-023-00240-w

35. Cohen-Shelly M, Attia ZI, Friedman PA, et al. Electrocardiogram screening for aortic valve stenosis using artificial intelligence. *Eur Heart J.* 2021;42(30):2885-2896. doi:10.1093/eurheartj/ehab153

36. Kwon JM, Lee SY, Jeon KH, et al. Deep learning-based algorithm for detecting aortic stenosis using electrocardiography. *J Am Heart Assoc.* 2020;9(7):e014717. doi:10.1161/JAHA.119.014717

37. Elias P, Poterucha TJ, Rajaram V, et al. Deep learning electrocardiographic analysis for detection of left-sided valvular heart disease. *J Am Coll Cardiol.* 2022;80(6):613-626. doi:10.1016/j.jacc.2022.05.029

38. Semsarian C, Ingles J, Maron MS, Maron BJ. New perspectives on the prevalence of hypertrophic cardiomyopathy. *J Am Coll Cardiol.* 2015;65(12):1249-1254. doi:10.1016/j.jacc.2015.01.019

39. Ko WY, Siontis KC, Attia ZI, et al. Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. *J Am Coll Cardiol.* 2020;75(7):722-733. doi:10.1016/j.jacc.2019.12.030

40. Siontis KC, Wieczorek MA, Maanja M, et al. Hypertrophic cardiomyopathy detection with artificial intelligence electrocardiography in international cohorts: an external validation study. *Eur Heart J Digit Health.* 2024;5(4):416-426. doi:10.1093/ehjdh/ztae029

41. Sammani A, Jansen M, de Vries NM, et al. Automatic identification of patients with unexplained left ventricular hypertrophy in electronic Health record data to improve targeted treatment and family screening. *Front Cardiovasc Med.* 2022;9:768847. doi:10.3389/fcvm.2022.768847

42. Coats CJ, Masri A, Nassif ME, et al. Dosing and safety profile of Aficamten in symptomatic obstructive hypertrophic cardiomyopathy: results from SEQUOIA-HCM. *J Am Heart Assoc.* 2024;13(15):e035993. doi:10.1161/JAHA.124.035993

43. Olivotto I, Oreziak A, Barriales-Villa R, et al. Mavacamten for treatment of symptomatic obstructive hypertrophic

cardiomyopathy (EXPLORER-HCM): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2020;396(10253):759-769. doi:10.1016/S0140-6736(20)31792-X

44. Force T, McDonagh TA, Metra M, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail*. 2022;24(1):4-131. doi:10.1002/ejhf.2333

45. Aljaroudi W, Alraies MC, Halley C, et al. Impact of progression of diastolic dysfunction on mortality in patients with normal ejection fraction. *Circulation*. 2012;125(6):782-788. doi:10.1161/CIRCULATIONAHA.111.066423

46. Sangha V, Nargesi AA, Dhingra LS, et al. Detection of left ventricular systolic dysfunction from electrocardiographic images. *Circulation*. 2023;148(9):765-777. doi:10.1161/circulationaha.122.062646

47. Klein CJ, Ozcan I, Attia ZI, et al. Electrocardiogram-artificial intelligence and immune-mediated necrotizing myopathy: predicting left ventricular dysfunction and clinical outcomes. *Mayo Clin Proc Innov Qual Outcomes*. 2022;6(5):450-457. doi:10.1016/j.mayocpiqo.2022.08.003

48. Sun JY, Qiu Y, Guo HC, et al. A method to screen left ventricular dysfunction through ECG based on convolutional neural network. *J Cardiovasc Electrophysiol*. 2021;32(4):1095-1102. doi:10.1111/jce.14936

49. van de Leur RR, Bos MN, Taha K, et al. Improving explainability of deep neural network-based electrocardiogram interpretation using variational auto-encoders(). *Eur Heart J Digit Health*. 2022;3(3):390-404. doi:10.1093/ehjdh/ztac038

50. Cho J, Lee B, Kwon JM, et al. Artificial intelligence algorithm for screening heart failure with reduced ejection fraction using electrocardiography. *ASAIO J*. 2021;67(3):314-321. doi:10.1097/MAT.0000000000001218

51. Kwon JM, Kim KH, Eisen HJ, et al. Artificial intelligence assessment for early detection of heart failure with preserved ejection fraction based on electrocardiographic features. *Eur Heart J Digit Health*. 2021;2(1):106-116. doi:10.1093/ehjdh/ztaa015

52. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25(1):70-74. doi:10.1038/s41591-018-0240-2

53. Attia ZI, Kapa S, Yao X, et al. Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction. *J Cardiovasc Electrophysiol*. 2019;30(5):668-674. doi:10.1111/jce.13889

54. Yagi R, Goto S, Katsumata Y, MacRae CA, Deo RC. Importance of external validation and subgroup analysis of artificial intelligence in the detection of low ejection fraction from electrocardiograms. *Eur Heart J Digit Health*. 2022;3(4):654-657. doi:10.1093/ehjdh/ztac065

55. König S, Hohenstein S, Nitsche A, et al. Artificial intelligence-based identification of left ventricular systolic dysfunction from 12-lead electrocardiograms: external validation and advanced application of an existing model. *Eur Heart J Digit Health*. 2024;5(2):144-151. doi:10.1093/ehjdh/ztad081

56. Byrne RA, Rossello X, Coughlan JJ, et al. 2023 ESC guidelines for the management of acute coronary syndromes: developed by the task force on the management of acute coronary syndromes of the European Society of Cardiology (ESC). *Eur Heart J*. 2023;44(38):3720-3826. doi:10.1093/eurheartj/ehad191

57. Gulati M, Levy PD, Mukherjee D, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain: executive summary: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. *Circulation*. 2021;144(22):e336-e367. doi:10.1161/CIR.0000000000001030

58. Chen KW, Wang YC, Liu MH, et al. Artificial intelligence-assisted remote detection of ST-elevation myocardial infarction using a mini-12-lead electrocardiogram device in prehospital ambulance care. *Front Cardiovas Med*. 2022;9(9):1001982. doi:10.3389/fcvm.2022.1001982

59. Chang KC, Hsieh PH, Wu MY, et al. Usefulness of multi-labelling artificial intelligence in detecting rhythm disorders and acute ST-elevation myocardial infarction on 12-lead electrocardiogram. *Eur Heart J Digit Health*. 2021;2(2):299-310. doi:10.1093/ehjdh/ztab029

60. Byrne RA, Rossello X, Coughlan JJ, et al. 2023 ESC guidelines for the management of acute coronary syndromes. *Eur Heart J Acute Cardiovasc Care*. 2024;13(1):55-161. doi:10.1093/ehjacc/zuad107

61. Herman R, Meyers HP, Smith SW, et al. International evaluation of an artificial intelligence-powered electrocardiogram model detecting acute coronary occlusion myocardial infarction. *Eur Heart J Digit Health*. 2024;5(2):123-133. doi:10.1093/ehjdh/ztad074

62. Liu WC, Lin CS, Tsai CS, et al. A deep learning algorithm for detecting acute myocardial infarction. *EuroIntervention*. 2021;17(9):765-773. doi:10.4244/EIJ-D-20-01155

63. Gustafsson S, Gedon D, Lampa E, et al. Development and validation of deep learning ECG-based prediction of myocardial infarction in emergency department patients. *Sci Rep*. 2022;12(1):19615. doi:10.1038/s41598-022-24254-x

64. Yao X, Rushlow DR, Inselman JW, et al. Artificial intelligence–enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat Med*. 2021;27(5):815-819. doi:10.1038/s41591-021-01335-4

65. Adedinsewo DA, Morales-Lara AC, Afolabi BB, et al. Artificial intelligence guided screening for cardiomyopathies in an obstetric population: a pragmatic randomized clinical trial. *Nat Med*. 2024;30(10):2897-2906. doi:10.1038/s41591-024-03243-9

66. Isogai T, Kamiya CA. Worldwide incidence of peripartum cardiomyopathy and overall maternal mortality. *Int Heart J*. 2019;60(3):503-511. doi:10.1536/ihj.18-729

67. Lin C-S, Liu W-T, Tsai D-J, et al. AI-enabled electrocardiography alert intervention and all-cause mortality: a pragmatic randomized clinical trial. *Nat Med*. 2024;30(5):1461-1470. doi:10.1038/s41591-024-02961-4

68. Lin C, Liu WT, Chang CH, et al. Artificial intelligence–powered rapid identification of ST-elevation myocardial infarction via electrocardiogram (ARISE) — a pragmatic randomized controlled trial. *NEJM AI*. 2024;1(7):AIoa2400190. doi:10.1056/AIoa2400190

69. Shah S. Northwestern tempus AI-enabled electrocardiography (NOTABLE) trial (NOTABLE). https://clinicaltrials.gov/study/NCT065115052024

70. Jung IH. Evaluating AI-Gatekeeper software in coronary artery stenosis screening: a multicenter RCT (AIGatekeeper). 2023. https://clinicaltrials.gov/study/NCT06178900

71. van de Sande D, Van Genderen ME, Smit JM, et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter. *BMJ Health Care Inform*. 2022;29(1):e100495. doi:10.1136/bmjhci-2021-100495

72. Nestor Maslej LF, Perrault R, Parli V, et al. "The AI Index 2024 Annual Report,". AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. 2024.

73. Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. *BMC Med*. 2023;21(1):70. doi:10.1186/s12916-023-02779-w

74. Young BSJ-J. The new ISO/IEC standard for automated ECG interpretation. *Heart*. 2021;2:410-418.

75. Administration USFD. Premarket Approval (PMA). 2024.

76. Directorate-General for Internal Market I, Entrepreneurship and SMEs. Internal market, industry, entrepreneurship and SMEs. 2024. https://single-market-economy.ec.europa.eu/single-market/ce-marking/manufacturers_en

77. Artificial Intelligence Act. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. COM(2021) 206 final. 2024, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

78. General data protection regulation (GDPR), regulation (EU) 2016/679. *Off J Eur Union*. 2016; 32016R0679:1-88 L 119/1.

79. Gadotti A, Rocher L, Houssiau F, Cretu AM, de Montjoye YA. Anonymization: the imperfect science of using data while preserving privacy. *Sci Adv*. 2024;10(29):eadn7053. doi:10.1126/sciadv.adn7053

80. Gu X, Sabrina F, Fan Z, Sohail S. A review of privacy enhancement methods for federated learning in healthcare systems. *Int J Environ Res Public Health*. 2023;20(15):6539. doi:10.3390/ijerph20156539

81. Balasubramaniam N, Kauppinen M, Rannisto A, Hiekkanen K, Kujala S. Transparency and explainability of AI systems: from ethical guidelines to requirements. *Inf Softw Technol*. 2023;159:107197. doi:10.1016/j.infsof.2023.107197

82. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1(5):206-215. doi:10.1038/s42256-019-0048-x

83. Chang CH, Lin CS, Luo YS, Lee YT, Lin C. Electrocardiogram-based heart age estimation by a deep learning model provides more information on the incidence of cardiovascular disorders. *Front Cardiovasc Med*. 2022;9:754909. doi:10.3389/fcvm.2022.754909

84. Diamant N, Di Achille P, Weng LC, et al. Deep learning on resting electrocardiogram to identify impaired heart rate recovery. *Cardiovasc Digit Health J*. 2022;3(4):161-170. doi:10.1016/j.cvdhj.2022.06.001

85. Safdar MF, Nowak RM, Pałka P. Pre-processing techniques and artificial intelligence algorithms for electrocardiogram (ECG) signals analysis: a comprehensive review. *Comput Biol Med*. 2024;170:107908. doi:10.1016/j.compbiomed.2023.107908

86. Goldberger A, Amaral L, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000;101(23):e215-e220.

87. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3(1):160035. doi:10.1038/sdata.2016.35

88. Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. *IEEE Eng Med Biol Mag*. 2001;20:45-50.

89. Collins GA-O, Moons KA-OX, Dhiman PA-O, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378

90. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open*. 2021;11(6):e047709. doi:10.1136/bmjopen-2020-047709

91. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med*. 2020;26(9):1364-1374. doi:10.1038/s41591-020-1034-x

92. Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomed Tech*. 1995;40:317-318.

93. Wagner P, Strodthoff N, Bousseljot R-D, et al. PTB-XL, a large publicly available electrocardiography dataset. *Sci Data*. 2020;7(1):154. doi:10.1038/s41597-020-0495-6

94. Jager F, Taddei A, Moody GB, et al. Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Med Biol Eng Comput*. 2003;41(2):172-182. doi:10.1007/BF02344885

95. Taddei A, Distante G, Emdin M, et al. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *Eur Heart J*. 1992;13:1164-1172.

96. Kalyakulina AI, Yusipov II, Moskalenko VA, et al. LUDB: a new open-access validation tool for electrocardiogram delineation algorithms. *IEEE Access*. 2020;8:186181-186190. doi:10.1109/ACCESS.2020.3029211

97. Lugovaya TS. Biometric human identification based on electrocardiogram. *[Master's thesis] Faculty of Computing Technologies and Informatics, Electrotechnical University "LETI", Saint-Petersburg, Russian Federation*. 2005.

98. Liu H, Chen D, Chen D, et al. A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements. *Sci Data*. 2022;9(1):272. doi:10.1038/s41597-022-01403-5