An Artificial Intelligence-Enabled ECG Algorithm to Identify Patients with Left Ventricular Systolic Dysfunction Presenting to the Emergency Department with Dyspnea

Running title: *Adedinsewo et al.; Evaluating dyspnea in the ED*

Demilade Adedinsewo, MD, MPH[1]; Rickey E. Carter, PhD[2]; Zachi Attia, MSc[3]; Patrick Johnson, BS[2]; Anthony H. Kashou, MD[3]; Jennifer L. Dugan, MD[4]; Michael Albus, MD[5]; Johnathan M. Sheele, MD, MPH[5]; Fernanda Bellolio, MD, MS[6]; Paul A. Friedman, MD[4]; Francisco Lopez-Jimenez, MD, MSc[4]; Peter A. Noseworthy, MD[4]

[1]Division of Cardiovascular Medicine, [2]Department of Health Sciences Research, [5]Department of Emergency Medicine, Mayo Clinic, FL; [3]Department of Medicine, [4]Division of Cardiovascular Medicine, [6]Department of Emergency Medicine, Mayo Clinic, MN

Correspondence:
Demilade Adedinsewo, MD, MPH
Division of Cardiovascular Diseases
Mayo Clinic
4500 San Pablo Rd, S
Jacksonville, FL 32224
Tel: 904-953-7278
Fax: 904-953-2911
Email: adedinsewo.demilade@mayo.edu

Journal Subject Terms: Cardiomyopathy; Electrocardiology (ECG); Echocardiography; Diagnostic Testing; Heart Failure

Abstract:

Background - Identification of systolic heart failure among patients presenting to the emergency department (ED) with acute dyspnea is challenging. The reasons for dyspnea are often multifactorial. A focused physical evaluation and diagnostic testing can lack sensitivity and specificity. The objective of this study was to assess the accuracy of an artificial intelligence-enabled electrocardiogram (AI-ECG) to identify patients presenting with dyspnea who have left ventricular systolic dysfunction (LVSD).

Methods - We retrospectively applied a validated AI-ECG algorithm for the identification of LVSD (defined as left ventricular ejection fraction ≤ 35%) to a cohort of patients aged ≥ 18 years who were evaluated in the ED at a Mayo Clinic site with dyspnea. Patients were included if they had at least one standard 12-lead ECG acquired on the date of the ED visit and an echocardiogram performed within 30 days of presentation. Patients with prior LVSD were excluded. We assessed the model performance using area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity.

Results - A total of 1,606 patients were included. Median time from ECG to echocardiogram was 1 day (Q1: 1, Q3: 2). The AI-ECG algorithm identified LVSD with an AUC of 0.89 (95% CI: 0.86 – 0.91) and accuracy of 85.9%. Sensitivity, specificity, negative predictive value, and positive predictive value were 74%, 87%, 97%, and 40%, respectively. To identify an ejection fraction < 50%, the AUC, accuracy, sensitivity, and specificity were 0.85 (95% CI: 0.83 – 0.88), 86%, 63%, and 91%, respectively. NT-Pro BNP alone at a cut-off of >800 identified LVSD with an AUC of 0.80 (95% CI: 0.76 – 0.84).

Conclusions - The ECG is an inexpensive, ubiquitous, painless test which can be quickly obtained in the ED. It effectively identifies LVSD in selected patients presenting to the ED with dyspnea when analyzed with AI and outperforms NT-Pro BNP.

Key words: ECG; dyspnea; heart failure; cardiomyopathy; artificial intelligence

Nonstandard Abbreviations and Acronyms

ACC: American College of Cardiology

AHA: American Heart Association

AI: Artificial intelligence

ARNI: angiotensin receptor–neprilysin inhibitor

AUC: Area under the receiver operating curve

BNP: B-type natriuretic peptide

NT-Pro BNP: N-terminal pro-B-type natriuretic peptide

CNN: Convolutional neural network

ECG: Electrocardiogram

ED: Emergency Department

EF: Ejection fraction

HFpEF: Heart failure with preserved ejection fraction

LVSD: Left ventricular systolic dysfunction

ROC: Receiver operating characteristic

Introduction

Dyspnea is a frequent complaint in the emergency department (ED) accounting for approximately 1.2 million annual ED visits in the United States [1]. Ascertaining the etiology of dyspnea can be challenging given the variability in clinical symptoms and presentation. Heart failure is a common cause of cardiac-related dyspnea among ED patients and accurate identification of these patients is essential for acute management and final disposition. Heart failure is also one of the most common diagnoses assigned to ED patients who become hospitalized [2].

Identification of left ventricular systolic dysfunction (LVSD) in the ED can effectively diagnose heart failure and appropriately guide management in the acute setting. A left ventricular ejection fraction (LVEF) ≤ 35% is associated with increased mortality in patients with and without heart failure [3-5].

We hypothesized that the application of an artificial intelligence-enabled electrocardiogram (AI-ECG) algorithm may provide for a rapid and inexpensive means to effectively identify LVSD in patients presenting to the ED with dyspnea. This AI-ECG algorithm has been previously demonstrated to be very effective in identification of LVSD [6]. To test this hypothesis, we performed a retrospective study of patients presenting to the Mayo Clinic ED with acute dyspnea.

Methods

The data, analytic methods, and study materials that support the findings of this study are available from the corresponding author upon reasonable request.

Study population

We identified adult patients 18 years and older (n=21,309) presenting to the ED at all Mayo Clinic sites (AZ, FL, MN) and Mayo Clinic Health Systems (MCHS) with a complaint of dyspnea between May 2018 and February 2019 who had at least 1 standard 10-second 12-lead ECG performed within 24 hours of ED visit. Reported dyspnea was identified based on ICD-9 codes: 786.0 (including 786.00 – 786.09) and ICD-10 codes: R06 (R06.00 – R06.9). For patients with multiple ED visits, the first ED visit was selected as the index visit. The first ECG performed within 24 hours of index ED visit was selected when multiple ECGs were available. From this cohort, we subsequently identified patients who had a comprehensive two-dimensional

echocardiogram performed within 30 days of index ED visit. All ECGs were acquired at a sampling rate of 500Hz using a GE Marquette ECG machine (Marquette, WI) and stored using the MUSE® ECG data management system (GE Healthcare, Chicago, IL). Quantitative data from echocardiography performed are recorded at the time of the acquisition in a Mayo Clinic developed, custom database (Echo Image Management System [EIMS], Rochester, MN). LVEF was estimated using standard methods recommended by the American Society of Echocardiography [7].

Exclusion

We excluded patients with a known prior diagnosis of systolic, diastolic, or unspecified heart failure using ICD 10 diagnosis codes I11, I13, I50, I27.22 and Z95.811 (n=4,812) within 2 years prior to index ED visit. We also excluded patients without an echocardiogram within 30 days of index ED visit (n=14,072), prior echocardiogram demonstrating left ventricular ejection fraction ≤ 35% (n=20), ECGs not performed on the date of ED visit (n=101), ECGs not readily accessible in MUSE® (n=444), and those without readily accessible echocardiogram accession numbers (n=254). Our final sample size included 1,606 patients (7.5% of patients meeting inclusion criteria) (Figure 1). The study was approved by the Mayo Clinic Internal Review Board including a waiver of informed consent.

Primary and secondary outcomes

Our primary outcome was the identification of patients with new LVSD defined as LVEF ≤35% within 30 days of the ED visit using a deep learning network for ECGs performed at the time of ED visit. Our secondary outcome was the identification of LVEF EF <50% within 30 days of the ED visit using the same deep learning network. For diagnostic accuracy assessment, the gold standard was LVEF as measured on a two-dimensional echocardiogram.

AI Model

We employed a previously described AI-ECG algorithm developed and validated for identification of LVEF ≤35% [6, 8] with no additional training or optimization. This algorithm employed a convolutional neural network (CNN) trained with Keras with a Tensorflow (Google, Mountain View, CA). Details of the algorithm derivation have been previously published [6, 8]. There was no overlap in the study data utilized and published in Nature Medicine.

Statistical Analyses

The primary global measure of model performance for this study was the area under the receiver operating characteristics curve (AUC) formed by modeling the AI-ECG algorithm prediction of the probability of LVSD in relationship to the clinically-determined diagnosis of LVSD within 30 days of the ED presentation. Routine measures of diagnostic performance based on dichotomized predictions (e.g., sensitivity, specificity) were obtained using a previously determined threshold of >=0.256 to indicate a positive screen. Ninety-five percent exact confidence intervals were calculated for all measures of diagnostic performance except for AUC. In this case, the large sample approximation of the DeLong method with optimization by Sun & Xu was used[9, 10]. The AI-ECG algorithm's performance was tested against BNP on subjects that had BNP measured at time of ED visit. Summaries for the AUC of BNP alone and in combination with the AI-ECG algorithm were developed. The incremental benefit of the AI-ECG prediction of LVSD was tested using DeLong's test.

In order to provide some general description about comorbidities the sample had at time of ED presentation, the following approach was taken. Using standardized code sets of ICD9 and ICD10 codes per diagnosis, the Mayo Clinic Unified Data Platform was queried for the presence of at least 1 code within the code set within 30 days post the ECG. If a single code was

found, the patient was considered positive for the condition. Statistical analyses were computed using R version 3.6.2.

Results

Study Population Characteristics

A total of 1606 patients were included. Overall, the median age of patients evaluated in the ED for dyspnea was 68 years, approximately half were female (47%), and the majority were white (91%). The median time to echocardiogram following ED visit was 1 day (Q1: 1, Q3: 2). Only 54% of patients had NT-Pro BNP levels assessed, 43% had high-sensitivity troponin values, and 63% had serum creatinine at the index ED visit (Table 1).

Left ventricular ejection fraction ≤ 35%

Utilization of an AI-ECG algorithm for identification of new LVSD among ED patients presenting with dyspnea achieved an area under the receiver operating characteristic curve (AUC) of 0.89 (95% CI: 0.86 – 0.91) with an accuracy of 85.9% (95% CI: 84.1% – 87.6%). Sensitivity, specificity, negative predictive value, and positive predictive value were 74%, 87%, 97%, and 40%, respectively (Figure 2).

Left ventricular ejection fraction < 50%

For our secondary outcome, the AI-ECG algorithm achieved an AUC of 0.85 (95% CI: 0.83 – 0.88) with an accuracy of 86% (95% CI: 84.2% – 87.7%). Sensitivity, specificity, negative predictive value, and positive predictive value were 63%, 91%, 92%, and 62%, respectively (Figure 3).

Sub-Populations

The AI-ECG algorithm appeared to have slightly better performance characteristics for identification of LVSD (EF ≤ 35%) in younger (AUC=0.91) and female (AUC 0.90) patients, albeit with less precision (Figure 4). Overall, the AI-ECG algorithm's diagnostic accuracy was similar across the subgroups examined. We also evaluated its performance for identification of EF < 50% and its performance was similar across patient sub-groups (Figure 5).

Comparing AI-ECG Vs. NT-Pro BNP in the ED

In a sub-sample of our patient population who had NT-Pro BNP values available (54%, n=866), NT-pro BNP alone at a cutoff value > 800 identified new LVSD (ejection fraction ≤ 35%) with an AUC of 0.80 (95% CI: 0.76 – 0.84) demonstrating a superior diagnostic value of a single AI-ECG algorithm in this patient cohort (p<0.0001). Addition of NT-Pro BNP to the AI-ECG algorithm added a marginal incremental value with improvement in AUC from 0.89 to 0.91 (p=0.091).

30- Day Clinical Outcomes

Patient with new LVSD identified by echocardiography were significantly more likely to be re-hospitalized with heart failure (32% Vs. 10%, p=<0.001). There was however no difference in 30-day all cause re-hospitalizations or repeat ED visits.

Discussion

In this study, we demonstrate that an AI-ECG algorithm can be an effective tool for rapid detection of LVSD in patients presenting to the ED with acute dyspnea with an AUC of 0.89. Our study provides evidence to support real world application of an AI-ECG algorithm in routine clinical practice.

This study is particularly important as its goal is to identify patients with significant depression in LV systolic function in the ED. Early identification of these patients provides a potential opportunity for linkage to essential cardiovascular care for appropriate management, follow-up diagnostic testing as appropriate, medication optimization and device based therapies such as implantable cardioverter-defibrillator (ICD) and cardiac resynchronization therapy (CRT)[4]. The ECG is inexpensive, ubiquitous, painless, quickly obtained, and can be performed with minimal training. The AHA/ACC guidelines identified dyspnea as the most common angina equivalent and recommend obtaining a 12-lead ECG in all adult patients presenting to the ED with dyspnea [11], which makes the ECG an appropriate screening tool. The cost of the ECG compared to an echocardiogram likely makes it more cost effective for screening patients in the ED for LVSD.

Although, not all patients with cardiac related dyspnea have LVSD. Some patients may present with dyspnea secondary to heart failure with preserved ejection fraction (HFpEF), angina presenting as dyspnea, or pulmonary edema due to acute valvular heart disease or hypertensive crisis. However, given the high mortality risk in patients with significant LVSD (ejection fraction ≤35%)[4], identifying this sub-group in the ED provides a unique avenue for early linkage to cardiovascular care and device based therapies.

Echocardiography is considered the ideal non-invasive test for evaluation of left ventricular systolic function [7] but this procedure is highly operator dependent, requires significant training, and may not be readily available in the ED [12]. Formal echocardiography requires review and interpretation by a trained cardiologist which takes considerable time and may not be feasible for making acute care decisions.

Studies have examined various point-of-care tests for differentiating between cardiac and non-cardiac causes of dyspnea in the ED. Natriuretic peptides are by far the most widely utilized and the only biomarkers recommended by the American College of Cardiology/ American Heart Association (ACC/AHA) heart failure guidelines (Class I A) for initial evaluation of heart failure in patients presenting with dyspnea [13]. The "Breathing Not Properly" study was the landmark clinical trial that established the utility of B-type natriuretic peptide (BNP) values in identifying patients presenting to the ED with heart failure related dyspnea [14]. The final diagnosis of heart failure in this study was adjudicated based on a combination of clinical symptoms, physical examination findings, and diagnostic tests reviewed by cardiologists. However, only 77% of those deemed to have heart failure related dyspnea had echocardiograms performed [14]. Several other studies have also demonstrated the utility of point of care BNP and N-terminal pro-B-type natriuretic peptide (NT-pro BNP) testing in the acute care setting [15-21]. However, the misclassification rate with BNP testing has been reported to be as high as 14 – 29% [2]. Multiple factors are also known to affect natriuretic peptide levels such as obesity, age, chronic kidney disease, hemodialysis, pulmonary hypertension, sepsis, chronic atrial fibrillation, and angiotensin receptor–neprilysin inhibitor (ARNI) [13, 22-27]. In addition, measurement of BNP has not been shown to have any effect on clinical outcomes or medications administered [13, 28]. Our study demonstrated the AI-ECG algorithm outperforms NT-Pro BNP alone for identifying patients with new LVSD (AUC 0.89 vs. 0.80).

Other rapid diagnostic modalities previously evaluated in the ED or acute care setting for evaluation of dyspnea include: chest radiograph [29], inferior vena cava diameter on ultrasound [30], lung ultrasound [29, 31-33], partial pressure of end-tidal CO2 [34], bioimpedance [35], plasma volume status [36] and focused cardiac ultrasound [37], all with associated limitations. Making a diagnosis of

acute heart failure in the ED using a combination of history, physical examination, chest radiograph, and ECG was noted to be discordant with the final diagnosis 25% of the time [2]. The use of cardiopulmonary ultrasound in the ED has been shown to be superior to a combination of clinical examination, NT-pro BNP, and chest radiograph for establishing the etiology of acute dyspnea [38]; however, performance of a cardiopulmonary ultrasound also requires additional training in ultrasonography. A meta-analysis of various studies using these rapid diagnostic modalities found bedside lung ultrasound and echocardiography to be the most useful tests for confirming acute heart failure while use of natriuretic peptides were more effective for diagnostic exclusion [2].

Clinical diagnoses made in the ED often remain in patient's medical records even if the final diagnosis changes with the availability of additional imaging tests or upon further evaluation by the admitting or consulting physician. An analysis of the National Hospital Discharge Survey noted the number of hospitalizations with any mention heart failure tripled between 1979 and 2004. However, only 30% of these patients had heart failure listed as the first or primary diagnosis [39], suggesting that heart failure may not be the main clinical diagnosis or may not be acute (i.e. chronic heart failure). Therefore, the additional information provided by an AI-ECG algorithm may assist in improving diagnostic accuracy and documentation in the emergency department and patient disposition. While not assessed in this work, we intend to study this prospectively. In addition, AI-ECG may be particularly helpful and potentially add incremental valuable information as a screening tool in community EDs without ready access to cardiologists or echocardiography.

We also observed that patients with LVSD were significantly more likely to be re-hospitalized with heart failure. These findings suggest that the use of an AI-ECG algorithm could

potentially identify patients at risk for repeat heart failure hospitalizations and provides a unique opportunity to implement specific interventions to prevent this while in the ED including early follow up with a cardiologist, initiation of guideline directed medical therapy and social work if needed.

A recent analysis of Medicare beneficiaries revealed a diagnosis of heart failure had the highest preventable healthcare spending among elderly patients in the acute care setting above bacterial pneumonia, urinary tract infections, and diabetes complications [40]. In the United States, health care costs for patients with heart failure are projected to increase from $20.9 billion in 2012 to $53.1 billion by 2030 [41]. As such, it is imperative that patients with heart failure related dyspnea are accurately and efficiently identified in the ED and appropriate therapies initiated early to potentially reduce associated health care costs due to readmissions related to delays in appropriate cardiovascular care.

Limitations

Our study utilized ICD diagnosis codes for excluding patients with prior heart failure as such, some patients may have been missed or inappropriately excluded. In addition, we only included patients who had a confirmatory echocardiogram performed within 30 days of index ED visit but may have inadvertently excluded patients with new heart failure who did not have a follow-up echocardiogram within 30 days or those who had an echocardiogram performed at a different facility. The omission of the confirmatory echocardiogram required to diagnosis LVSD may result in biased measures of positive and negative predictive value given the true underlying disease prevalence may be different from what was observed in the analysis.

Strengths

A major strength of this study is the ability to utilize an existing, rapid, non-invasive test- the ECG, to provide valuable additional information in the care of patients with dyspnea in the ED. This study adds to the growing body of literature demonstrating practical applications of AI based algorithms in the field of cardiovascular medicine such as rapid determination of implantable cardiac device type and model using images from a chest radiograph[42], heart failure mortality risk prediction[43] and prognostication using cardiopulmonary exercise testing in heart failure[44].

Conclusion

An AI-ECG algorithm was able to identify LVSD with high accuracy in patients presenting to the ED with dyspnea. The application of an AI-ECG algorithm in the ED could improve diagnostic accuracy, facilitate appropriate disposition and provide an avenue to identify high risk patients early and link them to appropriate cardiovascular care. Prospective studies are needed to further evaluate the effectiveness of this algorithm, practicality, effect on real-time improvement in diagnostic evaluation, cost-effectiveness, and association with long-term clinical outcomes.

Sources of Funding: None

Disclosures: None

References:

1. Rui P KK, Ashman JJ. National Hospital Ambulatory Medical Care Survey: 2016 Emergency Department Summary Tables. 2016.

2. Martindale JL, Wakai A, Collins SP, Levy PD, Diercks D, Hiestand BC, Fermann GJ, Desouza I, Sinert R. Diagnosing acute heart failure in the emergency department: a systematic review and meta-analysis. *Acad Emerg Med*. 2016;23:223-242.

3. Toma M, Ezekowitz JA, Bakal JA, O'Connor CM, Hernandez AF, Sardar MR, Zolty R, Massie BM, Swedberg K, Armstrong PW, et al. The relationship between left ventricular ejection fraction and mortality in patients with acute heart failure: insights from the ASCEND-HF Trial. *Eur J Heart Fail*. 2014;16:334-341.

4. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, et al. 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-e239.

5. Wehner GJ, Jing L, Haggerty CM, Suever JD, Leader JB, Hartzel DN, Kirchner HL, Manus JNA, James N, Ayar Z, et al. Routinely reported ejection fraction and mortality in clinical practice: where does the nadir of risk lie? *Eur Heart J*. 2019:ehz550.

6. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriquez-Sarano M, Noseworthy PA, Munger TM, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. 2019;25:70-74.

7. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr*. 2015;28:1-39.e14.

8. Attia ZI, Kapa S, Yao X, Lopez-Jimenez F, Mohan TL, Pellikka PA, Carter RE, Shah ND, Friedman PA, Noseworthy PA. Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction. *J Cardiovasc Electrophysiol*. 2019;30:668-674.

9. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-45.

10. Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters*. 2014;21:1389-1393.

11. Amsterdam EA, Wenger NK, Brindis RG, Casey DE, Jr., Ganiats TG, Holmes DR, Jr., Jaffe AS, Jneid H, Kelly RF, Kontos MC, et al. 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130:e344-e426.

12. Gil Martinez P, Mesado Martinez D, Curbelo Garcia J, Cadinanos Loidi J. Amino-terminal pro-B-type natriuretic peptide, inferior vena cava ultrasound, and biolectrical impedance analysis for the diagnosis of acute decompensated CHF. *Am J Emerg Med*. 2016;34:1817-22.

13. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Jr., Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Am Coll Cardiol*. 2017;70:776-803.

14. Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P, Omland T, Storrow AB, Abraham WT, Wu AH, et al. Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med*. 2002;347:161-7.

15. Dao Q, Krishnaswamy P, Kazanegra R, Harrison A, Amirnovin R, Lenert L, Clopton P, Alberto J, Hlavin P, Maisel AS. Utility of B-type natriuretic peptide in the diagnosis of congestive heart failure in an urgent-care setting. *J Am Coll Cardiol*. 2001;37:379-85.

16. Logeart D, Saudubray C, Beyne P, Thabut G, Ennezat P-V, Chavelas C, Zanker C, Bouvier E, Solal AC. Comparative value of Doppler echocardiography and B-type natriuretic peptide assay in the etiologic diagnosis of acute dyspnea. *J Am Coll Cardiol*. 2002;40:1794-1800.

17. McCullough PA, Nowak RM, McCord J, Hollander JE, Herrmann HC, Steg PG, Duc P, Westheim A, Omland T, Knudsen CW, et al. B-type natriuretic peptide and clinical judgment in emergency diagnosis of heart failure: analysis from Breathing Not Properly (BNP) Multinational Study. *Circulation*. 2002;106:416-22.

18. Knudsen CW, Omland T, Clopton P, Westheim A, Abraham WT, Storrow AB, McCord J, Nowak RM, Aumont M-C, Duc P. Diagnostic value of B-type natriuretic peptide and chest radiographic findings in patients with acute dyspnea. *Am J Med*. 2004;116:363-368.

19. Mueller C, Scholer A, Laule-Kilian K, Martina B, Schindler C, Buser P, Pfisterer M, Perruchoud AP. Use of B-type natriuretic peptide in the evaluation and management of acute dyspnea. *N Engl J Med*. 2004;350:647-654.

20. Green SM, Martinez-Rumayor A, Gregory SA, Baggish AL, O'Donoghue ML, Green JA, Lewandrowski KB, Januzzi JL. Clinical uncertainty, diagnostic accuracy, and outcomes in emergency department patients presenting with dyspnea. *Arch Intern Med*. 2008;168:741-748.

21. Ray P, Delerme S, Jourdain P, Chenevier-Gobeaux C. Differential diagnosis of acute dyspnea: the value of B natriuretic peptides in the emergency department. *Q J Med*. 2008;101:831-843.

22. Homsak E, Ekart R. Hemodiafiltration affects NT-proBNP but not ST2 serum concentration in end-stage renal disease patients. *Clin Biochem*. 2016;49:1159-1163.

23. Santos-Araujo C, Leite-Moreira A, Pestana M. Clinical value of natriuretic peptides in chronic kidney disease. *Nefrologia*. 2015;35:227-33.

24. Kim BJ, Hwang SJ, Sung KC, Kim BS, Kang JH, Lee MH, Park JR. Assessment of factors affecting plasma BNP levels in patients with chronic atrial fibrillation and preserved left ventricular systolic function. *Int J Cardiol*. 2007;118:145-50.

25. Tagore R, Ling LH, Yang H, Daw HY, Chan YH, Sethi SK. Natriuretic peptides in chronic kidney disease. *Clin J Am Soc Nephrol*. 2008;3:1644-51.

26. Kadri AN, Kaw R, Al-Khadra Y, Abuamsha H, Ravakhah K, Hernandez AV, Tang WHW. The role of B-type natriuretic peptide in diagnosing acute decompensated heart failure in chronic kidney disease patients. *Arch Med Sci*. 2018;14:1003-1009.

27. Madamanchi C, Alhosaini H, Sumida A, Runge MS. Obesity and natriuretic peptides, BNP and NT-proBNP: mechanisms and diagnostic implications for heart failure. *Int J Cardiol*. 2014;176:611-7.

28. Schneider H-G, Lam L, Lokuge A, Krum H, Naughton MT, Smit PDV, Bystrzycki A, Eccleston D, Federman J, Flannery G. B-type natriuretic peptide testing, clinical outcomes, and health services use in emergency department patients with dyspnea. *Ann Intern Med*. 2009;150:365-371.

29. Maw AM, Hassanin A, Ho PM, McInnes MDF, Moss A, Juarez-Colunga E, Soni NJ, Miglioranza MH, Platz E, DeSanto K, et al. Diagnostic Accuracy of Point-of-Care Lung Ultrasonography and Chest Radiography in Adults With Symptoms Suggestive of Acute Decompensated Heart Failure: A Systematic Review and Meta-analysis. *JAMA Netw Open*. 2019;2:e190703.

30. Yamanoglu A, Celebi Yamanoglu NG, Parlak I, Pinar P, Tosun A, Erkuran B, Akgur A, Satilmis Siliv N. The role of inferior vena cava diameter in the differential diagnosis of dyspneic patients; best sonographic measurement method? *Am J Emerg Med*. 2015;33:396-401.

31. Prosen G, Klemen P, Strnad M, Grmec Š. Combination of lung ultrasound (a comet-tail sign) and N-terminal pro-brain natriuretic peptide in differentiating acute heart failure from chronic obstructive pulmonary disease and asthma as cause of acute dyspnea in prehospital emergency setting. *Critical Care*. 2011;15:R114.

32. Pivetta E, Goffi A, Nazerian P, Castagno D, Tozzetti C, Tizzani P, Tizzani M, Porrino G, Ferreri E, Busso V, et al. Lung ultrasound integrated with clinical assessment for the diagnosis of acute decompensated heart failure in the emergency department: a randomized controlled trial. *Eur J Heart Fail*. 2019;21:754-766.

33. Buessler A, Chouihed T, Duarte K, Bassand A, Huot-Marchand M, Gottwalles Y, Penine A, Andre E, Nace L, Jaeger D, et al. Accuracy of Several Lung Ultrasound Methods for the Diagnosis of Acute Heart Failure in the ED: A Multicenter Prospective Study. *Chest*. 2019.

34. Klemen P, Golub M, Grmec Š. Combination of quantitative capnometry, N-terminal pro-brain natriuretic peptide, and clinical assessment in differentiating acute heart failure from

pulmonary disease as cause of acute dyspnea in pre-hospital emergency setting: study of diagnostic accuracy. *Croat Med J*. 2009;50:133-142.

35. Genot N, Mewton N, Bresson D, Zouaghi O, Francois L, Delwarde B, Kirkorian G, Bonnefoy-Cudraz E. Bioelectrical impedance analysis for heart failure diagnosis in the ED. *Am J Emerg Med*. 2015;33:1025-9.

36. Chouihed T, Rossignol P, Bassand A, Duarte K, Kobayashi M, Jaeger D, Sadoune S, Buessler A, Nace L, Giacomin G, et al. Diagnostic and prognostic value of plasma volume status at emergency department admission in dyspneic patients: results from the PARADISE cohort. *Clin Res Cardiol*. 2019;108:563-573.

37. Carlino MV, Paladino F, Sforza A, Serra C, Liccardi F, de Simone G, Mancusi C. Assessment of left atrial size in addition to focused cardiopulmonary ultrasound improves diagnostic accuracy of acute heart failure in the Emergency Department. *Echocardiography*. 2018;35:785-791.

38. Gallard E, Redonnet J-P, Bourcier J-E, Deshaies D, Largeteau N, Amalric J-M, Chedaddi F, Bourgeois J-M, Garnier D, Geeraerts T. Diagnostic performance of cardiopulmonary ultrasound performed by the emergency physician in the management of acute dyspnea. *Am J Emerg Med*. 2015;33:352-358.

39. Fang J, Mensah GA, Croft JB, Keenan NL. Heart failure-related hospitalization in the U.S., 1979 to 2004. *J Am Coll Cardiol*. 2008;52:428-34.

40. Figueroa JF, Joynt Maddox KE, Beaulieu N, Wild RC, Jha AK. Concentration of Potentially Preventable Spending Among High-Cost Medicare Subpopulations: An Observational Study. *Ann Intern Med*. 2017;167:706-713.

41. Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*. 2016;13:368-78.

42. Howard JP, Fisher L, Shun-Shin MJ, Keene D, Arnold AD, Ahmad Y, Cook CM, Moon JC, Manisty CH, Whinnett ZI, et al. Cardiac Rhythm Device Identification Using Neural Networks. *JACC Clin Electrophysiol*. 2019;5:576-586.

43. Adler ED, Voors AA, Klein L, Macheret F, Braun OO, Urey MA, Zhu W, Sama I, Tadel M, Campagnari C, et al. Improving risk prediction in heart failure using machine learning. *Eur J Heart Fail*. 2019:10.1002/ejhf.1628.

44. Hearn J, Ross HJ, Mueller B, Fan C-P, Crowdy E, Duhamel J, Walker M, Alba AC, Manlhiot C. Neural Networks for Prognostication of Patients With Heart Failure. *Circ Heart Fail*. 2018;11:e005193-e005193.

Table 1: Demographic and baseline characteristics of study population.

Characteristics	LVSD (No) n = 1442	LVSD (Yes) n = 164	Overall n = 1606	P value
Age, years*	67.9 (56.3, 77.6)	68.2 (58.4, 77.2)	67.9 (56.7, 77.5)	0.86
Sex, female	706 (49.0%)	53 (32.3%)	759 (47.3%)	**<0.001**
Race				0.19
- Black/African American	61 (4.2%)	8 (4.9%)	69 (4.3%)	
- White	1310 (90.8%)	143 (87.2%)	1453 (90.5%)	
- Other	71 (4.9%)	13 (7.9%)	84 (5.2%)	
Ethnicity				0.84
- Hispanic or Latino	36 (2.5%)	4 (2.4%)	40 (2.5%)	
- Not Hispanic or Latino	1357 (94.1%)	153 (93.4%)	1510 (94.0%)	
- Other	49 (3.4%)	7 (4.2%)	56 (3.5%)	
BMI (kg/m^2)*†	29.1 (24.9, 34.8)	29.0 (24.9, 35.2)	29.1 (24.9, 34.8)	0.57
Time between ECG and echocardiogram*	1.0 (1.0, 2.0)	1.0 (0.0, 1.0)	1.0 (1.0, 2.0)	**<0.001**
Serum creatinine at index ED visit* ‡	0.9 (0.8, 1.2)	1.1 (0.9, 1.4)	1.0 (0.8, 1.2)	**<0.001**
NT-Pro BNP at index ED visit* ‡	719.0 (233.5, 2188.5)	3808.0 (2009.0, 6771.5)	944.5 (278.2, 3100.0)	**<0.001**
High sensitivity troponin T at index ED visit* ‡	20.0 (12.0, 40.0)	26.5 (15.0, 61.8)	22.0 (13.0, 43.0)	**0.006**
History of myocardial infarction	303 (21.0%)	67 (40.9%)	370 (23.1%)	**<0.001**
Diabetes	400 (27.8%)	54 (32.9%)	454 (28.3%)	0.17
Peripheral artery disease	505 (35.0%)	62 (37.8%)	567 (35.3%)	0.49
Cerebrovascular disease	280 (19.4%)	28 (17.1%)	308 (19.2%)	0.53
Chronic kidney disease	362 (25.1%)	59 (36.0%)	421 (26.2%)	**0.004**
Chronic pulmonary disease	618 (42.9%)	54 (32.9%)	672 (41.9%)	**0.015**

P-values presented are from Fisher's Exact Test for count data and Wilcoxon Rank Sum tests for numeric variables
*Data summarized as median (25^{th}, 75^{th} percentile)
†Body mass index (BMI) data was unavailable for 194 patients
‡Overall, 63%, 54% and 43% of patients has serum creatinine, NT-Pro BNP and high sensitivity troponin values respectively

Figure Legends:

Figure 1: Patient flow diagram

Figure 2: Receiver operative characteristic (ROC) curve for identification of LVEF ≤ 35% among patients presenting to the ED with dyspnea

Figure 3: Receiver operating characteristic (ROC) curve for identification of LVEF < 50% among patients presenting to the ED with dyspnea.

Figure 4: Forest plot showing AI-ECG algorithm performance for identification of LVSD (EF ≤ 35%) stratified by age group and sex

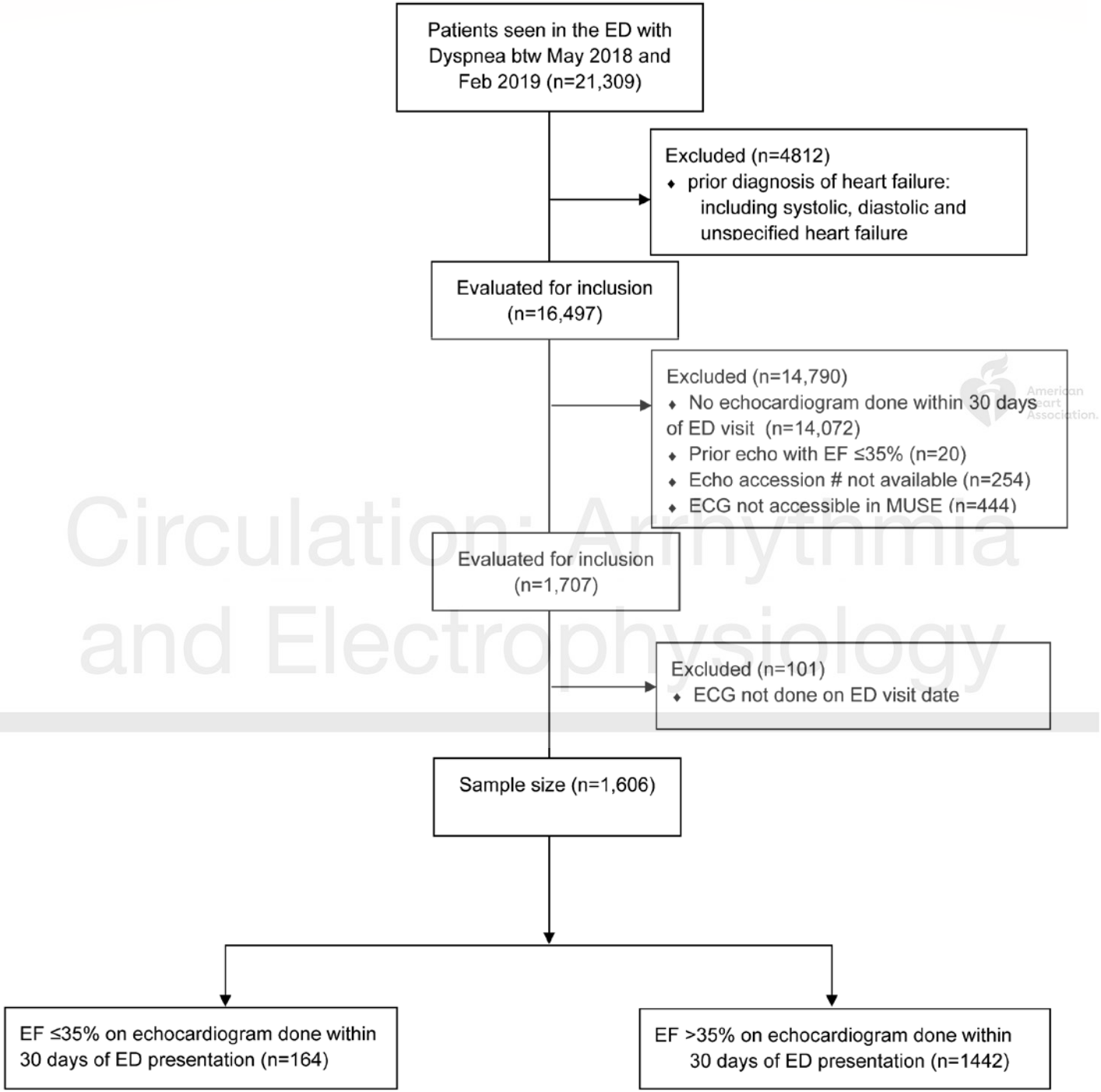
Figure 5: Forest plot showing AI-ECG algorithm performance for identification of LVSD (EF < 50%) stratified by age group and sex

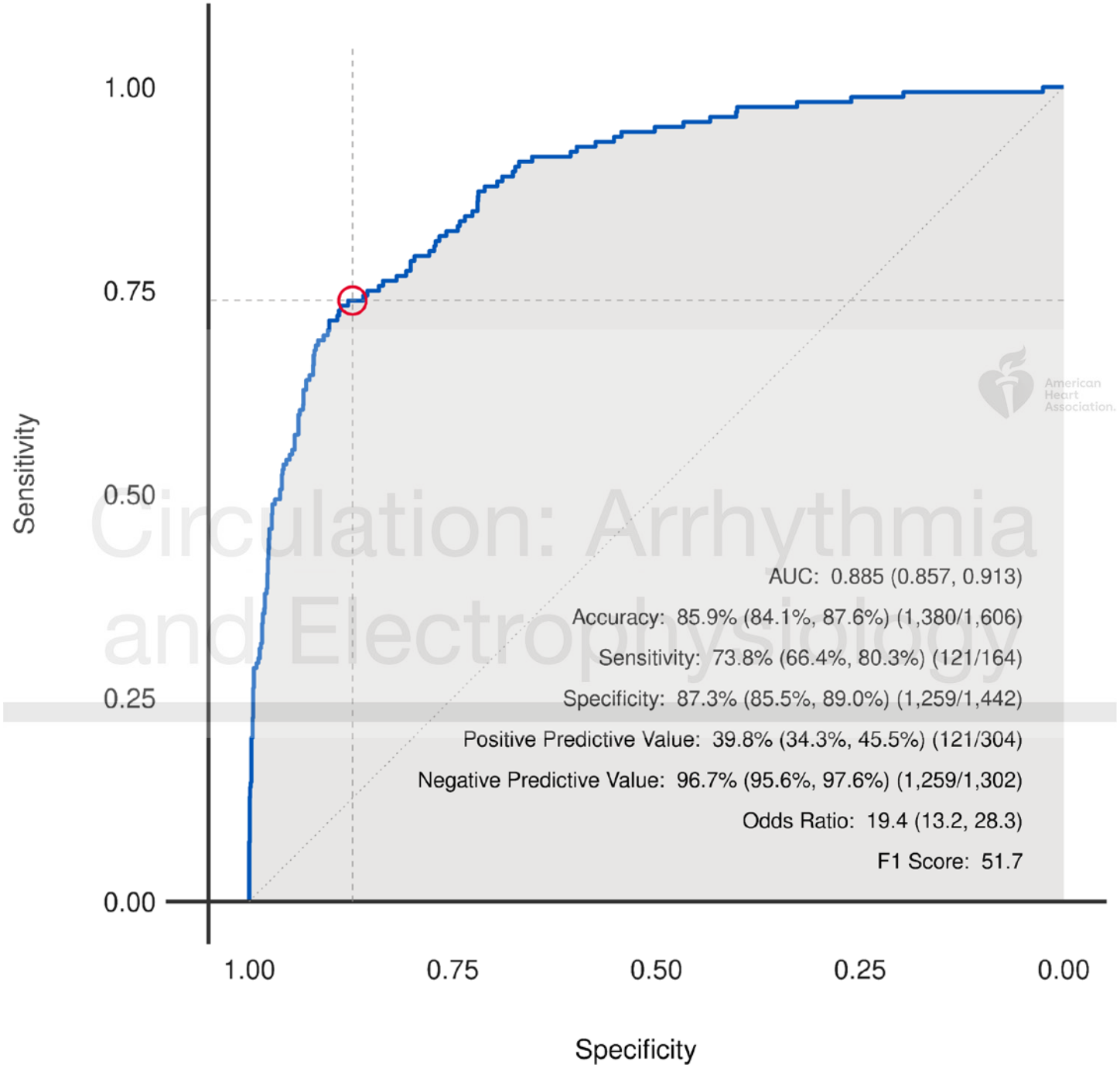
What is Known:

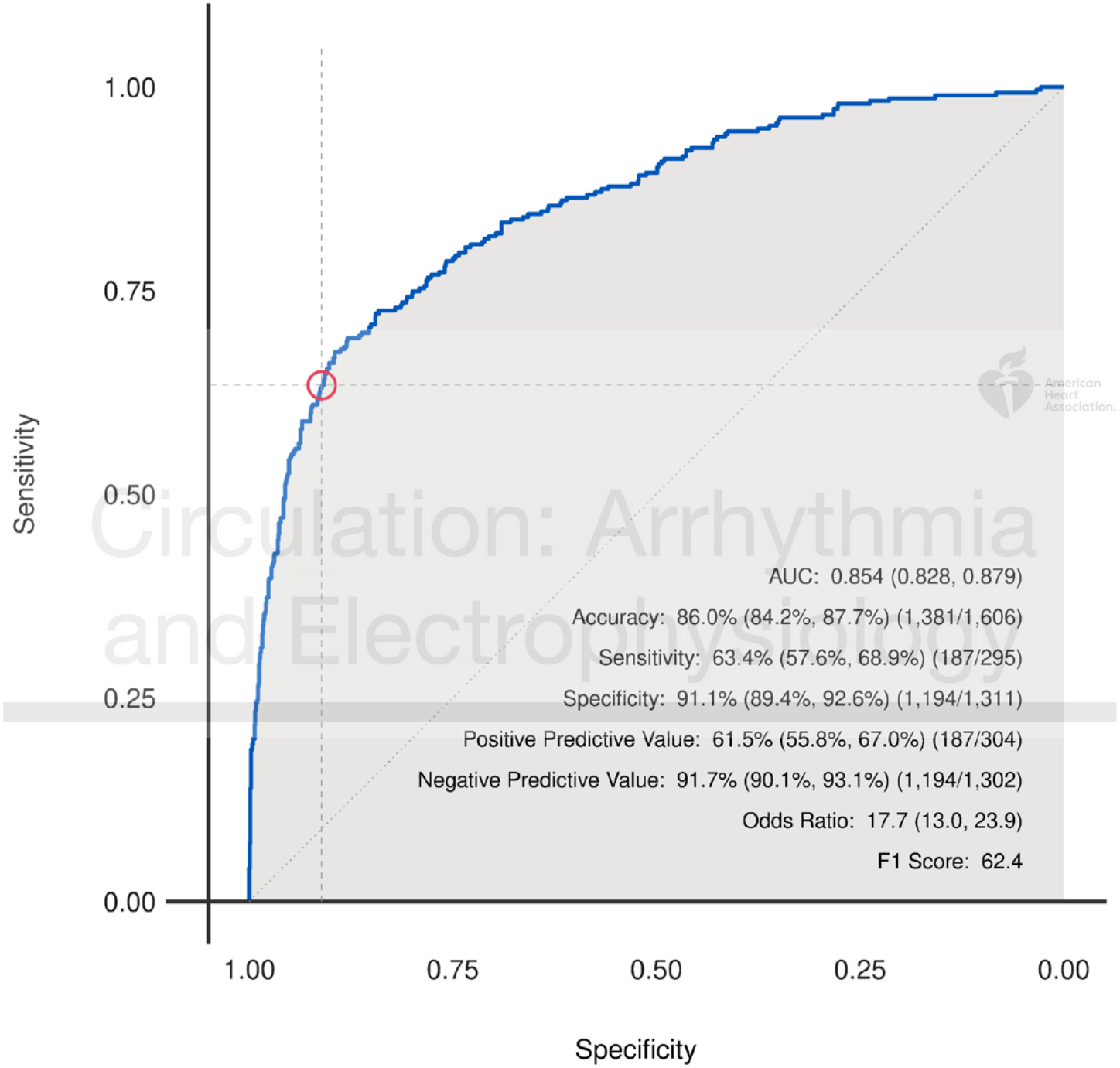
- Artificial Intelligence (AI) algorithms can predict left ventricular dysfunction using a 12-lead ECG
- The clinical applications of AI algorithms in routine clinical practice remain unclear

What the Study Adds:

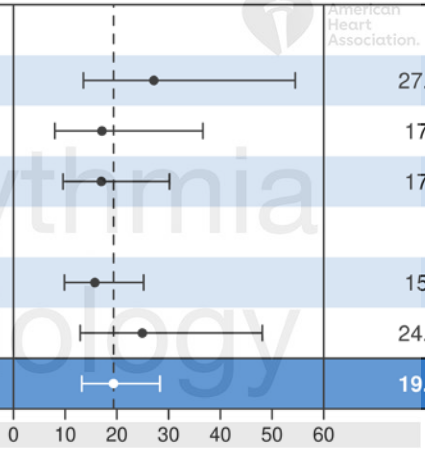
- An AI-enabled ECG algorithm applied retrospectively to a sample of patients evaluated in an acute care setting for dyspnea can reliably identify left ventricular dysfunction
- An AI-enabled ECG outperforms NT-Pro BNP in identifying left ventricular dysfunction in the emergency room
- Utilization of an AI-enabled ECG in the acute care setting is feasible



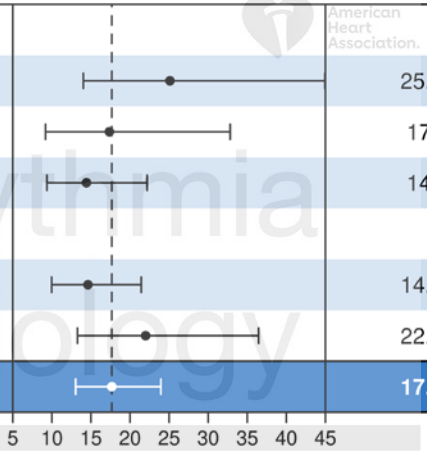




Group	AUC	Sensitivity	Specificity	Odds Ratio
Age Group				
[18,59]	0.909 (0.864, 0.955)	75.5% (61.7%, 86.2%)	89.8% (86.7%, 92.5%)	27.2 (13.5, 54.5)
(59,69]	0.875 (0.819, 0.932)	68.4% (51.3%, 82.5%)	88.8% (85.0%, 91.9%)	17.1 (8.0, 36.6)
(69,89]	0.871 (0.827, 0.916)	75.3% (63.9%, 84.7%)	84.8% (81.7%, 87.5%)	17.0 (9.6, 30.2)
Sex				
Male	0.869 (0.830, 0.907)	73.9% (64.7%, 81.8%)	84.8% (82.0%, 87.3%)	15.8 (9.9, 25.2)
Female	0.904 (0.863, 0.944)	73.6% (59.7%, 84.7%)	89.9% (87.5%, 92.1%)	24.9 (12.9, 48.1)
Overall	**0.885 (0.857, 0.913)**	**73.8% (66.4%, 80.3%)**	**87.3% (85.5%, 89.0%)**	**19.4 (13.2, 28.3)**



Group	AUC	Sensitivity	Specificity	Odds Ratio
Age Group				
[18,59]	0.893 (0.854, 0.931)	65.9% (54.8%, 75.8%)	92.9% (90.0%, 95.1%)	25.1 (14.0, 44.9)
(59,69]	0.842 (0.787, 0.896)	59.7% (47.0%, 71.5%)	92.1% (88.6%, 94.8%)	17.4 (9.2, 32.8)
(69,89]	0.832 (0.792, 0.873)	63.6% (55.2%, 71.5%)	89.2% (86.3%, 91.6%)	14.4 (9.4, 22.2)
Sex				
Male	0.834 (0.800, 0.869)	63.5% (56.3%, 70.2%)	89.4% (86.8%, 91.6%)	14.6 (10.0, 21.4)
Female	0.875 (0.834, 0.915)	63.3% (52.9%, 72.8%)	92.7% (90.5%, 94.6%)	22.0 (13.3, 36.4)
Overall	**0.854 (0.828, 0.879)**	**63.4% (57.6%, 68.9%)**	**91.1% (89.4%, 92.6%)**	**17.7 (13.0, 23.9)**



Graphic Abstract