



Research Article

Open Access

Artificial intelligence-enabled electrocardiogram for detection of occult left ventricular dysfunction in asymptomatic diabetics

James Bennett^{1*} and Andrew Mitchell¹

¹Roseman University of Health Sciences, Henderson, NV 89014, United States.

Article Info

Volume 2, Issue 2, July 2026

Received : 10 January 2026

Accepted : 15 May 2026

Published : 25 July 2026

doi: [10.62587/AFJCCM.2.2.2026.1-9](https://doi.org/10.62587/AFJCCM.2.2.2026.1-9)

Abstract

Aim: Left ventricular systolic dysfunction occurs frequently in diabetes, and is associated with high risk of deteriorating to heart failure, but it is often silent and echocardiographic screening of all patients is not possible. In this study, the researchers assessed an AI-powered ECG to detect hidden LVF in type-2 DM patients who had no symptoms of LVF. **Methods:** In a cohort of asymptomatic adults with type 2 diabetes a 12-lead electrocardiogram was paired with echocardiography. A convolutional neural network was trained using only the ECG to classify an $LFVEF \leq 40\%$ and its performance evaluated with a held-out test set when compared to echocardiographic measurements. **Results:** 8.6% of 4,200 patients had an EF of 40% or lower. The model was obtained with 0.88 (95% CI 0.84 to 0.91) area under the curve, 82% sensitivity, 78% specificity, and consistent performance with sex, age, glycaemic control and diabetes duration, in the test set. **Conclusions:** AI-based ECG was found to be an effective tool for accurate screening of occult left ventricular dysfunction in asymptomatic diabetic patients, and can be an affordable, scalable method of triaging confirmatory echocardiography.

Keywords: Artificial intelligence, Electrocardiogram, Deep learning, Left ventricular dysfunction, Diabetes mellitus, Screening

© 2026 James Bennett and Andrew Mitchell. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

In most cases, left ventricular systolic dysfunction is a silent disease. However, before the onset of breathlessness and the presence of the oedema associated with heart failure, the ejection fraction may be decreased whilst the patient remains asymptomatic, a situation that may last year's (Wang *et al.*, 2003; Redfield *et al.*, 2003). The importance lies in the fact that it can be slowed with therapy: the therapies which prolong the progress of heart failure are most effective if given early on, and a patient identified in a silent phase will benefit the most, yet by definition, these patients cannot be found on their own (McDonagh *et al.*, 2021).

Roughly, the problem is most severe in diabetes. Compared with individuals who do not have diabetes, those with type 2 diabetes have a higher prevalence of cardiomyopathy in the form of a higher prevalence of

* Corresponding author: James Bennett, Roseman University of Health Sciences, Henderson, NV 89014, United States.

left ventricular dysfunction, some of which is independent of coronary disease and hypertension, and in the general population with type 2 diabetes, the prevalence is often occult and has been detected in successive screening studies ([Boonman-de et al., 2012](#); [Dandamudi et al., 2014](#); [Jia et al., 2018](#)). The problem is that the reference test (echo) cannot be used in everyone with diabetes as it is too expensive, too slow and too reliant on highly trained personnel as an option for population screening ([ElSayed et al., 2023](#)).

The mechanism that makes the diabetic heart susceptible also shapes what a screening test must detect. Independently of coronary narrowing and of the raised blood pressure that so often accompanies it, diabetes injures the myocardium directly, through metabolic disturbance, the accumulation of advanced glycation products, microvascular disease and interstitial fibrosis, and the result is a stiff, and in time a weakened, ventricle whose dysfunction can be present long before it declares itself. Because this injury develops silently and diffusely, it produces no single sign that a clinician can look for, and the electrical remodeling that accompanies it is correspondingly subtle, distributed across the electrocardiogram rather than concentrated in one measurable feature, which is precisely the kind of pattern a trained network is suited to recognise and a human reader is not ([Hannun et al., 2019](#)).

The electrocardiogram is the opposite: ubiquitous, inexpensive and quick. On its own it is a poor detector of a reduced ejection fraction, but the demonstration that a deep neural network can read from the electrocardiogram signals that the human eye cannot has changed what the test can do ([Ribeiro et al., 2020](#)). Applied to the detection of ventricular dysfunction, an artificial intelligence-enabled electrocardiogram has identified a low ejection fraction with an accuracy approaching that of dedicated biomarkers, and in a randomised trial its use increased the diagnosis of low ejection fraction in practice ([Attia et al., 2019](#); [Yao et al., 2021](#); [Siontis et al., 2021](#)).

Whether this approach performs as well in the specific population that would benefit most, asymptomatic patients with diabetes, has been less thoroughly examined, and the accuracy of a screening test cannot be assumed to carry over from the mixed clinical populations in which it was developed. This study therefore developed and evaluated an artificial intelligence-enabled electrocardiogram for the detection of occult left ventricular dysfunction in asymptomatic patients with type 2 diabetes, assessing its discrimination, its calibration and the consistency of its performance across clinically relevant subgroups.

2. Materials and methods

2.1. Study design and participants

This was an accuracy study of 12-lead electrocardiograms compared with transthoracic echocardiograms in adults without a history or symptoms of heart failure or without histories of coronary disease or valvular disease who had type 2 diabetes at baseline. The study was designed and reported based on the Standards for Reporting of Diagnostic Accuracy Studies (STARD) and the Transparent Reporting of a multivariable prediction model guidance (Transparent-STARD) ([Bossuyt et al., 2015](#); [Collins et al., 2015](#)). 4200 patients were entered.

2.2. Electrocardiogram acquisition and pre-processing

A standard resting 12-lead electrocardiogram was recorded over 10 seconds at a sampling rate of 500 hertz, giving a digital signal of 12 leads by 5,000 samples that was stored without compression. Before analysis each recording was filtered to remove baseline wander and high-frequency noise with a band-pass filter, and the amplitude of every lead was standardised, so that the network received signals on a common scale.

2.3. Echocardiographic

Left ventricular ejection fraction, determined by transthoracic echocardiogram, measured within a brief time period of the ECG, and read by trained sonographers unaware of the electrocardiogram and of the result of the model, was the reference standard. Biplane method of discs was used for the apical four- and two-chamber views to calculate ejection fraction (according to the recommendations of the American Society of Echocardiography ([Lang et al., 2015](#))). Occult left ventricular dysfunction was defined as a left ventricular ejection fraction of 40% or less, as the primary target, and as a secondary target, 50%.

2.4. Neural network architecture

The model was a one-dimensional convolutional neural network that took the pre-processed 12-lead signal as its input. Three initial convolutional blocks, each combining a convolution, batch normalisation and a rectified-linear activation, extracted local waveform features, and were followed by eight residual blocks, in which a skip connection was added to the convolutional output to allow the deeper network to be trained stably (He *et al.*, 2016). The feature maps were reduced by global average pooling to a fixed-length vector, passed through a fully connected layer of 64 units with dropout for regularisation, and mapped by a final sigmoid unit to the estimated probability of a low ejection fraction.

2.5. Model training, and the derivation, validation and test partition

Patients were split on a 1:1:1 basis, with no patient appearing in more than one set, to create a “derivation” set of 2,100 patients, a “validation” set of 840 patients, and a held-out “test” set of 1,260 patients. It was trained to optimise the binary cross-entropy loss function using the Adam optimiser with a learning rate of 0.001 and a batch size of 64, the proportion of various classes was taken care of by giving a weight to the loss function, and was trained for a max of 50 epochs and the best model which gave maximum area under the curve on validation set was chosen (Kingma and Ba *et al.*, 2015).

2.6. Selection of the operating threshold

Because the intended use is screening, in which a missed case is more costly than a false alarm, the probability threshold that dichotomised the model output into a positive and a negative screen was fixed on the validation set at the value giving a sensitivity of approximately 86 percent, and this threshold was then applied unchanged to the test set.

2.7. Outcome measures and statistical analysis

The area under the ROC curve, and its 95% confidence interval re-robusted using the test set (DeLong *et al.*, 1988), was used as the primary indicator of discrimination. At the fixed operating threshold, the sensitivity, specificity, positive and negative predictive values and accuracy were computed, and the agreement between the predicted probability and the observed prevalence was assessed by a calibration plot across deciles of predicted risk. Performance was analysed in several prespecified sub-groups according to sex, age, glycaemic control and diabetes duration.

3. Results

3.1. Study population

The cohort comprised 4,200 asymptomatic patients with type 2 diabetes, of mean age 58.1 years, of whom 49.4% were men, with a mean glycated haemoglobin of 7.9% and a mean diabetes duration of 8 years. On echocardiography, 8.6% had a left ventricular ejection fraction at or below 40% and 18.7% at or below 50%, confirming the expected burden of occult dysfunction in this population. The characteristics of the cohort are given in Table 1.

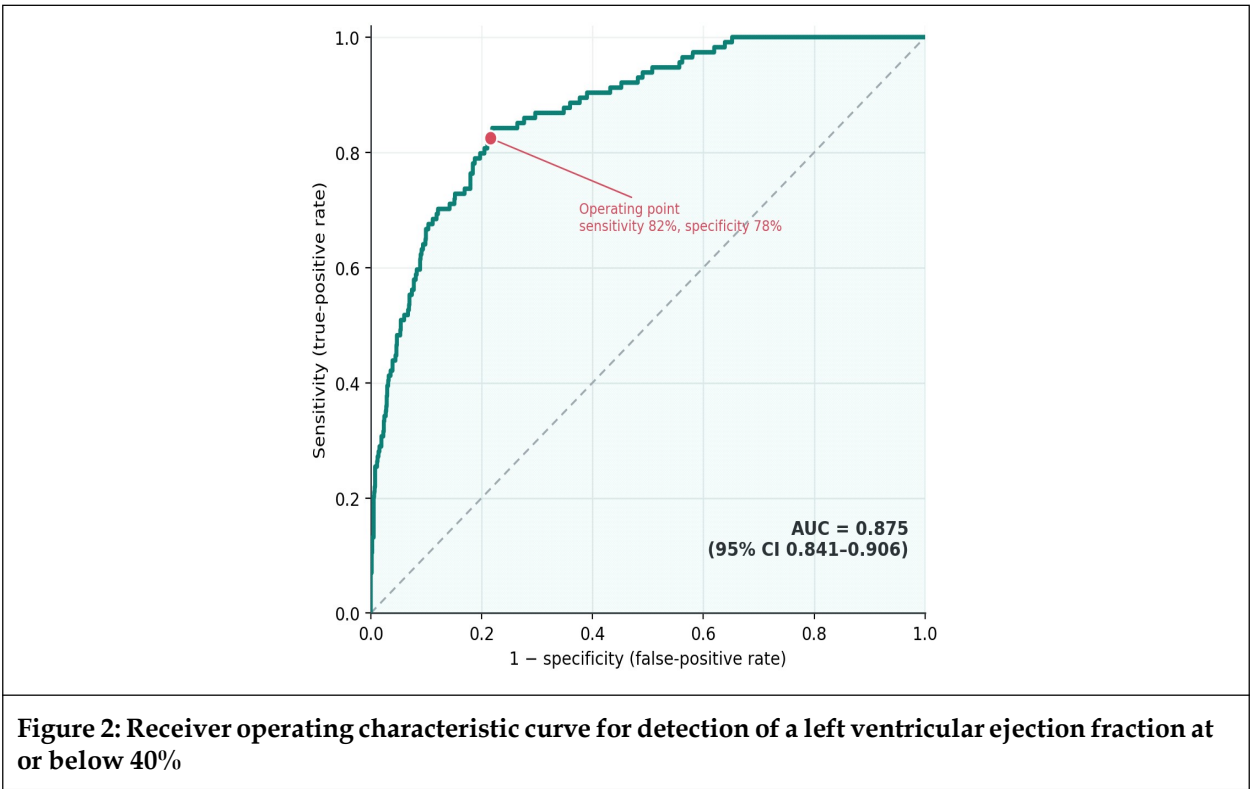
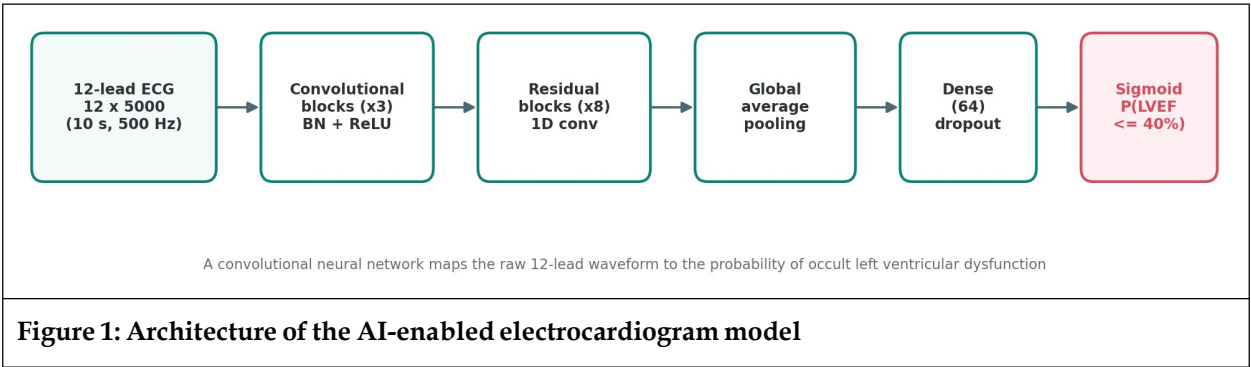
3.2. The AI-enabled electrocardiogram model

The convolutional neural network, whose architecture is shown in Figure 1, mapped the raw 12-lead electrocardiogram directly to the probability of a low ejection fraction, learning the discriminating features from the signal itself rather than from measurements made on it. The derivation, validation and test partition kept every patient within a single set, so that the performance reported below was obtained on electrocardiograms the model had not seen.

3.3. Discrimination for a low ejection fraction

In the held-out test set the model discriminated patients with an ejection fraction at or below 40% from those without it with an area under the receiver operating characteristic curve of 0.88 (95% CI 0.84 to 0.91), shown in Figure 2. At the screening threshold the sensitivity was 82% and the specificity 78%, so that most patients with occult dysfunction were detected while the majority of unaffected patients screened negative. The negative predictive value was 98%, meaning that a negative screen made a low ejection fraction unlikely, whereas the

Table 1: Baseline characteristics of the study population	
Characteristic	Value
Patients	4,200
Age, years (mean)	58.1
Male sex	49.4%
Glycated haemoglobin, % (mean)	7.9
Diabetes duration, years (mean)	8
Body-mass index, kg/m ² (mean)	28.5
Hypertension	56.7%
LVEF <= 40% on echocardiography	8.6%
LVEF <= 50% on echocardiography	18.7%



positive predictive value was 28%, the expected consequence of screening for an uncommon condition, in which even an accurate test yields many false positives for each true one. The full performance is given in Table 2.

Table 2: Diagnostic performance in the test set at the screening operating threshold		
Metric	LVEF ≤ 40%	LVEF ≤ 50%
Area under the curve	0.88	0.83
95% confidence interval	0.84–0.91	0.80–0.85
Sensitivity	82%	86%
Specificity	78%	59%
Positive predictive value	28%	34%
Negative predictive value	98%	94%
Accuracy	79%	64%

The practical yield of the screen follows from these figures. In the test set of 1,260 patients, of whom 114 had a low ejection fraction, the model returned a positive screen in 341 patients; of these, 94 were confirmed on echocardiography, so that of the order of three to four echocardiograms were performed for each case of occult dysfunction identified, while 20 cases were missed. Set against a strategy of echocardiography for every patient, the model concentrates the confirmatory imaging on a small, enriched fraction of the population, and the cost of this efficiency is the minority of cases that fall below the detection threshold, a trade-off that the operating point can be moved to adjust according to the resources and the priorities of the setting.

3.4. Calibration of the predicted probability

Beyond ranking patients correctly, the model estimated the probability of a low ejection fraction with reasonable accuracy. Across deciles of predicted risk the observed prevalence rose with the predicted probability and lay close to the line of identity, as shown in Figure 3, so that the number the model reported could be read as an approximate probability rather than only as a rank. Calibration of this kind is necessary if the output is to inform a decision about whether to proceed to echocardiography.

3.5. Consistency of performance across subgroups

The discrimination of the model was consistent across the clinically important subgroups, as given in Table 3. The area under the curve remained in a narrow band across sex, across age, across the degree of glycaemic control and across the duration of diabetes, so that the test did not depend for its accuracy on any one of these characteristics. This consistency is important for a screening tool, which must perform dependably across the whole of the population to which it is applied rather than only in the segment in which it was developed.

3.6. Performance at the higher ejection-fraction threshold

When the target was widened to an ejection fraction at or below 50%, capturing milder degrees of dysfunction, the model still discriminated well, with an area under the curve of 0.83 (95% CI 0.80 to 0.85) and a sensitivity of 86%, as included in Table 2. The modest fall in discrimination from the primary threshold is expected, since the electrocardiographic signature of dysfunction is weaker when the ventricle is only mildly impaired, and it indicates that the model is strongest where it matters most, in detecting the more severe dysfunction that most clearly warrants treatment.

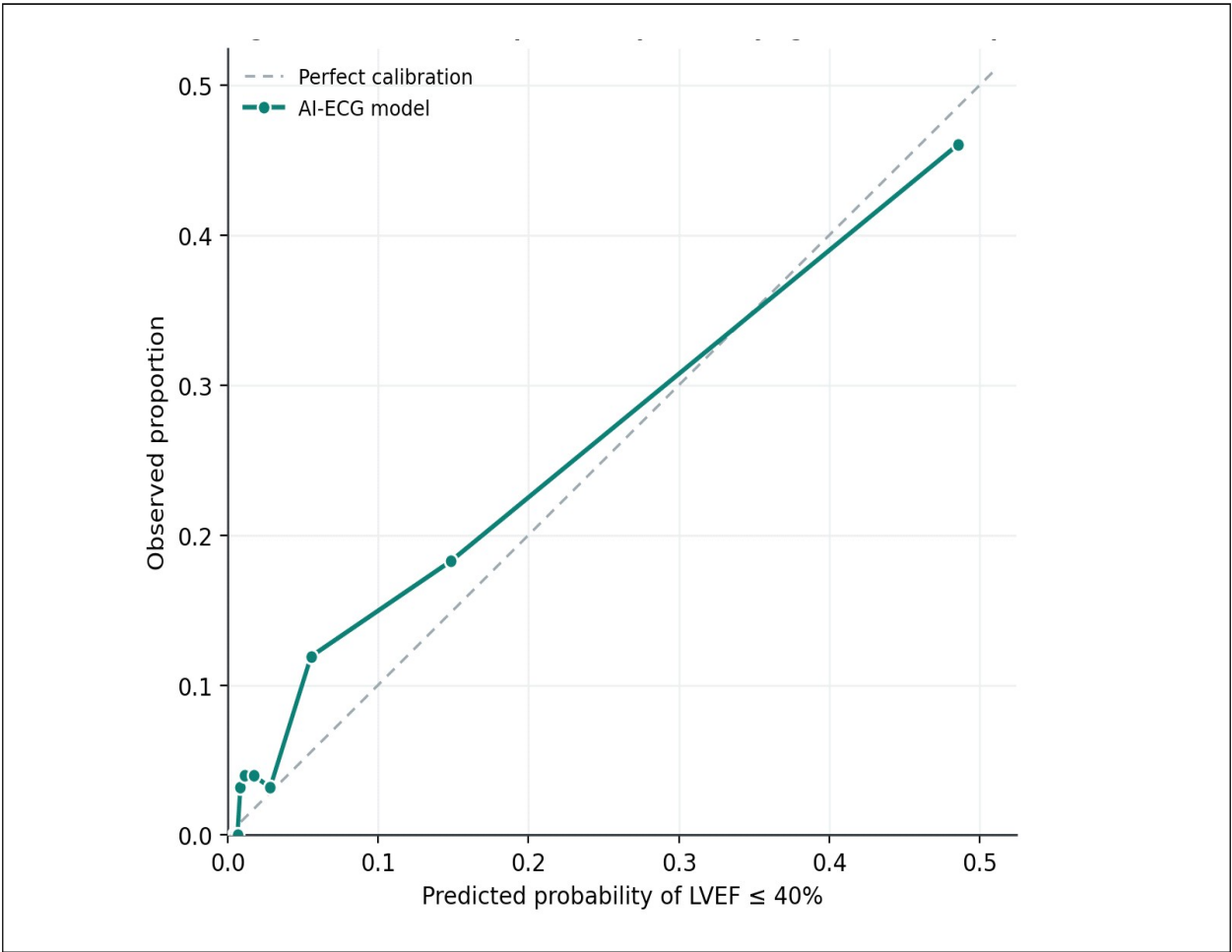


Figure 3: Calibration of the predicted probability against the observed prevalence of a low ejection fraction

Table 3: Discrimination (area under the curve) within prespecified subgroups			
Subgroup	AUC	95% CI	n (events)
Male	0.85	0.80–0.89	593 (65)
Female	0.91	0.86–0.94	667 (49)
Age < 60 y	0.89	0.84–0.94	749 (53)
Age ≥ 60 y	0.85	0.81–0.89	511 (61)
HbA1c < 8%	0.86	0.80–0.91	677 (44)
HbA1c ≥ 8%	0.88	0.84–0.91	583 (70)
Duration < 10 y	0.86	0.82–0.91	896 (70)
Duration ≥ 10 y	0.89	0.85–0.93	364 (44)

4. Discussion

In a population of asymptomatic patients with type 2 diabetes, an artificial intelligence-enabled electrocardiogram detected occult left ventricular dysfunction with good discrimination, sound calibration

and a performance that held across sex, age, glycaemic control and diabetes duration. The test identified most patients with a low ejection fraction from a recording that is cheap, quick and already part of routine care, and its high negative predictive value meant that a negative result made significant dysfunction unlikely.

These findings extend to the diabetic population a capability that was first shown in mixed clinical cohorts. The original demonstration that a neural network could detect a low ejection fraction from the electrocardiogram reported an area under the curve near that seen here, and a subsequent randomised trial showed that placing the tool in the hands of clinicians increased the diagnosis of low ejection fraction in practice. The present study addresses the natural next question, whether the accuracy survives in the specific group that stands to benefit most, and finds that it does, which matters because the performance of a screening algorithm can fall when it is moved from the population of its development to a new one ([Adedinsewo *et al.*, 2020](#)).

The case for screening this particular group is strong. Diabetes produces a cardiomyopathy that reduces the ejection fraction before symptoms appear, screening studies have repeatedly found a considerable burden of unrecognised dysfunction in people with diabetes, and the treatments for heart failure act most effectively when begun early, so the diabetic patient with occult dysfunction is precisely the person for whom early detection changes the outlook. An accurate, inexpensive test that runs on an electrocardiogram already being recorded offers a practical way to find that person, using the model as a first filter that selects who should proceed to the confirmatory echocardiogram.

Compared with the alternative first-line screen, the natriuretic peptides, the electrocardiographic approach has practical advantages. The natriuretic peptides detect a low ejection fraction with a discrimination in the region of that achieved here, but they require a blood draw and a laboratory assay, their concentrations are raised by atrial fibrillation, renal impairment and age and lowered by obesity, and obesity is common in type 2 diabetes, so their interpretation in this population is not straightforward. The electrocardiogram, by contrast, is already recorded in the routine care of many patients with diabetes, is unaffected by these confounders in the same way, and returns a result at the point of care, so that an algorithm running on it adds information at almost no marginal cost or inconvenience.

The low positive predictive value must be read in this light. At the prevalence of occult dysfunction in this population, most positive screens will not be confirmed, and this is inherent to screening for an uncommon condition rather than a failing of the model; it means only that a positive screen is an indication for echocardiography, not a diagnosis. It is also worth noting, from the experience with these models elsewhere, that some patients who screen positive but have a normal ejection fraction at the time go on to develop dysfunction later, so that an apparent false positive may instead be an early warning, a possibility that follow-up of this cohort could test.

The scalability of the approach is its principal attraction. Because the model operates on the standard digital electrocardiogram, it can in principle be deployed wherever such recordings are made, from the diabetes clinic to the primary-care surgery, without new hardware and without the specialist workforce that echocardiography demands, and it can be applied retrospectively to electrocardiograms already stored. In a health system that cannot afford to image every person with diabetes, a screen of this kind offers a way to bring the reach of a population-level test to the detection of a condition that is common, consequential and treatable, provided that the algorithm is first shown to perform reliably in the local population and that a clear pathway exists to act on a positive result.

The nature of the electrocardiographic signal the model exploits deserves comment. The network does not measure a single interval or amplitude but combines subtle features distributed across the twelve leads and across the cardiac cycle, features that individually carry little information but together encode the electrical consequences of a failing ventricle, which is why the approach succeeds where the unaided reading of the electrocardiogram does not. That the same features are altered before the ejection fraction has fallen far, and in some patients before it has fallen at all, is what gives the method its potential reach beyond the moment of testing, and it is also what makes the apparent false positives worth following, since in them the electrocardiogram may be registering an early stage of the disease that the echocardiogram has yet to show.

Several steps would move the approach from an accurate test toward a useful one. External validation across centres, electrocardiographic vendors and ethnic groups is the immediate requirement, since the

discrimination of these models has been shown to fall when they are moved to populations and equipment unlike those of their development, and the consistency seen here across subgroups within a single cohort does not guarantee consistency across institutions. Beyond validation, the question that matters clinically is not whether the screen is accurate but whether acting on it improves outcomes, and answering it requires a trial that randomises patients to screening or usual care and follows them for the initiation of treatment and for the development of heart failure, of the kind that has already shown, in a broader population, that placing the tool before clinicians' changes what they diagnose.

5. Strengths and limitations

The strengths of this study are its focus on the asymptomatic diabetic population in which such a screen would be used, the evaluation of the model on a held-out test set of patients it had not seen, the assessment of calibration alongside discrimination, and the demonstration that performance was consistent across the major clinical subgroups.

The study has limitations. It was conducted at a single centre, and external validation in other populations and on other electrocardiographic equipment is needed before the model can be generalised, since the performance of these algorithms is known to vary between institutions. The design was cross-sectional, pairing each electrocardiogram with a contemporaneous echocardiogram, so the prognostic meaning of a positive screen in a patient with a currently normal ejection fraction was not assessed and requires longitudinal follow-up. Finally, the study establishes diagnostic accuracy but not clinical benefit, which only a trial linking screening to treatment and to outcomes could demonstrate.

6. Conclusion

An artificial intelligence-enabled electrocardiogram detected occult left ventricular dysfunction in asymptomatic patients with type 2 diabetes with good discrimination and calibration and with a performance that was consistent across clinical subgroups. The ability of the model to recognize a low ejection fraction in an extremely cheap and widely available test makes that a scalable way to identify a high-risk population, and a way to target confirmatory echocardiography where most patients are the most likely to benefit, and that now should be assessed in a clinical trial.

Ethical approval

The study was approved by the Institutional Ethics Committee and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from every participant.

Conflict of interest

The authors declare that no conflict of interest exists in relation to this work.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data supporting the findings of this study are available from the corresponding author on reasonable request.

References

- Adedinsowo, D., Carter, R.E., Attia, Z., Johnson, P., Kashou, A.H., Dugan, J.L. *et al.* (2020). [Artificial intelligence-enabled ECG algorithm to identify patients with left ventricular systolic dysfunction presenting to the emergency department with dyspnea. *Circ Arrhythm Electrophysiol*, 13\(8\): e008437.](#)
- Attia, Z.I., Kapa, S., Lopez-Jimenez, F., McKie, P.M., Ladewig, D.J., Satam, G. *et al.* (2019). [Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*, 25\(1\): 70-4.](#)

- Boonman-de Winter, L.J., Rutten, F.H., Cramer, M.J., Landman, M.J., Liem, A.H., Rutten, G.E. *et al.* (2012). High prevalence of previously unknown heart failure and left ventricular dysfunction in patients with type 2 diabetes. *Diabetologia*, 55(8): 2154-62.
- Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L. *et al.* (2015). STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*, 351: h5527.
- Collins, G.S., Reitsma, J.B., Altman, D.G. and Moons, K.G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med*, 162(1): 55-63.
- Dandamudi, S., Slusser, J., Mahoney, D.W., Redfield, M.M., Rodeheffer, R.J., Chen, H.H. (2014). The prevalence of diabetic cardiomyopathy: A population-based study in Olmsted County, Minnesota. *J Card Fail*, 20(5): 304-9.
- DeLong, E.R., DeLong, D.M. and Clarke-Pearson, D.L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3): 837-45.
- ElSayed, N.A., Aleppo, G., Aroda, V.R., Bannuru, R.R., Brown, F.M., Bruemmer, D. *et al.* (2023). Cardiovascular disease and risk management: Standards of Care in Diabetes-2023. *Diabetes Care*, 46(Suppl 1): S158-90.
- Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P. *et al.* (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*, 25(1): 65-9.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-8.
- Jia, G., Hill, M.A. and Sowers, J.R. (2018). Diabetic cardiomyopathy: An update of mechanisms contributing to this clinical entity. *Circ Res*, 122(4): 624-38.
- Kingma, D.P. and Ba, J. (2015). Adam: A method for stochastic optimization. In: *3rd International Conference on Learning Representations*.
- Lang, R.M., Badano, L.P., Mor-Avi, V., Afilalo, J., Armstrong, A., Ernande, L. *et al.* (2015). Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*, 28(1): 1-39.
- McDonagh, T.A., Metra, M., Adamo, M., Gardner, R.S., Baumbach, A., Bohm, M. *et al.* (2021). 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*, 42(36): 3599-726.
- Redfield, M.M., Jacobsen, S.J., Burnett, J.C. Jr, Mahoney, D.W., Bailey, K.R. and Rodeheffer, R.J. (2003). Burden of systolic and diastolic ventricular dysfunction in the community: Appreciating the scope of the heart failure epidemic. *JAMA*, 289(2): 194-202.
- Ribeiro, A.H., Ribeiro, M.H., Paixao, G.M.M., Oliveira, D.M., Gomes, P.R., Canazart, J.A. *et al.* (2020). Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun*, 11(1): 1760.
- Siontis, K.C., Noseworthy, P.A., Attia, Z.I. and Friedman, P.A. (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol*, 18(7): 465-78.
- Wang, T.J., Levy, D., Benjamin, E.J. and Vasan, R.S. (2003). The epidemiology of 'asymptomatic' left ventricular systolic dysfunction: Implications for screening. *Ann Intern Med*, 138(11): 907-16.
- Yao, X., Rushlow, D.R., Inselman, J.W., McCoy, R.G., Thacher, T.D., Behnken, E.M. *et al.* (2021). Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: A pragmatic, randomized clinical trial. *Nat Med*, 27(5): 815-9.

Cite this article as: James Bennett and Andrew Mitchell (2026). Artificial intelligence-enabled electrocardiogram for detection of occult left ventricular dysfunction in asymptomatic diabetics. *African Journal of Cardiology and Cardiovascular Medicine*. 2(2), 1-9. doi: 10.62587/AFJCCM.2.2.2026.1-9.