medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Using Artificial Intelligence to Predict Heart Failure Risk from Single-lead Electrocardiographic Signals: A Multinational Assessment

Short Title: AI-enabled HF Risk Prediction Using Single-lead ECG

Lovedeep S Dhingra MBBS[1]*, Arya Aminorroaya MD, MPH[1]*, Aline Pedroso Camargos PhD[1], Akshay Khunte BS[2], Veer Sangha BS[1,3], Daniel McIntyre MD, PhD[4], Clara K Chow MBBS, PhD[4,5], Folkert W Asselbergs MD, PhD[6,7,8], Luisa CC Brant MD, PhD[9,10], Sandhi M Barreto MD, PhD[11], Antonio Luiz P Ribeiro MD, PhD[9,10], Harlan M Krumholz MD, SM[1,12,13], Evangelos K Oikonomou MD, DPhil[1], Rohan Khera MD, MS[1,12,14,15]

[1]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA
[2]Department of Computer Science, Yale University, New Haven, CT, USA
[3]Department of Engineering Science, University of Oxford, Oxford, UK
[4]Westmead Applied Research Centre, Faculty of Medicine and Health, The University of Sydney, Westmead, Australia
[5]Department of Cardiology, Westmead Hospital, Sydney, Australia
[6]Department of Cardiology, Amsterdam Cardiovascular Sciences, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands
[7]Institute of Health Informatics, University College London, London, UK
[8]The National Institute for Health Research University College London Hospitals Biomedical Research Centre, University College London, London, UK
[9]Department of Internal Medicine, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
[10]Telehealth Center and Cardiology Service, Hospital das Clínicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
[11]Department of Preventive Medicine, School of Medicine, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
[12]Center for Outcomes Research and Evaluation (CORE), Yale New Haven Hospital, New Haven, CT, USA
[13]Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA
[14]Section of Biomedical Informatics and Data Science, Yale School of Medicine, New Haven, CT, USA
[15]Section of Health Informatics, Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA

***** Contributed equally as co-first authors

Correspondence to:
Rohan Khera, MD, MS
195 Church Street, 6th Floor, New Haven, CT 06510
(203) 764-5885; rohan.khera@yale.edu; @rohan_khera

Word Count: 3,381

Keywords: Heart failure, Deep learning, Wearable Devices, Electrocardiograms, Risk Stratification, Predictive Modelling

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

KEY POINTS

Question: Can single-lead electrocardiogram (ECG) tracings predict heart failure (HF) risk?

Findings: We evaluated a noise-adapted artificial intelligence (AI) algorithm for single-lead ECGs as the sole input across multinational cohorts, spanning a diverse integrated US health system and large community-based cohorts in the UK and Brazil. A positive AI-ECG screen was associated with a 3- to 7-fold higher HF risk, independent of age, sex, and comorbidities. The AI model achieved incremental discrimination and improved reclassification for HF over the pooled cohort equations to prevent HF (PCP-HF).

Meaning: A noise-adapted AI model for single-lead ECG predicted the risk of new-onset HF, representing a scalable HF risk-stratification strategy for portable and wearable devices.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ABSTRACT

Importance: Despite the availability of disease-modifying therapies, scalable strategies for heart failure (HF) risk stratification remain elusive. Portable devices capable of recording single-lead electrocardiograms (ECGs) can enable large-scale community-based risk assessment.

Objective: To evaluate an artificial intelligence (AI) algorithm to predict HF risk from noisy single-lead ECGs.

Design: Multicohort study.

Setting: Retrospective cohort of individuals with outpatient ECGs in the integrated Yale New Haven Health System (YNHHS) and prospective population-based cohorts of UK Biobank (UKB) and Brazilian Longitudinal Study of Adult Health (ELSA-Brasil).

Participants: Individuals without HF at baseline.

Exposures: AI-ECG-defined risk of left ventricular systolic dysfunction (LVSD).

Main Outcomes and Measures: Among individuals with ECGs, we isolated lead I ECGs and deployed a noise-adapted AI-ECG model trained to identify LVSD. We evaluated the association of the model probability with new-onset HF, defined as the first HF hospitalization. We compared the discrimination of AI-ECG against the pooled cohort equations to prevent HF (PCP-HF) score for new-onset HF using Harrel's C-statistic, integrated discrimination improvement (IDI), and net reclassification improvement (NRI).

Results: There were 194,340 YNHHS patients (age 56 years [IQR, 41-69], 112,082 women [58%]), 42,741 UKB participants (65 years [59-71], 21,795 women [52%]), and 13,454 ELSA-Brasil participants (56 years [41-69], 7,348 women [55%]) with baseline ECGs. A total of 3,929 developed HF in YNHHS over 4.5 years (2.6-6.6), 46 in UKB over 3.1 years (2.1-4.5), and 31 in ELSA-Brasil over 4.2 years (3.7-4.5). A

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

positive AI-ECG screen was associated with a 3- to 7-fold higher risk for HF, and each 0.1 increment in the model probability portended a 27-65% higher hazard across cohorts, independent of age, sex, comorbidities, and competing risk of death. AI-ECG's discrimination for new-onset HF was 0.725 in YNHHS, 0.792 in UKB, and 0.833 in ELSA-Brasil. Across cohorts, incorporating AI-ECG predictions in addition to PCP-HF resulted in improved Harrel's C-statistic (Δ=0.112-0.114), with an IDI of 0.078-0.238 and an NRI of 20.1%-48.8% for AI-ECG vs. PCP-HF.

Conclusions and Relevance: Across multinational cohorts, a noise-adapted AI model with lead I ECGs as the sole input defined HF risk, representing a scalable portable and wearable device-based HF risk-stratification strategy.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

BACKGROUND

Accessible strategies for risk stratification for heart failure (HF) remain elusive despite the availability of evidence-based therapies that can effectively modify the trajectory of the disease.[1,2] Clinical scores to predict HF risk, including the pooled cohort equations to prevent HF (PCP-HF) and the Health ABC score,[3,4] require extensive clinical evaluation, including detailed history and physical exam, a 12-lead electrocardiogram (ECG), and other specialized testing.[3–8] These complex inputs limit their use, systematically excluding those without access to healthcare and preventing wider deployment for community-based screening.[7–9] Similarly, serum-based biomarkers such as N-terminal pro–B-type natriuretic peptide and high-sensitivity cardiac troponin, which portend a higher risk of HF when elevated, are limited by the need for blood draws, the need for storing blood samples, and frequent inaccessibility at the point-of-contact.[10–15] Thus, there is an unmet need for a simple, efficient, and scalable strategy for identifying those at the highest risk for developing HF in the community.

Given their increasing utility and ubiquity, portable and wearable devices capable of recording single-lead ECG tracings have been proposed as a platform for cardiovascular monitoring and screening.[16–19] Further, artificial intelligence (AI)-enhanced interpretation of ECGs (AI-ECG) has been shown to detect hidden cardiovascular disease signatures from single-lead ECGs, highlighting the potential of these devices for improving cardiovascular health.[20–24] However, these portable ECGs are prone to the introduction of noise during acquisition, which can limit the performance of AI models unless specialized measures are taken to ensure they are resilient to such noise.[20,25] Recently, we reported a novel approach for single-lead ECGs that incorporates random noising during model development, enabling the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

model to maintain consistent diagnostic performance across varying levels of real-world noises.[20] This initial deployment of our approach focused on developing a model that detects reduced left ventricular ejection fraction (LVEF) on single-lead ECG based on information from a concurrent echocardiogram, with the potential application of identifying subclinical left ventricular systolic dysfunction (LVSD). Recent studies also suggest that the AI-ECG signature for LVSD identifies other subtle markers of LV dysfunction, including abnormal LV strain and diastolic function, especially among those with a positive screen but preserved LVEF.[26] Of note, each of these is known to be associated with elevated HF risk.[27]

Given the increasing accessibility of single-lead ECGs, we tested the hypothesis that an AI model developed to detect the cross-sectional signature of LVSD from single-lead ECG tracings can identify the risk of developing HF. We longitudinally evaluated our approach in a cohort of individuals undergoing outpatient ECGs within a diverse US health system and two large population-based cohort studies in the UK and Brazil.

METHODS

Data Sources

We included three large cohorts spanning different countries and settings who had undergone an ECG: (i) individuals seeking outpatient care in the Yale New Haven Health System (YNHHS), a large and diverse healthcare system in the Northeastern US, including five independent hospitals and an outpatient medical network, (ii) participants in the UK Biobank, a nationwide UK-based cohort study, and (iii) participants in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil), the largest community-based cohort study from Brazil. While YNHHS included testing

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

and follow-up as a part of routine clinical care in an integrated health system, participants in UKB and ELSA-Brasil had detailed protocolized evaluation at baseline and comprehensive longitudinal follow-up. An overview of all data sources is included in the **eMethods**.

Study Population

In YNHHS, to approximate a cohort resembling a screening setting, we identified patients undergoing a 12-lead ECG in an outpatient health encounter during 2014-2023 without HF before the ECG. To account for the ECGs potentially being obtained as a part of a workup for HF, we included a 1-year blanking period from the first recorded encounter in the electronic health records (EHR) to identify those with prevalent HF (**eMethods; eFigure 1**). In YNHHS, 255,604 individuals had at least one outpatient ECG after the blanking period. We excluded 47,720 patients in the original model development population and 11,954 patients with prevalent HF. To further focus the cohort on those without established precursors of HF, we excluded 1,590 patients with LV dysfunction (LVEF under 50% or moderate/severe LV diastolic dysfunction) before the index ECG (**eFigure 2**).

In UKB, we included 42,366 participants who had undergone a 12-lead ECG during their imaging study visit in 2014-2020. We used the linkage of UKB with the EHR from the UK National Health Service to exclude 225 participants who had been hospitalized with a principal or secondary discharge diagnosis of HF before the ECG. In ELSA-Brasil, we included 13,739 participants who had undergone a 12-lead ECG during 2008-2010, excluding those with an HF diagnosis (N=227) or with an LVEF under 50% (N=58) on their baseline echocardiogram (**eFigure 2**).

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Study Outcomes

We defined the study outcome as new-onset HF characterized by HF hospitalizations. In YNHHS, this was defined as a hospitalization with an International Classification of Disease Tenth Revision – Clinical Modification (ICD-10-CM) code for HF as the principal discharge diagnosis (**eTable 1**). This approach was guided by the over 95% specificity of HF diagnosis codes, especially as the principal discharge diagnosis for a clinical diagnosis of HF.[28] Similarly, in UKB, we used the linked EHR to identify hospitalizations with HF as the principal diagnosis code. In ELSA-Brasil, HF was identified by in-person interview or telephonic surveillance for all hospitalizations, followed by independent medical record review and adjudication of HF hospitalizations by two cardiologists (**eMethods**).[29]

We further evaluated the association of AI-ECG probabilities with alternate definitions of new-onset HF, and composite outcomes, including (i) any hospitalization with a principal or secondary diagnosis code for HF, (ii) a subsequent echocardiogram with an LVEF under 50%, and (iii) a composite outcome of HF or all-cause death (**eMethods**). To evaluate the specificity of the AI-ECG-defined HF risk, we examined the risk of other cardiovascular conditions, including acute myocardial infarction (AMI), stroke hospitalizations, and all-cause mortality (**eTable 1**). A composite outcome of major adverse cardiovascular events (MACE) was defined as death or hospitalization for HF, AMI, or stroke.

Study Exposure

We defined the study exposure as the output probability of an AI-ECG model trained to detect a concurrent LVEF of less than 40% on lead I of a 12-lead ECG, representing the lead commonly captured by portable ECG devices.[20] This was

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

developed at the Yale New Haven Hospital (YNHH) using a novel approach of augmenting training data with random Gaussian noise (**eMethods**). The model achieved excellent discrimination (area under the receiver operating characteristic curve of 0.899 [95% CI, 0.888-0.909]; **eTable 2**) for detecting concurrent reduced LVEF across various noisy ECGs.

We deployed this established model without further development to lead I ECG signals across included individuals to obtain the LVSD probability, representing a continuous HF risk score. We defined a positive AI-ECG screen as a model output probability greater than 0.08, representing the probability threshold at which the model achieved a sensitivity of 90% for detecting LVSD during internal validation.[20] We further defined graded thresholds based on AI-ECG probabilities of 0-0.2, 0.2-0.4, 0.4-0.6, 0.6-0.8, and 0.8-1 to evaluate the association of a higher risk score with HF. Notably, while the model was developed for detecting the cross-sectional signature of LVSD using data from the YNHH alone, it was applied across all YNHHS sites and the population-based cohorts without any further development or fine-tuning for prediction of HF risk.

Study Covariates

We identified available demographic characteristics across cohorts, including age at the time of ECG, sex, and self-reported race and ethnicity. Comorbidities, including ischemic heart disease, hypertension, and type 2 diabetes mellitus, were defined using relevant EHR diagnosis codes in YNHHS and UKB (**eTable 1**). Obesity was defined as BMI ≥30 kg/m^2. In ELSA-Brasil, covariates were recorded at the baseline study visit.[30] Race was self-classified based on Brazil's National Bureau of Statistics definition and classified as White, Black, "Pardo", Asian, or Others.[30,31]

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Study Comparator

The PCP-HF score was used as the study comparator for predicting HF risk, representing a complex clinical risk model developed and validated with data from seven community-based cohorts.[3] The PCP-HF risk score uses demographics (age, sex, race), physical exam-based features (smoking status, BMI, systolic blood pressure), laboratory measurements (total cholesterol, high-density cholesterol, fasting blood glucose), medication history (antihypertensive medication use, antihyperglycemic medication use), and electrocardiographically-defined QRS duration. Across the study cohorts, these input features were determined using the EHR and/or study visits (**eMethods**).[32–35]

Statistical Analysis

We used age-, sex-, and comorbidity-adjusted Cox proportional hazard models with time-to-first HF event as the dependent variable and the AI-ECG-based screen results (positive/negative) or continuous model probability as the independent variable to evaluate the association of the model output with HF risk. Multi-outcome Fine-Gray subdistribution hazard models were used to account for the competing risk of death.[36]

The incremental discrimination of AI-ECG over PCP-HF for predicting time-to-HF hospitalization was reported as the difference in Harrel's C-statistics and 95% confidence interval (CI) using a one-shot nonparametric approach.[37] We calculated integrated discrimination improvement (IDI), categorical and continuous time-to-event net reclassification improvement (NRI),[38] and compared the net benefit of the AI-ECG model and the PCP-HF score across all probability thresholds, with further

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

details in **eMethods**.[39] We also evaluated the association of AI-ECG with HF, overall and across racial groups, after adjusting for the PCP-HF score in Cox models.

RESULTS

Study Population

From YNHHS, we included 194,340 individuals with a median age of 56 years (IQR, 41-69), comprising 112,082 (57.7%) women, 119,051 (61.3%) non-Hispanic White, 30,840 (15.9%) non-Hispanic Black, and 33,456 (17.2%) Hispanic. Over a median 4.5-year follow-up (IQR, 2.6-6.6), 3,929 (2.0%) had an HF hospitalization, 7,847 (4.0%) had an HF hospitalization or an LVEF below 50% on subsequent echocardiogram, and 10,747 (5.5%) died (**eTables 3-4**).

The 42,141 UKB participants had a median age of 65 years (IQR, 59-71), including 21,795 (51.7%) women, with 40,691 (96.6%) identifying as White and 304 (0.7%) as Black. Over a median follow-up of 3.1 years (IQR, 2.1-4.5), 46 (0.1%) had an HF hospitalization, and 346 (0.8%) died (**eTable 3**).

From ELSA-Brasil, the 13,454 participants had a median age of 51 years (IQR, 45-58), comprising 7,348 (54.6%) women, 6,920 (51.4%) adults identifying as White, 2,130 (15.8%) as Black, and 3,767 (28.0%) as “Pardo”. Over a median of 4.2 years (IQR, 3.7-4.5), 31 (0.2%) people developed HF, and 229 (1.7%) died.

Risk Stratification for New-Onset HF

In YNHHS, 43,559 (22.4%) patients screened positive on the AI model applied to the baseline single-lead ECG signal. A positive screen was associated with over 5-fold higher risk of developing HF (HR 5.07 [95% CI, 4.76-5.40]; **Table 1**). After accounting for differences in age and sex, a positive AI-ECG screen was associated

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

with a 3.3-fold higher risk of HF compared with a negative screen (adjusted HR [aHR], 3.32 [95% CI, 3.11-3.54]). The association remained statistically significant after accounting for differences in HF risk factors of prior ischemic heart disease, hypertension, type 2 diabetes, and obesity (aHR 2.81 [95% CI, 2.64-3.01]) and after additionally accounting for the competing risk of death (aHR of 2.73 [95% CI, 2.55-2.93]) (**Table 1**). The association of a positive screen with an elevated risk of HF was noted across YNHHS sites (**eTable 5**), demographic subgroups (**eTable 6**), and different HF definitions (**eTables 7-8**).

In UKB, 5,513 (13.1%) participants screened positive with the AI-ECG model. A positive AI-ECG screen portended a 7.5-fold higher hazard for developing HF (HR 7.52 [95% CI, 4.21-13.41]). After accounting for age, sex, HF risk factors, and the competing risk of death, screen-positive participants had a 5-fold higher risk of HF (aHR 5.02 [95% CI, 2.77-9.09]; **Table 1**).

In the ELSA-Brasil cohort, 1,928 (14.3%) participants had a positive AI-ECG screen, with a 9-fold higher risk for HF (age- and sex-adjusted HR 8.74 [95% CI, 4.13-18.48]) compared with screen-negative participants. This association was consistent even after accounting for the comorbidities and the competing risk of death (aHR 7.71 [95% CI, 3.62-16.46]; **Table 1**).

Risk Across Model Probability Increments

Across the YNHHS network, each 0.1 increment in the model output probability portended a 28% higher hazard of developing HF, adjusted for age, sex, comorbidities, and accounting for the competing risk of death (aHR 1.28 [95% CI, 1.26-1.30]; **Table 1**). Higher model probabilities were progressively associated with a higher risk of future HF across various probability bins, with consistent patterns

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

across the individual hospitals and the outpatient medical network (**Figure 2**, **eTables 5, 9-10**).

Across both UKB and ELSA-Brasil cohorts, a 0.1 increment in model probability was associated with 51% and 66% higher adjusted risk of HF (aHR 1.49 [95% CI, 1.36-1.63] and aHR 1.65 [95% CI, 1.46-1.87], respectively, **Table 1**), respectively. Model output probabilities of over 0.8 portended age- and sex-adjusted HRs of 13.68 (1.86-100.45) and 112.33 (25.75-489.96), at UKB and ELSA-Brasil, respectively, compared with those with probabilities under 0.2 (**Figure 2**, **eTable 9**).

Comparison with PCP-HF

The AI-ECG model had a discrimination based on Harrel's C-statistic of 0.725 (95% CI, 0.704-0.746) in YNHHS, compared with 0.634 (95% CI, 0.612-0.656) for PCP-HF score (p<0.001; **Table 2**). In UKB and ELSA-Brasil, the AI-ECG model's discrimination for HF was 0.792 (95% CI, 0.696-0.889) and 0.833 (95% CI, 0.720-0.946), respectively, which was not significantly different from PCP-HF (UKB: p=0.45; ELSA-Brasil: p=0.87). Across all cohorts, incorporating model probability with PCP-HF yielded a statistically significant improvement in discrimination over the use of PCP-HF alone (ΔC-statistic in YNHHS: 0.112 [95% CI, 0.091-0.134]; UKB: 0.114 [95% CI, 0.027-0.200]; ELSA-Brasil: 0.114 [95% CI, 0.034-0.194]; **Table 2**).

Compared with the PCP-HF score, the AI model had a positive IDI across all study cohorts, with an IDI of 0.078 (95% CI, 0.003-0.160) for YNHHS, 0.165 (95% CI, 0.075-0.252) for UKB, and 0.238 (95% CI, 0.107-0.369) for ELSA-Brasil (**Table 3**). The categorical NRI for AI-ECG compared with the PCP-HF score over the 0.08 threshold was 20.1% (95% CI, 1.48%-42.2%) in YNHHS, 29.1% (95% CI, 9.18%-49.4%) in UKB, and 48.8% (95% CI, 18.8%-49.4%) in ELSA-Brasil. The continuous

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

NRI was 44.2% (95% CI, 11.1%-77.9%) higher for AI-ECG than PCP-HF in YNHHS. Although continuous NRI was also numerically higher in UKB (15.5% [95% CI, -21.1% to 47.8%]) and ELSA-Brasil (7.7% [95% CI, -23.8% to 32.4%]), the differences were not statistically significant (**Table 3**). The AI-ECG model demonstrated consistent superior net benefit over PCP-HF across all probability thresholds in UKB and ELSA-Brasil. In YNHHS, the net benefit of AI-ECG was higher than that of PCP-HF in probability thresholds greater than 0.04, where the AI-ECG threshold is 0.08 (**eFigure 3**).

A positive AI-ECG screen was an independent predictor of HF risk after accounting for PCP-HF score, with a 3-fold hazard (aHR 3.29 [95% CI, 2.88-3.77]) in YNHHS, and 7- to 11-fold hazards in UKB (7.16 [95% CI, 3.67-13.98]) and ELSA-Brasil (11.27 [95% CI, 4.53-28.04]; **eFigure 4**), respectively, with consistent patterns across racial groups (**eFigure 5**).

Non-HF Cardiovascular Outcome Prediction

In YNHHS, a positive AI-ECG screen was associated with a modestly elevated risk of stroke and MACE (age- and sex-adjusted HRs: stroke, 1.17; MACE, 1.77; **eTables 11-12**) compared with a 3-fold increase in HF risk. In UKB and ELSA-Brasil, a positive screen portended a 1.5- to 4-fold hazard of stroke, death, and MACE compared with a 6- to 9-fold increase in HF risk (**eTable 11**).

DISCUSSION

Across clinically and geographically distinct cohorts, a noise-adapted deep learning model, trained to detect cross-sectional LVSD from only a lead I ECG, predicted the risk of future HF among individuals seeking outpatient care and community-dwelling

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

adults. Individuals with a positive AI-ECG screen had a 3- to 7-fold higher risk of developing HF compared with those with a negative screen, independent of known demographic and clinical risk factors. Higher AI-ECG probabilities were progressively associated with a higher risk of HF, with each 10% increment portending a 27-65% higher risk-adjusted hazard for HF across all cohorts. Further, the AI-ECG model demonstrated incremental discrimination, improved reclassification, and superior net benefit over the PCP-HF score. Therefore, our AI-based approach presents ideal characteristics for use as a non-invasive digital biomarker for elevated HF risk using a single-lead ECG.

Applications of deep learning for ECGs have demonstrated the ability to identify subtle signatures of structural heart disorders previously considered electrically silent,[40–48] with applications extending to detecting LVSD from single-lead tracings.[20,49–51] Recently, the US Food and Drug Administration also granted clearance to an AI tool using electronic stethoscope-based single-lead ECGs for cross-sectional LVSD detection.[52] Our study demonstrates that a noise-adapted AI-ECG model can predict the risk of new-onset HF using single-lead ECG tracings. Given the increasing accessibility of portable and wearable devices capable of acquiring ECG signals outside a clinical setting,[19,53,54] this approach can be applied widely to identify individuals at a high risk of HF.[54] While the ECGs acquired with these devices are often distorted by positioning and movement of electrodes or artifacts due to skeletal muscle contraction during acquisition,[25,55] our unique noise-adapted training approach can ensure reliable inference from these noisy ECGs.[20]

In this study, we opted for a specific definition of HF based on the principal discharge diagnosis code, a criterion that has demonstrated high specificity.[28] Nonetheless, the association of a positive screen with an elevated HF risk was

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

consistent across several sensitivity analyses that defined the condition differently in YNHHS and UKB, and in ELSA-Brasil where the outcomes were explicitly adjudicated. The robust performance across clinically and demographically distinct cohorts indicates that the model captures a predictive signature of HF across a spectrum of disease phenotypes independent of site-specific coding practices.[56,57] Moreover, the dose-dependent association of higher AI-ECG scores with progressively elevated HF risk enables graded risk stratification and risk-informed management. Notably, while a positive screen was also associated with a modestly elevated risk of other cardiovascular outcomes, including MACE, the predictive signature identified by the model was more specific for HF.

Our study has important implications for defining HF risk. While several clinical risk score- and serum biomarker-based strategies have been proposed to identify those at the highest risk for developing HF, these strategies often require extensive clinical evaluation and blood testing.[7,8,15] This limits their scope to patients with established access to healthcare services, preventing community-based application.[8,15,58,59] In contrast, our AI-based approach using single-lead ECG tracings can enable HF risk stratification outside clinical settings. The ability to use a single portable device to record ECGs for multiple individuals allows for the design of innovative and efficient community-based screening programs.[60,61] Successful health promotion strategies, such as targeted hypertension management in barbershops and cancer screening in churches across the US,[62,63] can be adapted to promote HF screening, especially among individuals who are traditionally less likely to seek preventive medical care.[58] The ease of use and the brief time required for ECG acquisition with these devices can enable a non-laboratory-based strategy, with the potential for integration into national-level non-communicable disease screening

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

programs globally, especially in low- and middle-income countries.[60,61,64] This scalability and potential community health benefits necessitate prospective clinical and cost-effectiveness assessments for AI-based HF risk stratification using portable and wearable devices.

Our study has certain limitations. First, waveforms extracted from the lead I of clinical ECGs may not be identical to those captured by portable devices. While our noise-augmentation approach has previously demonstrated sustained performance during testing on ECGs with real-world noises,[20] prospective validation of the model on portable- and wearable-acquired ECGs is necessary before deployment for HF screening in a community setting. Second, despite the wide geographic coverage of YNHHS, out-of-hospital clinical outcomes may not have been captured, thereby representing a lower risk of HF and other cardiovascular outcomes compared with the protocolized follow-up in UKB and ELSA-Brasil. Moreover, while we included only ECGs performed in an outpatient setting, the patients who underwent ECG testing were clinically selected, indicating an unmeasured potential risk profile of those who underwent a clinical ECG but had a negative AI-ECG screen. However, the controls in this setting all underwent ECG screens as well. Finally, while the approach offers an accessible strategy for identifying individuals at elevated risk of HF, it is not clear if this risk is modifiable. Nonetheless, a robust screening strategy can enable targeted management of known HF risk factors.

CONCLUSION

Across clinically and geographically distinct cohorts, we used a noise-resilient AI model with a lead I ECG tracing as the sole input to define the risk of future HF, with value over conventional risk scores in both performance and ease of deployment.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

With the increasing availability of single-lead ECGs on portable and wearable devices, this AI-ECG-based non-invasive digital biomarker can enable scalable stratification of HF risk across communities.

ACKNOWLEDGMENTS

Author Contributions

Drs Dhingra, Aminorroaya, and Khera had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version for submission.

Study concept and design: Dhingra, Lovedeep S; Aminorroaya, Arya; Oikonomou, Evangelos K; Khera, Rohan.

Acquisition, analysis, or interpretation of data: Dhingra, Lovedeep S; Aminorroaya, Arya; Pedroso Camargos, Aline; Khunte, Akshay; Sangha, Veer; McIntyre, Daniel; Chow, Clara K; Asselbergs, Folkert W; Brant, Luisa CC; M Barreto, Sandhi; Ribeiro, Antonio Luiz P; Krumholz, Harlan M; Oikonomou, Evangelos K; Khera, Rohan

Drafting of the manuscript: Dhingra, Lovedeep S; Aminorroaya, Arya; Pedroso Camargos, Aline.

Critical revision of the manuscript for important intellectual content: Khunte, Akshay; Sangha, Veer; McIntyre, Daniel; Chow, Clara K; Asselbergs, Folkert W; Brant, Luisa CC; M Barreto, Sandhi; Ribeiro, Antonio Luiz P; Krumholz, Harlan M; Oikonomou, Evangelos K; Khera, Rohan.

Statistical analysis: Dhingra, Lovedeep S; Aminorroaya, Arya.

Obtained funding: Khera, Rohan.

Administrative, technical, or material support: Aminorroaya, Arya; Dhingra, Lovedeep S; Pedroso Camargos, Aline; Khera, Rohan.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Study supervision: Khera, Rohan.

Conflict of Interest Disclosures

Dr. Khera is an Associate Editor of JAMA. Dr. Khera and Mr. Sangha are the coinventors of U.S. Provisional Patent Application No. 63/346,610, "Articles and methods for format-independent detection of hidden cardiovascular disease from printed electrocardiographic images using deep learning" and are co-founders of Ensight-AI. Dr. Khera receives support from the National Heart, Lung, and Blood Institute of the National Institutes of Health (under awards R01HL167858 and K23HL153775) and the Doris Duke Charitable Foundation (under award 2022060). He receives support from the Blavatnik Foundation through the Blavatnik Fund for Innovation at Yale. He also receives research support, through Yale, from Bristol-Myers Squibb, BridgeBio, and Novo Nordisk. In addition to 63/346,610, Dr. Khera is a coinventor of U.S. Pending Patent Applications WO2023230345A1, US20220336048A1, 63/484,426, 63/508,315, 63/580,137, 63/606,203, 63/619,241, and 63/562,335. Dr. Khera and Dr. Oikonomou are co-founders of Evidence2Health, a precision health platform to improve evidence-based cardiovascular care. Dr. Oikonomou is a co-inventor of the U.S. Patent Applications 63/508,315 & 63/177,117 and has been a consultant to Caristo Diagnostics Ltd (all outside the current work). Dr. Krumholz works under contract with the Centers for Medicare & Medicaid Services to support quality measurement programs. He is associated with research contracts through Yale University from Janssen, Kenvue, and Pfizer. In the past three years, Dr. Krumholz received options for Element Science and Identifeye and payments from F-Prime for advisory roles. He is a co-founder of and holds equity in Hugo Health, Refactor Health, and Ensight-AI. Dr. Ribeiro is supported in part by the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

National Council for Scientific and Technological Development - CNPq (grants 465518/2014-1, 310790/2021-2, 409604/2022-4 e 445011/2023-8). Dr. Brant is supported in part by CNPq (307329/2022-4). Dr. Asselbergs is supported by Heart4Data, which received funding from the Dutch Heart Foundation and ZonMw (2021-B015), and UCL Hospitals NIHR Biomedical Research Centre.

Funding/Support

Dr. Khera was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (under awards R01HL167858 and K23HL153775) and the Doris Duke Charitable Foundation (under award 2022060). Dr. Oikonomou was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (under award 1F32HL170592).

Role of the Funder/Sponsor

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer

The views expressed in this article are those of the authors and not necessarily any funders.

Data Sharing Statement

The analyzed de-identified data are available for the UK Biobank cohort from the UK Biobank's Access Management System, and for the ELSA-Brasil cohort upon a

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

reasonable request to the corresponding author. Individual-level data for the YNHHS cohort cannot be made available due to HIPAA regulations enforced by the Yale IRB. Additional supporting information (statistical/analytic code) is available upon request to the corresponding author.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

REFERENCES

1. Shahim B, Kapelios CJ, Savarese G, Lund LH. Global public health burden of heart failure: An updated review. *Card Fail Rev*. 2023;9. doi:10.15420/cfr.2023.05

2. Bozkurt B, Ahmad T, Alexander KM, et al. Heart failure epidemiology and outcomes statistics: A report of the heart failure society of America. *J Card Fail*. 2023;29(10):1412-1451.

3. Khan SS, Ning H, Shah SJ, et al. 10-Year Risk Equations for Incident Heart Failure in the General Population. *J Am Coll Cardiol*. 2019;73(19):2388-2397.

4. Butler J, Kalogeropoulos A, Georgiopoulou V, et al. Incident Heart Failure Prediction in the Elderly. *Circ Heart Fail*. 2008;1(2):125-133.

5. Pandey A, Khan MS, Patel KV, Bhatt DL, Verma S. Predicting and preventing heart failure in type 2 diabetes. *Lancet Diabetes Endocrinol*. 2023;11(8):607-624.

6. Kalogeropoulos A, Psaty BM, Vasan RS, et al. Validation of the Health ABC Heart Failure Model for Incident Heart Failure Risk Prediction. *Circ Heart Fail*. 2010;3(4):495-502.

7. Echouffo-Tcheugui JB, Greene SJ, Papadimitriou L, et al. Population Risk Prediction Models for Incident Heart Failure. *Circ Heart Fail*. 2015;8(3):438-447.

8. Nadarajah R, Younsi T, Romer E, et al. Prediction models for heart failure in the community: A systematic review and meta-analysis. *Eur J Heart Fail*. 2023;25(10):1724-1738.

9. Agarwal SK, Chambless LE, Ballantyne CM, et al. Prediction of Incident Heart Failure in General Practice. *Circ Heart Fail*. 2012;5(4):422-429.

10. Grewal J, McKelvie R, Lonn E, et al. BNP and NT-proBNP predict echocardiographic severity of diastolic dysfunction. *Eur J Heart Fail*. 2008;10(3):252-259.

11. Campbell DJ, Gong FF, Jelinek MV, et al. Prediction of incident heart failure by serum amino-terminal pro-B-type natriuretic peptide level in a community-based cohort. *Eur J Heart Fail*. 2019;21(4):449-459.

12. Yan I, Börschel CS, Neumann JT, et al. High-sensitivity cardiac troponin I levels and prediction of heart failure: Results from the BiomarCaRE consortium. *JACC Heart Fail*. 2020;8(5):401-411.

13. Kelder JC, Cowie MR, McDonagh TA, et al. Quantifying the added value of BNP in suspected heart failure in general practice: an individual patient data meta-analysis. *Heart*. 2011;97(12):959-963.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

14. Jia X, Al Rifai M, Hoogeveen R, et al. Association of Long-term Change in N-Terminal Pro-B-Type Natriuretic Peptide With Incident Heart Failure and Death. *JAMA Cardiol*. 2023;8(3):222-230.

15. Watson CJ, Gallagher J, Wilkinson M, et al. Biomarker profiling for risk of future heart failure (HFpEF) development. *J Transl Med*. 2021;19(1):61.

16. Al-Alusi MA, Khurshid S, Wang X, et al. Trends in Consumer Wearable Devices With Cardiac Sensors in a Primary Care Cohort. *Circ Cardiovasc Qual Outcomes*. 2022;15(7):e008833.

17. Gehr S, Russmann C. Shaping the future of cardiovascular medicine in the new era of wearable devices. *Nat Rev Cardiol*. 2022;19(8):501-502.

18. Aminorroaya Arya, Dhingra Lovedeep S., Nargesi Arash A., Oikonomou Evangelos K., Krumholz Harlan M., Khera Rohan. Use of Smart Devices to Track Cardiovascular Health Goals in the United States. *JACC: Advances*. 2023;2(7):100544.

19. Dhingra LS, Aminorroaya A, Oikonomou EK, et al. Use of Wearable Devices in Individuals With or at Risk for Cardiovascular Disease in the US, 2019 to 2020. *JAMA Netw Open*. 2023;6(6):e2316634.

20. Khunte A, Sangha V, Oikonomou EK, et al. Detection of left ventricular systolic dysfunction from single-lead electrocardiography adapted for portable and wearable devices. *NPJ Digit Med*. 2023;6(1):124.

21. Shankar SV, Oikonomou EK, Khera R. CarDS-Plus ECG Platform: Development and Feasibility Evaluation of a Multiplatform Artificial Intelligence Toolkit for Portable and Wearable Device Electrocardiograms. *medRxiv*. Published online October 3, 2023. doi:10.1101/2023.10.02.23296404

22. Mathews SM, Kambhamettu C, Barner KE. A novel application of deep learning for single-lead ECG classification. *Comput Biol Med*. 2018;99:53-62.

23. Dupulthys S, Dujardin K, Anné W, et al. Single-lead electrocardiogram Artificial Intelligence model with risk factors detects atrial fibrillation during sinus rhythm. *Europace*. 2024;26(2):euad354.

24. Gibson CM, Mehta S, Ceschim MRS, et al. Evolution of single-lead ECG for STEMI detection using a deep learning approach. *Int J Cardiol*. 2022;346:47-52.

25. Bläsing D, Buder A, Reiser JE, Nisser M, Derlien S, Vollmer M. ECG performance in simultaneous recordings of five wearable devices using a new morphological noise-to-signal index and Smith-Waterman-based RR interval comparisons. *PLoS One*. 2022;17(10):e0274994.

26. Oikonomou EK, Sangha V, Dhingra LS, et al. Artificial intelligence-enhanced risk stratification of cancer therapeutics-related cardiac dysfunction using electrocardiographic images. *bioRxiv*. Published online March 15, 2024:2024.03.12.24304047. doi:10.1101/2024.03.12.24304047

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

27. Haji K, Huynh Q, Wong C, Stewart S, Carrington M, Marwick TH. Improving the characterization of stage A and B heart failure by adding global longitudinal strain. *JACC Cardiovasc Imaging*. 2022;15(8):1380-1387.

28. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: A systematic review and meta-analysis. *PLoS One*. 2014;9(8):e104519.

29. Dicionários de variáveis e Data Books. Elsa. Published October 26, 2021. Accessed March 6, 2024. http://elsabrasil.org/pesquisadores/dicionarios-de-variaveis-e-data-books/

30. Schmidt MI, Duncan BB, Mill JG, et al. Cohort profile: Longitudinal study of adult health (ELSA-brasil). *Int J Epidemiol*. 2015;44(1):68-75.

31. IBGE. Accessed March 31, 2024. https://www.ibge.gov.br/en/home-eng.html

32. Littlejohns TJ, Sudlow C, Allen NE, Collins R. UK Biobank: opportunities for cardiovascular research. *Eur Heart J*. 2019;40(14):1158-1166.

33. Elliott P, Peakman TC, UK Biobank. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int J Epidemiol*. 2008;37(2):234-244.

34. Mill JG, Pinto K, Griep RH, et al. Medical assessments and measurements in ELSA-Brasil. *Rev Saude Publica*. 2013;47 Suppl 2:54-62.

35. Bensenor IM, Griep RH, Pinto KA, et al. Routines of organization of clinical tests and interviews in the ELSA-Brasil investigation center. *Rev Saude Publica*. 2013;47 Suppl 2:37-47.

36. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med*. 2017;36(27):4391-4400.

37. Kang L, Chen W, Petrick NA, Gallas BD. Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach. *Stat Med*. 2015;34(4):685-703.

38. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11-21.

39. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016;352:i6.

40. Sangha V, Nargesi AA, Dhingra LS, et al. Detection of Left Ventricular Systolic Dysfunction From Electrocardiographic Images. *Circulation*. Published online July 25, 2023. doi:10.1161/CIRCULATIONAHA.122.062646

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

41. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25(1):70-74.

42. Bjerkén LV, Rønborg SN, Jensen MT, Ørting SN, Nielsen OW. Artificial intelligence enabled ECG screening for left ventricular systolic dysfunction: a systematic review. *Heart Fail Rev*. Published online November 8, 2022. doi:10.1007/s10741-022-10283-1

43. König S, Hohenstein S, Nitsche A, et al. Artificial intelligence-based identification of left ventricular systolic dysfunction from 12-lead electrocardiograms: External validation and advanced application of an existing model. *Eur Heart J Digit Health*. doi:10.1093/ehjdh/ztad081

44. Acharya UR, Fujita H, Oh SL, et al. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. *Applied Intelligence*. 2019;49(1):16-27.

45. Sudarshan VK, Acharya UR, Oh SL, et al. Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2s of ECG signals. *Comput Biol Med*. 2017;83:48-58.

46. Dhingra LS, Sangha V, Aminorroaya A, et al. A multicenter evaluation of the impact of procedural and pharmacological interventions on deep learning-based electrocardiographic markers of hypertrophic cardiomyopathy. *bioRxiv*. Published online January 16, 2024. doi:10.1101/2024.01.15.24301011

47. Sangha V, Dhingra LS, Oikonomou EK, et al. Identification of hypertrophic cardiomyopathy on electrocardiographic images with deep learning. *medRxiv*. Published online December 28, 2023:2023.12.23.23300490. doi:10.1101/2023.12.23.23300490

48. Dhingra LS, Aminorroaya A, Sangha V, et al. Scalable risk stratification for heart failure using artificial intelligence applied to 12-lead electrocardiographic images: A multinational study. *bioRxiv*. Published online April 3, 2024. doi:10.1101/2024.04.02.24305232

49. Attia ZI, Harmon DM, Dugan J, et al. Prospective evaluation of smartwatch-enabled detection of left ventricular dysfunction. *Nat Med*. 2022;28(12):2497-2503.

50. Attia ZI, Dugan J, Rideout A, 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*. 2022;3(3):373-379.

51. Bachtiger P, Petri CF, Scott FE, 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*. 2022;4(2):e117-e125.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

52. U.S. Food & Drug Administration. 21 CFR 870.2380 Eko Low Ejection Fraction Tool (ELEFT). Published online March 28, 2024. https://www.accessdata.fda.gov/cdrh_docs/pdf23/K233409.pdf

53. Kumar S, Victoria-Castro AM, Melchinger H, et al. Wearables in cardiovascular disease. *J Cardiovasc Transl Res*. Published online September 9, 2022. doi:10.1007/s12265-022-10314-0

54. Singhal A, Cowie MR. The role of wearables in heart failure. *Curr Heart Fail Rep*. 2020;17(4):125-132.

55. Lacirignola F, Pasero E. Hardware design of a wearable ECG-sensor: Strategies implementation for improving CMRR and reducing noise. In: *2017 European Conference on Circuit Theory and Design (ECCTD)*. IEEE; 2017. doi:10.1109/ecctd.2017.8093244

56. Dhingra LS, Shen M, Mangla A, Khera R. Cardiovascular Care Innovation through Data-Driven Discoveries in the Electronic Health Record. *Am J Cardiol*. 2023;203:136-148.

57. Nargesi AA, Adejumo P, Dhingra L, et al. Automated identification of heart failure with reduced ejection fraction using deep learning-based natural language processing. *bioRxiv*. Published online September 11, 2023:2023.09.10.23295315. doi:10.1101/2023.09.10.23295315

58. Sandhu AT, Tisdale RL, Rodriguez F, et al. Disparity in the setting of incident heart failure diagnosis. *Circ Heart Fail*. 2021;14(8). doi:10.1161/circheartfailure.121.008538

59. Lee WC, Serag H, Ohsfeldt RL, et al. Racial disparities in type of heart failure and hospitalization. *J Immigr Minor Health*. 2019;21(1):98-104.

60. Shah K, Pandya A, Kotwani P, et al. Cost-effectiveness of portable electrocardiogram for screening cardiovascular diseases at a primary health center in Ahmedabad district, India. *Front Public Health*. 2021;9. doi:10.3389/fpubh.2021.753443

61. Chen W, Khurshid S, Singer DE, et al. Cost-effectiveness of Screening for Atrial Fibrillation Using Wearable Devices. *JAMA Health Forum*. 2022;3(8):e222419-e222419.

62. Victor RG, Lynch K, Li N, et al. A cluster-randomized trial of blood-pressure reduction in black barbershops. *N Engl J Med*. 2018;378(14):1291-1301.

63. McNeill LH, Reitzel LR, Escoto KH, et al. Engaging black churches to address cancer health disparities: Project CHURCH. *Front Public Health*. 2018;6. doi:10.3389/fpubh.2018.00191

64. Proietti M, Farcomeni A, Goethals P, et al. Cost-effectiveness and screening performance of ECG handheld machine in a population screening programme: The Belgian Heart Rhythm Week screening programme. *Eur J Prev Cardiol*. 2019;26(9):964-972.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

FIGURES

Figure 1. Study Overview. Abbreviations: BMI, Body Mass Index; BP; Blood Pressure; CNN, Convolutional Neural Network; ECG, Electrocardiogram; EF, Ejection Fraction; EHR, Electronic Health Record; ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; HDL, High-density Lipoprotein Cholesterol; HF, Heart Failure; LV, Left Ventricle; YNH, Yale New Haven Hospital; YNHHS, Yale New Haven Health System.

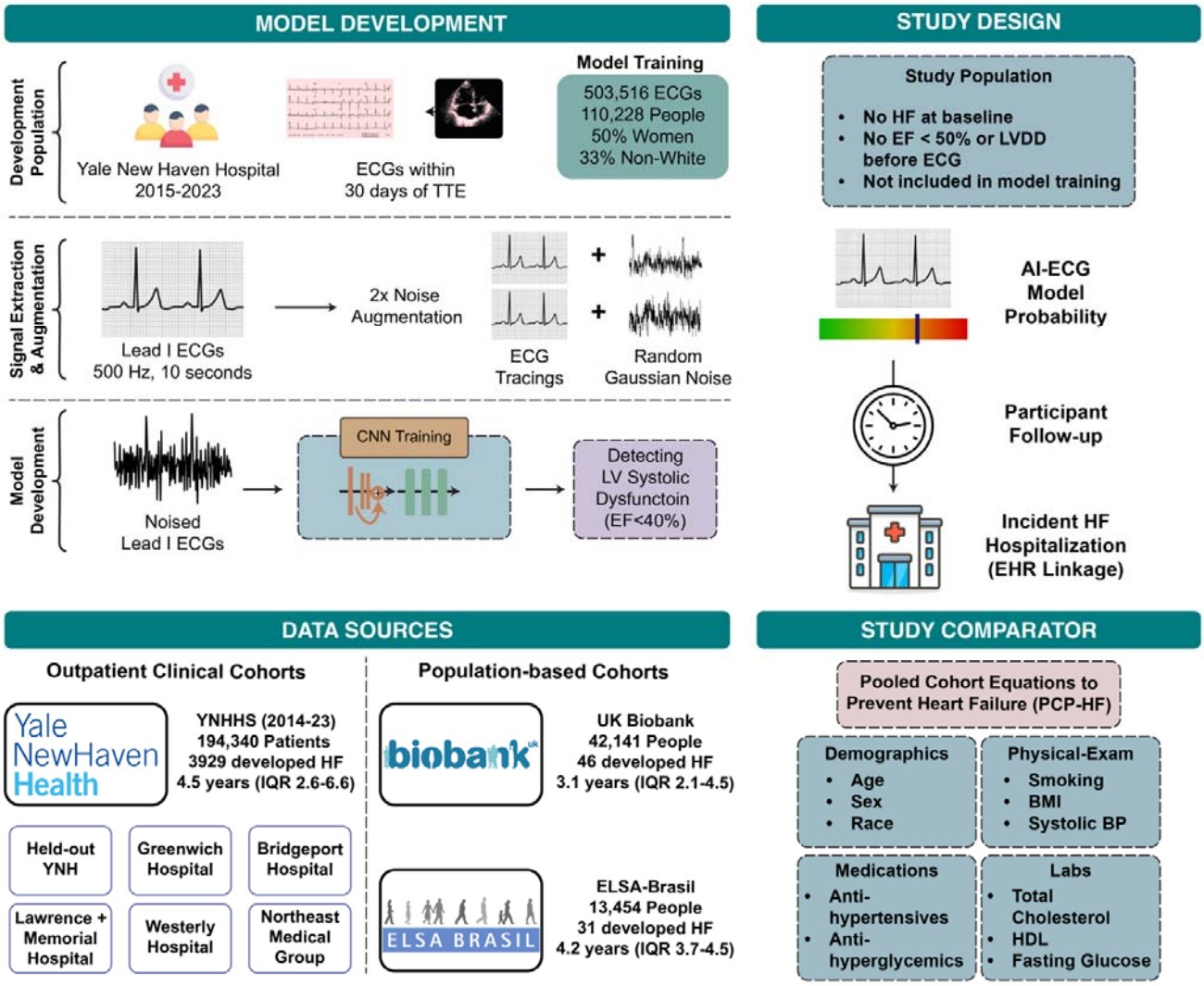
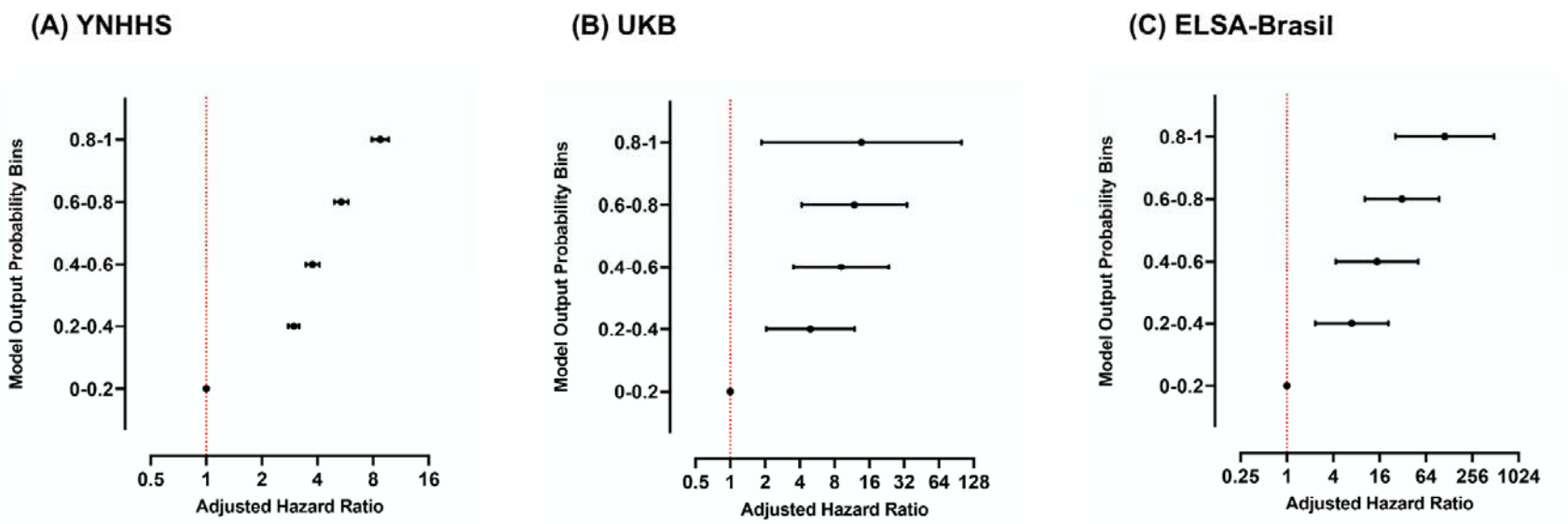


Figure 2. Age- and Sex-adjusted Hazard for Heart Failure across Model Output Probability Bins. Abbreviations: ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; UKB, UK Biobank; YNHHS, Yale New Haven Health System



medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

TABLES

Table 1. Cox Proportional Hazard Models for Predicting Heart Failure Risk Based on AI-ECG Probability. Abbreviations: ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; IHD, Ischemic Heart Disease; HTN, hypertension; T2DM, type-2 diabetes mellitus; UKB, UK Biobank; YNHHS, Yale New Haven Health System.

Predictive Model Inputs	YNHHS		UKB		ELSA-Brasil	
	Positive Screen	Per 0.1 Increment	Positive Screen	Per 0.1 Increment	Positive Screen	Per 0.1 Increment
AI-ECG Probability	5.07 (4.76-5.40)	1.45 (1.44-1.47)	7.52 (4.21-13.41)	1.55 (1.40-1.71)	11.11 (5.32-23.19)	1.83 (1.64-2.05)
AI-ECG Probability + Age + Sex	3.32 (3.11-3.54)	1.32 (1.30-1.33)	5.96 (3.32-10.68)	1.52 (1.37-1.68)	8.74 (4.13-18.48)	1.75 (1.56-1.97)
AI-ECG Probability + Age + Sex + IHD + HTN + T2DM + Obesity	2.81 (2.64-3.01)	1.28 (1.26-1.30)	5.02 (2.77-9.09)	1.49 (1.33-1.66)	7.71 (3.62-16.46)	1.72 (1.52-1.93)
AI-ECG Probability + Age + Sex and accounting for competing risk of death	3.22 (3.01-3.45)	1.30 (1.29-1.32)	5.91 (3.33-10.50)	1.51 (1.38-1.66)	8.67 (4.02-18.70)	1.74 (1.55-1.96)
AI-ECG Probability Age + Sex + IHD + HTN + T2DM + Obesity and accounting for competing risk of death	2.73 (2.55-2.93)	1.27 (1.25-1.28)	4.99 (2.81-8.87)	1.49 (1.36-1.63)	6.53 (2.91-14.67)	1.65 (1.46-1.87)

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 2. Comparison of Discrimination for AI-ECG Model Output Probability and Pooled Cohort equations to Prevent Heart Failure for Heart Failure. Abbreviations: ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; PCP-HF, Pooled Cohort equations to Prevent Heart Failure; UKB, UK Biobank; YNHHS, Yale New Haven Health System**.**

Covariates	YNHHS			UKB			ELSA-Brasil		
	Harrel's C-statistic	Marginal difference over Harrel's C-statistic for PCP-HF	P-value	Harrel's C-statistic	Marginal difference over Harrel's C-statistic for PCP-HF	P-value	Harrel's C-statistic	Marginal difference over Harrel's C-statistic for PCP-HF	P-value
PCP-HF	0.634 (0.612 - 0.656)	-	-	0.740 (0.646 - 0.834)	-	-	0.821 (0.748 - 0.893)	-	-
AI-ECG Model Output Probability	0.725 (0.704 - 0.746)	0.091 (0.063 - 0.120)	< 0.001	0.792 (0.696 - 0.889)	0.052 (-0.083 - 0.186)	0.45	0.833 (0.720 - 0.946)	0.012 (-0.144 - 0.168)	0.87
AI-ECG Model Output Probability + Age + Sex	0.726 (0.705 - 0.746)	0.092 (0.068 - 0.116)	< 0.001	0.836 (0.762 - 0.909)	0.096 (-0.106 - 0.202)	0.07	0.904 (0.844 - 0.964)	0.084 (-0.017 - 0.185)	0.10
AI-ECG Model Output Probability + PCP-HF	0.746 (0.727 - 0.766)	0.112 (0.091 - 0.134)	< 0.001	0.854 (0.785 - 0.923)	0.114 (0.027 - 0.200)	0.01	0.934 (0.903 - 0.966)	0.114 (0.034 - 0.194)	0.005

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Table 3. Integrated Discrimination Improvement and Categorical and Continuous Time-to-Event Net Reclassification Index of AI-ECG Model Output Probability over Pooled Cohort Equations to Prevent Heart Failure. Abbreviations: ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; IDI, Integrated Discrimination Improvement; NRI, Net Reclassification Index; PCP-HF, Pooled Cohort equations to Prevent Heart Failure; UKB, UK Biobank; YNHHS, Yale New Haven Health System**.**

Metric	YNHHS	UKB	ELSA-Brasil
IDI	0.078 (0.003 to 0.160)	0.165 (0.075 to 0.252)	0.238 (0.107 to 0.369)
Categorical NRI	0.201 (0.015 to 0.422)	0.291 (0.092 to 0.494)	0.488 (0.188 to 0.741)
Continuous NRI	0.442 (0.111 to 0.779)	0.155 (-0.211 to 0.478)	0.077 (-0.238 to 0.324)

medRxiv preprint doi: https://doi.org/10.1101/2024.05.27.24307952; this version posted May 27, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .