npj Digital Medicine

https://doi.org/10.1038/s41746-026-02485-w

Article in Press

Reimagining atrial fibrillation screening beyond age-based thresholds using AI

Received: 20 December 2025

Accepted: 14 February 2026

Cite this article as: Menon, T.P., Mahajan, A., Powell, D. Reimagining atrial fibrillation screening beyond age-based thresholds using AI. *npj Digit. Med.* (2026). https://doi.org/10.1038/s41746-026-02485-w

Tara P. Menon, Arjun Mahajan & Dylan Powell

Reimagining Atrial Fibrillation Screening Beyond Age-Based Thresholds Using AI

Tara P. Menon[1], Arjun Mahajan[2], Dylan Powell[3]*

[1]Virginia Tech Carilion School of Medicine, Roanoke, VA, USA
[2]Harvard Medical School, Boston, MA, USA
[3]Faculty of Health Sciences & Sport, University of Stirling, Stirling, United Kingdom

***Corresponding Author**: dylan.powell@stir.ac.uk

Abstract

Atrial fibrillation (AF) affects over 50 million people worldwide and carries substantial downstream morbidity, mortality, and cost. Yet many contemporary screening programs rely primarily on age thresholds -- an approach that is operationally simple but can be imprecise for identifying near-term risk. AI applied to handheld single-lead ECGs can predict incident AF with accuracy similar to established clinical risk scores, but real-world deployment remains limited by signal noise, workflow complexity, and unclear risk thresholds.

Main Text

Atrial fibrillation (AF) is a growing public health burden, affecting about 52.6 million people worldwide in 2021, and accounting for roughly 340,000 deaths. In the U.S. alone, patients with AF incur an additional $6,200 in healthcare costs per year compared with those without[1,2]. European Society of Cardiology guidelines, for example, recommend screening in adults aged ≥65 years during routine clinical encounters (that is, assessments performed 'opportunistically' at a regular visit rather than through a dedicated screening program)[3]. While this approach is easy to implement at scale, age thresholds can be an imprecise proxy for atrial substrate and near-term risk -- screening many people who will not develop AF soon while offering limited guidance for identifying higher-risk individuals under 65.

One illustration of this trade-off comes from VITAL-AF, a pragmatic primary-care trial embedding handheld single-lead ECG screening into routine visits for patients aged ≥65 years[4]. Over two years, incident AF was identified in only a small fraction of screened individuals (3.1%), highlighting how an age cut-off can capture many people who will not develop AF in the near term[4]. Importantly, because VITAL-AF enrolled only older adults, it does not resolve how best to approach risk identification in those under 65—where age thresholds offer little guidance and where earlier atrial remodeling may still matter for prevention.

Clinical risk prediction tools may offer a more targeted alternative. The Cohorts for Aging Research in Genomic Epidemiology Atrial Fibrillation (CHARGE-AF) model accurately predicts incident AF but requires calculation of 11 clinical variables, limiting its use in clinical practice[5]. As a result, despite guideline endorsement of risk-informed approaches, most clinicians fall back on the simple age threshold because it is practical, fast, and does not require additional data entry during routine visits. This gap between methodological rigor and real-world usability emphasizes the need for scalable screening strategies that balance predictive accuracy with clinical practicality.

The Promise of Handheld Electrocardiography

Handheld single-lead ECG devices (1L ECG) offer a potential solution to both the efficiency problem and the complexity of traditional risk scores[6]. These devices generate 30-second ECG recordings using finger electrodes, which are then uploaded via mobile apps to secure servers for automated analysis and review. The 2020 ESC guideline has accepted these devices as a Class I recommendation for opportunistic screening because of their diagnostic validity[3].

While these devices detect current AF, they have rarely been used to predict future risk. Handheld ECG devices differ from standard 12-lead ECGs in several ways, including sampling rates, greater susceptibility to noise, and less controlled recording environments. These factors have historically limited the ability to directly apply machine-learning models trained on 12-lead ECGs to handheld devices[7,8]. Nonetheless, large-scale smartwatch studies using optical photoplethysmography have demonstrated the potential of wrist-worn wearables for AF detection -- the Apple Heart Study, for example, enrolled over 400,000 participants and found that 34% of those with an irregular pulse notification had AF on subsequent ECG patch monitoring, with 84% of notifications concordant with AF[9].

Applying Transfer Learning to ECG Modalities

The discrepancy in modalities is addressed by Khurshid et al. using transfer learning, where a pre-trained convolutional neural network model developed using data from over 450,000 single-lead recordings extracted from 12-lead ECGs was applied to handheld 1L ECG recordings (Figure 1)[10]. The 1L ECG-AI model predicts incident atrial fibrillation using a multi-task architecture that processes thousands of data points per recording. A key practical innovation in this pipeline is the selection of the quietest 10-second window from each 30-second tracing, enabling the model to prioritize cleaner signal segments rather than learning from noise-contaminated recordings.

The transfer learning approach works remarkably well. Despite not having been trained on any data from handheld devices, the 1L ECG-AI model achieved performance similar to a model trained de novo on VITAL-AF data (AUROC 0.672 versus 0.667). Notably, fine-tuning the model on handheld-device data did not offer meaningful improvement, suggesting that broad pre-training on large ECG datasets captures features that generalize well across acquisition settings[10].

Clinical Performance and Risk Reclassification

Beyond technical generalizability across ECG modalities, a key question is whether this transfer learning approach improves clinical risk stratification in practice. In the complete VITAL-AF population, the 1L ECG-AI model combined with age and sex (1L ECG-AI AS) achieved an AUROC of 0.695 (95% CI 0.637–0.742), which was not significantly different from CHARGE-AF (AUROC 0.679, 95% CI 0.623–0.730)[10].

Notably, 1L ECG-AI AS facilitates risk-informed screening that enables substantial efficiency gains. Using a ≥3% two-year risk threshold, the authors report favorable reclassification compared with an age ≥65 strategy, with decision curve analysis supporting net benefit across clinically relevant threshold probabilities[10]. This framing makes the central point: the output is a probability that can be aligned to local resources, follow-up capacity, and patient preferences *if* thresholds are defined transparently.

The model demonstrates physiologic plausibility. Model-attention maps consistently emphasized the P-wave region -- an expected location for subtle signs of atrial stress -- suggesting the network is focusing on physiologically relevant patterns[11]. Performance is better for persistent than paroxysmal forms of AF (AUROC 0.717 versus 0.601), likely reflecting more prominent ECG manifestations in persistent disease[10].

Implementation Challenges and Limitations

Several barriers limit immediate clinical implementation. First, the computational requirements for model inference are modest and well suited to cloud-based deployment, but establishing robust signal-quality preprocessing pipelines and integrating these AI tools into existing clinical workflows may be challenging outside well-resourced health systems, particularly in smaller practices[10]. Signal quality also remains an issue: the VITAL-AF medical assistants were trained specifically for device operation, whereas consumers using these devices at home may obtain noisier tracings[4,12]. The confidence intervals in subgroup analyses, particularly for paroxysmal versus persistent AF, reflect these precision limitations[10].

Moreover, generalizability issues also persist. For example, the VITAL-AF trial population was mostly white (82.9%) from New England family practices, reducing confidence in how the model might perform in regions or populations with different AF prevalence, comorbidities, or typical ECG characteristics[4]. The algorithm has not been studied in younger populations or in patients with implanted devices[10]. The regulatory process for AI-assisted screening devices requires proof of clinical utility in addition to technical validation. Similar to other digital diagnostics, broader adoption depends on outcome studies that demonstrate real-world benefit beyond algorithmic accuracy alone[13].

Future Directions for Risk-Informed Screening

Despite these limitations, several technical and clinical advances could address these barriers. Appropriate risk thresholds will need to balance sensitivity and specificity -- higher risk thresholds increase specificity and positive predictive value, which may be more cost-efficient in resource-limited environments by reducing false-positive referrals, whereas lower risk thresholds increase sensitivity, which may be preferable in broader screening programs where the priority is to minimize missed cases of incident AF[10]. Khurshid et al. report better performance in younger and lower-comorbidity groups, suggesting potential for screening extension to individuals under 65 years[10]. Integration into electronic health records, which may allow automated calculation of AF risk during routine care, could make risk-informed screening easier to adopt.

Longer-term monitoring strategies need further exploration. While the two-year prediction interval corresponds well to screening intervals, accumulating serial 1L ECGs may detect changes in atrial substrate over time[10]. Continuous or intermittent ECG analysis using devices such as smartwatches could offer richer information for dynamic risk updating than single-timepoint screening, but it represents an “active” measurement that requires user engagement and is best suited for confirmatory testing and risk stratification[14]. In contrast, passive, always-on photoplethysmography on smartwatches supports continuous, population-scale surveillance, with ECG (whether handheld, phone-based, or smartwatch-based) reserved for rhythm confirmation and treatment decisions[15,16].

Conclusions

Age thresholds for AF screening can be considered somewhat crude instruments—they may catch too many people who will never develop the condition, while missing younger individuals already at risk. Khurshid et al. demonstrate that AI analysis of a handheld single-lead ECG can predict new cases of AF as effectively as traditional risk prediction scores, using just age, sex, and a 30-second rhythm strip[10]. Yet several challenges must be addressed before these technologies enter routine practice, including signal quality management, determination of appropriate risk thresholds, demonstration of clinical utility in diverse populations, and establishment of regulatory frameworks. Closing the gap between algorithmic potential and clinical adoption could shift AF screening from age-based thresholds to individualized, risk-based prevention.

Acknowledgements
Not applicable.

Competing interests
DP, TM, and AM declare no competing financial or non-financial interest. DP, TM, and AM are News and Views editors at npj Digital Medicine. TM, AM and DP played no role in the internal review or decision to publish this News and Views article.

Author Contributions
T.P.M. drafted the manuscript and created Figure 1. A.M. and D.P. contributed to critical review, revision, and editing of the manuscript. All authors approved the final version of the manuscript.

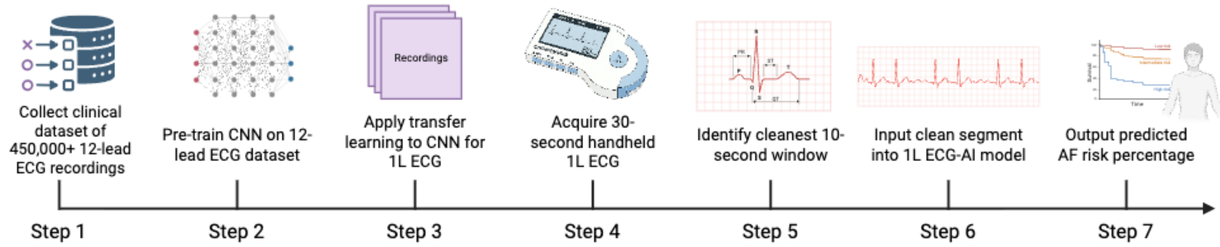
Funding Declaration
The authors received no financial support or grant for the research, authorship, and/or publication of this article.

References

1. Li, X. *et al.* Frontiers | Global burden and health inequality of atrial fibrillation/atrial flutter from 1990 to 2021. https://doi.org/10.3389/fcvm.2025.1585980 doi:10.3389/fcvm.2025.1585980.
2. See, C. *et al.* Healthcare Expenditure on Atrial Fibrillation in the United States. *JACC Adv.* **4**, 101716 (2025).
3. G, H. *et al.* 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *PubMed* https://pubmed.ncbi.nlm.nih.gov/32860505/ (2021).
4. Ashburner, J. M. *et al.* Design and rationale of a pragmatic trial integrating routine screening for atrial fibrillation at primary care visits: The VITAL-AF trial. *Am. Heart J.* **215**, 147–156 (2019).
5. A, A. *et al.* Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *PubMed* https://pubmed.ncbi.nlm.nih.gov/23537808/ (2013).
6. Bray, J. J. H. *et al.* Single-lead ECGs (AliveCor) are a feasible, cost-effective and safer alternative to 12-lead ECGs in community diagnosis and monitoring of atrial fibrillation. *BMJ Open Qual.* **10**, e001270 (2021).
7. Khurshid, S. *et al.* ECG-Based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. *Circulation* **145**, 122–133 (2022).

8. Metcalfe, J. Z. *et al.* Validation of a Handheld 6-Lead Device for QT Interval Monitoring in Resource-Limited Settings. *JAMA Netw. Open* **7**, e2415576 (2024).
9. Perez, M. V. *et al.* Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. *N. Engl. J. Med.* **381**, 1909–1917 (2019).
10. S, K. *et al.* Artificial intelligence-enabled analysis of handheld single-lead electrocardiograms to predict incident atrial fibrillation: an analysis of the VITAL-AF randomized trial. *PubMed* https://pubmed.ncbi.nlm.nih.gov/41299008/ (2025).
11. Impact of P-wave indices in prediction of atrial fibrillation—Insight from loop recorder analysis - Kreimer - 2021 - Annals of Noninvasive Electrocardiology - Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1111/anec.12854.
12. Mahajan, A., Heydari, K. & Powell, D. Wearable AI to enhance patient safety and clinical decision-making. *Npj Digit. Med.* **8**, 176 (2025).
13. Windecker, D. *et al.* Generalizability of FDA-Approved AI-Enabled Medical Devices for Clinical Use. *JAMA Netw. Open* **8**, e258052 (2025).
14. Barrera, N. *et al.* Accuracy of Smartwatches in the Detection of Atrial Fibrillation. *JACC Adv.* **4**, 102133 (2025).
15. Pan, Y. *et al.* Continuous atrial fibrillation monitoring using a wearable smartwatch: Using long-term Holter as reference. *Digit. Health* **11**, 20552076251314105 (2025).
16. Frontiers | Continuous 24-h Photoplethysmogram Monitoring Enables Detection of Atrial Fibrillation. https://www.frontiersin.org/journals/physiology/articles/10.3389/fphys.2021.778775/full.

Figure 1. Transfer learning pipeline enabling AF risk prediction from handheld single-lead ECGs



A schematic overview of the AI-enabled transfer learning pipeline used to predict atrial fibrillation (AF) risk from handheld single-lead (1L) ECG recordings. A convolutional neural network (CNN) is first pre-trained on over 450,000 clinical 12-lead ECGs (Steps 1–2). Transfer learning adapts the model for use with handheld 1L ECGs (Step 3). A 30-second 1L ECG is acquired (Step 4), the cleanest 10-second window is automatically identified (Step 5), and the segment is analyzed by the 1L ECG-AI model (Step 6) to produce an AF risk estimate (Step 7).