**SUPPLEMENTAL CONTENT**

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

Lovedeep S Dhingra MBBS, Arya Aminorroaya MD, MPH, Aline Pedroso Camargos PhD, Akshay Khunte BS, Veer Sangha BS, Daniel McIntyre MD, PhD, Clara K Chow MBBS, PhD, Folkert W Asselbergs MD, PhD, Luisa CC Brant MD, PhD, Sandhi M Barreto MD, PhD, Antonio Luiz P Ribeiro MD, PhD, Harlan M Krumholz MD, SM, Evangelos K Oikonomou MD, DPhil, Rohan Khera MD, MS

Table of Contents

[eMethods 3](#_Toc167647319)

[eFigure 1. Overview of Cohort Creation at the Yale New Haven Health System. 7](#_Toc167647320)

[eFigure 2. Consort Diagram for Study Cohorts. 8](#_Toc167647321)

[eFigure 3. Net Benefit of AI-ECG Model Output Probability and Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score Across Thresholds at (A) Yale New Haven Health System (B) UK Biobank (C) Brazilian Longitudinal Study of Adult Health. 9](#_Toc167647322)

[eFigure 4. Cumulative Hazard for Heart Failure Adjusted for Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score at (A) Yale New Haven Health System (B) UK Biobank (C) Brazilian Longitudinal Study of Adult Health. 10](#_Toc167647323)

[eFigure 5. Cumulative Hazard for Heart Failure Adjusted for Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score at Yale New Haven Health System. 11](#_Toc167647324)

[eTable 1. Model Performance for Cross-sectional Detection of Left Ventricular Systolic Dysfunction. 12](#_Toc167647325)

[eTable 2. International Classification of Disease Tenth Revision Codes for the Identification of Comorbidities and Outcomes. 13](#_Toc167647326)

[eTable 3. Population Characteristics of the Study Cohorts. 14](#_Toc167647327)

[eTable 4. Population Characteristics of the Yale New Haven Health System Sites. 15](#_Toc167647328)

[eTable 5. Model Performance for Predicting Heart Failure Risk. 16](#_Toc167647329)

[eTable 6. Model Performance for Prediction of Heart Failure Across Demographic Subgroups. 17](#_Toc167647330)

[eTable 7. Age- and Sex- Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure-related Outcomes. 18](#_Toc167647331)

[eTable 8. Age- and Sex- Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure-related Outcomes. 19](#_Toc167647332)

[eTable 9. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure across Model Output Probabilities 20](#_Toc167647333)

[eTable 10. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure across Model Output Probabilities at the Yale New Haven Health System Sites. 21](#_Toc167647334)

[eTable 11. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Non-Heart Failure Clinical Outcomes. 22](#_Toc167647335)

[eTable 12. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Non-Heart Failure Clinical Outcomes Across Yale New Haven Health System Sites. 23](#_Toc167647336)

# **eMethods**

**Data Sources**

The Yale New Haven Health System (YNHHS) is the largest referral center in southern New England and serves a diverse patient population. The YNHHS includes five hospitals, Yale New Haven Hospital, Bridgeport Hospital, Greenwich Hospital, Lawrence and Memorial Hospital, and Westerly Hospital, and a large network of community outpatient clinics, the Northeast Medical Group. The electronic health records (EHR) data was acquired during patient care at YNHHS using Epic and was extracted from the Clarity database.1,2

UK Biobank (UKB) is a prospective cohort of 502,468 community-dwelling adults aged 40-69 years recruited during 2006-2010.3 A group of these participants accepted to participate in the third or fourth UKB study visit during which the participants underwent 12-lead electrocardiograms (ECGs) in 2014-2021. The UKB dataset is linked with the national EHR from the UK National Health Service predating UKB enrollment, enabling access to EHR diagnosis codes.4,5 We used data from UKB under research application#71033.

 The Brazilian Longitudinal Study of Adult Health (ELSA-Brasil) study, a large multicenter prospective cohort study conducted in Brazil, enrolled,105 community-dwelling adults aged 35-74 years at their baseline visit during 2008-2010.6,7 These participants represent active and retired civil servants from six higher education and research institutions in Brazilian state capitals in three geographical regions of the country: Southeast (Belo Horizonte, Rio de Janeiro, São Paulo and Vitória), South (Porto Alegre) and Northeast (Salvador).8 The ELSA-Brasil study aimed to investigate the development and progression of chronic diseases and their determinants in the Brazilian adult population. Baseline data were collected using validated instruments, physical examinations, laboratory assessments, and imaging modalities.6 Additionally, all participants underwent protocolized 12-lead ECG and echocardiogram.6,7 To ascertain exposure status and to identify changes in baseline, ELSA-Brasil participants present for in-person follow-up visits every three to four years. Moreover, telephone interviews occur annually to obtain information on new diagnoses, hospitalization, and death with adjudicated clinical events based on expert medical record review.6

**Study Population**

In YNHHS, to identify patients with prevalent heart failure (HF) at the time of ECG, we identified the first recorded encounter for all patients within the EHR and followed for 1 year. Patients with prevalent HF based on either a diagnosis code for HF or an echocardiogram with left ventricular ejection fraction under 50% or left ventricular diastolic dysfunction (defined as “moderate” or “severe” left ventricular diastolic dysfunction) we excluded from the study. The baseline ECG for patients was defined as an outpatient ECG recorded after this 1-year blanking period to exclude prevalent HF (e**Figure 1**). The YNHHS cohort also excluded patients previously included in the development of the AI-ECG algorithm and a small proportion of individuals who opted out of research participation (<0.01% of all YNHHS patients).

**Study Exposure**

The model development population consisted of 503,516 ECGs from 110,228 unique patients (**Figure 1**). We used raw voltage data from 12-lead ECGs obtained as standard 10-second 12-lead ECGs at a sampling frequency of 500 Hz or 250 Hz and extracted the lead I waveforms. Median filtering was conducted by subtracting a one-second median filter from the acquired signals to eliminate baseline drift. To incorporate noise during the model development, we isolated four distinct noises from a 5-minute random Gaussian noise within four frequency ranges of 3-12 Hz, 12-50 Hz, 50-100 Hz, and 100-150 Hz, each corresponding to the frequency range of a specific type of real-world noise.9 Each ECG in the training set was included twice with different random noises signal-to-noise ratios.

**Study Outcomes and Covariates**

In ELSA-Brasil, HF was identified either by in-person interview or the annual telephonic surveillance and investigated by a designated committee that contacted health providers and requested copies of medical records for all hospitalizations. After investigation, the cardiovascular events were adjudicated by an independent review of two cardiologists. A third senior cardiologist defined the event in case of disagreement.10 HF was identified from hospitalization records, based on the presence of a clinical diagnosis of HF, with the individual receiving pharmacological therapy for HF, in addition to any of the following: (1) pulmonary congestion on chest X-ray, (2) reduced ejection fraction or systolic dysfunction observed on cardiac imaging, or (3) preserved ejection fraction with evidence of moderate to severe diastolic dysfunction.

Information about all-cause death was available in the YNHHS EHR, with in-hospital mortality data supplemented from the Connecticut death index to improve capturing out-of-hospital patient mortality. Similarly, information about mortality was available in UKB via linkage to the EHR and the UK national death registries. Information about death in the ELSA-Brasil study was recorded via telephonic surveillance and confirmed using the national mortality database and death certificates.

**Study Comparator**

The pooled cohort equations to prevent HF (PCP-HF) represent sex- and race-specific equations for estimating the 10-year risk of incident HF. To align with the score development, across cohorts, the PCP-HF score was calculated for White and Black individuals between 30 and 80 years of age with complete documentation of the score covariates. The calculated 10-year risk score was adjusted based on the length of follow-up for each individual to estimate the risk of HF over the study period.

 In YNHHS, PCP-HF features were extracted from the EHR. Body mass index (BMI), systolic blood pressure, and laboratory measurements closest to and within two years of the ECG acquisition date were used for calculation. In UKB, the demographic features were identified from the baseline visit. Blood pressure measurement and smoking status assessment were conducted at the time of ECG acquisition. Laboratory values were measured in the first and second study visits, while ECGs were recorded in third and fourth visits.11 We used the laboratory values closest to the ECG acquisition for the calculation of PCP-HF score. History of hypertension and diabetes were defined using ICD diagnosis codes from the linked EHR and self-reported use of anti-hypertensive and anti-hyperglycemic medications was recorded at the time of ECG acquisition.4 In ELSA-Brasil, all PCP-HF features, including the ECG recording, were captured at the baseline visit using established study protocols.12,13

**Statistical Analysis**

Integrated discrimination improvement (IDI) was calculated as the difference between the improvements in the average predicted probabilities for those with and without the outcome across the AI-ECG and the PCP-HF probabilities. Categorical net reclassification improvement (NRI) was calculated for the 0.08 threshold of the AI-ECG model. Net benefit evaluates true positives while accounting for the potential for increased false positives, ranging from 0-1 with higher values showing greater benefit. This was calculated using the following formula:

$$Net Benefit = \left(\frac{True Positive}{Number}\right)-\left(\frac{Probability Threshold}{1-Probability Threshold}\right)×\left(\frac{False Positive}{Number}\right)$$

Categorical variables were reported as counts (percentages), and continuous variables as median (interquartile range [IQR]).All statistical tests were 2-sided with a level of significance set at 0.05. compareC package in R was used for calculating and comparing Harrel's C-statistics.14 All analyses were conducted using a combination of Python 3.11.2 and R version 4.2.0. The Yale Institutional Review Board approved the study protocol and waived the need for informed consent as the study involves secondary analysis of pre-existing data.

**Online References**

1. Chishtie J, Sapiro N, Wiebe N, et al. Use of Epic Electronic health record system for health care research: Scoping review. *J Med Internet Res*. 2023;25(1):e51003.

2. 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.

3. Palmer LJ. UK Biobank: bank on it. *Lancet*. 2007;369(9578):1980-1982.

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

5. Raisi-Estabragh Z, Petersen SE. Cardiovascular research highlights from the UK Biobank: opportunities and challenges. *Cardiovasc Res*. 2020;116(1):e12-e15.

6. Aquino EML, Barreto SM, Bensenor IM, et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. *Am J Epidemiol*. 2012;175(4):315-324.

7. 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.

8. Aquino EML, Araujo MJ, Almeida M da CC, et al. Recrutamento de participantes no Estudo Longitudinal de Saúde do Adulto. *Rev Saúde Pública*. 2013;47:10-18.

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

10. Barreto SM, Ladeira RM, Bastos M do SCB de O, et al. ELSA-Brasil strategies for outcome identification, investigation and ascertainment. *Rev Saude Publica*. 2013;47 Suppl 2:79-86.

11. 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.

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

13. 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.

14. 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.

**eFigure 1. Overview of Cohort Creation at the Yale New Haven Health System.** Abbreviations: ECG, Electrocardiograms; EF, Ejection Fraction; HF, Heart Failure; YNHHS, Yale New Haven Health System.

****

**eFigure 2. Consort Diagram for Study Cohorts.** Abbreviations: ECG, Electrocardiogram; ELSA Brasil, Brazilian Longitudinal Study of Adult Health; HF, Heart Failure; YNHHS, Yale New Haven Health System



**eFigure 3. Net Benefit of AI-ECG Model Output Probability and Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score Across Thresholds at (A) Yale New Haven Health System (B) UK Biobank (C) Brazilian Longitudinal Study of Adult Health.** Abbreviations: AI-ECG, Artificial Intelligence-enhanced Electrocardiography; 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

****

**eFigure 4. Cumulative Hazard for Heart Failure Adjusted for Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score at (A) Yale New Haven Health System (B) UK Biobank (C) Brazilian Longitudinal Study of Adult Health.** Abbreviations: aHR, Adjusted Hazard Ratio; ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; UKB, UK Biobank; YNHHS, Yale New Haven Health System



**eFigure 5. Cumulative Hazard for Heart Failure Adjusted for Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) Risk Score at Yale New Haven Health System.** Abbreviations: aHR, Adjusted Hazard Ratio



**eTable 1. International Classification of Disease Tenth Revision Codes for the Identification of Comorbidities and Outcomes.** Abbreviations: ICD-10-CM, International Classification of Disease Tenth Revision Clinical Modification Codes.

|  |  |
| --- | --- |
| **Condition** | **ICD-10-CM codes** |
| **Heart Failure** | 'I11.0','I13.0','I13.2','I50','I50.0','I50.1','I50.9','Z95.81','I09.81' |
| **Acute Myocardial Infarction** | 'I21', 'I22', 'I23', 'I24.0', 'I24.8', 'I24.9' |
| **Stroke** | 'G45','G45.0','G45.1','G45.2','G45.3','G45.4','G45.8','G45.9','I63','I63.0','I63.1','I63.2','I63.3','I63.4','I63.5','I63.8','I63.9','I64','I65','I65.0','I65.1','I65.2','I65.3','I65.8','I65.9','I66','I66.0','I66.1','I66.2','I66.3','I66.4','I66.8','I66.9','I67.2','I69.3','I69.4' |
| **Type 2 Diabetes Mellitus** | 'E11','E11.0','E11.1','E11.2','E11.3','E11.4','E11.5','E11.6','E11.7','E11.8','E11.9','O24.1' |
| **Hypertension** | 'I10','I11','I11.0','I11.9','I12','I12.0','I12.9','I13','I13.0','I13.1','I13.2','I13.9','I67.4','O10','O10.0','O10.1','O10.2','O10.3','O10.9','O11' |
| **Ischemic Heart Disease** | 'I20', 'I20.0', 'I20.8', 'I20.9', 'I21', 'I21.0', 'I21.1', 'I21.2', 'I21.3', 'I21.4', 'I21.9', 'I21.X', 'I22', 'I22.0', 'I22.1', 'I22.8', 'I22.9', 'I23', 'I23.0', 'I23.1', 'I23.2', 'I23.3', 'I23.4', 'I23.5', 'I23.6', 'I23.8', 'I24', 'I24.0', 'I24.1', 'I24.8', 'I24.9', 'I25', 'I25.0', 'I25.1', 'I25.2', 'I25.5', 'I25.6', 'I25.8', 'I25.9', 'Z95.1', 'Z95.5' |

**eTable 2. Model Performance for Cross-sectional Detection of Left Ventricular Systolic Dysfunction.** Abbreviations: AUPRC, Area Under the Precision Recall Curve; AUROC, Area Under the Receiver Operating Characteristic Curve; NPV, Negative Predictive Value; PPV, Positive Predictive Value.

|  |  |
| --- | --- |
| **Metric** | **Value** |
| **Number of Individuals in Held-out Test Set** | 11203 |
| **Prevalence of Left Ventricular Systolic Dysfunction** | 8.9% |
| **Model Probability Threshold** | 0.08 |
| **AUROC** | 0.899 (0.888-0.909) |
| **AUPRC** | 0.451 (0.417-0.487) |
| **Sensitivity** | 0.904 |
| **Specificity** | 0.563 |
| **PPV** | 0.169 |
| **NPV** | 0.984 |
| **F1 Score** | 0.286 |

**eTable 3. Population Characteristics of the Study Cohorts.** Abbreviations: AMI, acute myocardial infarction; ECG, Electrocardiogram; ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; HF, heart failure; IQR, Interquartile Range; LVEF, Left Ventricular Ejection Fraction; UKB, UK Biobank; YNHHS, Yale New Haven Health System

|  |  |  |  |
| --- | --- | --- | --- |
| **Characteristic** | **YNHHS** | **UKB** | **ELSA-Brasil** |
| **Number** | 194340 | 42141 | 13454 |
| **Age at ECG, Median [IQR]** | 56 [41,69] | 65 [59,71] | 51 [45,58] |
| **Female Sex, N (%)** | 112082 (57.7) | 21795 (51.7) | 7348 (54.6) |
| **Race/Ethnicity,****N (%)** | **White** | 119051 (61.3) | 40691 (96.6) | 6920 (51.4) |
| **Black** | 30840 (15.9) | 304 (0.7) | 2130 (15.8) |
| **Hispanic** | 33456 (17.2) | - | - |
| **Asian** | 3573 (1.8) | 600 (1.4) | - |
| **Other** | 2172 (1.1) | 546 (1.3) | 637 (4.7) |
| **Brazilian “Pardo”** | - | - | 3767 (28.0) |
| **Missing** | 5248 (2.7) | - | - |
| **Death, N (%)** | 10747 (5.5) | 346 (0.8) | 229 (1.7) |
| **Follow-up Time, Years; Median [IQR]** | 4.5 [2.6,6.6] | 3.1 [2.1,4.5] | 4.2 [3.7, 4.5] |
| **Positive Screens, N (%)** | 43559 (22.4) | 5513 (13.1) | 1928 (14.3) |
| **Primary HF hospitalization during follow-up, N (%)** | 3929 (2.0) | 46 (0.1) | 31 (0.2) |
| **Primary HF hospitalization or an echocardiogram with LVEF < 50% during follow-up, N (%)** | 7847 (4.0) | - | - |
| **Any HF hospitalization during follow-up, N (%)** | 14355 (7.4) | 231 (0.5) | - |
| **Any HF hospitalization or an echocardiogram with LVEF < 50% during follow-up, N (%)** | 16380 (8.4) | - | - |
| **Primary AMI hospitalization during follow-up, N (%)** | 375 (0.2) | 208 (0.5) | 60 (0.4) |
| **Primary Stroke hospitalization during follow-up, N (%)** | 3360 (1.7) | 210 (0.5) | 59 (0.4) |
| **Major Adverse Cardiovascular Events during follow-up, N (%)** | 16618 (8.6) | 768 (1.8) | 338 (2.5) |

**eTable 4. Population Characteristics of the Yale New Haven Health System Sites.** Abbreviations: AMI, acute myocardial infarction; ECG, Electrocardiogram; HF, heart failure; IQR, Interquartile Range; LVEF, Left Ventricular Ejection Fraction; NEMG, Northeast Medical Group; L&M, Lawrence and Memorial Hospital; YNH, Yale New Haven Hospital

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Characteristic** | **YNH** | **Bridgeport** | **Greenwich** | **L&M** | **Westerly** | **NEMG** |
| **Number** | 96929 | 32696 | 17980 | 19250 | 4683 | 22802 |
| **Age at ECG, Median [IQR]** | 54.4 [38.8,67.2] | 51.6 [37.5,64.2] | 56.2 [42.8,70.6] | 58.4 [43.6,71.1] | 65.9 [54.2,76.3] | 64.2 [54.3,73.8] |
| **Female Sex, N (%)** | 55887 (57.7) | 19868 (60.8) | 10542 (58.6) | 11333 (58.9) | 2534 (54.1) | 11918 (52.3) |
| **Race/Ethnicity,****N (%)** | **White** | 57644 (59.5) | 12614 (38.6) | 12277 (68.3) | 13716 (71.3) | 4333 (92.5) | 18467 (81.0) |
| **Black** | 18392 (19.0) | 8313 (25.4) | 978 (5.4) | 1740 (9.0) | 83 (1.8) | 1334 (5.9) |
| **Hispanic** | 15015 (15.5) | 10373 (31.7) | 3598 (20.0) | 2737 (14.2) | 122 (2.6) | 1611 (7.1) |
| **Asian** | 2044 (2.1) | 351 (1.1) | 512 (2.8) | 267 (1.4) | 35 (0.7) | 364 (1.6) |
| **Other** | 1141 (1.2) | 314 (1.0) | 138 (0.8) | 377 (2.0) | 58 (1.2) | 144 (0.6) |
| **Missing** | 2693 (2.8) | 731 (2.2) | 477 (2.7) | 413 (2.1) | 52 (1.1) | 882 (3.9) |
| **Death, N (%)** | 5198 (5.4) | 1660 (5.1) | 869 (4.8) | 1200 (6.2) | 409 (8.7) | 1411 (6.2) |
| **Follow-up Time, Years; Median [IQR]** | 4.9 [2.7,6.9] | 4.8 [3.1,6.6] | 5.3 [3.6,6.8] | 3.4 [2.0,4.7] | 3.3 [1.5,4.6] | 4.7 [3.0,6.7] |
| **Positive Screens, N (%)** | 20958 (21.6) | 7087 (21.7) | 3333 (18.5) | 4673 (24.3) | 1422 (30.4) | 6086 (26.7) |
| **Primary HF hospitalization during follow-up, N (%)** | 1515 (1.6) | 636 (1.9) | 355 (2.0) | 559 (2.9) | 252 (5.4) | 612 (2.7) |
| **Primary HF hospitalization or an echocardiogram with LVEF < 50% during follow-up, N (%)** | 3003 (3.1) | 1252 (3.8) | 517 (2.9) | 1031 (5.4) | 361 (7.7) | 1683 (7.4) |
| **Any HF hospitalization during follow-up, N (%)** | 6526 (6.7) | 2192 (6.7) | 1026 (5.7) | 1517 (7.9) | 659 (14.1) | 2435 (10.7) |
| **Any HF hospitalization or an echocardiogram with LVEF < 50% during follow-up, N (%)** | 7271 (7.5) | 2492 (7.6) | 1105 (6.1) | 1766 (9.2) | 714 (15.2) | 3032 (13.3) |
| **Primary AMI hospitalization during follow-up, N (%)** | 51 (0.1) | 49 (0.1) | 17 (0.1) | 39 (0.2) | 22 (0.5) | 197 (0.9) |
| **Primary Stroke hospitalization during follow-up, N (%)** | 1225 (1.3) | 556 (1.7) | 386 (2.1) | 536 (2.8) | 169 (3.6) | 488 (2.1) |
| **Major Adverse Cardiovascular Events during follow-up, N (%)** | 7335 (7.6) | 2627 (8.0) | 1466 (8.2) | 2065 (10.7) | 734 (15.7) | 2391 (10.5) |

**eTable 5. Model Performance for Predicting Heart Failure Risk.** Abbreviations: IHD, Ischemic Heart Disease; HTN, hypertension; NEMG, Northeast Medical Group; L&M, Lawrence and Memorial Hospital; T2DM, type-2 diabetes mellitus; YNH, Yale New Haven Hospital

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model**  | **Covariates** | **YNH** | **Bridgeport** | **Greenwich** | **L&M** | **Westerly** | **NEMG** |
| **Positive Screen** | **Per 0.1 Increment** | **Positive Screen** | **Per 0.1 Increment** | **Positive Screen** | **Per 0.1 Increment** | **Positive Screen** | **Per 0.1 Increment** | **Positive Screen** | **Per 0.1 Increment** | **Positive Screen** | **Per 0.1 Increment** |
| **Cox Proportional Hazard Model** | Model Probability | 5.04 (4.55-5.58) | 1.47 (1.44-1.50) | 5.20 (4.45-6.09) | 1.46 (1.42-1.50) | 5.68 (4.61-7.00) | 1.53 (1.47-1.59) | 4.39 (3.71-5.20) | 1.43 (1.39-1.48) | 4.24 (3.29-5.48) | 1.34 (1.28-1.40) | 4.39 (3.73-5.17) | 1.37 (1.33-1.41) |
| **Cox Proportional Hazard Model** | Model Probability + Age + Sex | 3.52 (3.17-3.91) | 1.34 (1.32-1.37) | 3.36 (2.85-3.95) | 1.33 (1.29-1.37) | 2.86 (2.30-3.55) | 1.30 (1.24-1.36) | 2.87 (2.41-3.42) | 1.31 (1.27-1.36) | 3.14 (2.42-4.07) | 1.24 (1.18-1.30) | 3.02 (2.55-3.57) | 1.27 (1.23-1.31) |
| **Cox Proportional Hazard Model** | Model Probability + Age + Sex + IHD + HTN + T2DM + Obesity | 3.02 (2.71-3.36) | 1.31 (1.28-1.33) | 2.76 (2.34-3.25) | 1.28 (1.24-1.32) | 2.50 (2.01-3.10) | 1.25 (1.19-1.31) | 2.41 (2.02-2.87) | 1.28 (1.24-1.33) | 2.89 (2.22-3.75) | 1.22 (1.16-1.28) | 2.63 (2.23-3.12) | 1.24 (1.21-1.28) |
| **Fine-Gray Subdistribution Hazard Model** | Model Probability + Age + Sex + Competing Risk of Death | 3.38 (3.09-3.69) | 1.32 (1.30-1.34) | 3.30 (2.92-3.73) | 1.31 (1.28-1.34) | 2.88 (2.45-3.37) | 1.28 (1.25-1.32) | 3.01 (2.63-3.45) | 1.30 (1.27-1.33) | 3.04 (2.47-3.74) | 1.23 (1.18-1.27) | 2.96 (2.49-3.52) | 1.26 (1.22-1.30) |
| **Fine-Gray Subdistribution Hazard Model** | Model Probability Age + Sex + IHD + HTN + T2DM + Obesity + Competing Risk of Death | 2.93 (2.62-3.27) | 1.30 (1.27-1.32) | 2.71 (2.28-3.21) | 1.27 (1.23-1.31) | 2.39 (1.90-3.00) | 1.24 (1.18-1.30) | 2.32 (1.94-2.78) | 1.26 (1.22-1.31) | 2.86 (2.19-3.72) | 1.20 (1.14-1.27) | 2.57 (2.16-3.06) | 1.23 (1.19-1.27) |

**eTable 6. Model Performance for Prediction of Heart Failure Across Demographic Subgroups.** Abbreviations: CI, Confidence Interval; ELSA-Brasil, Brazilian Longitudinal Study of Adult Health; HF, Heart Failure; UKB, UK Biobank; YNHHS, Yale New Haven Health System

|  |  |  |  |
| --- | --- | --- | --- |
| **Subgroup** | **YNHHS** | **UKB** | **ELSA-Brasil** |
| **Total Number of Individuals at Risk** | **Number of HF Events** | **Age- and Sex- Adjusted Cox Proportional Hazard Ratios (95% CI)** | **Total Number of Individuals at Risk** | **Number of HF Events** | **Age- and Sex- Adjusted Cox Proportional Hazard Ratios (95% CI)** | **Total Number of Individuals at Risk** | **Number of HF Events** | **Age- and Sex- Adjusted Cox Proportional Hazard Ratios (95% CI)** |
| **Age < 65** | 131829 (67.8) | 1107 | 4.96 (4.40-5.60) | 20802 (49.4) | 9 | 6.30 (1.68-23.57) | 12038 (89.4) | 21 | 11.41 (4.55-28.64) |
| **Age ≥ 65** | 62511 (32.2) | 2822 | 2.79 (2.58-3.01) | 21345 (50.6) | 37 | 5.86 (3.06-11.23) | 1416 (10.6) | 10 | 4.69 (1.31-16.71) |
| **Female** | 112082 (57.7) | 2013 | 2.94 (2.69-3.21) | 21795 (51.7) | 11 | 3.76 (1.10-12.88) | 7348 (54.6) | 11 | 10.21 (3.88-26.91) |
| **Male** | 82258 (42.3) | 1916 | 3.8 (3.45-4.19) | 20346 (48.3) | 35 | 6.92 (3.52-13.60) | 6106 (45.4) | 20 | 6.76 (2.03-22.47) |
| **Race/Ethnicity** | **White** | 119051 (61.3) | 2942 | 3.03 (2.81-3.27) | 40691 (96.6) | 46 |  | 6920 (51.4) | 15 | 10.17 (3.40-30.46) |
| **Black** | 30840 (15.9) | 500 | 4.35 (3.60-5.26) | 304 (0.7) | 0 | - | 2130 (15.8) | 9 | 11.80 (2.38-58.47) |
| **Hispanic** | 33456 (17.2) | 359 | 4.40 (3.54-5.47) | 0 | - | - | - | - | - |
| **Asian** | 3573 (1.8) | 36 | 4.69 (2.35-9.37) | 600 (1.4) | 0 | - | - | - | - |
| **Other** | 2172 (1.1) | 28 | 2.87 (1.34-6.17) | 546 (1.3) | 0 | - | 637 (4.7) | 0 | - |
| **Brazilian “Pardo”** | - | - | - | - | - | - | 3767 (28.0) | 7 | 3.69 (0.81-16.8) |
| **Missing** | 5248 (2.7) | 64 | 2.38 (1.43-3.97) | 0 | - | - | - | - | - |

**eTable 7. Age- and Sex- Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure-related Outcomes.** Abbreviations: AMI, acute myocardial infarction; ELSA Brasil, Brazilian Longitudinal Study of Adult Health; HF, heart failure; LVEF, Left Ventricular Ejection Fraction; UKB, UK Biobank; YNHHS, Yale New Haven Health System

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **YNHHS** | **UKB** | **ELSA-Brasil** |
| **Primary HF Hospitalization** | 3.32 (3.11-3.54) | 5.96 (3.32-10.68) | 8.74 (4.13-18.48) |
| **Primary HF Hospitalization or an Echocardiogram with LVEF < 50%** | 3.88 (3.70-4.06) | - | - |
| **Primary HF Hospitalization or Death** | 1.98 (1.91-2.05) | 1.91 (1.51-2.41) | 2.59 (1.99-3.35) |
| **Any Hospitalization with HF**  | 2.43 (2.35-2.52) | 3.55 (2.72-4.64) | - |
| **Any Hospitalization with HF or an Echocardiogram with LVEF < 50%** | 2.62 (2.54-2.71) | - | - |
| **Any HF Hospitalization or Death** | 2.07 (2.01-2.13) | 2.18 (1.81-2.63) | - |

**eTable 8. Age- and Sex- Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure-related Outcomes.** Abbreviations: AMI, acute myocardial infarction; HF, heart failure; NEMG, Northeast Medical Group; LVEF, Left Ventricular Ejection Fraction; L&M, Lawrence and Memorial Hospital; YNH, Yale New Haven Hospital

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Outcome** | **YNH** | **Bridgeport** | **Greenwich** | **L&M** | **Westerly** | **NEMG** |
| **Primary HF Hospitalization** | 3.52 (3.17-3.91) | 3.36 (2.85-3.95) | 2.86 (2.3-3.55) | 2.87 (2.41-3.42) | 3.14 (2.42-4.07) | 3.02 (2.55-3.57) |
| **Primary HF Hospitalization or an Echocardiogram with LVEF < 50%** | 3.81 (3.54-4.11) | 3.64 (3.24-4.09) | 2.80 (2.34-3.35) | 3.42 (3.01-3.89) | 2.76 (2.23-3.42) | 4.83 (4.35-5.37) |
| **Primary HF Hospitalization or Death** | 1.90 (1.81-2.00) | 1.85 (1.70-2.02) | 1.99 (1.76-2.24) | 2.18 (1.97-2.42) | 1.76 (1.49-2.07) | 2.09 (1.90-2.30) |
| **Any Hospitalization with HF**  | 2.42 (2.31-2.55) | 2.47 (2.26-2.69) | 2.20 (1.94-2.50) | 2.65 (2.38-2.94) | 1.95 (1.67-2.28) | 2.27 (2.09-2.46) |
| **Any Hospitalization with HF or an Echocardiogram with LVEF < 50%** | 2.54 (2.42-2.66) | 2.56 (2.36-2.77) | 2.22 (1.96-2.50) | 2.79 (2.53-3.07) | 1.93 (1.66-2.24) | 2.87 (2.66-3.08) |
| **Any HF Hospitalization or Death** | 2.03 (1.95-2.11) | 2.01 (1.87-2.15) | 1.97 (1.78-2.18) | 2.28 (2.09-2.48) | 1.73 (1.51-1.98) | 2.06 (1.92-2.21) |

**eTable 9. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure across Model Output Probabilities.** Abbreviations: ELSA Brasil, Brazilian Longitudinal Study of Adult Health; UKB, UK Biobank; YNHHS, Yale New Haven Health System

|  |  |  |  |
| --- | --- | --- | --- |
| **Model output probability bins** | **YNHHS** | **UKB** | **ELSA-Brasil** |
| **0-0.2** | Reference | Reference | Reference |
| **0.2-0.4** | 2.9 (2.66-3.17) | 4.94 (2.05-11.93) | 6.98 (2.34-20.82) |
| **0.4-0.6** | 3.52 (3.14-3.95) | 9.11 (3.52-23.57) | 14.83 (4.32-50.95) |
| **0.6-0.8** | 4.95 (4.37-5.6) | 11.87 (4.15-33.9) | 31.17 (10.27-94.66) |
| **0.8-1.0** | 7.12 (6.04-8.39) | 13.68 (1.86-100.45) | 112.33 (25.75-489.96) |

**eTable 10. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Heart Failure across Model Output Probabilities at the Yale New Haven Health System Sites.** Abbreviations: NEMG, Northeast Medical Group; L&M, Lawrence and Memorial Hospital; YNH, Yale New Haven Hospital

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model output probability bins** | **YNH** | **Bridgeport** | **Greenwich** | **L&M** | **Westerly** | **NEMG** |
| **0-0.2** | Reference | Reference | Reference | Reference | Reference | Reference |
| **0.2-0.4** | 3.56 (3.09-4.10) | 2.31 (1.82-2.93) | 2.66 (1.98-3.58) | 2.3 (1.79-2.96) | 3.00 (2.18-4.11) | 2.39 (1.92-2.98) |
| **0.4-0.6** | 3.70 (3.06-4.46) | 3.91 (2.92-5.23) | 3.74 (2.56-5.47) | 3.59 (2.71-4.76) | 2.26 (1.35-3.80) | 2.98 (2.23-3.97) |
| **0.6-0.8** | 5.83 (4.80-7.07) | 5.49 (3.99-7.56) | 4.22 (2.66-6.69) | 4.42 (3.19-6.11) | 3.89 (2.44-6.20) | 3.73 (2.70-5.14) |
| **0.8-1.0** | 9.21 (7.16-11.85) | 6.16 (3.99-9.50) | 5.71 (2.81-11.59) | 9.47 (6.37-14.07) | 2.55 (1.04-6.23) | 4.96 (3.28-7.48) |

**eTable 11. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Non-Heart Failure Clinical Outcomes.** Abbreviations: AMI, acute myocardial infarction; ELSA Brasil, Brazilian Longitudinal Study of Adult Health; HF, heart failure; UKB, UK Biobank; YNHHS, Yale New Haven Health System

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **YNHHS** | **UKB** | **ELSA-Brasil** |
| **Primary AMI Hospitalization** | 1.12 (0.90-1.40) | 1.23 (0.85-1.76) | 3.00 (1.78-5.08) |
| **Primary Stroke Hospitalization** | 1.17 (1.09-1.26) | 1.54 (1.10-2.15) | 3.86 (2.28-6.51) |
| **All-cause Death** | 0.99 (0.95-1.03) | 1.58 (1.22-2.05) | 2.18 (1.65-2.88) |
| **Major Adverse Cardiovascular Events** | 1.77 (1.71-1.83) | 1.62 (1.37-1.93) | 2.68 (2.14-3.36) |

**eTable 12. Age- and Sex-Adjusted Cox Proportional Hazard Models for the Prediction of Non-Heart Failure Clinical Outcomes Across Yale New Haven Health System Sites.** Abbreviations: AMI, acute myocardial infarction; HF, heart failure; NEMG, Northeast Medical Group; L&M, Lawrence and Memorial Hospital; YNH, Yale New Haven Hospital

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Outcome** | **YNH** | **Bridgeport** | **Greenwich** | **L&M** | **Westerly** | **NEMG** |
| **Primary AMI Hospitalization** | 1.16 (0.72-1.86) | 1.45 (0.92-2.29) | 3.20 (1.62-6.29) | 2.03 (1.27-3.27) | 1.19 (0.56-2.54) | 0.74 (0.54-1.01) |
| **Primary Stroke Hospitalization** | 1.18 (1.06-1.31) | 1.23 (1.08-1.39) | 1.20 (1.02-1.41) | 1.15 (1.00-1.32) | 0.98 (0.75-1.27) | 1.02 (0.84-1.24) |
| **All-cause Death** | 1.08 (1.04-1.13) | 1.14 (1.08-1.21) | 1.23 (1.13-1.33) | 1.14 (1.05-1.24) | 1.08 (0.93-1.26) | 0.99 (0.89-1.10) |
| **Major Adverse Cardiovascular Events** | 1.93 (1.87-2.00) | 1.92 (1.83-2.02) | 1.93 (1.79-2.07) | 2.09 (1.96-2.23) | 1.76 (1.56-1.97) | 1.69 (1.56-1.84) |