



Mini Review

Copyright © Pavel Antiperovitch

# Should we Trust Artificial Intelligence to Read Our ECGs?

**Nan Shi MD and Pavel Antiperovitch MD FRCPC\***

Department of Medicine, Division of Cardiology, Western University, London Health Sciences Centre, London, Canada

\*Corresponding author: Pavel Antiperovitch MD FRCPC, Department of Medicine, Division of Cardiology University Hospital, London, Canada.

To Cite This Article: Nan Shi MD and Pavel Antiperovitch MD FRCPC\*. Should we Trust Artificial Intelligence to Read Our ECGs?. Am J Biomed Sci & Res. 2024 21(1) AJBSR.MS.ID.002794, DOI: 10.34297/AJBSR.2024.21.002794

Received: 📅 December 19, 2023; Published: 📅 January 03, 2024

## Abstract

In the 21st century, Artificial Intelligence (AI) is being increasingly studied and developed for medical applications, including Electrocardiogram (ECG) interpretation. Despite promising results, there is no consensus in the medical community regarding the reliability of AI or its potential to become standard of care when it comes to ECG interpretation. This review aims to examine the evolution of AI in ECG interpretation, starting from basic QRS identification to surpassing Cardiologists in diagnosing arrhythmias. Additionally, we explore the role of AI in diagnosing diseases outside of arrhythmia and its ability to predict response to cardiac treatments. Finally, we discuss the challenges AI must overcome before it can move beyond the realm of research and become widely accepted as the standard of care.

**Keywords:** Artificial intelligence, Electrocardiogram, Machine learning, Deep learning

## Introduction

Since the advent of the first single-lead ECG in 1887, it has been the most common diagnostic test for arrhythmia [1]. While ECG interpretation is a mandatory part of any medical curriculum, it is a challenging to do so accurately, with great inter-clinician variability [2]. Since the 1950s, efforts have been made to automate ECG interpretation through computers [3]. However, as of today computer interpretations of ECGs are only used as an adjunct, not substitute, for physician interpretation given the rates of error [4]. No doubt, computers still struggle to accurately interpret ECGs, especially certain arrhythmias.

In the past decade, the field of Artificial Intelligence (AI) has experienced a meteoric rise [5]. The use of AI in diagnostic testing involves computational extraction of a pattern from a training dataset and applying it to make predictions about unseen data. Machine Learning (ML) is a subset of AI where the computer fits data from a training set into statistical models, without any external human definition, by minimizing prediction error ("cost function") compared to traditional probability and statistics, especially in non-linear models. Deep Learning (DL) is a subset of ML that mimics the human nervous system through linear functions ("nodes")

arranged in series ("layers"). During training, each node is thought to represent a simple feature of the training data; there are many layers in a DL model hence the term "deep". The main advantage of DL compared to ML lies in its ability to process non-numerical data such as images [6]. This makes DL an especially attractive option to study ECGs.

## Arrhythmia Diagnosis

Attempts to study arrhythmia via AI date back to last century. In 1993, *Edenbrandt, et al.* [7]. used 500 ECG ST segments to train an artificial neural network to identify ST changes such as ST elevation. The network was able to characterize ST elevation much more accurately than conventional criteria but fell short when compared to an experienced cardiologist [7]. More promising results were delivered in 2007 when *Yu, et al.* used wavelet transformation and probabilistic neural networks to classify ECG beats, including bundle branch block and premature atrial/ventricular rhythms, with 99.65% accuracy. However, the study has drawbacks including a small sample size of 23 ECGs and lack of comparison with conventional criteria or cardiologist interpretation [8].



A large-scale study was done in Brazil in 2018 using 1,557,415 ECGs in Brazil were used to train a deep neural network to detect 6 different rhythms. The model's accuracy was represented by the F1 score. The DL model was able to outperform emergency residents and final year medical students. It even performed better than cardiology residents in 5 out of 6 rhythms [9]. In a landmark study by Hannun, *et al.* [10], a deep neural network was trained with 91,232 single-lead ECGs to identify rhythms including normal sinus, atrial fibrillation, AV block, and ventricular tachycardia. The gold standard was a panel of expert cardiologists who agree on a diagnosis. The model performed extremely well, with an AUC of 0.97, which illustrates almost perfect arrhythmia identification. This model was found to perform similarly to individual cardiologists in diagnostic accuracy and made similar mistakes as cardiologists [10]. This study confirmed that our current DL models can attain the accuracy of Cardiologists for certain arrhythmia diagnoses.

### Identification of Biosignature for Disease

In addition to interpreting ECG rhythms, DL models are also used to identify the biological signature of silent diseases such as asymptomatic cardiomyopathy. At the Mayo Clinic, a convolutional neural network was trained to identify patients with reduced heart function, defined as left ventricular ejection fraction (LVEF)  $\leq 35\%$ , based on ECGs alone. The model was tested on 52,870 patients, with sensitivity and specificity over 85%. Notably, the "false positives" identified by the neural network were found 4.1 times more likely to develop cardiomyopathy in the following 5 years compared to true negatives [11]. This revealed that the model identified an ECG biosignature of early cardiomyopathy, which can be used to predict a drop in ejection fraction, allowing initiation of early preventative treatment. Being able to predict a reduced ejection fraction based on a readily available test such as ECG can transform the way we practice medicine, since ECGs are far more available than echocardiograms [12]. This application opens the door for ECG to become a cheap, accessible tool for asymptomatic patient LVEF screening and monitoring.

### Prediction of Treatment Response

DL models can also be used to predict outcomes as well as responses to specific patient interventions. One such challenge is selecting patients who would benefit from a cardiac device implantation called Cardiac Resynchronization Therapy (CRT). This is a special pacemaker that can pace both sides of the heart to provide synchronous contraction [13]. However, identifying patients who would benefit from this is challenging, as nearly 1/3 of patients do not respond to CRT, with response usually defined as echocardiographic evidence of reverse remodelling of the heart such as recovery of LVEF [14]. In 2023, Wouters, *et al.* from Utrecht used a DL model predict response to CRT using a pre-procedure ECG. In fact, investigators were also able to predict the risk of requiring a heart transplant, left ventricular assistive device, and death in heart failure patients. They used an DL-based autoencoder that transforms ECGs into 21 numbers that are used in the algorithm to calculate

patient outcomes. The model significantly outperformed current methods of screening patients for CRT therapy [15]. Investigators also examined how this model was able to make its predictions and found that the ST segment in addition to the QRS holds a lot of prognostic information. This showcases the potential of DL to turn ECG into a treatment-response prediction tool.

### Challenges to Overcome

The operationalization of deep learning models in ECG interpretation poses several challenges at a full scale. Many of the models published face limitations in achieving the performance necessary for practical clinical applications. Take, for instance, a model designed to screen ECGs for left ventricular ejection fraction, which exhibited an alarming tendency to overcall reduced heart function in 66% of cases. Additional techniques can be used to boost performance to a clinically acceptable range. To enhance the reliability and applicability of neural networks, it is crucial to undergo external validation using independently generated datasets as a key developmental step. Explain ability also emerges as a paramount factor in the development of deep learning tools. Investigators must employ techniques like LIME and SHAP to understand the decision-making process of the model. Without a clear understanding of how the neural network operates and which features it relies on, clinicians are unlikely to place trust in its predictions. Each DL model intended for practical use must confront and resolve these pivotal challenges.

### Conclusion

The short answer to the question of whether we can place trust in AI to read our ECGs is yes, but not yet. The field of AI has made significant strides, having reached the performance and sometimes outperforming medical professionals in certain applications. However, for it to be a standard of care, relying solely on AI as the ECG reader requires extensive validation across multiple centres in different countries. Such models must also pass the hurdles of regulatory approval, although similar technologies for arrhythmia detection are already getting approved by the Food & Drug Administration [16]. For now, clinicians should be comfortable with employing AI as a more advanced version of existing computerized interpretations of ECGs - as an adjunct and reference. Aside from ECG interpretation, AI can creatively turn the cheap, accessible test that is the ECG into a screening tool for conditions that often require more expensive or invasive testing for prediction of patient outcomes and responsiveness to treatment. The future of electrophysiology, cardiology, and medicine will be forever, if not already changed by AI.

### Acknowledgements

None.

### Conflict of Interest

None.

## References

1. Breen C, Bond R, Finlay D (2019) A clinical decision support tool to assist with the interpretation of the 12-lead electrocardiogram. *Health Informatics J* 25(1): 51-61.
2. Breen CJ, Kelly GP, Kernohan WG (2022) ECG interpretation skill acquisition: A review of learning, teaching and assessment. *J Electrocardiol* 73: 125-128.
3. Taback L, Marden E, Mason HL, Pipberger HV (1959) Digital recording of electrocardiographic data. *Ire Trans Biomed Electron.* 6(3): 167-171.
4. Schläpfer J, Wellens HJ (2017) Computer-Interpreted Electrocardiograms: Benefits and Limitations. *J Am Coll Cardiol* 70(9): 1183-1192.
5. Somani S, Russak AJ, Richter F, Zhao S, Vaid A, et al. (2021) Deep learning and the electrocardiogram: review of the current state-of-the-art. *Eurospace* 23(8): 1179-1191.
6. Chen D, Liu S, Kingsbury P, Sohn S, Storlie CB, et al. (2019) Deep learning and alternative learning strategies for retrospective real-world clinical data. *NPJ Digit Med* 2(1): 43.
7. Edenbrandt L, Devine B, Macfarlane PW (1993) Classification of electrocardiographic ST-T segments-human expert vs artificial neural network. *Eur Heart J* 14(4): 464-468.
8. Yu S, Chen Y (2007) Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. *Pattern Recognit Lett* 28(10): 1142-1150.
9. Ribeiro AL, Paixão GMM, Gomes PR, Ribeiro MH, Ribeiro AH, et al. (2019) Tele-electrocardiography and bigdata: The CODE (Clinical Outcomes in Digital Electrocardiography) study. *J Electrocardiol* 57(S): 75-78.
10. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, et al. (2019) Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med* 25(1): 65-69.
11. Attia ZI, Kapa S, Lopez Jimenez F, McKie PM, Ladewig DJ, et al. (2019) Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med* 25(1): 70-74.
12. McDonagh TA, McDonald K, Maisel AS (2008) Screening for asymptomatic left ventricular dysfunction using B-type natriuretic Peptide. *Congest Heart Fail* 14(4S): 5-8.
13. McAlister FA, Ezekowitz J, Hooton N, Vendermeer B, Spooner et al. (2007) Cardiac resynchronization therapy for patients with left ventricular systolic dysfunction: a systematic review. *JAMA* 297(22): 2502-2514.
14. Nakai T, Ikeya Y, Kogawa R, Otsuka N, Wakamatsu Y, et al. (2021) What Are the Expectations for Cardiac Resynchronization Therapy? A Validation of Two Response Definitions. *J Clin Med* 10(3): 514.
15. Wouters PC, van de Leur RR, Vessies MB, Van Stipdonk AMW, Ghossein MA, et al. (2023) Electrocardiogram-based deep learning improves outcome prediction following cardiac resynchronization therapy. *Eur Heart J* 44(8): 680-692.
16. (2023) U.S. Food & Drug Administration. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices.