Article

Artificial intelligence guided screening for cardiomyopathies in an obstetric population: a pragmatic randomized clinical trial

Check for updates

A list of authors and their affiliations appears at the end of the paper

Nigeria has the highest reported incidence of peripartum cardiomyopathy worldwide. This open-label, pragmatic clinical trial randomized pregnant and postpartum women to usual care or artificial intelligence (AI)-guided screening to assess its impact on the diagnosis left ventricular systolic dysfunction (LVSD) in the perinatal period. The study intervention included digital stethoscope recordings with point of-care AI predictions and a 12-lead electrocardiogram with asynchronous AI predictions for LVSD. The primary end point was identification of LVSD during the study period. In the intervention arm, the primary end point was defined as the number of identified participants with LVSD as determined by a positive AI screen, confirmed by echocardiography. In the control arm, this was the number of participants with clinical recognition and documentation of LVSD on echocardiography in keeping with current standard of care. Participants in the intervention arm had a confirmatory echocardiogram at baseline for AI model validation. A total of 1,232 (616 in each arm) participants were randomized and 1,195 participants (587 intervention arm and 608 control arm) completed the baseline visit at 6 hospitals in Nigeria between August 2022 and September 2023 with follow-up through May 2024. Using the AI-enabled digital stethoscope, the primary study end point was met with detection of 24 out of 587 (4.1%) versus 12 out of 608 (2.0%) patients with LVSD (intervention versus control odds ratio 2.12, 95% CI 1.05–4.27; $P = 0.032$). With the 12-lead AI-electrocardiogram model, the primary end point was detected in 20 out of 587 (3.4%) versus 12 out of 608 (2.0%) patients (odds ratio 1.75, 95% CI 0.85–3.62; $P = 0.125$). A similar direction of effect was observed in prespecified subgroup analysis. There were no serious adverse events related to study participation. In pregnant and postpartum women, AI-guided screening using a digital stethoscope improved the diagnosis of pregnancy-related cardiomyopathy. ClinicalTrials.gov registration: NCT05438576

✉ e-mail: adedinsewo.demilade@mayo.edu

In the United States, cardiomyopathy is a leading cause of maternal mortality and the number one cause of death in the postpartum period. Its incidence is estimated to be 1 in 2,000 (refs. 1,2) and as high as 1 in ~700 among African American women with milder forms likely going undetected in the absence of screening[2]. In Nigeria, the incidence is reported to be 1 in 96 deliveries[3,4], the highest reported worldwide[5]. Cardiomyopathy occurring during pregnancy and postpartum is challenging to diagnose due to similarities between heart failure symptoms and those related to physiologic changes of pregnancy[6,7]. This leads to a delay in diagnosis and consequently, adverse maternal outcomes[6]. While cardiomyopathy can occur de novo during pregnancy, hemodynamic changes of pregnancy can also unmask previously undiagnosed or asymptomatic LVSD[8].

AI-enabled electrocardiograms (ECGs) have shown effectiveness in identifying multiple cardiovascular pathologies[9–12], including detection of low left ventricular ejection fraction (LVEF)[13,14]. A retrospective study (area under the curve (AUC) = 0.89)[15] and a pilot prospective study among pregnant and postpartum women in the United States showed AI-based screening to be effective (AUC = 1.00 using a 12-lead ECG and 0.98 using a digital stethoscope)[16] in identifying pregnancy-related LVSD with LVEF < 45%. Other retrospective studies in the United States and the Republic of Korea have also shown good performance of separate AI-ECG models for detecting perinatal LVSD[17,18]. Based on these previous studies, we surmised that this new technology has the potential to enhance screening and identification of LVSD in the peripartum period; however, it remains unknown whether AI-guided screening improves cardiomyopathy detection in obstetric patients beyond the current standard of care. To address this question, we conducted an open-label, randomized, pragmatic clinical trial among pregnant and postpartum women to evaluate whether AI-guided screening (using a digital stethoscope and 12-lead ECG) improves the diagnosis of pregnancy-related LVSD in an obstetric population in Nigeria compared to usual care.

Results

Sample characteristics
A total of 1,232 participants were randomized and enrolled between 15 August 2022 and 28 September 2023 across six sites in Nigeria. Of these, 1,195 (587 in the intervention arm and 608 in the control arm) completed baseline assessments following randomization and were included in a modified intention-to-treat (mITT) analysis as prespecified in the study protocol with follow-up through 15 May 2024 (Fig. 1). For participants in the control arm, baseline assessments included performance of a 12-lead ECG (with AI-based predictions for age and sex), and for those in the intervention arm, baseline assessments involved performance of a 12-lead ECG (with AI-based predictions for age, sex and LVSD), digital stethoscope recordings (with AI-based predictions for LVSD) and an echocardiogram at baseline (for validation of AI-based LVSD prediction results) (Fig. 2). To avoid inflating the detection rate for LVSD based on the protocolized echocardiogram at baseline in the intervention arm, a positive AI screen was required to be present to count the echocardiogram findings for the purpose of the primary end point. Given in-built data quality checks embedded as part of the US Food and Drug Administration (FDA)-cleared 12-lead ECG and digital stethoscope AI algorithms, recorded data deemed to be of poor quality in relation to the AI prediction results were conservatively assumed to be negative predictions (n = 7 out of 587 and n = 7 out of 587 respectively).
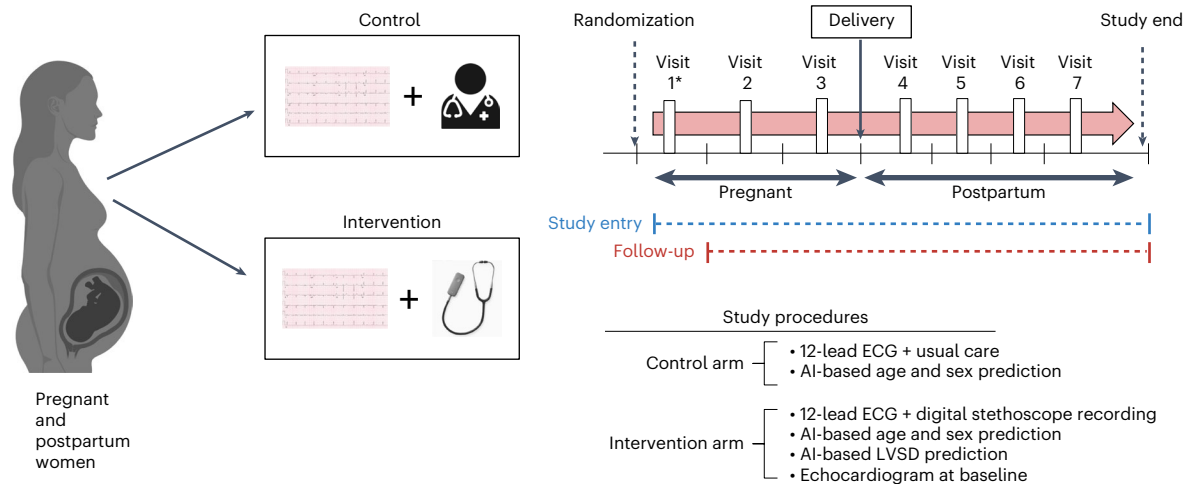
Study participant characteristics are displayed in Table 1. The median age was 31 years and all women identified as Black. Fifty-five percent were of Yoruba ethnicity, 28% Hausa, 11% Igbo and 6% were from other ethnic groups. At baseline, 150 (12.5%) women had been diagnosed with a hypertensive disorder of pregnancy (which includes chronic hypertension, gestational hypertension, pre-eclampsia and eclampsia) during the index pregnancy.

Primary outcome
Digital stethoscope. At study entry in the mITT analysis set, 22 cases (3.7%) of LVSD (LVEF < 50%) were identified in the intervention arm (positive point-of-care AI prediction for LVSD, maximum prediction across all recording locations confirmed with echocardiography) compared to 11 (1.8%) in the control arm; odds ratio 2.11, 95% CI 1.02–4.40; P = 0.041. At the study end, three additional LVSD cases were identified with two in the intervention arm and one in the control arm, resulting in a total of 24 cases (4.1%) of LVSD (LVEF < 50%) identified in the intervention arm compared to 12 (2.0%) in the control arm; odds ratio 2.12,
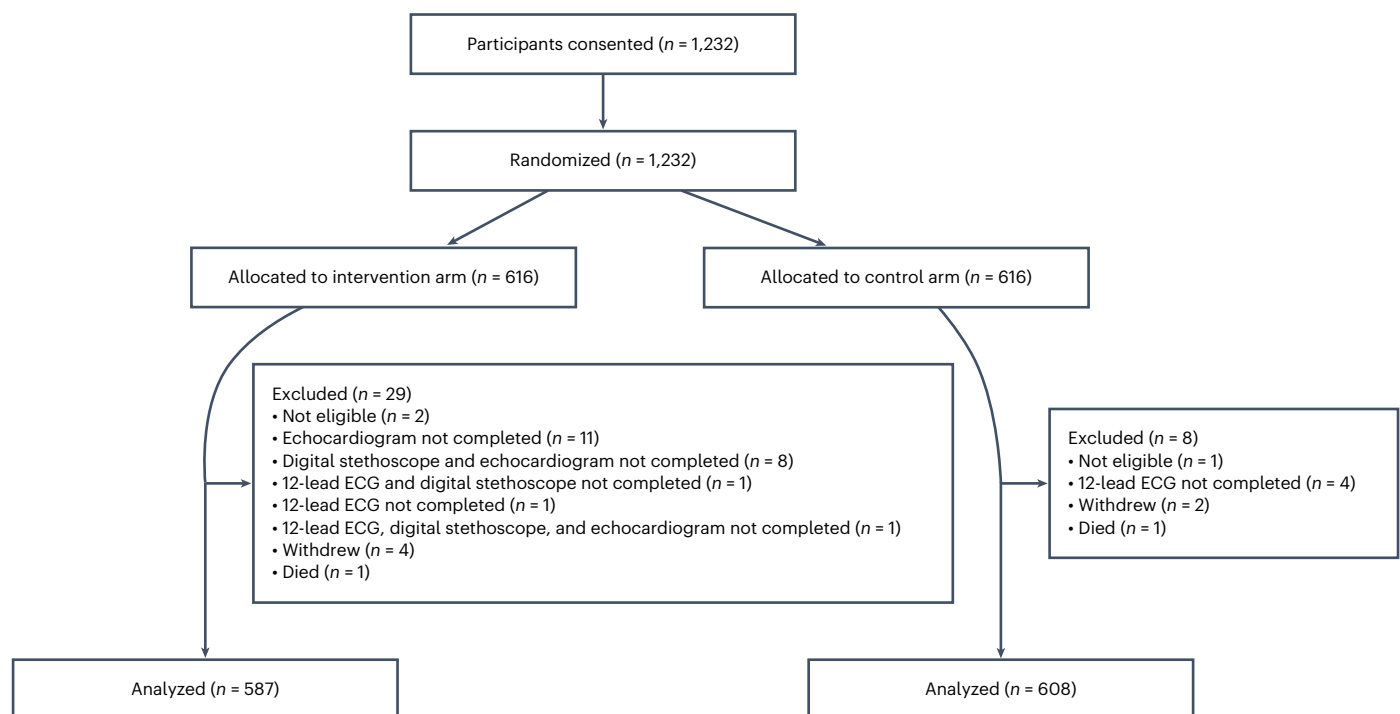
Fig. 1 | Consort diagram. This diagram shows study participant flow in the screening for peripartum cardiomyopathies in Nigeria study (SPEC-AI Nigeria).

Flowchart:
- Participants consented (n = 1,232)
- Randomized (n = 1,232)
  - Allocated to intervention arm (n = 616)
    - Excluded (n = 29)
      - Not eligible (n = 2)
      - Echocardiogram not completed (n = 11)
      - Digital stethoscope and echocardiogram not completed (n = 8)
      - 12-lead ECG and digital stethoscope not completed (n = 1)
      - 12-lead ECG not completed (n = 1)
      - 12-lead ECG, digital stethoscope, and echocardiogram not completed (n = 1)
      - Withdrew (n = 4)
      - Died (n = 1)
    - Analyzed (n = 587)
  - Allocated to control arm (n = 616)
    - Excluded (n = 8)
      - Not eligible (n = 1)
      - 12-lead ECG not completed (n = 4)
      - Withdrew (n = 2)
      - Died (n = 1)
    - Analyzed (n = 608)

Fig. 2 | Summary of study design. This figure summarizes the study design, interventions, study-related visits and key study procedures. Participants could enter the study at any time point during pregnancy or postpartum (up to 12 months). As such, each individual participant could have up to seven visits if they enter the study in the first trimester of pregnancy and fewer depending on the time of study entry. AI-based screening was performed up to seven times during the study period, including during each trimester of pregnancy (first trimester, <14 weeks; second trimester, 14 to <28 weeks; and third trimester, 28 to <42 weeks and post-term) (up to three ECGs), between delivery and 6 weeks, between 6 weeks and 3 months, between 3 and 5 months, and between 5 and 12 months postpartum (up to four ECGs). Only participants in the intervention arm had a baseline echocardiogram as well as a simultaneous portable ECG recorded at each time point. AI-based prediction for LVSD using the digital stethoscope was available in real time at the point of care and 12-lead AI-ECG predictions for LVSD were provided asynchronously, usually within 1 week of ECG acquisition. *, visit 1 can vary for each participant depending on the time point at study entry in relation to delivery.

95% CI 1.05–4.27; $P = 0.032$ (Table 2). The estimated number needed to screen (NNS) to detect one additional case of LVSD was 47.

The 36 cases of LVSD were primarily from the northern region of Nigeria (86%, 31 out of 36). As a sensitivity analysis, the primary end point was analyzed with adjustment for site. This analysis yielded consistent results; odds ratio 2.25, 95% CI 1.09–4.66; $P = 0.029$. To address the potential limitation of the mITT analysis, a full intention-to-treat (ITT) analysis for the primary end point using the digital stethoscope was conducted to include all 1,232 individuals randomized. To be conservative, all initially excluded patients were included and assumed to have normal LVEF. The point estimate for the intervention effect remained stable and statistically significant for the primary outcome measured over the 18-month study period showing an unadjusted odds ratio of 2.04 (95% CI 1.01–4.12, $P = 0.042$) and site-adjusted odds ratio of 2.13 (95% CI 1.03–4.41, $P = 0.041$).

12-lead ECG. *US FDA-cleared AI-ECG algorithm.* At study entry, 18 cases (3.1%) of LVSD (LVEF < 50%) were identified in the intervention arm (positive 12-lead ECG AI prediction for LVSD, confirmed with echocardiography) compared to 11 (1.8%) in the control arm; odds ratio 1.72, 95% CI 0.80–3.67; $P = 0.158$. At the study end, three additional LVSD cases were identified with two in the intervention arm and one in the control arm, resulting in a total of 20 cases (3.4%) of LVSD (LVEF < 50%) identified in the intervention arm compared to 12 (2.0%) in the control arm; odds ratio 1.75, 95% CI 0.85–3.62; $P = 0.125$ (Table 2). Although a numerically higher number of LVSD cases were identified with the 12-lead AI-ECG screening, this did not reach statistical significance.

Original Mayo Clinic AI-ECG algorithm. At study entry, 14 cases (2.4%) of LVSD (LVEF < 50%) were identified in the intervention arm (positive 12-lead AI-ECG prediction for LVSD, confirmed with echocardiography) compared to 11 (1.8%) in the control arm; odds ratio 1.33, 95% CI 0.60–2.94; $P = 0.487$. At the study end, five additional LVSD cases were identified with four in the intervention arm and one in the control arm resulting in a total of 18 cases (3.1%) of LVSD (LVEF < 50%) identified in the intervention arm compared to 12 (2.0%) in the control arm; odds ratio 1.57, 95% CI 0.75–3.29; $P = 0.227$ (Table 2). Similar to the

US FDA-cleared model, a numerically higher number of LVSD cases were identified with this AI model but this did not reach statistical significance.

Secondary outcomes

Digital stethoscope performance by subgroups. Evaluation of the primary outcome within prespecified subgroups were consistent with a tendency toward improved detection of LVSD in the intervention arm compared to the control arm across all prespecified subgroups (age group, ethnicity, region, presence of hypertensive disorders and pregnancy/postpartum status). A summary figure is provided in Fig. 3a.

12-Lead ECG performance by subgroups. Subgroup analysis using the US FDA-cleared 12-lead AI-ECG algorithm also showed consistent results across the prespecified subgroups except for age group, where the odds ratio for identifying LVSD among those younger than 30 years was 0.9 (95% CI 0.4–2.5) and 4.2 (95% CI 1.2–15.0) for those aged 30 years and older (Fig. 3b). This may suggest improved performance of the model in older compared to younger patients. We also provide a summary figure for the subgroup analysis using the original Mayo Clinic 12-lead AI-ECG algorithm (Fig. 3c) and the direction of effect was in keeping with the US FDA-cleared 12-lead AI-ECG algorithm.

Effectiveness of the AI-stethoscope for LVSD detection. Among participants in the intervention arm, the digital stethoscope (maximum prediction across all locations recorded) had an AUC of 0.976 (95% CI 0.953–0.998) for detection of LVEF < 50%, whereas for detection of LVEF < 40%, AUC was 0.985 (95% CI 0.974–0.996) using the diagnostic quality recordings (Supplementary Fig. 1). Table 3 provides a comprehensive examination of the diagnostic performance of the digital stethoscope AI model across the various recording locations and a range of ejection fraction thresholds at study entry for only the diagnostic quality recordings. Supplementary Table 1 provides a sensitivity analysis for the digital stethoscope's diagnostic performance results, including the recordings identified by the software as being of poor quality. Of note, only one case of LVSD identified on echocardiography ($n = 23$) was incorrectly classified, corresponding to a sensitivity of

Table 1 | Demographic and clinical characteristics of the study sample at baseline

Characteristic	n	Median (Q1– Q3) or n (%) of patients	
		Intervention arm (*n*=587)	Control arm (*n*=608)
Age, years	1,195	31 (27–35)	31 (26–35)
Race (Black)	1,195	587 (100.0%)	608 (100.0%)
Ethnicity	1,195		
Hausa		163 (27.8%)	174 (28.6%)
Igbo		61 (10.4%)	68 (11.2%)
Other		35 (6.0%)	41 (6.7%)
Yoruba		328 (55.9%)	325 (53.5%)
Status at entry	1,195		
Pregnant		423 (72.1%)	451 (74.2%)
Postpartum		164 (27.9%)	157 (25.8%)
Time point of pregnancy/ postpartum	1,195		
First trimester		37 (6.3%)	50 (8.2%)
Second trimester		150 (25.6%)	173 (28.5%)
Third trimester		238 (40.5%)	229 (37.7%)
Time of delivery or up to 6 weeks after delivery		116 (19.8%)	132 (21.7%)
Greater than 6 weeks and up to 3 months after delivery		23 (3.9%)	9 (1.5%)
Greater than 3 months and up to 5 months after delivery		7 (1.2%)	7 (1.2%)
Greater than 5 months and up to 12 months after delivery		16 (2.7%)	8 (1.3%)
Weight, kg	1,195	70 (60–81)	70 (60–83)
Height, cm	1,195	161 (157–165)	161 (157–165)
Systolic blood pressure, mm Hg	1,194	110 (100–120)	110 (100–120)
Diastolic blood pressure, mm Hg	1,195	70 (60–80)	70 (60–80)
Resting heart rate, bpm	1,195	87 (80–95)	88 (80–95)
Hemoglobin (g dl⁻¹)	1,046	11 (10–12)	11 (10–12)
Hematocrit (%)	1,088	33 (30–35)	33 (30–35)
Blood type	1,134		
A		98 (17.7%)	90 (15.5%)
B		106 (19.2%)	120 (20.7%)
AB		29 (5.2%)	28 (4.8%)
O		320 (57.9%)	343 (59.0%)
Hemoglobin genotype	1,136		
AA		402 (72.7%)	436 (74.8%)
AS		131 (23.7%)	124 (21.3%)
SS		5 (0.9%)	9 (1.5%)
Sc		1 (0.2%)	4 (0.7%)
Other		14 (2.5%)	10 (1.7%)
Infectious screen			
HIV	1,195	7 (1.2%)	6 (1.0%)
Hepatitis C	1,195	3 (0.5%)	7 (1.2%)
Syphilis	1,195	4 (0.7%)	3 (0.5%)
Hepatitis B	1,195	9 (1.5%)	6 (1.0%)
Urinalysis positive for protein	1,154	61 (10.8%)	64 (10.9%)

HIV, human immunodeficiency virus.

95.7% (22 out of 23) based on the maximum AI prediction in all recording locations. We also noted some differences in performance based on recording location (angled, V2 or handheld) with the highest individual performance for detection of LVEF < 50% across all metrics seen in the angled position (Table 3). While using the maximum prediction (any positive screen across all recording locations) increased sensitivity for detection of LVEF < 50%, this resulted in a decrease in the positive predictive value. Between the angled and maximum prediction, the positive predictive value was lowered from 30.8% to 18.0% while increasing the sensitivity from 87.0% to 95.7%. The false positive rate was also higher for the maximum prediction (100 out of 580, 17%) compared to the angled position (45 out of 562, 8%).

Effectiveness of the 12-lead AI-ECG for LVSD detection. Among participants in the intervention arm, the US FDA-cleared 12-lead AI-ECG algorithm had an AUC of 0.928 (95% CI 0.875–0.981) for detection of LVEF < 50%, whereas for detection of LVEF < 40%, the AUC was 0.928 (95% CI 0.865–0.990; Supplementary Fig. 2). Table 3 shows the diagnostic performance of the US FDA-cleared 12-lead AI-ECG model and the original Mayo Clinic 12-lead AI-ECG model results across a range of ejection fraction thresholds. For the original Mayo Clinic 12-lead AI-ECG model, the AUC values were 0.892 (95% CI 0.825–0.960) and 0.921 (95% CI 0.864–0.979) for detection of LVEF < 50% and <40%, respectively (Supplementary Fig. 3).

Other prespecified (exploratory) outcomes
A summary of other prespecified (exploratory) outcomes is shown in Table 2. There was no statistically significant difference in the number of composite adverse outcomes between groups (100 out of 587 versus 104 out of 608, OR 1.00, 95% CI 0.74–1.35 P = 0.976). Composite adverse outcomes include systolic heart failure, diastolic heart failure, gestational hypertension, pre-eclampsia, eclampsia, gestational diabetes, valvular heart disease, atrial arrhythmias, sustained ventricular arrhythmias and any other reported pregnancy-related complication. We also found no statistically significant difference in the number of composite adverse cardiovascular outcomes between groups (56 out of 587 versus 53 out of 608, OR 1.10, 95% CI 0.74–1.64, P = 0.621). Composite adverse cardiovascular outcomes include diastolic heart failure, gestational hypertension, pre-eclampsia, eclampsia, valvular heart disease, atrial arrhythmias and sustained ventricular arrhythmias.

There were 15 (1.3%) deaths (all-cause mortality) in the overall study cohort, 12 out of 587 (2%) in the intervention arm and 3 out of 608 (0.5%) in the control arm (hazard ratio 4.20, 95% CI 1.18–14.87). Of these, 5 out of 12 (42%) in the intervention arm and 3 out of 3 (100%) in the control arm were attributed to cardiovascular causes, showing a numerically higher but a statistically insignificant difference in CVD-related mortality between the two groups (hazard ratio 1.75, 95% CI 0.42–7.33) (Table 2).

Safety. There were no serious adverse events reported in relation to study participation as defined in the study protocol (Supplementary Information). Five participants reported skin irritation due to placement of adhesive ECG electrodes on the skin, one in the intervention arm and four in the control arm.

Discussion
The SPEC-AI Nigeria randomized trial evaluated AI-guided screening (using a digital stethoscope and standard 12-lead ECG) compared to a control group among pregnant and postpartum women in Nigeria and showed: (1) AI-guided screening with a digital stethoscope doubled the diagnosis of pregnancy-related cardiomyopathy when compared to usual obstetric care, identifying women who may have otherwise remained undiagnosed with this condition. (2) We demonstrate a high prevalence of LVSD, in an obstetric population in Nigeria that supports the need for screening. (3) We found that an AI-enabled digital

Table 2 | Primary and other prespecified outcomes

Outcome	Intervention (n=587)	Control (n=608)	Effect estimate (95% CI)	P value
Primary outcome				
AI-enabled digital stethoscope	24/587	12/608	2.12 (1.05, 4.27)	0.032
AI-enabled 12-lead ECG (US FDA-cleared)	20/587	12/608	1.75 (0.85, 3.62)	0.125
AI-enabled 12-lead ECG (Mayo model)	18/587	12/608	1.57 (0.75, 3.29)	0.227
Other prespecified (exploratory) outcomes				
Composite adverse events[a]	100/587	104/608	1.00 (0.74, 1.35)	0.975
Composite cardiovascular events[b]	56/587	53/608	1.10 (0.74, 1.64)	0.621
Diastolic heart failure	2/587	1/608	2.08 (0.19, 22.95)	0.543
Pre-eclampsia	21/587	19/608	1.15 (0.61, 2.16)	0.664
Eclampsia	6/587	4/608	1.56 (0.44, 5.55)	0.490
Valvular heart disease[c]	1/587	1/608	1.04 (0.06, 16.60)	0.980
All-cause mortality[d]	12/587	3/608	4.20 (1.18, 14.87)	0.026
Cardiovascular mortality[d]	5/587	3/608	1.75 (0.42, 7.33)	0.442

The odds ratio and 95% large sample CI was estimated using a logistic regression model and statistical significance was assessed with a Pearson chi-squared test at the $\alpha = 0.05$ level of significance (two-sided). No adjustments for multiple comparison were performed. [a]Composite adverse events include systolic heart failure, diastolic heart failure, heart failure hospitalization, any cardiomyopathy, hypertensive disorders of pregnancy, valvular heart disease, atrial arrhythmias, sustained ventricular arrhythmias and any other reported pregnancy-related complication. [b]Composite cardiovascular events include diastolic heart failure, gestational hypertension, pre-eclampsia, eclampsia, valvular heart disease, atrial arrhythmias and sustained ventricular arrhythmias. [c]This captures clinically recognized valvular heart disease complicating pregnancy or a new diagnosis with clinical impact during index pregnancy. [d]Cox-proportional hazards regression analysis was performed and hazard ratios are reported with 95% CI.

stethoscope that analyzes single-lead ECG and phonocardiogram recordings, as well as an AI-enabled 12-lead ECG detected the presence of cardiomyopathy with high sensitivity, specificity and negative predictive value. Only one case of LVSD was incorrectly classified by the digital stethoscope at study entry.
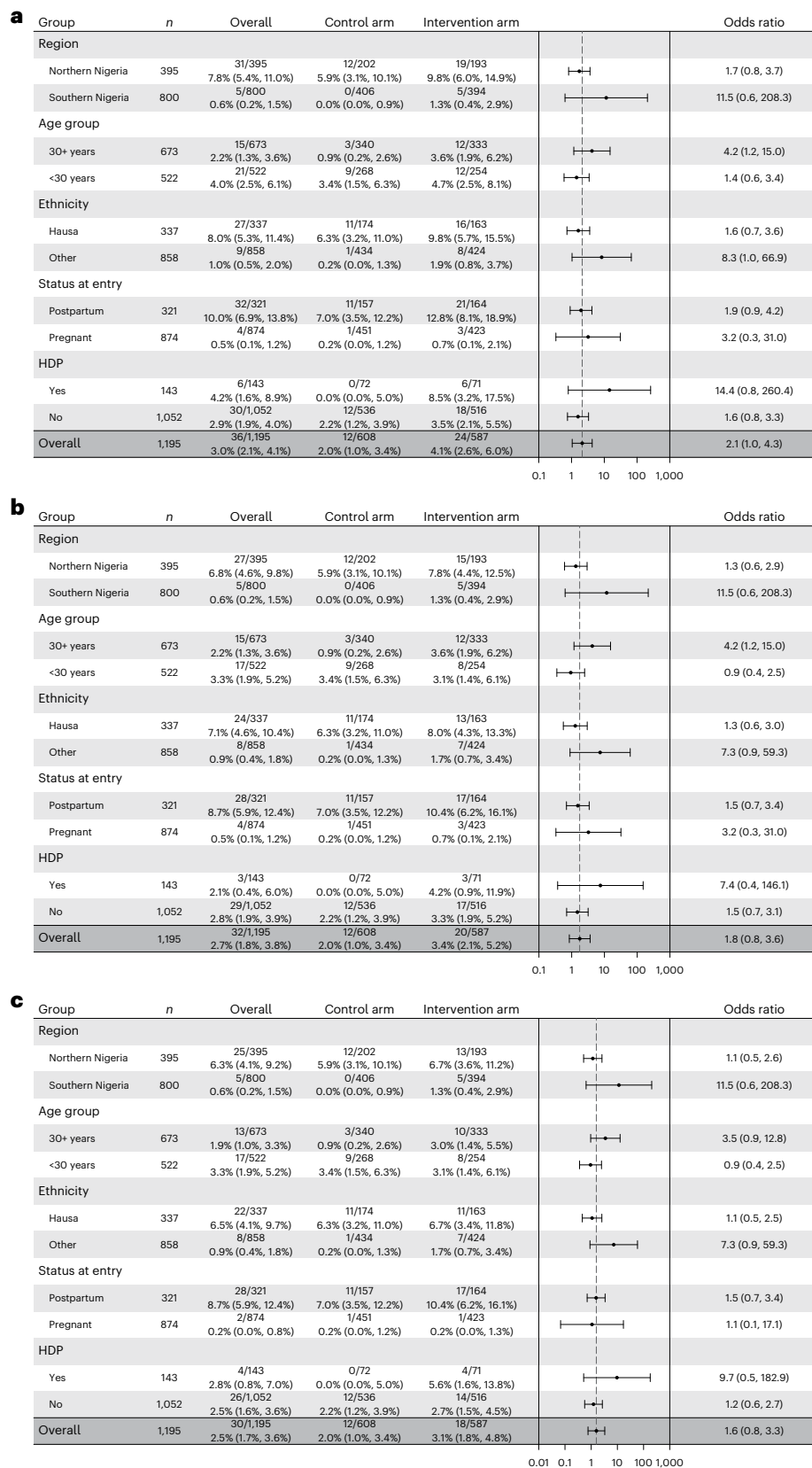
We completed a large prospective trial evaluating an AI-guided intervention for cardiomyopathy screening in an obstetric population and the results are consistent with findings from a pilot, prospective study conducted by members of our team in a sample of 100 pregnant and postpartum individuals and the EAGLE trial, which evaluated an AI-enabled 12-lead ECG in the primary care patient population. In the pilot prospective study, we showed that an AI-enabled 12-lead ECG had perfect discrimination for detection of LVEF < 50% and the digital stethoscope had an AUC of 0.97 with 100% sensitivity[16]. The EAGLE trial showed that AI-guided screening, using a standard 12-lead ECG, improved the detection of asymptomatic LVSD with a 32% increase in the odds of detection[14], whereas our study showed a doubling of the odds of LVSD detection. A large, prospective observational study in the United Kingdom has also demonstrated the effectiveness of an AI-enabled digital stethoscope in detecting LVSD at the point of care in a primary care population[19] and small, prospective US-based studies showed similar effectiveness[16,20]. The key advantages of this screening modality in a low-to-middle-income country such as Nigeria include the following: portability of equipment (allowing it to be transported easily across clinical locations and settings), battery-powered (so it does not rely on availability of a stable electric power supply), enclosed system (reduced likelihood of electromagnetic interference on recorded ECG signals) and real-time availability of AI predictions at the point of care. It also has the potential to improve risk stratification before cardiology referral given the shortage of cardiovascular specialists in the country; as of 2023, there were approximately 450 cardiologists in Nigeria in a population of 213.4 million people[21].

We show in this study, that the NNS to identify one new case of perinatal LVSD is 47. Taken in context with other studies in the obstetric population, the NNS to prevent one case of preterm pre-eclampsia using the Fetal Medicine Foundation triple test is 250 (ref. 22), the NNS to prevent one case of postpartum depression using the Edinburgh Postnatal Depression Scale is 25 (ref. 23), the NNS to detect one additional case of severe congenital heart disease using fetal echocardiography is 436 (ref. 24) and the NNS to detect one spontaneous preterm delivery using cervical length measured with transvaginal ultrasound ranges from 161 to 1,018 for high- and low-risk pregnancies, respectively[25]. This emphasizes the powerful benefit of using AI for LVSD screening in this population. The result of this study also validates the effectiveness of the AI-enabled digital stethoscope and 12-lead ECG for cardiomyopathy screening in pregnancy, postpartum, a predominantly Black population and in low-resource obstetric settings and supports the hypothesis that digital technologies have the potential to narrow the disparity gap in cardiovascular care.

The maternal mortality ratio unfortunately remains very high in Nigeria (estimated at 1,047 per 100,000 live births in 2020) and in many low-to-middle-income countries[26]. Nigeria had the highest number of maternal deaths in 2020 (approximately 82,000), accounting for 28.5% of all global maternal deaths[26]. In the current study, the observed mortality rate was 1.3% and 53% of these were from cardiovascular causes; however, the study sample is not necessarily representative of the broader obstetric population. In Nigeria, the high maternal mortality is attributed to inadequate access to timely and affordable health services. Despite these challenges, the availability and uptake of mobile and smart phone technologies has grown exponentially. In fact, Nigeria and Ethiopia are estimated to become the fastest growing economies for mobile phone and mobile internet technology in Africa by 2025 (ref. 27). This highlights an opportunity to leverage digital and mobile technologies to transform healthcare in Nigeria and other African countries.

Peripartum cardiomyopathy, a unique form of pregnancy-related LVSD believed to occur de novo in late pregnancy and the early postpartum period[28], is a leading cause of maternal mortality particularly in the postpartum period[29,30], and its true incidence remains unknown due to lack of routine screening[2]. Previously reported estimates from Nigeria note an incidence rate of ~1 in 100 live births. In this study, the total number of cardiomyopathy cases identified during an 18-month time frame was 37 out of 1,195 (3.1%), higher than previously reported, although this study did not attempt to exclude patients with previous diagnosis of peripartum cardiomyopathy or known LVSD before pregnancy. Studies have shown that delays in the diagnosis of cardiomyopathy during the peripartum period is associated with poorer outcomes[6] as such, it is imperative that we are able to identify cardiac dysfunction early so that appropriate care can be initiated to reduce associated adverse maternal and infant outcomes. More recently, one study demonstrated increased risk of adverse outcomes in subsequent pregnancies among women with a previous history of peripartum cardiomyopathy[31]. This supports the need to screen women across the perinatal continuum, including before conception. A low-cost screening intervention for effective and early identification of pregnancy-related cardiomyopathy has the potential to reduce adverse maternal and infant outcomes given that these are closely linked[32]. In fact, maternal death is known to be a large driver of infant mortality with up to a 46-fold higher risk of infant death within the first month[33]. Although in the present study, we showed a significantly higher all-cause mortality (largely from renal failure,

a

Group	n	Overall	Control arm	Intervention arm	Odds ratio
Region					
Northern Nigeria	395	31/395 7.8% (5.4%, 11.0%)	12/202 5.9% (3.1%, 10.1%)	19/193 9.8% (6.0%, 14.9%)	1.7 (0.8, 3.7)
Southern Nigeria	800	5/800 0.6% (0.2%, 1.5%)	0/406 0.0% (0.0%, 0.9%)	5/394 1.3% (0.4%, 2.9%)	11.5 (0.6, 208.3)
Age group					
30+ years	673	15/673 2.2% (1.3%, 3.6%)	3/340 0.9% (0.2%, 2.6%)	12/333 3.6% (1.9%, 6.2%)	4.2 (1.2, 15.0)
<30 years	522	21/522 4.0% (2.5%, 6.1%)	9/268 3.4% (1.5%, 6.3%)	12/254 4.7% (2.5%, 8.1%)	1.4 (0.6, 3.4)
Ethnicity					
Hausa	337	27/337 8.0% (5.3%, 11.4%)	11/174 6.3% (3.2%, 11.0%)	16/163 9.8% (5.7%, 15.5%)	1.6 (0.7, 3.6)
Other	858	9/858 1.0% (0.5%, 2.0%)	1/434 0.2% (0.0%, 1.3%)	8/424 1.9% (0.8%, 3.7%)	8.3 (1.0, 66.9)
Status at entry					
Postpartum	321	32/321 10.0% (6.9%, 13.8%)	11/157 7.0% (3.5%, 12.2%)	21/164 12.8% (8.1%, 18.9%)	1.9 (0.9, 4.2)
Pregnant	874	4/874 0.5% (0.1%, 1.2%)	1/451 0.2% (0.0%, 1.2%)	3/423 0.7% (0.1%, 2.1%)	3.2 (0.3, 31.0)
HDP					
Yes	143	6/143 4.2% (1.6%, 8.9%)	0/72 0.0% (0.0%, 5.0%)	6/71 8.5% (3.2%, 17.5%)	14.4 (0.8, 260.4)
No	1,052	30/1,052 2.9% (1.9%, 4.0%)	12/536 2.2% (1.2%, 3.9%)	18/516 3.5% (2.1%, 5.5%)	1.6 (0.8, 3.3)
Overall	1,195	36/1,195 3.0% (2.1%, 4.1%)	12/608 2.0% (1.0%, 3.4%)	24/587 4.1% (2.6%, 6.0%)	2.1 (1.0, 4.3)

0.1   1   10   100   1,000

b

Group	n	Overall	Control arm	Intervention arm	Odds ratio
Region					
Northern Nigeria	395	27/395 6.8% (4.6%, 9.8%)	12/202 5.9% (3.1%, 10.1%)	15/193 7.8% (4.4%, 12.5%)	1.3 (0.6, 2.9)
Southern Nigeria	800	5/800 0.6% (0.2%, 1.5%)	0/406 0.0% (0.0%, 0.9%)	5/394 1.3% (0.4%, 2.9%)	11.5 (0.6, 208.3)
Age group					
30+ years	673	15/673 2.2% (1.3%, 3.6%)	3/340 0.9% (0.2%, 2.6%)	12/333 3.6% (1.9%, 6.2%)	4.2 (1.2, 15.0)
<30 years	522	17/522 3.3% (1.9%, 5.2%)	9/268 3.4% (1.5%, 6.3%)	8/254 3.1% (1.4%, 6.1%)	0.9 (0.4, 2.5)
Ethnicity					
Hausa	337	24/337 7.1% (4.6%, 10.4%)	11/174 6.3% (3.2%, 11.0%)	13/163 8.0% (4.3%, 13.3%)	1.3 (0.6, 3.0)
Other	858	8/858 0.9% (0.4%, 1.8%)	1/434 0.2% (0.0%, 1.3%)	7/424 1.7% (0.7%, 3.4%)	7.3 (0.9, 59.3)
Status at entry					
Postpartum	321	28/321 8.7% (5.9%, 12.4%)	11/157 7.0% (3.5%, 12.2%)	17/164 10.4% (6.2%, 16.1%)	1.5 (0.7, 3.4)
Pregnant	874	4/874 0.5% (0.1%, 1.2%)	1/451 0.2% (0.0%, 1.2%)	3/423 0.7% (0.1%, 2.1%)	3.2 (0.3, 31.0)
HDP					
Yes	143	3/143 2.1% (0.4%, 6.0%)	0/72 0.0% (0.0%, 5.0%)	3/71 4.2% (0.9%, 11.9%)	7.4 (0.4, 146.1)
No	1,052	29/1,052 2.8% (1.9%, 3.9%)	12/536 2.2% (1.2%, 3.9%)	17/516 3.3% (1.9%, 5.2%)	1.5 (0.7, 3.1)
Overall	1,195	32/1,195 2.7% (1.8%, 3.8%)	12/608 2.0% (1.0%, 3.4%)	20/587 3.4% (2.1%, 5.2%)	1.8 (0.8, 3.6)

0.1   1   10   100   1,000

c

Group	n	Overall	Control arm	Intervention arm	Odds ratio
Region					
Northern Nigeria	395	25/395 6.3% (4.1%, 9.2%)	12/202 5.9% (3.1%, 10.1%)	13/193 6.7% (3.6%, 11.2%)	1.1 (0.5, 2.6)
Southern Nigeria	800	5/800 0.6% (0.2%, 1.5%)	0/406 0.0% (0.0%, 0.9%)	5/394 1.3% (0.4%, 2.9%)	11.5 (0.6, 208.3)
Age group					
30+ years	673	13/673 1.9% (1.0%, 3.3%)	3/340 0.9% (0.2%, 2.6%)	10/333 3.0% (1.4%, 5.5%)	3.5 (0.9, 12.8)
<30 years	522	17/522 3.3% (1.9%, 5.2%)	9/268 3.4% (1.5%, 6.3%)	8/254 3.1% (1.4%, 6.1%)	0.9 (0.4, 2.5)
Ethnicity					
Hausa	337	22/337 6.5% (4.1%, 9.7%)	11/174 6.3% (3.4%, 11.0%)	11/163 6.7% (3.4%, 11.8%)	1.1 (0.5, 2.5)
Other	858	8/858 0.9% (0.4%, 1.8%)	1/434 0.2% (0.0%, 1.3%)	7/424 1.7% (0.7%, 3.4%)	7.3 (0.9, 59.3)
Status at entry					
Postpartum	321	28/321 8.7% (5.9%, 12.4%)	11/157 7.0% (3.5%, 12.2%)	17/164 10.4% (6.2%, 16.1%)	1.5 (0.7, 3.4)
Pregnant	874	2/874 0.2% (0.0%, 0.8%)	1/451 0.2% (0.0%, 1.2%)	1/423 0.2% (0.0%, 1.3%)	1.1 (0.1, 17.1)
HDP					
Yes	143	4/143 2.8% (0.8%, 7.0%)	0/72 0.0% (0.0%, 5.0%)	4/71 5.6% (1.6%, 13.8%)	9.7 (0.5, 182.9)
No	1,052	26/1,052 2.5% (1.6%, 3.6%)	12/536 2.2% (1.2%, 3.9%)	14/516 2.7% (1.5%, 4.5%)	1.2 (0.6, 2.7)
Overall	1,195	30/1,195 2.5% (1.7%, 3.6%)	12/608 2.0% (1.0%, 3.4%)	18/587 3.1% (1.8%, 4.8%)	1.6 (0.8, 3.3)

0.01   0.1   1   10   100   1,000

Fig. 3 | Forest plots showing primary outcome stratified by subgroups.
a, Performance of the digital stethoscope. **b**, Performance of the US FDA-cleared 12-lead AI-ECG algorithm. **c**, Performance of the original Mayo Clinic 12-lead AI-ECG algorithm in detecting the primary outcome within each prespecified subgroup. Data in the columns are presented as frequencies and percentages with 95% exact CI in parenthesis. The column with error bars represents odds ratio estimates depicted as a black dot and the error bar represents the large sample 95% CI around the odds ratio estimate. The odds ratios and 95% large sample CI were estimated using logistic regression. HDP, hypertensive disorder of pregnancy (includes chronic hypertension, gestational hypertension, pre-eclampsia and eclampsia).

Table 3 | Secondary outcomes: AI-enabled digital stethoscope and 12-lead ECG performance in the intervention arm at study entry (n=587)

Outcome	Model (threshold)	n/N	AUC	Sensitivity	Specificity	Positive predictive value	Negative predictive value
LVEF ≤35%	Angled (0.43)	562/587	0.972 (0.940, 1.000)	94.1% (71.3%, 99.9%) 16/17	91.0% (88.3%, 93.3%) 496/545	24.6% (14.8%, 36.9%) 16/65	99.8% (98.9%, 100.0%) 496/497
	V2 (0.43)	564/587	0.960 (0.927, 0.993)	93.8% (69.8%, 99.8%) 15/16	87.4% (84.3%, 90.1%) 479/548	17.9% (10.4%, 27.7%) 15/84	99.8% (98.8%, 100.0%) 479/480
	Handheld (0.43)	556/587	0.973 (0.953, 0.994)	66.7% (38.4%, 88.2%) 10/15	96.9% (95.0%, 98.2%) 524/541	37.0% (19.4%, 57.6%) 10/27	99.1% (97.8%, 99.7%) 524/529
	Max prediction (0.43)	580/587	0.983 (0.971, 0.996)	100.0% (80.5%, 100.0%) 17/17	81.3% (77.9%, 84.5%) 458/563	13.9% (8.3%, 21.4%) 17/122	100.0% (99.2%, 100.0%) 458/458
	Mayo 12-lead (0.256)	587/587	0.919 (0.855, 0.983)	41.2% (18.4%, 67.1%) 7/17	97.9% (96.4%, 98.9%) 558/570	36.8% (16.3%, 61.6%) 7/19	98.2% (96.8%, 99.2%) 558/568
	Revised 12-lead (0.56)	580/587	0.957 (0.904, 1.000)	85.7% (57.2%, 98.2%) 12/14	96.8% (95.0%, 98.1%) 548/566	40.0% (22.7%, 59.4%) 12/30	99.6% (98.7%, 100.0%) 548/550
LVEF <40%	Angled (0.43)	562/587	0.976 (0.949, 1.000)	95.0% (75.1%, 99.9%) 19/20	91.5% (88.8%, 93.7%) 496/542	29.2% (18.6%, 41.8%) 19/65	99.8% (98.9%, 100.0%) 496/497
	V2 (0.43)	564/587	0.962 (0.934, 0.990)	94.7% (74.0%, 99.9%) 18/19	87.9% (84.9%, 90.5%) 479/545	21.4% (13.2%, 31.7%) 18/84	99.8% (98.8%, 100.0%) 479/480
	Handheld (0.43)	556/587	0.972 (0.952, 0.992)	64.7% (38.3%, 85.8%) 11/17	97.0% (95.2%, 98.3%) 523/539	40.7% (22.4%, 61.2%) 11/27	98.9% (97.5%, 99.6%) 523/529
	Max prediction (0.43)	580/587	0.985 (0.974, 0.996)	100.0% (83.2%, 100.0%) 20/20	81.8% (78.3%, 84.9%) 458/560	16.4% (10.3%, 24.2%) 20/122	100.0% (99.2%, 100.0%) 458/458
	Mayo 12-lead (0.256)	587/587	0.921 (0.864, 0.979)	45.0% (23.1%, 68.5%) 9/20	98.2% (96.8%, 99.2%) 557/567	47.4% (24.4%, 71.1%) 9/19	98.1% (96.6%, 99.0%) 557/568
	Revised 12-lead (0.56)	580/587	0.928 (0.865, 0.990)	76.5% (50.1%, 93.2%) 13/17	97.0% (95.2%, 98.2%) 546/563	43.3% (25.5%, 62.6%) 13/30	99.3% (98.1%, 99.8%) 546/550
LVEF <45%	Angled (0.43)	562/587	0.963 (0.932, 0.995)	87.0% (66.4%, 97.2%) 20/23	91.7% (89.0%, 93.8%) 494/539	30.8% (19.9%, 43.4%) 20/65	99.4% (98.2%, 99.9%) 494/497
	V2 (0.43)	564/587	0.960 (0.933, 0.987)	90.9% (70.8%, 98.9%) 20/22	88.2% (85.2%, 90.8%) 478/542	23.8% (15.2%, 34.3%) 20/84	99.6% (98.5%, 99.9%) 478/480
	Handheld (0.43)	556/587	0.943 (0.905, 0.982)	55.0% (31.5%, 76.9%) 11/20	97.0% (95.2%, 98.3%) 520/536	40.7% (22.4%, 61.2%) 11/27	98.3% (96.8%, 99.2%) 520/529
	Max prediction (0.43)	580/587	0.976 (0.953, 0.998)	95.7% (78.1%, 99.9%) 22/23	82.0% (78.6%, 85.1%) 457/557	18.0% (11.7%, 26.0%) 22/122	99.8% (98.8%, 100.0%) 457/458
	Mayo 12-lead (0.256)	587/587	0.892 (0.825, 0.960)	43.5% (23.2%, 65.5%) 10/23	98.4% (97.0%, 99.3%) 555/564	52.6% (28.9%, 75.6%) 10/19	97.7% (96.1%, 98.8%) 555/568
	Revised 12-lead (0.56)	580/587	0.928 (0.875, 0.981)	70.0% (45.7%, 88.1%) 14/20	97.1% (95.4%, 98.4%) 544/560	46.7% (28.3%, 65.7%) 14/30	98.9% (97.6%, 99.6%) 544/550
LVEF <50%	Angled (0.43)	562/587	0.963 (0.932, 0.995)	87.0% (66.4%, 97.2%) 20/23	91.7% (89.0%, 93.8%) 494/539	30.8% (19.9%, 43.4%) 20/65	99.4% (98.2%, 99.9%) 494/497
	V2 (0.43)	564/587	0.960 (0.933, 0.987)	90.9% (70.8%, 98.9%) 20/22	88.2% (85.2%, 90.8%) 478/542	23.8% (15.2%, 34.3%) 20/84	99.6% (98.5%, 99.9%) 478/480
	Handheld (0.43)	556/587	0.943 (0.905, 0.982)	55.0% (31.5%, 76.9%) 11/20	97.0% (95.2%, 98.3%) 520/536	40.7% (22.4%, 61.2%) 11/27	98.3% (96.8%, 99.2%) 520/529
	Max prediction (0.43)	580/587	0.976 (0.953, 0.998)	95.7% (78.1%, 99.9%) 22/23	82.0% (78.6%, 85.1%) 457/557	18.0% (11.7%, 26.0%) 22/122	99.8% (98.8%, 100.0%) 457/458
	Mayo 12-lead (0.256)	587/587	0.892 (0.825, 0.960)	43.5% (23.2%, 65.5%) 10/23	98.4% (97.0%, 99.3%) 555/564	52.6% (28.9%, 75.6%) 10/19	97.7% (96.1%, 98.8%) 555/568
	Revised 12-lead (0.56)	580/587	0.928 (0.875, 0.981)	70.0% (45.7%, 88.1%) 14/20	97.1% (95.4%, 98.4%) 544/560	46.7% (28.3%, 65.7%) 14/30	98.9% (97.6%, 99.6%) 544/550

The results provided in this table are based on baseline assessments only where all participants in the intervention arm had a confirmatory echocardiogram carried out for validation of AI model performance. All results include ECGs/digital stethoscope recordings acquired on the same day as the echocardiogram. The angled, V2 and handheld models are based on the digital stethoscope recordings. The max prediction is derived as the maximum model output across the three recording positions (any positive prediction). The model and operating cutoff point of 0.43 were developed to detect an LVEF <40%. If a specific recording position was not captured or was of insufficient quality to be analyzed, the corresponding AI prediction was not included, resulting in a different effective sample size for each digital stethoscope recording location. The Mayo 12-lead model is the original Mayo Clinic LVEF model as reported in ref. 13. This model and operating point was trained to detect LVEF ≤35% at an operating cutoff point of 0.256. This model does not include any quality checks on the algorithm and provides a result for all available ECGs (n=587). The revised 12-lead model is based on a modified version of original Mayo Clinic LVEF model (US FDA-cleared version), trained to detect an LVEF <40% at an operating cutoff point of 0.56. This model had an in-built ECG data quality check. ECGs deemed to be of insufficient quality were thus not analyzed and a corresponding AI prediction was not generated resulting in an effective sample size of n=580.

infections or other unknown causes) in the screening arm compared to the control arm. Cardiovascular mortality was also numerically higher (5 versus 3) in the intervention arm but this did not reach statistical significance. The screening intervention was not tailored to influence physician practice or clinical care plans and treatment recommendations were at the discretion of the managing physician. As such, the direct impact of AI screening on treatment interventions, healthcare costs and utilization, mortality and other clinical outcomes will require further exploration and additional studies. Potential explanations or hypothesis for the observed higher mortality in the intervention arm include (1) a differential observation of mortality in the intervention arm due to increased contact with healthcare services, particularly among those diagnosed with cardiomyopathy who were referred for cardiovascular care. (2) Mortality ascertainment in this study may be prone to error as there are no national death registries and so this involved medical record review for in-hospital mortality (6 out of 15, 40% of reported deaths; 4 in the intervention arm and 2 in the control arm) and contacting patients and relatives via phone (9 out of 15, 60% of reported deaths). Cultural norms and practices related to providing news of death to unknown or unfamiliar parties may have impacted the reporting of death status by relatives to the study staff. Although the rates of attrition in the study were similar across study groups, the differential reporting of mortality status may be different in the control arm, which may not have had as much of a relationship with the study teams due to limited interactions and study-related testing, whereas those in the intervention arm likely developed a closer relationship with members of the study teams due to additional testing as well as increased interactions with the cardiology teams following the detection of LVSD.

Nevertheless, the use of AI-guided screening as an initial step has the potential not only to improve cardio-obstetric clinical care but also provides a huge opportunity for large population-based studies to assess the incidence and prevalence of cardiac dysfunction in the peripartum period in a safe manner.

The key strengths of this study include a large sample size and the enrollment of a predominantly Black, ethnically and regionally diverse obstetric population in Nigeria. The study results are based on the mITT analysis as specified in the statistical analysis plan. This approach was deemed reasonable as not obtaining the testing required to generate the associated AI prediction limits the ability to obtain or interpret a confirmatory echocardiogram in the intervention arm. In addition, we evaluated an ITT approach for the primary end point analysis using the digital stethoscope and demonstrated that an ITT analysis would not alter the study results and conclusions.

The key limitations of this study were introduced by the pragmatic clinical trial design and enrolling study participants at teaching hospitals with a licensed cardiologist and echocardiography capabilities. Two-thirds of study participants (66%) were either in the third trimester or postpartum at the time of study entry. While this represents the time frame at which peripartum cardiomyopathy is likely to develop, it limits the number of follow-up visits that the participants would be expected to receive as part of their routine clinical care. Only 61% of participants completed a second study visit, with rapidly decreasing percentages through the seventh visit. This attrition is due to a combination of factors, including variation in pregnancy time point at study entry, frequency of scheduled routine care visits postpartum and loss to follow-up. The trial was also designed to allow for up to seven visits and as such, it did not require completion of all seven visits to be compliant as would be expected in a standard explanatory clinical trial[34]. The limited follow-up resulted in similarity in the primary end point (detection over 18 months) compared to the baseline detection rates, as shown in the results.

Regarding the limitation of potential referral bias due to inclusion of hospital systems that had echocardiography capabilities, LVSD prevalence seen at these tertiary centers may not be reflective of the general obstetric patient population in Nigeria. Also, the actual frequency of LVSD in the control arm remains unknown, as echocardiograms were only performed in this group at the discretion of the managing physician. Although it is assumed that the distribution of LVSD is likely to be similar due to randomization, this was supported by a similar distribution of positive 12-lead AI-ECG predictions for LVSD, based on the US FDA-cleared algorithm, in the intervention and control arm (analyzed following study completion), 5.2% versus 7.3% respectively ($P = 0.136$); however, this cannot be confirmed as echocardiograms were not mandated among all participants in the control arm but only performed at the discretion of the managing physician (a disadvantage of the trial design). In addition, we are unable to estimate the potential impact of out-of-pocket costs on obtaining an echocardiogram in the control arm and the socioeconomic status of individual study participants was not assessed as part of the study. As such, not receiving an echocardiogram in the control arm may be influenced by a combination of (1) not being considered clinically necessary by the managing physician; (2) LVSD was suspected but an echocardiogram was not performed due to an out-of-pocket cost to the patient; or (3) the need to return to the clinic for an echocardiogram if it was recommended.

Finally, an important limitation in the design of this study was the selected cutoff for determination of the primary outcome, where LVSD was defined as LVEF < 50%. This did not exactly match categorizations used during model derivation. In particular, the original 12-lead model developed by the Mayo Clinic was trained to detect LVEF ≤ 35% (ref. 13) and this model was retrained to detect LVEF < 40% (US FDA-cleared version)[35] and again retrained using single leads to detect LVEF < 40% for use with the Eko digital stethoscope[36]. Table 3 shows a general pattern of results that model performance was improved with closer alignment of the study's LVSD categorization with the original model specifications.

We demonstrate in this study, the differential impact of the different AI algorithms used on LVSD detection rates in an obstetric patient population. Although all AI models had very good AUC values (values of 0.872 for the original Mayo Clinic 12-lead AI-ECG model, 0.923 for the US FDA-cleared 12-lead AI-ECG model, and 0.975 for the digital stethoscope model for detection of LVEF < 50%), we see variations in specific diagnostic performance metrics based on the preselected thresholds at model derivation, particularly sensitivity. These findings suggest that it might be important to carefully evaluate AI models in specific patient populations before deployment and considerations made for possible fine-tuning of the predetermined thresholds to suit specific patient populations and clinical environments.

In conclusion, among pregnant and postpartum women, AI-guided screening with the use of a digital stethoscope was associated with an increase in the diagnosis of cardiomyopathy associated with LVSD compared to usual care. This intervention has the potential to improve cardio-obstetric care by reducing delays in the diagnosis of a life-threatening but treatable condition.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-03243-9.

References

1. Goli, R. et al. Genetic and phenotypic landscape of peripartum cardiomyopathy. *Circulation* **143**, 1852–1862 (2021).
2. Gunderson, E. P. et al. Epidemiology of peripartum cardiomyopathy: incidence, predictors, and outcomes. *Obstet. Gynecol.* **118**, 583–591 (2011).
3. Karaye, K. M. et al. Clinical features and outcomes of peripartum cardiomyopathy in nigeria. *J. Am. Coll. Cardiol.* **76**, 2352–2364 (2020).

4. Karaye, K. M. et al. Incidence, clinical characteristics, and risk factors of peripartum cardiomyopathy in Nigeria: results from the PEACE Registry. *ESC Heart Fail.* **7**, 235–243 (2020).

5. Isogai, T. & Kamiya, C. A. Worldwide incidence of peripartum cardiomyopathy and overall maternal mortality. *Int. Heart J.* **60**, 503–511 (2019).

6. Hameed, A. B. et al. Pregnancy-related cardiovascular deaths in California: beyond peripartum cardiomyopathy. *Am. J. Obstet. Gynecol.* **213**, 379.e1–10 (2015).

7. Sliwa, K., Bauersachs, J., Arany, Z., Spracklen, T. F. & Hilfiker-Kleiner, D. Peripartum cardiomyopathy: from genetics to management. *Eur. Heart J.* **42**, 3094–3102 (2021).

8. Davis, M. B. et al. Team-based care of women with cardiovascular disease from pre-conception through pregnancy and postpartum: JACC Focus Seminar 1/5. *J. Am. Coll. Cardiol.* **77**, 1763–1777 (2021).

9. Attia, Z. I. 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* **394**, 861–867 (2019).

10. Ko, W. Y. et al. Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. *J. Am. Coll. Cardiol.* **75**, 722–733 (2020).

11. Cohen-Shelly, M. et al. Electrocardiogram screening for aortic valve stenosis using artificial intelligence. *Eur. Heart J.* **42**, 2885–2896 (2021).

12. Elias, P. et al. Deep learning electrocardiographic analysis for detection of left-sided valvular heart disease. *J. Am. Coll. Cardiol.* **80**, 613–626 (2022).

13. Attia, Z. I. et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat. Med.* **25**, 70–74 (2019).

14. Yao, X. et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat. Med.* **27**, 815–819 (2021).

15. Adedinsewo, D. A. et al. Detecting cardiomyopathies in pregnancy and the postpartum period with an electrocardiogram-based deep learning model. *Eur. Heart J. Digit. Health* **2**, 586–596 (2021).

16. Adedinsewo, D. et al. Artificial intelligence based screening for cardiomyopathy in an obstetric population: a pilot study. *Cardiovasc. Digit. Health J.* **5**, 132–140 (2024).

17. Karabayir, I. et al. Development and validation of an electrocardiographic artificial intelligence model for detection of peripartum cardiomyopathy. *Am. J. Obstet. Gynecol. MFM* **6**, 101337 (2024).

18. Lee, Y. et al. An artificial intelligence electrocardiogram analysis for detecting cardiomyopathy in the peripartum period. *Int. J. Cardiol.* **352**, 72–77 (2022).

19. Bachtiger, P. et al. Point-of-care screening for heart failure with reduced ejection fraction using artificial intelligence during ECG-enabled stethoscope examination in London, UK: a prospective, observational, multicentre study. *Lancet Digit. Health* **4**, e117–e125 (2022).

20. Attia, Z. I. et al. Automated detection of low ejection fraction from a one-lead electrocardiogram: application of an AI algorithm to an electrocardiogram-enabled digital stethoscope. *Eur. Heart J. Digit. Health* **3**, 373–379 (2022).

21. Ogah, O. S., Orimolade, O. A. & Jinadu, T. O. Cardiovascular diseases in nigeria: current status, threats, and opportunities. *Circulation* **148**, 1441–1444 (2023).

22. Chaemsaithong, P., Sahota, D. S. & Poon, L. C. First trimester preeclampsia screening and prediction. *Am. J. Obstet. Gynecol.* **226**, S1071–S1097.e2 (2022).

23. Beck, A. et al. Screening for depression among the general adult population and in women during pregnancy or the first-year postpartum: two systematic reviews to inform a guideline of the canadian task force on preventive health care. *Syst. Rev.* **11**, 176 (2022).

24. Bellavance, S., Cardinal, M.-P., Gobeil, L., Roy-Lacroix, M.-E. & Dallaire, F. The mathematical limitations of fetal echocardiography as a screening tool in the setting of a normal second-trimester ultrasound. *CJC Open.* **3**, 987–993 (2021).

25. Wikström, T. et al. Effect of second-trimester sonographic cervical length on the risk of spontaneous preterm delivery in different risk groups: a prospective observational multicenter study. *Acta Obstet. Gynecol. Scand.* **100**, 1644–1655 (2021).

26. *Trends in maternal mortality 2000 to 2020: estimates by* WHO, UNICEF, *UNFPA, World Bank Group and UNDESA/Population Division*. 2023. https://www.who.int/publications/i/item/9789240068759 (WHO, 2023).

27. *The Mobile Economy: Sub Saharan Africa 2022.* https://www.gsma.com/mobileeconomy/wp-content/uploads/2022/10/The-Mobile-Economy-Sub-Saharan-Africa-2022.pdf (GMSA, 2022).

28. Davis, M. B., Arany, Z., McNamara, D. M., Goland, S. & Elkayam, U. Peripartum cardiomyopathy: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **75**, 207–221 (2020).

29. Petersen, E. E. et al. Vital signs: pregnancy-related deaths, United States, 2011–2015, and strategies for prevention, 13 states, 2013–2017. *MMWR Morb. Mortal. Wkly Rep.* **68**, 423–429 (2019).

30. MacDorman, M. F., Thoma, M., Declcerq, E. & Howell, E. A. Racial and ethnic disparities in maternal mortality in the united states using enhanced vital records, 2016-2017. *Am. J. Public Health* **111**, 1673–1681 (2021).

31. Pachariyanon, P., Bogabathina, H., Jaisingh, K., Modi, M. & Modi, K. Long-term outcomes of women with peripartum cardiomyopathy having subsequent pregnancies. *J. Am. Coll. Cardiol.* **82**, 16–26 (2023).

32. *Maternal Mortality.* https://www.who.int/news-room/fact-sheets/detail/maternal-mortality (WHO, 2024).

33. Moucheraud, C. et al. Consequences of maternal mortality on infant and child survival: a 25-year longitudinal analysis in Butajira Ethiopia (1987–2011). *Reprod. Health* **12**, S4 (2015).

34. Le-Rademacher, J., Gunn, H., Yao, X. & Schaid, D. J. Clinical trials overview: from explanatory to pragmatic clinical trials. *Mayo Clin. Proc.* **98**, 1241–1253 (2023).

35. *Low Ejection Fraction AI-ECG Algorithm: Reduced Ejection Fraction Machine Learning-based Notification Software.* https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K232699 (US FDA, 2024).

36. *Eko Low Ejection Fraction Tool (ELEFT): Reduced Ejection Fraction Machine Learning-based Notification Software.* https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K233409 (US FDA, 2024).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2905

Demilade A. Adedinsewo ®[1] ✉, Andrea Carolina Morales-Lara[1], Bosede B. Afolabi ®[2], Oyewole A. Kushimo ®[3], Amam C. Mbakwem[3], Kehinde F. Ibiyemi ®[4], James Ayodele Ogunmodede[5], Hadijat Olaide Raji ®[4], Sadiq H. Ringim ®[6], Abdullahi A. Habib[7], Sabiu M. Hamza ®[6], Okechukwu S. Ogah[8], Gbolahan Obajimi[9], Olugbenga Oluseun Saanu[9], Olusoji E. Jagun[10], Francisca O. Inofomoh ®[11], Temitope Adeolu[11], Kamilu M. Karaye ®[12], Sule A. Gaya[13], Isiaka Alfa[12], Cynthia Yohanna[14], K. L. Venkatachalam[1], Jennifer Dugan[15], Xiaoxi Yao ®[15,16], Hanna J. Sledge[17], Patrick W. Johnson[17], Mikolaj A. Wieczorek[17], Zachi I. Attia ®[15], Sabrina D. Phillips[1], Mohamad H. Yamani[1], Yvonne Butler Tobah ®[18], Carl H. Rose[18], Emily E. Sharpe ®[19], Francisco Lopez-Jimenez[15], Paul A. Friedman ®[15], Peter A. Noseworthy ®[15], Rickey E. Carter[17] & on behalf of the SPEC-AI Nigeria Investigators*

[1]Department of Cardiovascular Medicine, Mayo Clinic, Jacksonville, FL, USA. [2]Department of Obstetrics and Gynaecology, College of Medicine and Centre for Clinical Trials, Research and Implementation Science, University of Lagos, Lagos, Nigeria. [3]Cardiology Unit, Department of Medicine, Lagos University Teaching Hospital, Lagos, Nigeria. [4]Department of Obstetrics & Gynaecology, University of Ilorin Teaching Hospital, Ilorin, Nigeria. [5]Department of Medicine, University of Ilorin, Ilorin, Nigeria. [6]Department of Medicine, Rasheed Shekoni Specialist Hospital, Dutse, Nigeria. [7]Department of Obstetrics and Gynaecology, Rasheed Shekoni Specialist Hospital, Dutse, Nigeria. [8]Department of Medicine, University of Ibadan, Oyo, Nigeria. [9]Department of Obstetrics and Gynaecology, University College Hospital Ibadan, Oyo, Nigeria. [10]Department of Obstetrics and Gynaecology, Olabisi Onabanjo University Teaching Hospital, Sagamu, Nigeria. [11]Cardiology Unit, Department of Medicine, Olabisi Onabanjo University Teaching Hospital, Sagamu, Nigeria. [12]Department of Medicine, Bayero University and Aminu Kano Teaching Hospital, Kano, Nigeria. [13]Department of Obstetrics and Gynaecology, Bayero University and Aminu Kano Teaching Hospital, Kano, Nigeria. [14]Lakeside Healthcare at Yaxley, the Health Centre, Peterborough, United Kingdom. [15]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA. [16]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA. [17]Department of Quantitative Health Sciences, Mayo Clinic, Jacksonville, FL, USA. [18]Department of Obstetrics and Gynecology, Mayo Clinic, Rochester, MN, USA. [19]Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN, USA. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: adedinsewo.demilade@mayo.edu

on behalf of the SPEC-AI Nigeria Investigators

Demilade A. Adedinsewo[1], Andrea Carolina Morales-Lara[1], Bosede B. Afolabi[2], Oyewole A. Kushimo[3], Amam C. Mbakwem[3], Kehinde F. Ibiyemi[4], James Ayodele Ogunmodede[5], Hadijat Olaide Raji[4], Sadiq H. Ringim[6], Abdullahi A. Habib[7], Sabiu M. Hamza[6], Okechukwu S. Ogah[8], Gbolahan Obajimi[9], Olugbenga Oluseun Saanu[9], Olusoji E. Jagun[10], Francisca O. Inofomoh[11], Temitope Adeolu[11], Kamilu M. Karaye[12], Sule A. Gaya[13], Isiaka Alfa[12], Cynthia Yohanna[14], K. L. Venkatachalam[1], Jennifer Dugan[15], Xiaoxi Yao[15,16], Hanna J. Sledge[17], Patrick W. Johnson[17], Mikolaj A. Wieczorek[17], Zachi I. Attia[15], Sabrina D. Phillips[1], Mohamad H. Yamani[1], Yvonne Butler Tobah[18], Carl H. Rose[18], Emily E. Sharpe[19], Francisco Lopez-Jimenez[15], Paul A. Friedman[15], Peter A. Noseworthy[15] & Rickey E. Carter[17]

Methods

Trial design and oversight

The Screening for PEripartum Cardiomyopathies with Artificial Intelligence in Nigeria (SPEC-AI Nigeria) trial was an investigator-initiated, pragmatic, multicenter, open label, randomized clinical trial in Nigeria designed to evaluate AI-guided ECG-based screening compared to clinical ECGs alone (control) in identifying cardiomyopathy associated with LVSD among pregnant and postpartum women receiving obstetric care. The study was approved by the Mayo Clinic institutional review board as well as local ethics research committees at all participating sites in Nigeria. The trial was registered before study start (ClinicalTrials.gov registration number NCT05438576). Details of the trial design and methods have been previously published[37] and the study protocol is provided in the Supplementary Information. The CONSORT-AI Extension[38] guideline was utilized in reporting study design and results.

This trial was funded by the Mayo Clinic (Centers for Digital Health and Community Health and Engagement Research) and in part by the Mayo Clinic Building Interdisciplinary Research Careers in Women's Health (BIRCWH) Program funded by the National Institutes of Health (NIH) (grant no. K12 AR084222) and the Mayo Clinic's Center for Clinical and Translational Sciences (grant no. UL1 TR002377). Portable ECGs, phonocardiogram recordings and AI predictions using the Eko DUO digital stethoscope were extracted by the Eko Health team and sent to the coordinating center for analysis. The 12-lead ECGs were analyzed using proprietary low ejection fraction AI-ECG algorithms[13,35]. The US FDA-cleared 12-lead AI-ECG algorithm and digital stethoscope AI algorithm had in-built ECG data quality checks[35,36]. All authors participated in data interpretation and paper review for intellectual content.

Study participants

Inclusion criteria were female, aged 18–49 years, pregnant or within 12 months postpartum, and currently receiving obstetric care (prenatal and postpartum) at six participating hospitals in Nigeria: Aminu Kano Teaching Hospital, Lagos University Teaching Hospital, Olabisi Onabanjo University Teaching Hospital, Rasheed Shekoni Specialist Hospital, University College Hospital Ibadan and University of Ilorin Teaching Hospital (Supplementary Fig. 4). Study participant age and sex were confirmed by self-report at the time of screening and enrollment. All participating sites had a licensed cardiologist (at least two at each site) and echocardiography capabilities to ensure that access to a cardiovascular specialist and echocardiograms was available to all study participants at the discretion of the managing physician.

Exclusion criteria were complex congenital heart disease (single ventricle physiology or substantial intracardiac shunts with cardiac structural abnormalities), notable cardiac conduction abnormalities (including but not limited to complete heart block and or presence of a pacemaker) and inability to provide informed consent. All study participants provided written or oral informed consent in accordance with local ethics research committee approvals.

Randomization

Following informed consent, study participants were randomized in a 1:1 fashion to either AI-guided screening for cardiomyopathy (intervention arm) or a standard 12-lead clinical ECG in addition to usual care as dictated by the managing physician (control arm). Randomization was performed in real time using dynamic minimization with the study site as a stratification factor, through a web-based application (iMedidata).

Trial procedures

All study participants had a standard 12-lead ECG performed at the time of enrollment and AI predictions were generated for age and sex estimation and provided to the study teams following receipt of digital ECG files (asynchronously, and not at point of care). The control arm received standard 12-lead ECGs and AI predictions for age and sex[39] (attention control) to control for the potential benefit that getting a

clinical ECG test might introduce to the study. Those assigned to the intervention arm received in addition, portable ECGs recorded with a digital stethoscope with binary AI predictions for LVSD (positive or negative), 12-lead AI-ECG binary prediction for LVSD (positive or negative), as well as a confirmatory echocardiogram at baseline for validation of AI model performance.

The digital stethoscope was used to record 15-s ECGs and phonocardiograms in two locations on the chest: V2 (placed vertically at the left sternal border) and angled (across the left upper chest at an angle) (Supplementary Fig. 5), in addition to a 15-s ECG recording in the handheld position. AI predictions (positive or negative) for the presence of cardiomyopathy based on the digital stethoscope recording were available to the study team in real time through a mobile app. The 'maximum' prediction (any positive prediction across all recording locations) was selected for determination of the primary end point based on its alignment with standard cardiac examination procedures using a stethoscope, which involves multiple auscultation points where abnormal findings in at least one location would be sufficient for a physician to consider additional testing. As such, this approach would mirror how this tool would be used if it was deployed in clinical practice.

The 12-lead ECGs were acquired in a standard fashion as with routine clinical care, in a supine or semi-recumbent position for 10 s on standard ECG paper, at a sampling rate of 500 Hz using a GE Marquette 2000 ECG machine (GE Healthcare). Raw 12-lead ECG files were extracted from the machines in .xml formats and uploaded to a secure cloud-based file share portal. This was downloaded by the coordinating center staff, analyzed and AI prediction results were provided to the study teams subsequently, usually within 1 week or less following receipt of the ECG files.

Following the return of 12-lead AI predictions to the study teams, the site investigators were advised to obtain a repeat echocardiogram for newly positive AI results (based on the original Mayo Clinic AI Algorithm) according to the study protocol; however, given the digital stethoscope AI prediction for low ejection fraction was made available to the study investigators in real time at the point of care, obtaining a repeat echocardiogram at any time point remained at the discretion of the managing physician.

AI algorithms

The digital stethoscope and the 12-lead ECG algorithms were based on a convolutional neural network, trained on over 100,000 adults[13], and independent of all data collected in this study. The original Mayo Clinic 12-lead AI-ECG algorithm was developed to detect LVEF ≤ 35% (ref. 13) and this model has been evaluated in a clinical trial among patients receiving primary care in the United States[14]. The US FDA-cleared 12-lead AI-ECG algorithm is a modification of the original Mayo Clinic 12-lead model and was retrained to detect an LVEF < 40% (ref. 35). The US FDA-cleared model had an in-built ECG data quality check. As such ECGs deemed to be of insufficient quality did not have AI predictions generated (algorithm version lvef_v2.2.0). To be conservative when AI predictions were not generated, this was assumed to be a negative screen.

The stethoscope modifications to the original Mayo Clinic 12-lead ECG algorithm were to allow for use with a single-lead ECG at an LVEF threshold of <40% (refs. 19,20); however, before the study start, the stethoscope's AI model was further refined to incorporate data from a single-lead ECG in addition to phonocardiogram recordings for detection of LVEF < 40% and this model, previously evaluated in a pilot study among obstetric patients in the United States[16], was used in this study (algorithm version ELEFT 7.2.0). ECG and phonocardiogram recording data quality checks were also built into the stethoscope device and poor-quality recordings with unreliable AI prediction results were assumed to be negative.

Variables and echocardiograms

Demographic and clinical variables were obtained from study participants and medical charts and entered onto a secure online REDCap

database v.14.0.36. Echocardiogram images were acquired by sonographers, cardiology residents or consultant cardiologists according to each site's standard clinical and/or research practices. Images were interpreted at each hospital site by a licensed cardiologist. All echocardiogram images with a reported low ejection fraction in addition to a sample of those with normal ejection fraction were uploaded to a secure file share portal and reviewed at the coordinating center by D.A.A. In cases where ejection fraction estimates by the local site and the coordinating center were discordant or in cases where images were of poor quality to make an adequate assessment of left ventricular ejection, a repeat echocardiogram was performed at the site.

Outcomes

The primary outcome was identification of cardiomyopathy, defined as LVEF < 50% based on two-dimensional echocardiography. For the control arm, the number of participants with clinical recognition and documentation of LVSD (based on echocardiography) was used for the determination of the primary outcome, in keeping with current standard of care. The primary outcome in the intervention group was the number of participants with LVSD, as determined by a positive AI screen for LVSD (for the digital stethoscope, this was any positive prediction across all recording locations, 'maximum prediction' and for the 12-lead ECG, this was any positive prediction if more than one ECG of sufficient quality was performed at a unique study-related encounter), confirmed by echocardiography at the time of ECG acquisition. For the primary outcome in the intervention arm, a positive AI screen was required. In a few patients ($n = 4$ for the digital stethoscope and $n = 5$ for the US FDA-cleared 12-lead ECG model), either or a combination of the baseline echocardiogram and AI screening modality could not be completed on the same date as the initial study encounter, although these were subsequently acquired or repeated on a different date. To avoid bias that could result from repeated ECG/digital stethoscope acquisitions on different dates, the first ECG/digital stethoscope recording acquired for the participant were used to determine results comparing the two treatment groups. Any incomplete 'first' AI screening test or those deemed to be of poor diagnostic quality were considered negative for determination of the primary end point. As such, if the AI screen was negative, not computed due to incomplete testing on the date of the initial encounter or of insufficient quality and the protocolized echocardiogram showed LVEF < 50%, the patient case was not counted as a positive detection of LVSD.

Secondary outcomes included AI model performance across prespecified subgroups (age group, ethnicity, region, presence of hypertensive disorders and pregnancy/postpartum status) and the effectiveness of the AI intervention in identifying LVEF < 45%, <40% and ≤35% within the intervention arm at baseline. For the analysis of the AI algorithms' performances in the intervention group, ECGs/digital stethoscope recordings of sufficient quality acquired on the same date as the baseline echocardiogram were used. ECGs/digital stethoscope recordings were repeated on the date of the echocardiogram if all study-related testing could not be completed on the same date as the initial study encounter. Other prespecified (exploratory) outcomes included composite adverse outcomes (including systolic heart failure, diastolic heart failure, gestational hypertension, pre-eclampsia, eclampsia, gestational diabetes, valvular heart disease, atrial arrhythmias, sustained ventricular arrhythmias and any other reported pregnancy-related complication), composite cardiovascular outcomes (including diastolic heart failure, gestational hypertension, pre-eclampsia, eclampsia, valvular heart disease, atrial arrhythmias and sustained ventricular arrhythmias) and all-cause mortality[37].

Study participants were followed up through 12 months postpartum or study end depending on which was earlier. As much as possible, follow-up visits were timed to correspond with subsequently scheduled clinic visits or hospital encounters, and home visits were also conducted at specific sites. Given that participants were allowed to enter the study at any time point between pregnancy and 12 months

postpartum in addition to the pragmatic study design, the time contributed to the study varied based on this, with a maximum of up to seven visits for participants enrolled in the first trimester of pregnancy (Fig. 2). Details of the proposed study follow-up visits at prespecified intervals have been published[37] and the study protocol is available in the Supplementary Information. Due to the minimal risk nature of this study, a data safety monitoring board was not established.

Statistical analysis

We assumed that the prevalence of LVEF < 50% would be 4% in the intervention arm and 1% in the control group[4,37], which yielded an estimate of 848 (424 per group) women to achieve 80% power at an α of 0.05. The estimates were rounded up to 500 per group to account for uncertainties in the calculations. In April 2023, an amendment to the protocol was submitted to increase the total study sample size to 1,200 to allow for each of the six sites to enroll up to 200 participants. The decision to increase the sample size was made blinded to all clinical data collected in the study at the time of the modification. The final analysis set excluded participants who did not complete the baseline visit (for the control arm, 12-lead ECG not obtained and for the intervention arm, 12-lead ECG, at least one digital stethoscope recording and or echocardiogram was not obtained), died before baseline testing or declined to participate (withdrew consent) following randomization, resulting in the mITT analysis set.

Descriptive statistics were calculated by group for all variables to evaluate whether balance was achieved with randomization. The odds ratio and 95% large sample CI was estimated using a logistic regression model and statistical significance was assessed with a Pearson chi-squared test. Two unplanned analyses were added to the study during the peer review process. (1) A full ITT analysis was generated by assuming any excluded participants, regardless of the reason (see Fig. 1 for reasons) were negative for the detection of LVSD. Thus, the denominators were increased to $n = 616$ for each treatment arm, while the overall number of events remained the same as the mITT analysis. The ITT analysis was otherwise conducted identically to the mITT described above. (2) A logistic regression model was fit that included site as a factor in the model. For this, the Wald test for the treatment effect obtained from the logistic regression was used for assessing statistical significance of the site-adjusted models.

Within the intervention arm, patients had echocardiograms performed at baseline to validate the performance of the AI algorithm as part of the study protocol. As such, standard measures of diagnostic performance along with their respective 95% CI were reported according to the STARD criteria. Receiver operating characteristic curves were estimated. The threshold for binary classification of the AI results utilized the value determined at model development. Additional details on the statistical analysis are provided in the statistical analysis plan in the Supplementary Information.

Statistical analyses were performed using R v.4.1.2. All tests were two sided and $P < 0.05$ was considered statistically significant.

Inclusion and ethics statement

This study was jointly designed and conducted with investigators from Mayo Clinic (Jacksonville, FL and Rochester, MN, USA) and all participating institutions in Nigeria. Members of the research team include physicians and researchers with expertise in adult cardiology, obstetrics and gynecology, maternal–fetal medicine, biostatistics and data science. Key team members are part of an international research consortium born out of mutual interest in investigating and designing interventions to address adverse cardio-obstetric outcomes among women in Nigeria. Enrollment and follow-up strategies were developed and modified by Nigerian investigators at each site to align with local contexts, preferences and participant convenience. All team members discussed and agreed on specific data types collected, data ownership and publication authorship. Roles and responsibilities of team

members and all other study staff were defined and agreed upon by the team before study start. Before and during the study, education and training support was provided to local researchers in Nigeria. This included training and support for study-related devices and procedures, presentation/lectures at national cardiovascular conferences/meetings and educational symposiums in Nigeria. Given the results of the study may result in a new cardiovascular diagnosis, which may have been clinically unrecognized, plans were put in place to facilitate cardiology referral for management as appropriate. This was carried out by ensuring that each study site was led by an obstetrician/gynecologist and cardiologist as co-principal investigators. Study participants also received the results of all testing carried out along with interpretation. Previous studies conducted by local investigators in Nigeria informed the study design and these are included as citations in this paper.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The underlying data supporting the findings of this study can be made available to clinical investigators and researchers upon request. Written requests for data sharing including an analysis plan will be required before approval. These requests will be individually assessed in consultation with the study team leads and co-investigators as appropriate. If other investigators are interested in performing additional analyses, these requests can be made to the corresponding author (D.A.A.) and analyses will be performed in collaboration with the Mayo Clinic. In all cases, any data and materials to be shared will be released via a Material Transfer Agreement. Individual-level data will be available and data sharing will ensure that the rights and privacy of individuals participating in the research always remains protected. The anticipated time frame to respond to initial data requests is 1 month.

Code availability

The AI algorithms used in this paper have been previously published and have recently received US FDA clearance[35,36]. The code itself cannot be shared because it is proprietary intellectual property that has been licensed to Anumana and Eko Health. The US FDA-cleared 12-lead AI-ECG algorithm can be accessed from Anumana and the digital stethoscope AI algorithm through Eko Health.

References

37. Adedinsewo, D. A. et al. Screening for peripartum cardiomyopathies using artificial intelligence in Nigeria (SPEC-AI Nigeria): clinical trial rationale and design. *Am. Heart J.* **261**, 64–74 (2023).

38. Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat. Med.* **26**, 1364–1374 (2020).

39. Attia, Z. I. et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. *Circ. Arrhythm. Electrophysiol.* **12**, e007284 (2019).

Acknowledgements

Author contributions

D.A.A., R.E.C. and P.A.N. initially designed the trial and provided study oversight. D.A.A. and A.C.M.L. performed site monitoring, all co-authors at participating institutions performed data acquisition, interpretation and entry. R.E.C., P.W.J., M.A.W. and H.J.S. performed data analysis. All authors contributed to study conceptualization, design and execution. D.A.A. wrote the initial paper draft and all co-authors contributed substantially to the writing and revising of the paper for key intellectual content and approved the final version for publication.

Competing interests

Additional information

nature portfolio

Corresponding author(s): Demilade Adedinsewo

Last updated by author(s): Aug 1, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ECGs were recorded with commercially available devices (GE MAC 2 000 and Eko DUO Digital Stethoscope). Commercially available AI algorithms were also utilized in this study: The Eko Low Ejection Fraction Tool -ELEFT (software version - ELEFT 7.2.0), and Anumana's Low Ejection Fraction AI-ECG Algorithm (software version - lvef_v2.2.0). Both algorithms have now received FDA clearance and are classified as reduced ejection fraction machine learning-based notification software.
Data analysis	The AI algorithms listed above were used to analyze the ECG and phonocardiogram data. All other statistical analyses were performed using R version 4.1.2. Data were collected and stored using REDCap version 14.0.36.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The underlying data supporting the findings of this study can be made available to clinical investigators and researchers upon request. Written requests for data sharing including an analysis plan will be required prior to approval. These requests will be individually assessed in consultation with the study team leads and co-investigators as appropriate. If other investigators are interested in performing additional analyses, these requests can be made to the corresponding author (DAA) and analyses will be performed in collaboration with the Mayo Clinic. In all cases, any data, and materials to be shared will be released via a Material Transfer Agreement. Individual-level data will be available and data sharing will ensure that the rights and privacy of individuals participating in the research always remains protected. Anticipated time frame to respond to initial data requests is 1 month.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Information on only one sex (female) was collected as part of this study as the trial aimed to study a condition that occurs exclusively in pregnant and postpartum females.
Population characteristics	Study participants were female, aged 18-49 years, pregnant or within 12 months postpartum, and receiving obstetric care (prenatal and postpartum) at 6 participating hospitals in Nigeria
Recruitment	All enrolling sites in Nigeria identified and enrolled consecutive consenting participants from outpatient, inpatient settings, or other hospital encounter. Potential participants were approached by members of the study team at the time of a scheduled clinic/hospital encounter (following initial screening procedures for eligibility) and information regarding the study was provided. If they agree to participate, consent was obtained in accordance with study protocol. Following informed consent, randomization was performed using an online-based tool and then study related testing and documentation completed based on assigned arm. (1) All participating sites were tertiary centers (teaching hospitals with a licensed cardiologist and echocardiography capabilities) as such, LVSD prevalence seen may not be reflective of the general obstetric patient population in Nigeria. (2) All sites communicated in and utilized English consent forms and documents being the official language spoken across the country. All sites were also approved to utilize oral consent and in situations were a potential participant did not speak English, the study description was communicated by the study teams in the patient's preferred language or in pidgin English in keeping with standard procedures used in providing clinical care. Study staff were fluent in the predominant non-English language (or dialect) at their respective site (i.e., Yoruba in the Southern locations and Hausa in the Northern locations). (3) Due to cultural and religious norms, participation of women in the Northern sites was influenced by the need to obtain permissions from their husbands and those who did not give permission for their wives to be enrolled in the study were deemed ineligible as these women also did not provide informed consent in these circumstances. Other than listed above, no additional forms of selection bias is assumed to be present.
Ethics oversight	The study was approved by the Mayo Clinic institutional review board as well as local ethics research committees at all participating sites in Nigeria: Aminu Kano Teaching Hospital, Lagos University Teaching Hospital, Olabisi Onabanjo University Teaching Hospital, Rasheed Shekoni Specialist Hospital, University College Hospital Ibadan, and University of Ilorin Teaching Hospital)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We assumed the prevalence of LVEF <50% would be 4% in the intervention arm and 1% rate in the control group4,34, which yielded an estimate of 848 (424/per group) women to achieve 80% power at an alpha of 0.05. The estimates were rounded up to 500/group to account for the uncertainties in the calculations. In April 2023, an amendment to the protocol was submitted to increase the total study sample size to

	1200 to allow for each of the six sites to enroll up to 200 participants. The decision to increase the sample size was made blinded to all clinical data collected in the study at time of the modification.
Data exclusions	The final analysis set excluded participants who did not complete the baseline visit (i.e., for the control arm, ECG not completed and for the intervention arm, ECG and echocardiogram not completed) or declined to participate (withdrew consent) following randomization resulting in a modified intention to treat analysis set.
Replication	- All echocardiogram images with a reported low ejection fraction in addition to a sample of those with normal ejection fraction were uploaded to a secure file share portal and reviewed at the coordinating center by DAA. These echocardiograms were over read at the coordinating center and verified to be consistent with clinical interpretations at the enrolling site. In cases where ejection fraction estimates by the local site and the coordinating center were discordant or in cases where images were of poor quality to make an adequate assessment of left ventricular ejection, a repeat echocardiogram was performed at the site and reviewed at the coordinating center. These attempts at replication of echocardiogram results were deemed successful. - Three different AI algorithms were evaluated in this study. These algorithms have been previously developed and 2 have received US FDA approval. No new algorithms were developed as part of this study. As such, no replications or additional tests for reproducibility was performed.
Randomization	Study participants were randomized in a 1:1 fashion to either AI-guided ECG-based screening for cardiomyopathy (intervention arm) or a standard 12-lead clinical ECG in addition to usual care as dictated by the managing physician (control arm). Randomization was performed in real-time using dynamic minimization with the study site as a stratification factor, through a web-based application (iMedidata).
Blinding	No blinding was done. This was designed as an open-label, pragmatic, randomized clinical trial.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT05438576
Study protocol	Supplemental appendix and published study design manuscript available at https://pubmed.ncbi.nlm.nih.gov/36966922/
Data collection	Participants were randomized between August 15, 2022 and September 28, 2023, at 6 participating hospitals in Nigeria: Aminu Kano Teaching Hospital, Lagos University Teaching Hospital, Olabisi Onabanjo University Teaching Hospital, Rasheed Shekoni Specialist Hospital, University College Hospital Ibadan, and University of Ilorin Teaching Hospital
Outcomes	The primary study endpoint was identification of LVSD during the study period. In the intervention arm, this was the number of identified participants with LVSD, as determined by a positive AI-ECG screen, confirmed by echocardiography at the time of ECG acquisition. In the control arm, this was the number of participants with clinical recognition and documentation of LVSD confirmed by echocardiography in keeping with current standard of care. All participants in the intervention arm had a confirmatory echocardiogram at baseline for validation of AI model effectiveness; however, for the detection of LVSD at baseline, a positive AI screen was required to be considered a detection. Secondary endpoints were evaluation of the primary endpoint within prespecified subgroups, and effectiveness of the AI-ECG at various LVEF cut points in the intervention arm