



Published in final edited form as:

Circ Heart Fail. 2020 November ; 13(11): e007462. doi:10.1161/CIRCHEARTFAILURE.120.007462.

Predictive Accuracy of Heart Failure-Specific Risk Equations in an Electronic Health Record-Based Cohort

Aakash Bavishi, MD^a, Matthew Bruce, MD^a, Hongyan Ning, MD,MS^b, Priya M. Freaney, MD^c, Peter Glynn, MD^a, Faraz S. Ahmad, MD, MS^c, Clyde W. Yancy, MD, ScM^c, Sanjiv J. Shah, MD^c, Norrina B. Allen, PhD^b, Suma X. Vupputuri, PhD^d, Laura J. Rasmussen-Torvik, PhD^b, Donald M. Lloyd-Jones, MD, ScM^{b,c}, Sadiya S. Khan, MD, MSc^{b,c}

^aDepartment of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

^bDepartment of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

^cDivision of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

^dKaiser Permanente, Mid-Atlantic Permanente Research Institute, Rockville MD

Abstract

Background: Guidelines recommend identification of individuals at risk for heart failure (HF). However, implementation of risk-based prevention strategies requires validation of HF-specific risk scores in diverse, real-world cohorts. Therefore, our objective was to assess the predictive accuracy of the Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) within a primary prevention cohort derived from the electronic health record (EHR).

Methods: We retrospectively identified patients between the ages of 30–79 years in a multi-center integrated healthcare system, free of cardiovascular disease, with available data on HF risk factors, and at least 5 years of follow-up. We applied the PCP-HF tool to calculate sex and race-specific 5-year HF risk estimates. Incident HF was defined by *International Classification of Diseases* codes. We assessed model discrimination and calibration, comparing predicted and observed rates for incident HF.

Results: Among 31,256 eligible adults, mean age was 51.4 years, 57% were women and 11% Black. Incident HF occurred in 568 patients (1.8%) over 5-year follow-up. The modified PCP-HF model for 5-year risk prediction of HF had excellent discrimination in white men (c-statistic 0.82, 95% confidence interval [0.79, 0.86]) and women (0.82, [0.78, 0.87]) and adequate discrimination in Black men (0.69, [0.60, 0.78]) and women (0.69, [0.52, 0.76]). Calibration was fair in all race-sex subgroups ($\chi^2 < 20$).

Address for Correspondence: Sadiya S. Khan, MD, MS, FACC, Assistant Professor of Medicine, Division of Cardiology, Department of Medicine and Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 N. Lake Shore Drive, 14-002, Chicago, IL 60611, 312-503-2515, s-khan-1@northwestern.edu.

Disclosures: Dr. Shah has received research grants from Actelion, AstraZeneca, Corvia, and Novartis; and has served as a consultant/advisory board/steering committee member for Abbott, Actelion, AstraZeneca, Amgen, Bayer, Boehringer-Ingelheim, Cardiora, Coridea, CVRx, Eisai, Ionis, Ironwood, Merck, MyoKardia, Novartis, Pfizer, Sanofi, Tenax, and United Therapeutics. All other authors have no disclosures to report.

Conclusions: A novel sex and race-specific risk score predicts incident HF in a “real-world,” EHR-based cohort. Integration of HF risk into the EHR may allow for risk-based discussion, enhanced surveillance, and targeted preventive interventions to reduce the public health burden of HF.

Subject Terms:

Heart Failure; Prevention

Heart failure (HF) remains a significant public health problem with an estimated prevalence of 6.2 million people in the United States (US).¹ Given the aging population and the growing prevalence of obesity and diabetes mellitus (DM), the prevalence of HF is estimated to exceed 8 million in the US by the year 2030 with total costs projected to surpass \$70 billion.² Despite significant advances in treatment strategies over the past 2 decades, HF remains a leading cause of death with recent trends demonstrating increasing age-adjusted mortality rates and widening disparities between black and white patients since 2011.³ American College of Cardiology/American Heart Association (ACC/AHA) HF guidelines highlight the need for improved identification of high-risk patients and implementation of targeted primary prevention strategies given these troubling trends.⁴

We recently published a new tool to predict incident HF: the Pooled Cohort equations to Prevent Heart Failure (PCP-HF), which was derived in five and validated in two population-based cohorts representative of the general US population (N=33,010).⁵ Cardiovascular risk prediction models, including the PCP-HF model, have been typically derived and validated in longitudinal cohort studies.⁶ However, the performance of predictive models derived in longitudinal studies may vary when implemented in a real-world clinical population due to differences in population characteristics and biases associated with cohort selection, inclusion of healthy volunteers, and participation incentives.^{6, 7}

The use of data from the electronic health record (EHR) representing a real-world clinical setting offers an opportunity to validate the PCP-HF model in a population that could potentially demonstrate the greatest benefit from risk-based primary prevention of HF. This validation study could also facilitate the implementation of the PCP-HF tool into clinical practice and guide the primary prevention efforts in health systems. Risk estimates could be automatically calculated within the EHR in the form of clinical decision support tools, bypassing the need for manual data entry in web-based tools or apps.

Therefore, we sought to construct a diverse, contemporary EHR-based primary prevention cohort and examine the predictive accuracy of the PCP-HF model in this “real world” patient sample to assess readiness for translation and implementation in clinical practice.

Methods

The data that support the findings of the study are available from the corresponding author upon reasonable request.

Data Sources

The original PCP-HF model and modified equations were derived in the Cardiovascular Lifetime Risk Pooling Project (LRPP) and detailed description of the LRPP cohorts and individual participant-level data harmonization have been previously published.⁸ For our external validation cohort, we used data from the Northwestern Medicine Enterprise Data Warehouse (NMEDW), which houses comprehensive demographic, laboratory, and prescription data on patients as well as claims for inpatient or outpatient diagnoses and procedures. This study was approved by the Institutional Review Board of Northwestern University.

