



Published in final edited form as:

Mayo Clin Proc. 2021 October ; 96(10): 2576–2586. doi:10.1016/j.mayocp.2021.02.029.

## Artificial Intelligence-Augmented Electrocardiogram Detection of Left Ventricular Systolic Dysfunction in the General Population

Anthony H. Kashou, M.D.<sup>a,\*</sup>, Jose R. Medina-Inojosa, M.D, M.Sc<sup>b,\*</sup>, Peter A. Noseworthy, M.D.<sup>b</sup>, Richard J. Rodeheffer, M.D.<sup>b</sup>, Francisco Lopez-Jimenez, M.D., M.B.A.<sup>b</sup>, Itzhak Zachi Attia, M.S.<sup>b</sup>, Suraj Kapa, M.D.<sup>b</sup>, Christopher G. Scott, M.S.<sup>c</sup>, Alexander T. Lee, BS<sup>c</sup>, Paul A. Friedman, M.D.<sup>b</sup>, Paul M. McKie, M.D.<sup>b</sup>

<sup>a</sup>Department of Medicine, Mayo Clinic, Rochester, MN, USA

<sup>b</sup>Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA

<sup>c</sup>Division of Biostatistics, Mayo Clinic, Rochester, MN, USA

### Abstract

**Objective:** To validate an artificial intelligence-augmented electrocardiogram (AI-ECG) algorithm for the detection of preclinical left ventricular systolic dysfunction (LVSD) in a large community-based cohort.

**Methods:** We identified a randomly selected community-based cohort of 2,041 subjects age 45 years in Olmsted County, Minnesota. All participants underwent a study echocardiogram and ECG. We first assessed the performance of the AI-ECG to identify LVSD (ejection fraction 40%). After excluding participants with clinical heart failure, we further assessed the AI-ECG to detect preclinical LVSD among all patients (n = 1,996) and in a high-risk subgroup (n = 1,348). Next we modelled an imputed screening program for preclinical LVSD detection where a positive AI-ECG triggered an echocardiogram. Finally, we assessed the ability of the AI-ECG to predict future LVSD. Participants were enrolled between January 1, 1997 and September 30, 2000 and LVSD surveillance was performed for 10-years after enrollment.

**Results:** For detection of LVSD in the total population (prevalence 2.0%), the area under the receiver operating curve (AUC) for AI-ECG was 0.97 (sensitivity 90%, specificity 92%); in the high-risk subgroup (prevalence 2.7%), the AUC was 0.97 (sensitivity 92%, specificity 93%). In an imputed screening program, identification of one preclinical LSVD case would require 88.3 AI-ECGs and 8.7 echocardiograms in the total population and 65.7 AI-ECGs and 5.5

**Corresponding Author:** Paul M. McKie, M.D., 200 First Street SW Rochester, MN 55905, Phone: 507-284-8896, Fax: 507-255-2550, mckie.paul@mayo.edu.

\* Authors contributed equally.

**Disclosures:** Mayo Clinic has licensed the underlying technology to EKO, a maker of digital stethoscopes with embedded ECG electrodes. Mayo Clinic may receive financial benefit from the use of this technology, but at no point will Mayo Clinic benefit financially from its use for the care of patients at Mayo Clinic. P.A.F., F.L.-J., S.K., and Z.I.A. may also receive financial benefit from this agreement.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

echocardiograms in the high-risk subgroup. The unadjusted hazard ratio for a positive AI-ECG for incident LVSD over 10-years was 2.31 (95% confidence interval: 1.32, 4.05;  $p = 0.004$ ).

**Conclusions:** AI-ECG can identify preclinical LVSD in the community and warrants further study as a screening tool for preclinical LVSD.

### Keywords

Artificial intelligence; electrocardiogram; left ventricular systolic dysfunction; systolic heart failure

---

## INTRODUCTION

Preclinical left ventricular systolic dysfunction (LVSD) is common yet under recognized in the general population (1,2). Despite the lack of symptoms, preclinical LVSD is associated with increased cardiovascular morbidity and mortality, and ultimately progresses to symptomatic heart failure (2,3). Treatment of preclinical LVSD can improve cardiac function, delay onset of symptomatic heart failure, and improve morbidity and mortality (4). Hence, there is a clear rationale for an early identification strategy.

Currently, there are no pragmatic strategies to identify preclinical LVSD in the community. Standard ECG screening for LVSD is ineffective (5). Screening with traditional echocardiograms is cost prohibitive. Portable handheld ultrasound holds promise, but requires training and expertise (6). While amino-terminal pro-B-type natriuretic peptide can detect LVSD, it requires an invasive blood draw and ultimately has insufficient performance to justify widespread implementation (7–9). Thus, there is an unmet need for a non-invasive screening tool for preclinical LVSD in the community.

In a recent retrospective study among patients who underwent concomitant clinical ECG and echocardiogram assessment, we derived and validated an artificial intelligence-augmented electrocardiogram (AI-ECG) algorithm that had favorable performance characteristics for the detection of LVSD (10). However, the ability of AI-ECG to detect LVSD in a randomly selected, community-based cohort and assess its predictive value for development of future LVSD remains unclear. In this study, we assess the predictive value of AI-ECG for the identification of both current and future LVSD in a large, well-characterized community-based cohort.

## METHODS

This analysis was conceived, performed, and funded by Mayo Clinic and the National Institute of Health with no industry support. Data was only used from participants who consented to the use of their anonymized records for research. The Mayo Clinic Foundation Institutional Review Board approved this study.

### Study Population and Medical Record Review

Our study population included a random age- and sex-stratified sample of Olmsted County residents over the age of 45 years as of January 1, 1997, who were participating in

the Rochester Epidemiology Project (11,12). None of the participants in this study were included in the previously published work to derive the AI-ECG algorithm (10). Participants were enrolled between January 1, 1997 and September 30, 2000. Of the 4,203 residents invited, 2,041 participated. Analysis of the medical records from 500 randomly selected residents who did not participate revealed similar age and sex distribution to that observed in participants with a similar prevalence of cardiovascular comorbidities (13). Participants underwent electrocardiography, phlebotomy, and echocardiography as part of the study and not for any specific clinical indication; all testing was performed on the same day. As previously described, medical records were reviewed to determine if a participant had coronary artery disease, prior myocardial infarction, hypertension, heart failure, or diabetes mellitus (14). Each participant's medical records were reviewed to determine if any diagnosis of heart failure had been made. If so, the participant's medical record was reviewed to determine whether the documented clinical information fulfilled Framingham criteria for the diagnosis of heart failure (15). Participants with a heart failure diagnosis in the medical record, which was validated using Framingham criteria, were labelled as having clinical heart failure regardless of the ejection fraction. Participants without a heart failure diagnosis but with systolic dysfunction (ejection fraction < 40%) by echocardiography were considered to have preclinical LVSD. The designation of preclinical LVSD does not imply that the participant would definitely develop heart failure or did not have symptoms, only that the participant had not sought evaluation or had not completed an evaluation that resulted in a diagnosis of heart failure.

The total population without a Framingham validated heart failure diagnosis included 1,996 participants. The high-risk subgroup included 1,391 participants with at least one of the following comorbidities: hypertension, coronary artery disease, prior or current smoking history, or diabetes mellitus (16,17). The high-risk subgroup without a Framingham validated heart failure diagnosis included 1,348 participants.

### AI-ECG Algorithm

We previously reported the application of a deep learning, convolutional neural network (CNN) to the standard 12-lead ECG enabled detection of LVSD with a receiver operated characteristic (ROC) area under the curve (AUC) of 0.93 (10). ECGs were performed with a Marquette ECG machine (GE Healthcare) and then stored with MUSE data management for retrieval. ECGs were supplied to the CNN as a matrix of digital raw voltages. Each lead was sampled at 500 Hz for 10 seconds resulting in a matrix of 12×5000 samples of raw voltages (12 leads, 10 seconds, 500 samples per seconds). The CNN used the Keras framework with a Tensorflow (Google) backend and Python capable of detecting and extracting very subtle patterns based on convolutional layers (18). In the previously published report (10), an internal validation and training set was used to optimize the network architecture and hyperparameters associated with each aspect of the model specification. Thereafter, the optimal network was selected and a ROC was created using a validation set with AUC as a primary assessment of network strength. Thresholds for the probability of having a low ejection fraction were subsequently chosen based on sensitivity and specificity. The CNN was then applied to the test data to assess its ability to predict low ejection fraction at various ejection fraction cutoff points. The thresholds to predict various ejection fraction cutoff

points which were defined in the previously published validation report (10) were used in the current study. No participants in the current study were used to derive the AI-ECG CNN.

### **Echocardiogram**

All echocardiograms were performed by one of three registered diagnostic cardiac sonographers with the same echocardiographic instrument (HP-2500) using standard methods as recommended by the American Society of Echocardiography and were interpreted by a single echocardiologist. Left ventricular ejection fraction was measured by M-mode using the modified Quinones formula, quantitative 2-D (BiPlane Simpsons), and semiquantitative 2-D (visual estimate) as previously described (1). For the purpose of this study, the ejection fraction value used in the models was the first available from a standard hierarchical sequence: biplane approach using the Simpson method, 2-D method, or M-mode; in the absence of any of these, the reported visually-estimated ejection fraction was used.

### **Heart Failure Incidence**

Employing the resources of the Rochester Epidemiology Project, passive surveillance for incident HF (outpatient or hospitalized diagnosis) was performed for 10-years from the time of the initial ECG and echocardiogram. More specifically, incident HF was retrieved using International Classification of Diseases (ICD-9/-10) codes 428.X and I50.X, an approach that has been validated previously (19). Each case of incident HF was reviewed independently and the ejection fraction at the time of the diagnosis recorded using the methods reported above by investigators masked to the AI-ECG outcome.

### **Modelling an Imputed Preclinical LVSD Screening Program**

We modelled an imputed screening program for preclinical LVSD in the total population without heart failure symptoms and in the high-risk subgroup. In this imputed screening program, a positive AI-ECG would trigger a reflex echocardiogram. We calculated how many AI-ECGs and echocardiograms would be required to identify one case of preclinical LVSD as well as the percentage of preclinical LVSD missed with such a screening program. The number of AI-ECGs required was calculated by taking the population of interest and dividing it by the number of LVSD cases in that population. The number of ECHOs required was calculated by taking the number of positive screens within the population of interest and dividing it by the number of true positives in that population.

### **Statistical Analysis**

Data was summarized using mean and standard deviation for continuous variables, which were normally distributed or using median and quartiles when the distribution was not normally distributed. Categorical variables were summarized using number and percentage. Logistic regression methods were used to test the association between ECG probability and an ejection fraction  $\geq 40\%$ . ROC curves were plotted and AUC, sensitivity, and specificity were subsequently estimated. The thresholds to predict various ejection fraction cutoff points which were defined in the previously published validation report (10) were used in the current study. These thresholds were then applied to the cohort with preclinical

LVSD to summarize implications for screening using AI-ECG. The positive likelihood ratio (sensitivity / [1-specificity]) and negative likelihood ratio ([1-sensitivity] / specificity) were calculated for these optimal cutoffs. We plotted the cumulative incidence of LVSD during follow-up while accounting for the competing risk of mortality and tested for the association with a false positive AI-ECG using Cox proportional hazard regression. Hazard ratios with 95% confidence intervals [CI] from unadjusted analyses, after adjustment for age and gender, and in a multivariable model including age, gender, and traditional risk factors for incident LVSD (history of hypertension, diabetes, and coronary artery disease) were obtained. SAS version 9.4 and JMP version 14.1 (Cary, NC) were used for statistical analyses.

## RESULTS

### Study Population Characteristics

The baseline characteristics of the total study population and high-risk subgroup are reported in Table 1. There were 40 and 37 participants with an ejection fraction  $\leq 40\%$  in the total study population and high-risk subgroup, respectively. Of the participants with an ejection fraction  $\leq 40\%$ , 22 in the total population and 20 in the high-risk subgroup had preclinical LVSD.

### Detection of LVSD

Table 2 summarizes the AUC, sensitivity, and specificity for the detection of a reduced ejection fraction ( $\leq 35\%$ ,  $\leq 40\%$ , and  $\leq 50\%$ ) in the total study population and subgroups. The AUC for an ejection fraction  $\leq 40\%$  in the total population was 0.97. Similar results were observed in the high-risk and age greater than 65 years subgroups, as well as for females and males. Figure 1 demonstrates the ROC for the detection of an ejection fraction  $\leq 40\%$  by AI-ECG in the total population and subgroups.

### Detection of Preclinical LVSD

We next assessed the ability of AI-ECG to identify preclinical LVSD (ejection fraction  $\leq 35\%$ ,  $\leq 40\%$ , and  $\leq 50\%$ ). All participants with Framingham criteria clinical heart failure regardless of the ejection fraction ( $n = 45$ ) were excluded from the following analyses, leaving 1,996 study participants of which 98 had an ejection fraction  $\leq 50\%$ , 22 had an ejection fraction  $\leq 40\%$ , and 9 had an ejection fraction  $\leq 35\%$ . Table 3 summarizes the AUC, sensitivity, and specificity for the detection of preclinical LVSD in the total study population and subgroups. The AUC for a preclinical LVSD (ejection fraction  $\leq 40\%$ ) in the total population was 0.97. Figure 2 demonstrates the ROC for the detection of preclinical LVSD (ejection fraction  $\leq 40\%$ ) in the total population and subgroups.

The prevalence of preclinical LVSD (ejection fraction  $\leq 40\%$ ), positive and negative likelihood ratios, as well as the percentage of participants needing a screening echocardiogram based upon an abnormal AI-ECG are shown in Table 4. For the identification of preclinical LVSD (ejection fraction  $\leq 40\%$ ) in the general population ( $n = 22$ , prevalence 1.1%), 8.4% of participants had a positive AI-ECG. For the identification of preclinical LVSD (ejection fraction  $\leq 40\%$ ) in the high-risk subgroup ( $n = 20$ , prevalence

1.5%), 7.5% of participants had a positive AI-ECG. The AI-ECG true positive rate was 11.3% in the total population and 17.8% in the high-risk subgroup.

### Modelling an Imputed Preclinical LVSD Screening Program

In an imputed screening program, 90.7 AI-ECGs and 8.8 echocardiograms in the total population and 67.4 AI-ECGs and 5.6 echocardiograms in the high-risk subgroup would need to be performed in order to identify each case of preclinical LVSD (ejection fraction  $< 40\%$ ). The number of negative echocardiograms (i.e., false positives) and the percentage of participants with preclinical LVSD who would be missed by AI-ECG screening (i.e., false negatives) are shown in Table 4.

### AI-ECG Predictive Value for Incident LVSD

Lastly, we assessed the ability of the AI-ECG algorithm to predict future LVSD (ejection fraction  $< 40\%$ ). For this analysis, we excluded participants with validated heart failure ( $n = 45$ ) or preclinical LVSD (ejection fraction  $< 40\%$ ) ( $n = 22$ ) at baseline. Of the remaining 1,974 participants without validated HF and an ejection fraction  $> 40\%$ , 149 had a positive AI-ECG (false positive) and 1,825 had a negative AI-ECG (true negative). Clinical characteristics and echocardiographic features of participants with a false positive AI-ECG are presented in the supplemental table. Over 10 years, of the participants with a positive AI-ECG without validated heart failure and an ejection fraction  $> 40\%$ , 14 developed incident LVSD (ejection fraction  $< 40\%$ ) compared to 92 participants with a negative AI-ECG at baseline (Figure 3). The five-year and 10-year incidence of LVSD were 3.5% and 9.4%, respectively, among those with a positive AI-ECG versus 1.6% and 5.0%, respectively, among those with negative AI-ECG. The unadjusted hazard ratio (95% CI) for a positive AI-ECG for incident LVSD (ejection fraction  $< 40\%$ ) was 2.31 (1.32, 4.05;  $p = 0.004$ ) and was 1.49 (0.79, 2.56;  $p = 0.25$ ) when adjusted for traditional heart failure clinical risk factors (age, gender, hypertension, coronary artery disease, and diabetes).

## DISCUSSION

In this study, we demonstrate for the first time that an AI-ECG algorithm can detect LVSD in a community-based cohort. The AI-ECG algorithm performs similarly well in individuals without heart failure symptoms and in a high-risk subgroup.

In a retrospective cohort of patients undergoing concomitant clinical ECG and echocardiograms, we previously demonstrated that AI-ECG improves the diagnostic yield for LVSD with an AUC of 0.93, sensitivity of 86.3%, and specificity of 85.7% for participants with an ejection fraction  $< 35\%$  (10). The current study was prospectively performed in a randomly selected community-based cohort where all participants underwent an echocardiogram and ECG as part of the study and not for any specific clinical indication. This well characterized, randomly selected community-based cohort enabled assessment of the AI-ECG algorithm for the detection of preclinical LVSD and its potential role as screening tool for LVSD in the community. In this cohort the AI-ECG algorithm's ability to detect LVSD was comparable, if not improved, to our initial analysis. To detect an ejection fraction  $< 40\%$ , the AI-ECG showed an AUC, sensitivity, and specificity of 0.97, 90.0%,



and 91.8%, respectively. Differences in the study populations may explain the higher AUC for the detection of LVSD in this study as compared to our initial report. Specifically, the current study was performed in a randomly selected, community-based cohort with the ECG and echocardiogram performed on the same day whereas the previous study included participants with a clinical ECG and echocardiogram pair within a two-week time period as well as those with transient LVSD, including tachycardia-mediated or stress-induced cardiomyopathies.

Clinical risk factors (20), the standard 12-lead ECG (5,21,22), plasma natriuretic peptides (5,22–24), and echocardiography – both standard and handheld ultrasound devices (6) – have been evaluated and trialed as potential strategies to detect LVSD. Diagnostic accuracy, cost effectiveness, invasive blood draws, and operator dependency limit the utility and justification of these approaches. AI-ECG’s ability to detect preclinical LVSD (ejection fraction < 40%) in the community provides hope for its future use as a potential screening tool. In an imputed screening protocol using AI-ECG in this study’s cohort, 8.4% of total population and 7.5% of participants in the high-risk subgroup would have a positive AI ECG. Of those with a positive AI-ECG, 11.5% of the total population and 18.0% of the high-risk subgroup would have preclinical LVSD (ejection fraction < 40%). To identify one individual with preclinical LVSD (ejection fraction < 40%) in the total population, 90.7 ECGs and 8.7 echocardiograms were performed. In the high-risk subgroup, 67.4 ECGs and 5.5 echocardiograms were performed to identify one individual with preclinical LVSD.

Participants with a false positive AI-ECG (i.e., positive AI-ECG but (ejection fraction < 40%) had a greater than 2-fold increase in incident clinical LVSD (ejection fraction < 40%) in unadjusted analysis. This suggests the CNN may be detecting early, subclinical, metabolic, or structural abnormalities. Indeed, the prevalence of clinical risk factors for heart failure was higher in subjects with a false positive ECG compared to a true negative ECG (Supplemental Table). The prevalence of incident LVSD was low, thereby limiting the power of the analysis and resulting in loss of statistical significance after adjusting for traditional clinical risk factors. Whether this group would benefit from serial screening or medical therapy to prevent the development of ventricular dysfunction remains unknown.

Our study represents the first prospective evaluation of the AI-ECG application in an unselected, community-based population. The prospective study design of the cohort, including protocol-based echocardiograms and electrocardiograms reduces the chances of error and bias. Despite these strengths, our results should be interpreted with caution due to the low prevalence of LVSD with an ejection fraction < 40%. We chose an ejection fraction cutoff value of 40% because it is a well-established guideline threshold for initiation medical therapy even in preclinical LVSD and because of the limited numbers of participants with an ejection fraction < 35%.

The standard ECG is a painless, non-invasive, inexpensive 10-second test available at most medical centers. Our current work demonstrates the potential utility of an AI-ECG algorithm as a screening tool for identifying preclinical LVSD. Furthermore, the ability to now obtain clinically useful AI-ECG results with mobile-based electrodes may eventually allow application of these technologies in nonclinical environments (25,26). Multicenter,

large community-based validation studies are required to further evaluate the effectiveness, accuracy, and economic implications of AI-ECG for the detecting preclinical LVSD. If AI-ECG proves to be an accurate and cost-effective screening tool, it would be important to assess whether its clinical application translates to improved patient outcomes.

## CONCLUSION

We demonstrate that an AI-ECG model can identify LVSD in a community population and warrants further study as a screening tool to detect preclinical LVSD in the general population.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Sources of Funding:

NIH grant NHLBI RO1-5502. This study was made possible by the Rochester Epidemiology Project (grant number R01-AG034676).

## ABBREVIATIONS

<b>AI-ECG</b>	artificial intelligence-augmented electrocardiogram
<b>AUC</b>	area under the curve
<b>CI</b>	confidence interval
<b>CNN</b>	convolutional neural network
<b>HR</b>	hazard ratio
<b>LVSD</b>	left ventricular systolic dysfunction
<b>ROC</b>	receiver operated characteristic

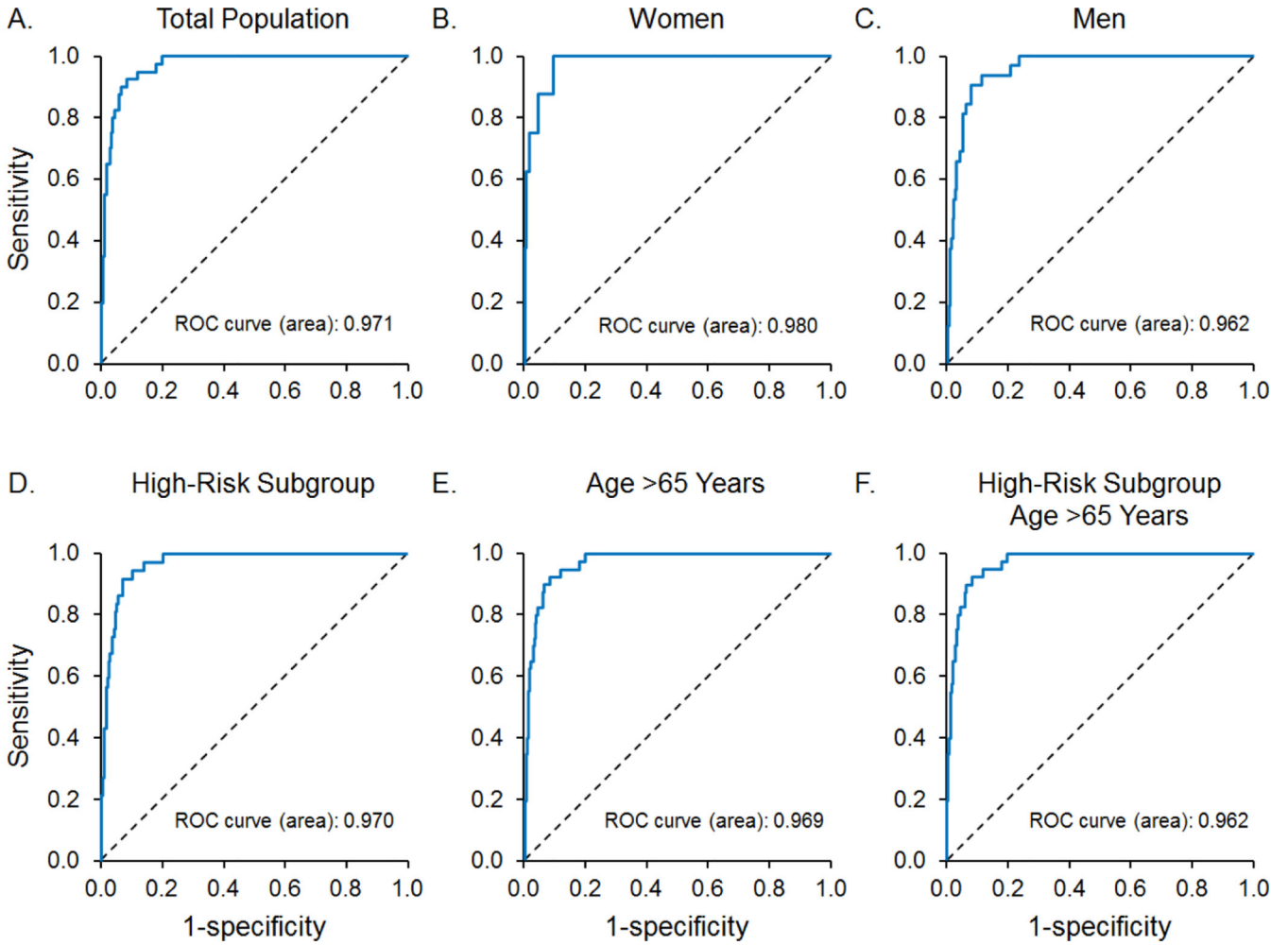
## REFERENCES

1. Redfield MM, Jacobsen SJ, Burnett JC Jr, Mahoney DW, Bailey KR, Rodeheffer RJ. Burden of systolic and diastolic ventricular dysfunction in the community: appreciating the scope of the heart failure epidemic. *JAMA*. 2003;289:194–202. [PubMed: 12517230]
2. Wang TJ, Evans JC, Benjamin EJ, Levy D, LeRoy EC, Vasan RS. Natural history of asymptomatic left ventricular systolic dysfunction in the community. *Circulation*. 2003;108:977–982. [PubMed: 12912813]
3. Echouffo-Tcheugui JB, Erqou S, Butler J, Yancy CW, Fonarow GC. Assessing the risk of progression from asymptomatic left ventricular dysfunction to overt heart failure: a systematic overview and meta-analysis. *JACC Heart Fail*. 2016;4:237–248. [PubMed: 26682794]
4. Pfeffer MA, Braunwald E, Moye LA, Basta L, Brown EJ, Cuddy TE, Davis BR, Geltman EM, Goldman S, Flaker GC, Klein M, Lamas G, Packer M, Rouleau J, Rouleau JL, Rutherford J, Wetheime JH, Hawkin CM. Effect of captopril on mortality and morbidity in patients with left-ventricular dysfunction after myocardial-infarction. Results of the survival and ventricular enlargement trial. The SAVE Investigators. *N Engl J Med*. 1992;327:669–677. [PubMed: 1386652]

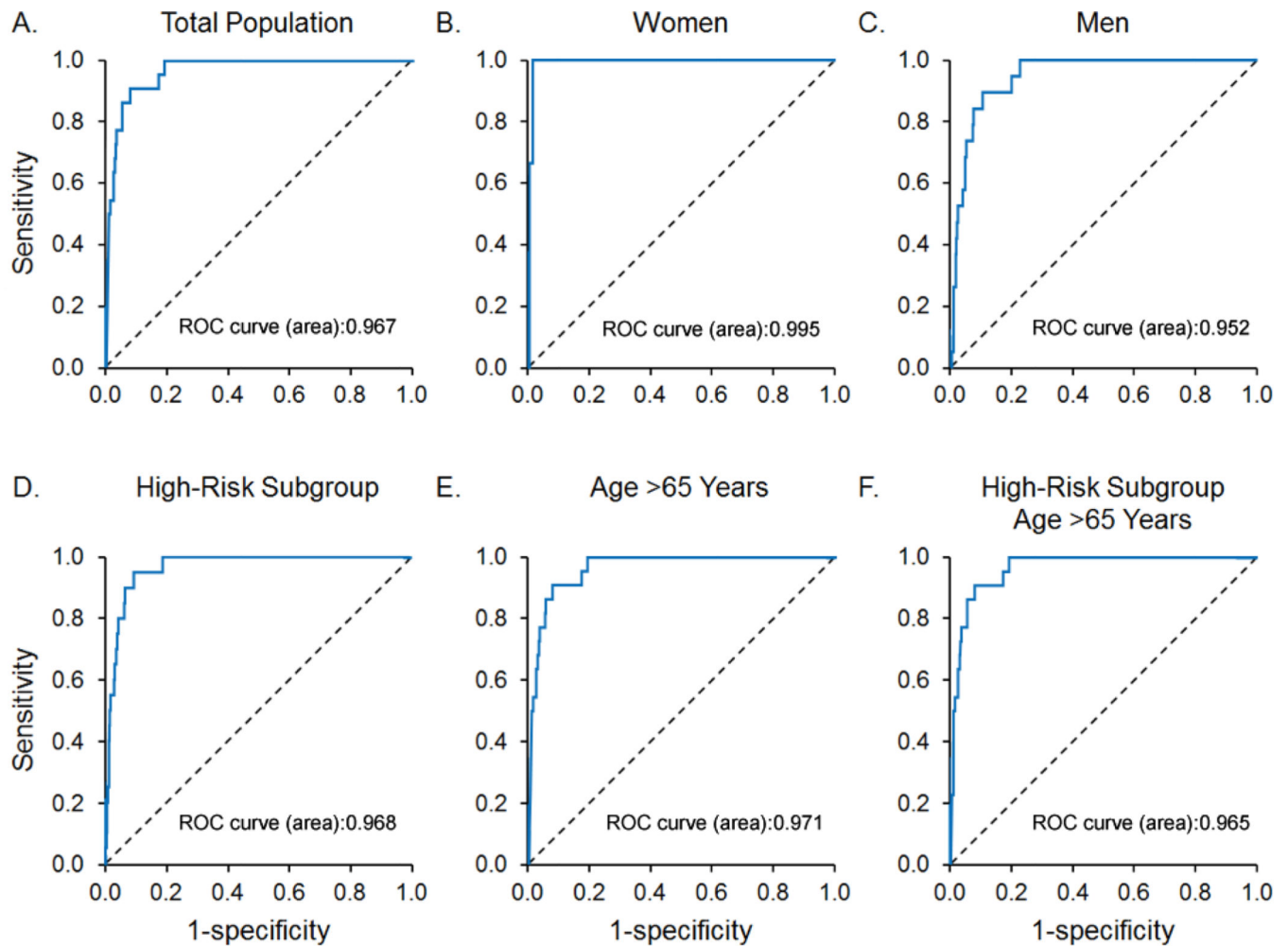


5. Davenport C, Cheng EY, Kwok YT, Lai AH, Wakabayashi T, Hyde C, Connock M. Assessing the diagnostic test accuracy of natriuretic peptides and ECG in the diagnosis of left ventricular systolic dysfunction: a systematic review and meta-analysis. *Br J Gen Pract*. 2006;56:48–56. [PubMed: 16438815]
6. Atherton JJ. Screening for left ventricular systolic dysfunction: is imaging a solution? *JACC Cardiovasc Imaging*. 2010;3:421–428. [PubMed: 20394904]
7. Costello-Boerrigter LC, Boerrigter G, Redfield MM, Rodeheffer RJ, Urban LH, Mahoney DW, Jacobsen SJ, Heublein DM, Burnett JC Jr. Amino-terminal pro-B-type natriuretic peptide and B-type natriuretic peptide in the general community: determinants and detection of left ventricular dysfunction. *J Am Coll Cardiol*. 2006;47:345–353. [PubMed: 16412859]
8. Redfield MM, Rodeheffer RJ, Jacobsen SJ, Mahoney DW, Bailey KR, Burnett JC Jr. Plasma brain natriuretic peptide to detect preclinical ventricular systolic or diastolic dysfunction: a community-based study. *Circulation*. 2004;109:3176–3181. [PubMed: 15184280]
9. Vasan RS, Benjamin EJ, Larson MG, Leip EP, Wang TJ, Wilson PW, Levy D. Plasma natriuretic peptides for community screening for left ventricular hypertrophy and systolic dysfunction: the Framingham heart study. *JAMA*. 2002;288:1252–1259. [PubMed: 12215132]
10. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriguez-Sarano M, Noseworthy PA, Munger TM, Asirvatham SJ, Scott CG, Carter RE, Friedman PA. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25:70–74. [PubMed: 30617318]
11. Melton LJ, 3rd. History of the Rochester Epidemiology Project. *Mayo Clin Proc*. 1996;71:266–274. [PubMed: 8594285]
12. Rocca WA, Yawn BP, St. Sauver JL, Grossardt BR, Melton LJ III. History of the Rochester Epidemiology Project: Half a Century of Medical Records Linkage in a US Population. *Mayo Clin Proc* 2012;87(12):1202–1213. [PubMed: 23199802]
13. Jacobsen SJ, Mahoney DW, Redfield MM, Bailey KR, Burnett JC Jr, Rodeheffer RJ. Participation bias in a population-based echocardiography study. *Ann Epidemiol*. 2004;14:579–584. [PubMed: 15350958]
14. McKie PM, Rodeheffer RJ, Cataliotti A, Martin FL, Urban LH, Mahoney DW, Jacobsen SJ, Redfield MM, Burnett JC Jr. Amino-terminal pro-B-type natriuretic peptide and B-type natriuretic peptide: biomarkers for mortality in a large community-based cohort free of heart failure. *Hypertension*. 2006;47:874–880. [PubMed: 16585413]
15. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. *N Engl J Med*. 1971;285:1441–1446. [PubMed: 5122894]
16. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2013. 62:e147–239.
17. McKie PM, Cataliotti A, Lahr BD, Martin FL, Redfield MM, Bailey KR, Rodeheffer RJ, Burnett JC Jr. The prognostic value of N-terminal pro-B-type natriuretic peptide for death and cardiovascular events in healthy normal and stage A/B heart failure subjects. *J Am Coll Cardiol*. 2010;55:2140–2147. [PubMed: 20447539]
18. van Rossum G. Python Tutorial, Technical Report CS-R9526 (CWI, Amsterdam, 1995).
19. Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, Jacobsen SJ. Trends in heart failure incidence and survival in a community-based population. *JAMA*. 2004;292:344–500. [PubMed: 15265849]
20. Vasan RS, Benjamin EJ, Larson MG, Leip EP, Wang TJ, Wilson PW, Levy D. Plasma natriuretic peptides for community screening for left ventricular hypertrophy and systolic dysfunction: the Framingham Heart Study. *JAMA*. 2002;288:1252–1259. [PubMed: 12215132]
21. Olesen LL, Andersen A. ECG as a first step in the detection of left ventricular systolic dysfunction in the elderly. *ESC Heart Failure*. 2016;3:44–52. [PubMed: 27774266]

22. Ewald B, Ewald D, Thakkestian A, Attia J. Meta-analysis of B type natriuretic peptide and N-terminal pro B natriuretic peptide in the diagnosis of clinical heart failure and population screening for left ventricular systolic dysfunction. *Intern Med J.* 2008;38:101–113. [PubMed: 18290826]
23. Latour-Perez J, Coves-Orts FJ, Abad-Terrado C, Abaira V, Zamora J. Accuracy of B-type natriuretic peptide levels in the diagnosis of left ventricular dysfunction and heart failure: a systematic review. *Eur J Heart Fail.* 2006;18:390–399. [PubMed: 16305826]
24. Betti I, Castelli G, Barchielli A, Belgini C, Boscherini V, De Luca L, Messeri G, Gheorghide M, Maisel A, Zuppiroli A. The role of N-terminal PRO-brain natriuretic peptide and echocardiography for screening asymptomatic left ventricular dysfunction in a population at high risk for heart failure. The PROBE-HF study. *J Card Fail.* 2009;15:377–384. [PubMed: 19477397]
25. Yasin OZ, Attia Z, Dillon JJ, DeSimone CV, Sapir Y, Dugan J, Somers VK, Ackerman MK, Asirvatham DJ, Scott CG, Bennet KE, Ladewig DJ, Sadot D, Geva AB, Friedman PA. Noninvasive blood potassium measurement using signal-processed, single-lead ecg acquired from a handheld smartphone. *J Electrocardiol.* 50:620–625. [PubMed: 28641860]
26. Attia ZI, Dugan J, Maidens J, Rideout A, Lopez-Jimenez F, Noseworthy PA, Asirvatham S, Pellikka PA, Ladewig DJ, Satam G, Pham S, Venkatraman S, Friedman P, Kapa S. Prospective Analysis of Utility of Signals From an Ecg-Enabled Stethoscope to Automatically Detect a Low Ejection Fraction Using Neural Network Techniques Trained From the Standard 12-Lead Ecg. 2019;140:A13447.

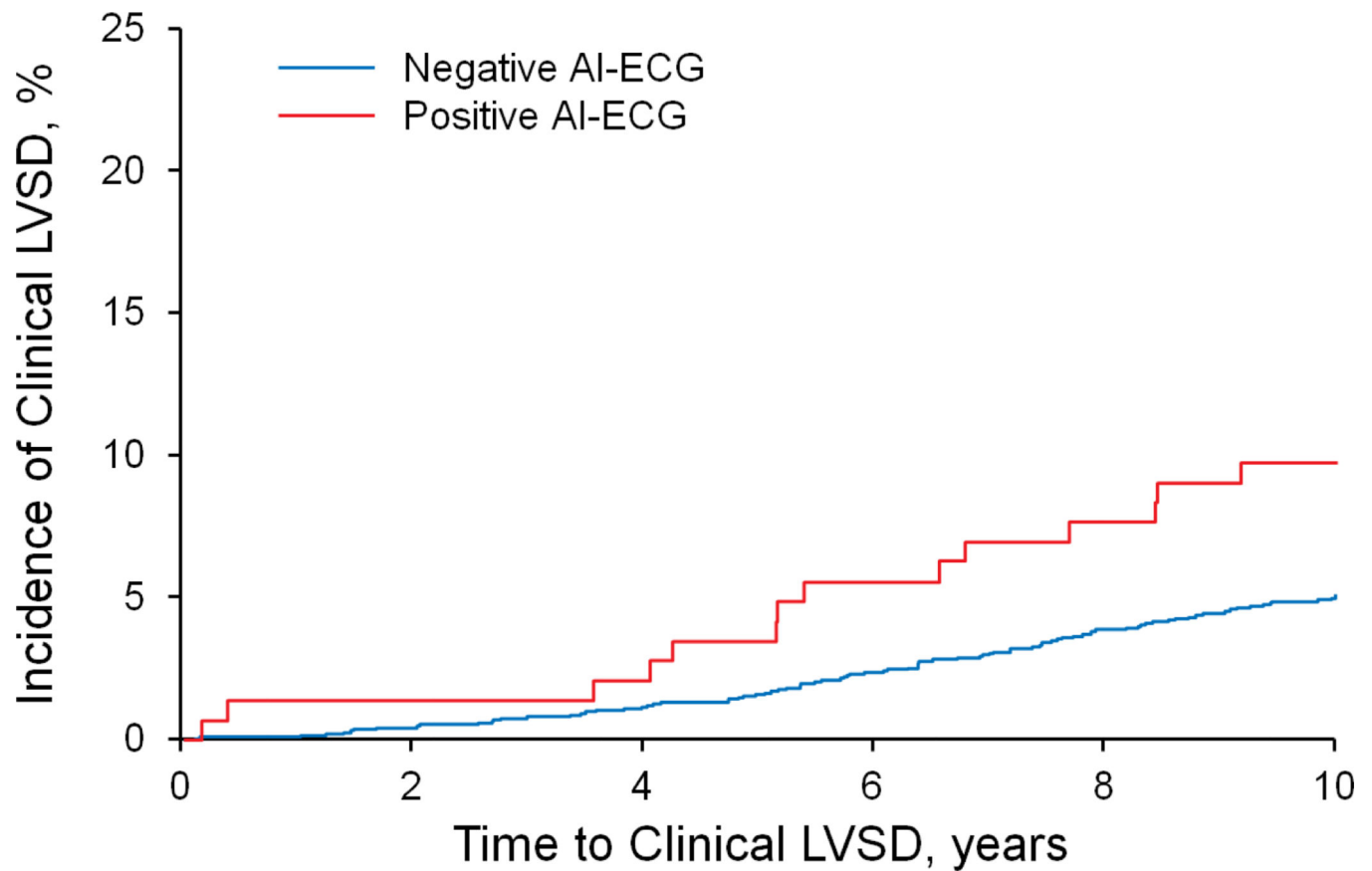


**Figure 1:** The receiver operator curves of AI-ECG to identify subjects with an ejection fraction 40% in the A) total population (n = 40), B) women (n = 8), C) men (n = 32), D) high-risk subgroup (n = 37), E) subjects over the age of 65 years (n = 27), and F) subjects in the high-risk subgroup over the age of 65 years (n = 26). *Abbreviations: AI-ECG = artificial intelligence-augmented electrocardiogram, ROC = receiver operator curves.*



**Figure 2:**

The receiver operator curves of AI-ECG to identify subjects with preclinical left ventricular systolic dysfunction (ejection fraction < 40%) in the A) total population (n = 22) ( B) women (n = 3), C) men (n = 19) D) high-risk subgroup (n = 20), E) subjects over the age of 65 years (n = 15), and F) subjects in the high-risk subgroup over the age of 65 years (n = 14). ). Abbreviations: AI-ECG = artificial intelligence-augmented electrocardiogram, ROC = receiver operator curves.



**Figure 3:**

Kaplan-Meier curve for unadjusted future clinical left ventricular systolic dysfunction (defined as clinical heart failure and an ejection fraction  $\leq$  40%) according to AI-ECG result among subjects without clinical heart failure and with an ejection fraction  $>$  40% at the time of study enrollment. Those with a positive AI-ECG screen had over two-fold risk of heart failure at 10-years (hazards ratio = 2.31 [1.32, 4.05],  $p = 0.004$ ). *Abbreviations: AI-ECG, artificial intelligence-augmented electrocardiogram; LVSD, left ventricular systolic dysfunction.*

**Table 1:**

Baseline characteristics of the study population and high-risk subgroup.

	Total Population		High-Risk Subgroup	
	Including Validated Heart Failure <sup>c</sup> n = 2041	Excluding Validated Heart Failure <sup>c</sup> n = 1996	Including Validated Heart Failure <sup>c</sup> n = 1391	Excluding Validated Heart Failure <sup>c</sup> n = 1348
Women gender, n (%)	1058 (52%)	1040 (52%)	633 (46%)	617 (46%)
Age, years, mean (SD)	63 (11)	62.5 (10)	63.5 (10)	63 (10)
Age, 65 years, n (%)	817 (40%)	784 (39%)	601 (43%)	569 (42%)
BMI, mean (SD)	28 (5)	28 (5)	29 (5)	29 (5)
Estimated GFR <sup>b</sup> , mean (SD)	84 (23)	85 (23)	83 (24)	84 (24)
NT-proBNP, median (Q1, Q3)	68 (28, 146)	66 (28, 140)	72 (28, 167)	69 (27, 158)
Diabetes, n (%)	151 (7%)	138 (7%)	151 (11%)	138 (10%)
Hypertension, n (%)	575 (28%)	546 (27%)	575 (41%)	546 (41%)
Coronary artery disease, n (%)	249 (12%)	215 (11%)	249 (18%)	215 (16%)
Smoking (current or previous), n (%)	1020 (50%)	994 (50%)	1020 (74%)	994 (74%)
Clinical heart failure <sup>c</sup> , n (%)	45 (2.2%)	0 (0%)	43 (3.1%)	0 (0%)
Ejection fraction 50%, n (%)	123 (6%)	98 (5%)	109 (8%)	86 (6%)
Ejection fraction 40%, n (%)	40 (2%)	22 (1%)	37 (3%)	20 (2%)
Ejection fraction 35%, n (%)	20 (1%)	9 (1%)	19 (1%)	8 (1%)

High-risk subgroup was defined as the presence of one or more of the following: diabetes, hypertension, coronary artery disease, or current/previous smoking history.

<sup>a</sup> Abbreviations: BMI = body mass index, GFR = glomerular filtration rate, NT-proBNP = N-terminal pro-B-type natriuretic peptide, SD = standard deviation

<sup>b</sup> Glomerular filtration rate calculated using the Modification of Diet in Renal Disease (MDRD) equation.

<sup>c</sup> Clinical heart failure regardless of the ejection fraction confirmed by trained nurse abstractors using Framingham criteria.



**Table 2:**

Test characteristics for AI-ECG for the detection of left ventricular systolic dysfunction.

	EF (%)	N	AUC	Sensitivity	Specificity
Total population, n = 2041	35	20 (1.0%)	0.981	100.0	93.5
	40	40 (2.0%)	0.971	90.0	91.9
	50	123 (6.0%)	0.880	76.4	87.9
High-risk, n = 1391	35	19 (1.4%)	0.974	100.0	91.9
	40	37 (2.7%)	0.970	91.9	92.9
	50	109 (7.8%)	0.884	78.9	86.6
Men, n = 983	35	16 (1.6%)	0.967	100.0	91.2
	40	32 (3.3%)	0.962	90.6	92.4
	50	91 (9.3%)	0.878	73.6	92.7
Women, n = 1058	35	4 (0.4%)	0.996	100.0	99.4
	40	8 (0.8%)	0.980	87.5	90.7
	50	32 (3.0%)	0.866	75.0	87.0
Age >65 years, n = 817	35	13 (1.6%)	0.976	100.0	89.4
	40	27 (3.3%)	0.969	92.6	87.2
	50	74 (9.1%)	0.907	87.8	81.7
High risk, Age >65 years, n = 601	35	12 (2.0%)	0.971	100.0	87.6
	40	26 (4.3%)	0.962	92.3	89.4
	50	70 (11.6%)	0.897	77.1	89.5

<sup>a</sup> Abbreviations: AI-ECG = artificial intelligence-augmented electrocardiogram, AUC = area under the curve, EF = ejection fraction;

Test characteristics for AI-ECG for the detection of preclinical left ventricular systolic dysfunction

**Table 3:**

	EF (%)	N	AUC	Sensitivity	Specificity
Total population, n = 1996	35	9 (0.5%)	0.978	100.0	94.3
	40	22(1.1%)	0.967	86.4	92.5
	50	98 (4.9%)	0.857	71.4	87.6
High risk, n = 1348	35	8 (0.6%)	0.969	100.0	93.1
	40	20 (1.5%)	0.968	90.0	90.8
	50	86 (6.4%)	0.861	74.4	86.5
Men, n = 956	35	7 (0.7%)	0.961	100.0	92.3
	40	19 (2.0%)	0.952	84.2	90.0
	50	75 (7.8%)	0.858	68.0	93.3
Women, n = 1040	35	2 (0.2%)	0.998	100.0	99.8
	40	3 (0.3%)	0.995	100.0	98.8
	50	23 (2.2%)	0.829	65.2	87.6
Age > 65 years, n = 784	35	6 (0.8%)	0.973	100.0	91.0
	40	15 (1.9%)	0.971	93.3	88.4
	50	56 (7.1%)	0.892	75.0	91.6
High risk, Age > 65 years, n = 569	35	5 (0.9%)	0.965	100.0	89.7
	40	14 (2.5%)	0.965	92.9	86.7
	50	53 (9.3%)	0.880	84.9	79.5

<sup>a</sup> Abbreviations: AI-ECG = artificial intelligence-augmented electrocardiogram, AUC = area under the curve, EF = ejection fraction

AI-ECG for the detection of preclinical left ventricular systolic dysfunction and the implications for screening <sup>b</sup>

Table 4:

	LVSD prevalence % (n)	+ LR	- LR	+ Screen % (n)	TP % (n)	FP % (n)	TN % (n)	FN % (n)	LVSD missed <sup>c</sup> % (n)
Total population, n = 1996	1.1 (22)	11.4	0.15	8.4 (168)	1.0 (19)	7.5 (149)	91.4 (1825)	0.2 (3)	13.6 (3)
High risk, n = 1348	1.5 (20)	14.4	0.11	7.5 (101)	1.3 (18)	6.2 (83)	92.4 (1245)	0.1 (2)	10.0 (2)
Men, n = 956	2.0 (19)	12.3	0.17	8.4 (80)	1.7 (16)	6.7 (64)	91.3 (873)	0.3 (3)	15.8 (3)
Women, n = 1040	0.3 (3)	11.4	0.00	9.0 (94)	0.3 (3)	8.8 (91)	91.0 (946)	0 (0)	0 (0)
Age > 65 years, n = 784	1.9 (15)	8.1	0.08	13.1 (103)	1.8 (14)	11.4 (89)	86.7 (680)	0.1 (1)	6.7 (1)
High risk, Age > 65 years, n = 569	2.5 (14)	10.3	0.08	11.1 (63)	2.3 (13)	8.8 (50)	88.8 (505)	0.2 (1)	7.1 (1)

<sup>a</sup> Abbreviations: AI-ECG = artificial intelligence-augmented electrocardiogram, ECHO = echocardiogram, FN = false negative, FP = false positive, LR = likelihood ratio, LVSD = left ventricular systolic dysfunction, - = negative, + = positive; TN = true negative, TP = true positive.

<sup>b</sup> Participants with clinical heart failure regardless of the ejection fraction confirmed by trained nurse abstractors using Framingham criteria were excluded from this analysis. A positive screen is defined as a positive AI-ECG. A true positive (TP) is defined as positive AI-ECG and positive ECHO. A false positive (FP) is defined as a positive AI-ECG and negative ECHO. A true negative (TN) is defined as a negative AI-ECG and negative ECHO. A false negative (FN) is defined as a negative AI-ECG and positive ECHO.

<sup>c</sup> LVSD missed = number of participants with LVSD and a negative AI-ECG; percentage = [(number of participants with LVSD with negative AI-ECG) / total number of participants with LVSD] x 100.