The PCP-HF Model

Details of the PCP-HF model have been published elsewhere.⁵ The equation includes the following variables to predict risk of HF: age, gender, race, current smoking, body mass index (BMI), systolic blood pressure (SBP) (treated or untreated), hypertension treatment, fasting glucose (treated or untreated), DM treatment, total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), and QRS duration. For the purposes of the current EHR-based validation study, a modified version of the PCP-HF equations without QRS duration were derived using the original Cardiovascular LRPP cohorts. This was necessitated by the high rate of missingness in electrocardiogram data (78%) in the EHR-based cohort. As electrocardiograms are not routinely indicated for screening, there were significant differences in risk factor profiles among patients who received an electrocardiogram compared with those who did not have an electrocardiogram (Online Table I). All coefficients were not significantly different compared with the original PCP-HF coefficients (Online Table II). We additionally internally validated the modified PCP-HF equations in the Cardiovascular LRPP and demonstrated good to excellent discrimination (Online Table III). The modified PCP-HF equations were then scaled to provide 5-year risk estimates, consistent with prior validation studies performed to assess model performance of the pooled cohort equations for atherosclerotic cardiovascular disease (ASCVD).^{9, 10}

Population Selection

Through the NMEDW, we identified patients aged 30–79 years who were seen in the outpatient setting for preventive care between January 1, 2005 and December 31, 2013 with available data on all risk factors and at least 5 years of follow-up. A 5-year look-back period prior to January 1, 2005 was used to assess key exclusion criteria. Patients were excluded based on the presence of *International Classification of Disease 9th and 10th revision* (ICD 9–10) codes located in the record (encounter diagnosis, problem list, medical history, and billing) for any of the following diagnoses at or prior to the baseline encounter: HF, coronary artery disease (CAD), peripheral arterial disease, stroke/cerebrovascular disease, and presence of pacemaker or implantable cardioverter-defibrillator to be consistent with the criteria used in the original derivation cohort (Online Table IV).

Exposure Assessment

Age, race, sex, current smoking status, SBP, and BMI were obtained from the baseline office visit encounter. Patients were stratified by race alone as white, Black, and non-white/non-

Black (patients who self-identified their race as non-white, non-Black, unknown race, or declined to disclose race). The primary analysis was restricted to white and Black adults (N=27,438) and supplemental analyses were performed in the non-white/non-Black adults (N=3818) based on the original derivation cohort participant demographics.

The following laboratory data were collected from patients within 1 year of the baseline outpatient encounter: random morning glucose, TC and HDL-C. If multiple laboratory values were present, the value closest to the baseline encounter was selected. DM status and hypertension status was assessed by ICD 9–10 codes at the baseline outpatient encounter (Online Table IV). Medication status for DM or hypertension was determined by the presence of ICD 9–10 codes for DM or hypertension and the presence of DM or anti-hypertension medications on home medication and prescription order lists from the baseline outpatient encounter.

Outcome Ascertainment and Adjudication

A new diagnosis of HF was identified if patients had at least 1 inpatient admission or at least 2 outpatient ambulatory encounters with a HF ICD 9–10 diagnostic code during follow-up. This was based on similarly published definitions of HF using EHR data that have demonstrated a positive predictive value of >95%.^{11–13} The classification of HF by ICD 9–10 codes in general has been reported to be highly specific and have good sensitivity in prior publications (pooled sensitivity 75.3% [95% CI 74.7–75.9] and pooled specificity was 96.8% in a meta-analysis of 11 studies between 1999–2009).^{14, 15} We evaluated the validity of our HF definition in a subset of cases. Two physician adjudicators independently reviewed the health record on a randomly selected subset of cases (n=100) and controls (n=50) using a validated protocol from the Multi-Ethnic Study of Atherosclerosis for diagnosis of probable or definite HF for adjudication (requiring symptoms, HF medication for probable HF and additional objective clinical criteria for definite HF).¹⁶ A third physician reviewer was used to resolve any disagreements. We combined probable and definite HF into a single HF end point for these analyses. Vital status in follow-up was also ascertained.

Statistical Analysis

Predicted 5-year risk of HF from race and sex-specific equations were applied to each patient and absolute risk was calculated in white and Black men and women for the primary analysis. In supplemental analyses, predicted HF risk was calculated for the non-Black/non-white group using the coefficients for white men and women as has been recommended in the ACC/AHA Primary Prevention Guidelines for the application of the pooled cohort equations for ASCVD.^{17, 18} Predicted HF risk was also calculated in white and Black men and women with available QRS data using the original PCP-HF equations scaled to provide 5- year risk estimates. Model performance of the modified and original PCP-HF tool was then evaluated using the Harrell C statistics with less than 0.60, 0.60 to 0.70, 0.70 to 0.80, and greater than 0.80 defined a priori based on prior publications as inadequate, adequate, acceptable, and excellent discrimination levels, respectively.¹⁹ Model calibration was evaluated by the Greenwood-Nam-D'Agostino (GND) approach with adequate calibration

defined a priori as $\chi^2 < 20$.¹⁹ All statistical analyses were performed with the use of SAS statistical software version 9.2 (SAS institute) and R version 3.1.2.

Results

Baseline Characteristics

Through the NMEDW, 31,256 individuals were identified (27,438 white and Black, men and women; and 3,818 non-white and non-Black men and women). Mean age was 51.4 years, 57% of individuals were women, 11% Black, and 12% non-white/non-Black. The baseline characteristics of the EHR cohort is shown in Table 1 for white and Black men and women and in Online Table V for non-white and non-Black men and women. Table 2 compares baseline characteristics of patients by level of predicted 5-year HF risk as determined by the modified PCP- HF model (without QRS). A greater proportion of Black men had higher predicted risk of HF. Rates of smoking, DM, and hypertension treatment were greater in those with higher predicted risk of HF.

HF Incidence and Adjudication

Incident HF event occurred in 568 patients (1.5%) over a 5-year follow-up period corresponding to 3.6 events per 1,000 patient years (Table 3 and Online Table VI). In the random subset with physician adjudication of HF events, the sensitivity of ICD-9 and 10 definitions of HF was high with 81% of cases meeting criteria. This was similar to recently published adjudication rates within our EHR and other administrative databases.^{14, 20, 21} None of the random subset of 50 controls evaluated had adjudicated HF.

Performance of HF Risk Prediction Model in Black and White Men and Women

We assessed model performance comparing predicted and observed rates for incident HF defined by established criteria using ICD codes. Performance of the modified PCP-HF model in white and Black patients is shown in Table 3 and Figure 1. Discrimination was excellent in white men (c-statistic 0.82, 95% confidence interval [0.79, 0.86]) and women (0.82, [0.78, 0.87]) and adequate in Black men (0.69, [0.60, 0.78]) and women (0.69, [0.52, 0.76]). Calibration was good in all race-sex subgroups based on pre-defined criteria with $\chi^2 < 20$. Comparison of predicted vs observed probability of incident HF is shown in Figure 2. Across all deciles, predicted HF risk closely mirrored observed HF risk in white men and white women. In Black men and Black women, there was generally an underestimation of HF risk with predicted HF risk consistently lower than observed HF risk with the exception of a few deciles (Figure 2).

Performance of HF Risk Prediction Model in Non-Black, Non-White Men and Women

Performance of the modified PCP-HF model in non-Black, non-white men and women is shown in Online Table VI and Online Figure I. Discrimination was excellent in non-Black, non-white men (c-statistic 0.80, 95% confidence interval [0.70, 0.90]) and women (c-statistic 0.90, 95% CI [0.86, 0.95]). Calibration was good in both groups. In non-Black, non-white men, predicted and observed HF risk were similar throughout most deciles with predicted HF risk lower than observed HF risk at higher deciles (Online Figure II). In non-

Black, non-white women, predicted and observed HF risk were similar throughout most deciles with differences at highest deciles (Online Figure II).

Performance of Original HF Risk Prediction Model in Black and White Men and Women with QRS Data

Performance of the original PCP-HF model in Black and white men and women with available QRS data is shown in Online Figure III. Discrimination was acceptable in white men (c-statistic 0.74, 95% confidence interval [0.67, 0.80]) and women (0.73, [0.65, 0.81]) and adequate in Black men (0.64, [0.54, 0.75]) and women (0.63, [0.54, 0.72]). Calibration was good in all race-sex groups.

Discussion

In this study, we validated race and sex-specific risk prediction equations for incident HF using an EHR cohort of adults without baseline cardiovascular disease. The modified PCP-HF tool studied here excluded QRS duration from the prediction model give current United States Preventative Task Force guidelines that appropriately recommend against routine screening electrocardiograms.²² Overall, the modified PCP-HF model had excellent discrimination in white adults and adequate discrimination in Black adults with good calibration in all race-sex subgroups based on pre-defined published thresholds, including in non-white, non-Black populations who remain understudied in cardiovascular risk prediction. In comparing model performance in patients with and without QRS data, the addition of QRS did not improve discrimination in any of the race-sex groups.

The critical question of how risk scores derived from longitudinal cohorts perform when applied to the EHR must be answered before integration and implementation of HF-specific risk prediction scores into the EHR. Rana *et al.* showed that the ACC/AHA pooled cohort equations for ASCVD risk had good predictive accuracy despite slight over-estimation of 5-year risk in non-diabetic adults when applied to a large, cotemporary EHR population.²³ Wolfson *et al.* showed that both the Framingham Risk Score and pooled cohort equations could produce accurate risk predictions in an EHR cohort for ASCVD.²⁴ Our study demonstrates that HF-specific risk prediction equations derived from multiple longitudinal cohorts can also be applied to a large, contemporary EHR population with adequate to excellent discrimination and good calibration.

A reliable, valid, easily integrated risk prediction score is only important if it leads to early, meaningful intervention to have an impact on curtailing the increasing prevalence of HF and rising burden of costs.² DM, hypertension, and smoking are modifiable risk factors for HF that can be targeted through the EHR in tandem with a risk score output.^{25–28} The customizability of the EHR to deliver “nudges” or automated electronic reminders offers the opportunity to improve rates of risk factor control.^{29, 30} In addition to aggressive risk factor modification such as intensive blood pressure lowering, patients specifically with DM identified at high risk of HF may be preferred candidates for emerging therapies in primary prevention of HF such as sodium-glucose cotransporter 2 inhibitors (SGLT2i) after an individualized patient-clinician risk benefit discussion. While use of SGLT2i are currently indicated in all patients with DM, data are now emerging for their use in prevalent HF with

reduced ejection fraction in the absence of DM, but are not currently available for their use in patients at high risk of HF who do not have DM.³¹

A HF risk prediction score could also be utilized for targeted, sequential risk stratification. The St. Vincent's Screening to Prevent Heart Failure Study trial demonstrated the utility of brain natriuretic peptide (BNP) screening and collaborative cardiovascular care in patients at risk for HF.³² A risk prediction score could help generate a threshold at which a BNP is reflexively ordered through the EHR in those patients at increased risk as population-wide BNP screening is unlikely to be cost-effective. Asymptomatic left ventricular dysfunction is an under-recognized condition with an estimated prevalence between 2% and 8% in the general community.³³ Despite a strong association between asymptomatic left ventricular dysfunction and development of HF, there is significant debate regarding utility and selection of patients from the general population who would benefit from screening echocardiography due to cost.³⁴ An integrated HF-specific risk prediction score could utilize clinical data in the EHR to identify high-risk patients who would likely benefit from screening echocardiography. Individuals with high risk imaging features (left atrial enlargement, abnormal global longitudinal strain, abnormal E/e' ratio, and left ventricular hypertrophy)³⁵ could benefit from frequent surveillance imaging and biomarker testing, aggressive risk factor targeting, and consideration of novel medical therapies.

There are multiple possible reasons why model discrimination was worse in Black men and women compared with white men and women. There was a consistent underestimation of HF risk seen in Black men and women, with higher discordance between observed and predicted incident HF particularly noted in the lower deciles of risk. This may have been due to selection bias as we relied on data from clinical encounters. Numerous studies have demonstrated racial disparities in primary care utilization with Black adults being less likely than white adults to have access to and interactions with a primary care provider.^{36, 37} This may lead to less routine general well-care visits in Black adults and a higher burden of comorbidities and cumulative exposure to risk factors at initial encounter, which was reflected in our EHR cohort. In addition, the smaller absolute sample of Black patients (N=3285) may have also been a contributing factor to discordance between observed and predicted incident HF noted in lower deciles. Inherent limitations of the PCP-HF model should also be considered, such as the lack of integration of other comorbidities, like chronic kidney disease as well as upstream social factors and root causes of health disparities (e.g. structural and systemic racism).³⁸ The Coronary Artery Risk Development in Young Adults study showed significantly higher rates of incident HF before 50 years of age in Black adults and renal insufficiency and socioeconomic status were found to be independent risk factors for incident HF.^{39, 40} Implementation of race-based risk equations, such as the PCP-HF tool, would require awareness of the potential for systematic underestimation of risk in Black patients as well as other vulnerable populations.⁴¹ Analogous to the use of a "CPR" or "calculate, personalize, reclassify" framework applied in the risk-based prevention of ASCVD that integrates the use of coronary artery calcium to refine risk estimation for individual patients, additional testing with BNP and echocardiography could be useful to improve HF risk assessment and inform personalized clinical decision-making for HF prevention.

Limitations of the current study include the use of data from the EHR, which may have lower accuracy and reproducibility than data collected from well-monitored longitudinal cohorts, but reflects the clinical data that is utilized in every day patient care and therefore is of value.⁷ Certain requirements for inclusion (e.g. specific laboratory values and 5 year follow-up) may have led to the selection of a cohort that was not truly representative of the general US population.⁴² While the proportion of patients requiring treatment for hypertension in our EHR cohort exceeded hypertension prevalence in the general US population,⁴³ the 5 year incidence of HF was 3.6 events per 1000 patient years which is similar to published findings from cohort studies including Framingham and Olmstead County.^{44,45, 46} The use of ICD 9–10 codes for a clinical diagnosis of HF also has inherent limitations although the validity of ICD codes for inpatient diagnosis of HF has been well-established and our physician-based adjudication was consistent with prior studies.^{47, 48} Finally, we do not integrate risk prediction for HF subtypes as there are no actionable differences in preventive strategies to prevent HF with preserved compared with reduced ejection fraction and prior risk models have shown similar contribution of underlying risk factors to both.⁴⁹

Conclusions

In summary, we present an analysis in a contemporary, clinical, population, validating a novel HF-specific risk estimation tool (PCP-HF) that provides sex- and race-specific estimates of 5-year risk of incident HF using risk factor data readily available in the primary care setting. Opportunities to promote aggressive risk factor control (intensive blood pressure lowering), targeted DM management with SGLT2i, and sequential risk stratification with imaging and biomarkers in those at greatest risk as assessed by the PCP-HF tool may allow for greater reduction in burden of HF in the community.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding:

Research reported in this publication was supported, in part, by the National Institutes of Health's National Center for Advancing Translational Sciences, Grant Number KL2TR001424 and the American Heart Association (#19TPA34890060) to SSK. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Sanjiv J. Shah is supported by grants from the National Institutes of Health (NIH; R01 HL107577, R01 HL127028, and R01 HL140731) and the American Heart Association (AHA; #16SFRN28780016 and #15CVGPPSD27260148).

Non-standard Abbreviations and Acronyms

ACC/AHA	American College of Cardiology/American Heart Association
ASCVD	atherosclerotic cardiovascular disease
BMI	body mass index
BNP	brain natriuretic peptide

CAD	coronary artery disease
DM	diabetes mellitus
EHR	electronic health record
GND	Greenwood-Nam-D'Agostino
HDL-C	high density lipoprotein cholesterol
HF	heart failure
ICD 9 –10	International Classification of Disease 9 th and 10 th revision
LRPP	Lifetime Risk Pooling Project
NMEDW	Northwestern Medicine Enterprise Data Warehouse
PCP-HF	Pooled Cohort equations to Prevent Heart Failure
SGLT2i	sodium-glucose cotransporter 2 inhibitors
SBP	systolic blood pressure
TC	total cholesterol
US	United States

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, de Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jimenez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JH, Alger HM, Wong SS and Muntner P. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*. 2017;135:e146–e603. [PubMed: 28122885]
2. Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, Ikonomidis JS, Khavjou O, Konstam MA and Maddox TM. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circulation: Heart Failure*. 2013;6:606–619. [PubMed: 23616602]
3. Glynn P, Lloyd-Jones DM, Feinstein MJ, Carnethon M and Khan SS. Disparities in Cardiovascular Mortality Related to Heart Failure in the United States. *Journal of the American College of Cardiology*. 2019;73:2354–2355. [PubMed: 31072580]
4. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr., Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM, Hollenberg SM, Lindenfeld J, Masoudi FA, McBride PE, Peterson PN, Stevenson LW and Westlake C. 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. *Circulation*. 2017;136:e137–e161. [PubMed: 28455343]
5. Khan SS, Ning H, Shah SJ, Yancy CW, Carnethon M, Berry JD, Mentz RJ, O'Brien E, Correa A, Suthahar N, de Boer RA, Wilkins JT and Lloyd-Jones DM. 10-Year Risk Equations for Incident Heart Failure in the General Population. *Journal of the American College of Cardiology*. 2019;73:2388–2397. [PubMed: 31097157]

6. Goldstein BA, Navar AM and Pencina MJ. Risk Prediction With Electronic Health Records: The Importance of Model Validation and Clinical Context. *JAMA cardiology*. 2016;1:976–977. [PubMed: 27732706]
7. Hersh WR, Weiner MG, Embi PJ, Logan JR, Payne PR, Bernstam EV, Lehmann HP, Hripcsak G, Hartzog TH, Cimino JJ and Saltz JH. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical care*. 2013;51:S30–7. [PubMed: 23774517]
8. Wilkins JT, Karmali KN, Huffman MD, Allen NB, Ning H, Berry JD, Garside DB, Dyer A and Lloyd-Jones DM. Data Resource Profile: The Cardiovascular Disease Lifetime Risk Pooling Project. *Int J Epidemiol*. 2015;44:1557–64. [PubMed: 26275450]
9. Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, Ballantyne CM and Go AS. Accuracy of the Atherosclerotic Cardiovascular Risk Equation in a Large Contemporary, Multiethnic Population. *Journal of the American College of Cardiology*. 2016;67:2118–2130. [PubMed: 27151343]
10. Muntner P, Colantonio LD, Cushman M, Goff DC Jr., Howard G, Howard VJ, Kissela B, Levitan EB, Lloyd-Jones DM and Safford MM. Validation of the atherosclerotic cardiovascular disease Pooled Cohort risk equations. *Jama*. 2014;311:1406–1415. [PubMed: 24682252]
11. Lee WY, Capra AM, Jensvold NG, Gurwitz JH and Go AS. Gender and risk of adverse outcomes in heart failure. *The American journal of cardiology*. 2004;94:1147–1152. [PubMed: 15518609]
12. McKee PA, Castelli WP, McNamara PM and Kannel WB. The natural history of congestive heart failure: the Framingham study. *N Engl J Med*. 1971;285:1441–6. [PubMed: 5122894]
13. Go AS, Lee WY, Yang J, Lo JC and Gurwitz JH. Statin Therapy and Risks for Death and Hospitalization in Chronic Heart Failure. *Jama*. 2006;296:2105–2111. [PubMed: 17077375]
14. McCormick N, Lacaille D, Bhole V and Avina-Zubieta JA. Validity of heart failure diagnoses in administrative databases: a systematic review and meta-analysis. *PLoS One*. 2014;9:e104519. [PubMed: 25126761]
15. Szeto HC, Coleman RK, Gholami P, Hoffman BB and Goldstein MK. Accuracy of computerized outpatient diagnoses in a Veterans Affairs general medicine clinic. *Am J Manag Care*. 2002;8:37–43. [PubMed: 11814171]
16. Steverson AB, Pawlowski AE, Schneider D, Nannapaneni P, Sanders JM, Achenbach CJ, Shah SJ, Lloyd-Jones DM and Feinstein MJ. Clinical characteristics of HIV-infected patients with adjudicated heart failure. *European journal of preventive cardiology*. 2017;24:1746–1758. [PubMed: 28945100]
17. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC, Sorlie P, Stone NJ and Wilson PWF. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk. *Circulation*. 2014;129:S49–S73. [PubMed: 24222018]
18. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, Himmelfarb CD, Khera A, Lloyd-Jones D, McEvoy JW, Michos ED, Miedema MD, Muñoz D, Smith SC, Virani SS, Williams KA, Yeboah J and Ziaeian B. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*. 2019;140:e596–e646. [PubMed: 30879355]
19. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*. 2010;121:1768–77. [PubMed: 20404268]
20. Feinstein MJ, Steverson AB, Ning H, Pawlowski AE, Schneider D, Ahmad FS, Sanders JM, Sinha A, Nance RM, Achenbach CJ, Christopher Delaney JA, Heckbert SR, Shah SJ, Hanna DB, Hsue PY, Bloomfield GS, Longenecker CT, Crane HM and Lloyd-Jones DM. Adjudicated Heart Failure in HIV-Infected and Uninfected Men and Women. *J Am Heart Assoc*. 2018;7:e009985. [PubMed: 30571387]
21. Prasada S, Rivera A, Nishtala A, Pawlowski AE, Sinha A, Bundy JD, Chadha SA, Ahmad FS, Khan SS, Achenbach C, Palella FJ Jr., Ramsey-Goldman R, Lee YC, Silverberg JI, Taiwo BO, Shah SJ, Lloyd-Jones DM and Feinstein MJ. Differential Associations of Chronic Inflammatory Diseases With Incident Heart Failure. *JACC Heart Fail*. 2020; 8(6): 489–498. [PubMed: 32278678]

22. Force UPST. Screening for Cardiovascular Disease Risk With Electrocardiography: US Preventive Services Task Force Recommendation Statement. *Jama*. 2018;319:2308–2314. [PubMed: 29896632]
23. Rana JS, Tabada GH, Solomon MD, Lo JC, Jaffe MG, Sung SH, Ballantyne CM and Go AS. Accuracy of the Atherosclerotic Cardiovascular Risk Equation in a Large Contemporary, Multiethnic Population. *Journal of the American College of Cardiology*. 2016;67:2118–2130. [PubMed: 27151343]
24. Wolfson J, Vock DM, Bandyopadhyay S, Kottke T, Vazquez-Benitez G, Johnson P, Adomavicius G and O'Connor PJ. Use and Customization of Risk Scores for Predicting Cardiovascular Events Using Electronic Health Record Data. *Journal of the American Heart Association*. 2017;6.4; e003670. [PubMed: 28438733]
25. Djoussé L, Driver JA and Gaziano JM. Relation between modifiable lifestyle factors and lifetime risk of heart failure. *Jama*. 2009;302:394–400. [PubMed: 19622818]
26. Kenchaiah S, Evans JC, Levy D, Wilson PW, Benjamin EJ, Larson MG, Kannel WB and Vasan RS. Obesity and the risk of heart failure. *New England Journal of Medicine*. 2002;347:305–313. [PubMed: 12151467]
27. Kenchaiah S, Sesso HD and Gaziano JM. Body mass index and vigorous physical activity and the risk of heart failure among men. *Circulation*. 2009;119:44–52. [PubMed: 19103991]
28. Suskin N, Sheth T, Negassa A and Yusuf S. Relationship of current and past smoking to mortality and morbidity in patients with left ventricular dysfunction. *Journal of the American College of Cardiology*. 2001;37:1677–1682. [PubMed: 11345383]
29. Linder JA, Rigotti NA, Schneider LI, Kelley JHK, Brawarsky P and Haas JS. An Electronic Health Record–Based Intervention to Improve Tobacco Treatment in Primary Care: A Cluster-Randomized Controlled Trial. *Archives of internal medicine*. 2009;169:781–787. [PubMed: 19398690]
30. Bertsimas D, Kallus N, Weinstein AM and Zhuo YD. Personalized diabetes management using electronic medical records. *Diabetes care*. 2016;dc160826.
31. McMurray JJV, Solomon SD, Inzucchi SE, Køber L, Kosiborod MN, Martinez FA, Ponikowski P, Sabatine MS, Anand IS, B Iohánková J, Böhm M, Chiang C-E, Chopra VK, de Boer RA, Desai AS, Diez M, Drozd J, Dukát A, Ge J, Howlett JG, Katova T, Kitakaze M, Ljungman CEA, Merkely B, Nicolau JC, O'Meara E, Petrie MC, Vinh PN, Schou M, Tereshchenko S, Verma S, Held C, DeMets DL, Docherty KF, Jhund PS, Bengtsson O, Sjöstrand M and Langkilde A-M. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. *New England Journal of Medicine*. 2019;381:1995–2008. [PubMed: 31535829]
32. Ledwidge M, Gallagher J, Conlon C, Tallon E, O'Connell E, Dawkins I, Watson C, O'Hanlon R, Birmingham M, Patle A, Badabhagn MR, Murtagh G, Voon V, Tilson L, Barry M, McDonald L, Maurer B and McDonald K. Natriuretic peptide-based screening and collaborative care for heart failure: the STOP-HF randomized trial. *Jama*. 2013;310:66–74. [PubMed: 23821090]
33. Wang TJ, Levy D, Benjamin EJ and Vasan RS. The epidemiology of “asymptomatic” left ventricular systolic dysfunction: implications for screening. *Annals of internal medicine*. 2003;138:907–16. [PubMed: 12779301]
34. Roger VL. Asymptomatic Left Ventricular Dysfunction. To Screen or Not to Screen? 2016;4:249–251. [PubMed: 26874384]
35. Yang H, Negishi K, Wang Y, Nolan M, Saito M and Marwick TH. Echocardiographic screening for non-ischaemic stage B heart failure in the community. *Eur J Heart Fail*. 2016;18:1331–1339. [PubMed: 27813300]
36. Arnett MJ, Thorpe RJ Jr., Gaskin DJ, Bowie JV and LaVeist TA. Race, Medical Mistrust, and Segregation in Primary Care as Usual Source of Care: Findings from the Exploring Health Disparities in Integrated Communities Study. *Journal of urban health : bulletin of the New York Academy of Medicine*. 2016;93:456–467. [PubMed: 27193595]
37. Gaskin DJ, Arbelaez JJ, Brown JR, Petras H, Wagner FA and Cooper LA. Examining racial and ethnic disparities in site of usual source of care. *Journal of the National Medical Association*. 2007;99:22–30. [PubMed: 17304965]

38. Kottgen A, Russell SD, Loehr LR, Crainiceanu CM, Rosamond WD, Chang PP, Chambless LE and Coresh J. Reduced Kidney Function as a Risk Factor for Incident Heart Failure: The Atherosclerosis Risk in Communities (ARIC) Study. *Journal of the American Society of Nephrology*. 2007;18:1307–1315. [PubMed: 17344421]
39. Bibbins-Domingo K, Pletcher MJ, Lin F, Vittinghoff E, Gardin JM, Arynchyn A, Lewis CE, Williams OD and Hulley SB. Racial differences in incident heart failure among young adults. *New England Journal of Medicine*. 2009;360:1179–1190. [PubMed: 19297571]
40. Ingelsson E, Lind L, Arnlov J and Sundstrom J. Socioeconomic factors as predictors of incident heart failure. *Journal of cardiac failure*. 2006;12:540–5. [PubMed: 16952788]
41. Akwo EA, Kabagambe EK, Harrell FE Jr, Blot WJ, Bachmann JM, Wang TJ, Gupta DK and Lipworth L. Neighborhood deprivation predicts heart failure risk in a low-income population of blacks and whites in the southeastern United States. *Circulation: Cardiovascular Quality and Outcomes*. 2018;11:e004052. [PubMed: 29317456]
42. Goldstein BA, Navar AM, Pencina MJ and Ioannidis J. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *Journal of the American Medical Informatics Association*. 2017;24:198–208. [PubMed: 27189013]
43. Fryar CD, Ostchega Y, Hales CM, Zhang G and Kruszon-Moran D. Hypertension Prevalence and Control Among Adults: United States, 2015–2016. NCHS data brief. 2017:1–8.
44. Bui AL, Horwich TB and Fonarow GC. Epidemiology and risk profile of heart failure. *Nature reviews Cardiology*. 2011;8:30–41. [PubMed: 21060326]
45. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, Murabito JM and Vasani RS. Long-term trends in the incidence of and survival with heart failure. *New England Journal of Medicine*. 2002;347:1397–1402. [PubMed: 12409541]
46. Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP and Jacobsen SJ. Trends in heart failure incidence and survival in a community-based population. *Jama*. 2004;292:344–350. [PubMed: 15265849]
47. Goyal A, Norton CR, Thomas TN, Davis RL, Butler J, Ashok V, Zhao L, Vaccarino V and Wilson PW. Predictors of incident heart failure in a large insured population: a one million person-year follow-up study. *Circulation Heart failure*. 2010;3:698–705. [PubMed: 20798277]
48. Huang HD, Turner M and Lederle F. Accuracy of ICD-9 codes for identifying acute heart failure hospitalizations. *Circulation: Cardiovascular Quality and Outcomes*. 2015;8:A320–A320.
49. Ho JE, Enserro D, Brouwers FP, Kizer JR, Shah SJ, Psaty BM, Bartz TM, Santhanakrishnan R, Lee DS, Chan C, Liu K, Blaha MJ, Hillege HL, van der Harst P, van Gilst WH, Kop WJ, Gansevoort RT, Vasani RS, Gardin JM, Levy D, Gottdiener JS, de Boer RA and Larson MG. Predicting Heart Failure With Preserved and Reduced Ejection Fraction: The International Collaboration on Heart Failure Subtypes. *Circulation Heart failure*. 2016;9:10.1161/CIRCHEARTFAILURE.115.003116 e003116. [PubMed: 27266854]

Clinical Impact

What is new?

- Using the recent published Pooled Cohort Equations to Prevent Heart Failure, we demonstrate for the first time the validation of heart failure-specific risk equations in a large, contemporary real-world cohort.
- We demonstrate that heart failure-specific risk equations derived from multiple longitudinal cohorts can be accurately applied to a primary prevention cohort derived from the Electronic Health Record.

What are Clinical Implications?

- Using clinical data readily available in an ambulatory setting, primary care physicians and cardiologists can utilize the PCP-HF to estimate race and sex specific risk of incident heart failure to guide personalized prevention.
- The PCP-HF can easily be integrated into electronic health records to identify patients who may benefit from more intensive measures to reduce risk of heart failure.
- Physicians should be aware of the limitations of the PCP-HF, including the potential for underestimation of risk in minorities and other vulnerable patient populations.

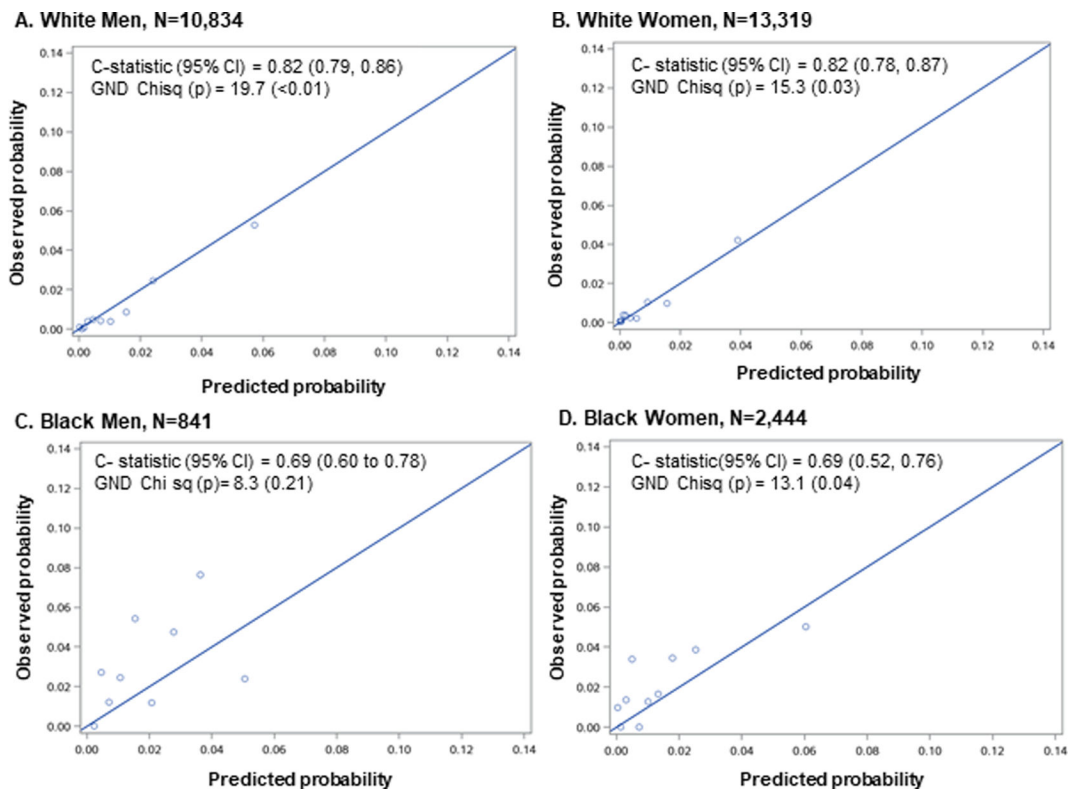


Figure 1.

Calibration and discrimination statistics of the modified Pooled Cohort Equation to Prevent Heart Failure (PCP-HF) model when applied to a contemporary, diverse electronic health record-based cohort stratified by A) White Men, B) White Women, C) Black Men and D) Black Women. In each race-sex group, patients were categorized into deciles based on predicted probability of incident heart failure. Deciles were collapsed as needed if <2 heart failure events occurred. Harrel's C index for discrimination and Greenwood-Nam D'Agostino chi square statistic for calibration were used. CI= confidence interval.

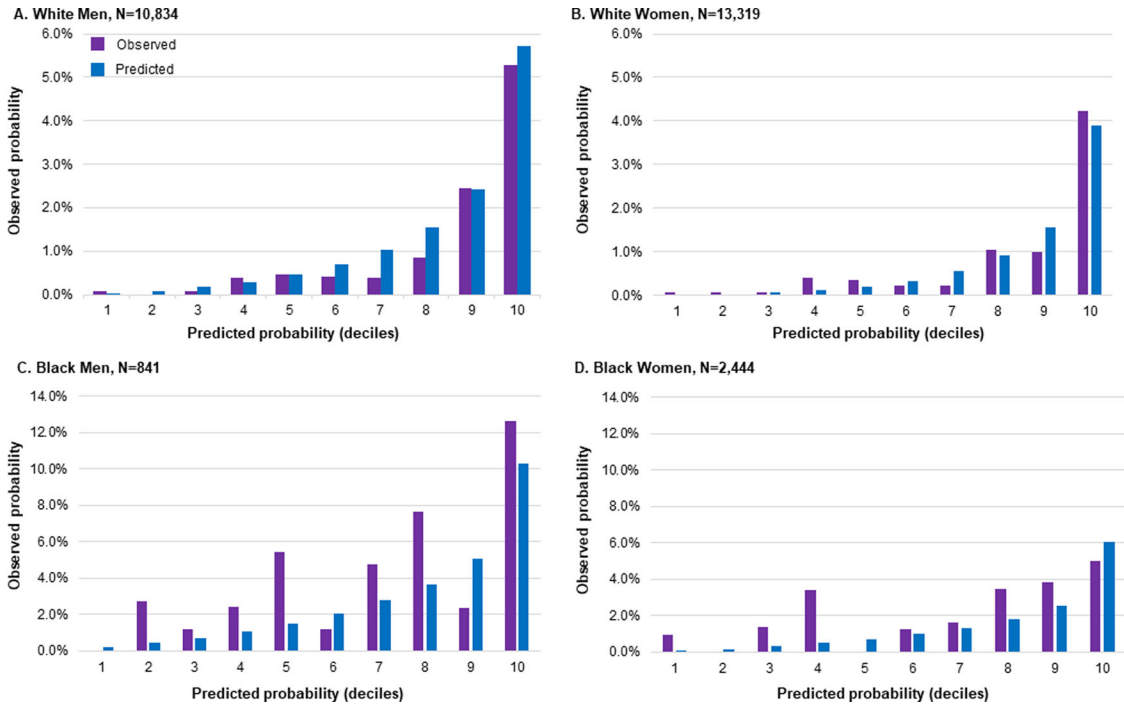


Figure 2. Predicted vs observed probability of incident diagnosis of heart failure in a contemporary, diverse electronic health record-based cohort stratified by A) White Men, B) White Women, C) Black Men, D) Black Women. Predicted probability estimated using modified Pooled Cohort Equation to Prevent Heart Failure (PCP-HF) model. In each race-sex group, patients were categorized into deciles based on predicted probability of incident heart failure. Observed probability was defined by established criteria using *Internal Classification of Disease* codes.

Table 1.

Baseline Patient Characteristics of the Electronic Health Record-Based Primary Prevention Cohort

	Black		White	
	Men N=841	Women N=2444	Men N=10,834	Women N= 13,319
Mean age, years (SD)	50.2 (10.8)	51.4 (11.3)	51.3 (10.7)	52.5 (11.4)
Current smoking, n (%)	119 (14)	247 (10)	1220 (11)	1204 (9)
Diabetes, n (%)	214 (6)	472 (13)	1207 (33)	1224 (33)
Diabetes treatment, n (%)	183 (6)	371 (12)	1060 (35)	948 (32)
Mean casual glucose, mg/dL (SD)	111 (48)	101 (37)	103 (29)	97 (23)
Mean systolic blood pressure, mm Hg (SD)	132 (17)	129 (18)	128 (16)	124 (17)
Hypertension treatment, n (%)	578 (69)	1717 (70)	5139 (47)	6032 (45)
Mean BMI, kg/m² (SD)	30.3 (5.6)	31.4 (6.8)	29.5 (5.1)	28.2 (6.5)
Mean total cholesterol, mg/dL (SD)	185 (40)	189 (37)	191 (37)	197 (36)
Mean HDL cholesterol, mg/dL (SD)	45 (13)	54 (15)	45 (12)	58 (15)

* Abbreviations: BMI, Body Mass Index; HDL, High Density Lipoprotein

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Baseline Characteristics of the Electronic Health Record-Based Primary Prevention Cohort Stratified by Predicted 5-Year Heart Failure Risk Category

	0–1% (n=21,858)	1–2.5% (n=5,830)	2.5–5% (n=2,509)	>5% (n=1,059)
Mean age, years (SD)	46.9 (8.6)	59.9 (8.5)	64.9 (9.6)	66.4 (9.4)
White, n (%)	17253 (79)	4350 (75)	1847 (74)	703 (67)
Black, n (%)	1623 (7)	975 (17)	440 (18)	247 (23)
Non-white, non-Black, n (%)	2982 (14)	505 (9)	222 (9)	109 (10)
Women, n (%)	13419 (61)	2970 (51)	1145 (46)	377 (36)
Current smoking, n (%)	1623 (7)	786 (19)	454 (18)	278 (26)
Diabetes, n (%)	755 (4)	753 (13)	763 (30)	738 (70)
Diabetes treatment, n (%)				
Mean casual glucose, mg/dL (SD)	94 (17)	105 (27)	118 (42)	159 (71)
Mean systolic blood pressure, mm Hg (SD)	122 (15)	133 (16)	138 (18)	141 (19)
Hypertension treatment, n (%)	7159 (33)	4595 (79)	2313 (92)	1028 (97)
Mean BMI, kg/m²	27.7 (5.5)	31.1 (6.2)	31.8 (6.4)	32.6 (6.8)
Mean total cholesterol, mg/dL (SD)	195 (35)	194 (38)	188 (40)	182 (42)
Mean HDL cholesterol, mg/dL (SD)	53 (15)	50 (15)	48 (14)	43 (13)

* Abbreviations: BMI, Body Mass Index; HDL, High Density Lipoprotein

Table 3.

Discrimination and Calibration Statistics of the Modified Pooled Cohort equations to Prevent HF (PCP-HF without QRS) Risk Equations in the Electronic Health Record-Based Primary Prevention Cohort

	White		Black	
	Men	Women	Men	Women
Total N	10834	13319	841	2444
Events	161	186	58	106
Events per 1000 person-years	3.0	2.8	13.8	8.7
C statistics (95% CI)	0.82 (0.79, 0.86)	0.82 (0.78, 0.87)	0.69 (0.60, 0.78)	0.69 (0.52, 0.76)
GND Chi-square, P value	19.7 (0.01)	15.3 (0.03)	8.3 (0.21)	13.1 (0.04)

* Abbreviations: CI, Confidence Interval; GND, Greenwood Nam D'Agostino

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript