Predicting adults likely to develop heart failure using readily available clinical information

Tova M. Bergsten, Robert Donnino, Binhuan Wang, Andrew Nicholson, Yixin Fang, Sundar Natarajan

PII: S0091-7435(19)30354-8

DOI: https://doi.org/10.1016/j.ypmed.2019.105878

Reference: YPMED 105878

To appear in: *Preventive Medicine*

Received date: 8 February 2019

Revised date: 21 August 2019

Accepted date: 23 October 2019

Please cite this article as: T.M. Bergsten, R. Donnino, B. Wang, et al., Predicting adults likely to develop heart failure using readily available clinical information, *Preventive Medicine*(2019), https://doi.org/10.1016/j.ypmed.2019.105878

Predicting Adults Likely to Develop Heart Failure Using Readily Available Clinical Information

An analysis of heart failure incidence using the NHEFS

Tova M. Bergsten[a], Robert Donnino[a,b], Binhuan Wang[a,b], Andrew Nicholson[a], Yixin Fang[c], Sundar Natarajan[a,b,*]

[a]VA New York Harbor Healthcare System, New York, NY; [b]New York University School of Medicine, New York, NY; [c]New Jersey Institute of Technology, Newark, NJ

Corresponding Author: Sundar Natarajan, 423 East 23rd Street, Room 15160-N, New York, NY 10010. Telephone # (212) 951-3287, Fax # (212) 951-3350, E-mail: Sundar.Natarajan@med.nyu.edu

Running head: Predicting Heart Failure Incidence from Clinical Information

Word count for text: (excluding abstract, tables, figures and references): 3418

Word count for abstract: 249

Number of tables: 3

Number of figures: 2

Abstract

BACKGROUND: Heart failure is a heavy burden to the health care system in the United States. Once heart failure develops, the quality of life and longevity are dramatically affected. It is critical to prevent it. We evaluated the predictive ability of readily available clinical information to identify those likely to develop heart failure.

METHODS: We used a CART model to determine the top predictors for heart failure incidence using the NHANES Epidemiologic Follow-up Study (NHEFS). The identified predictors were hypertension, diabetes, obesity, and myocardial infarction (MI). We evaluated the relationship between these variables and incident heart failure by the product-limit method and Cox models. All analyses incorporated the complex sample design to provide population estimates.

RESULTS: We analyzed data from 14,407 adults in the NHEFS. Participants with diabetes, MI, hypertension, or obesity had a higher incidence of heart failure than those without risk factors, with diabetes and MI being the most potent predictors. Individuals with multiple risk factors had a higher incidence of heart failure as well as a higher hazard ratio than those with just one risk factor. Combinations that included diabetes and MI had the highest incidence rates of heart failure per 1000 person years and the highest hazard ratios for incident heart failure.

CONCLUSIONS: Having diabetes, MI, hypertension or obesity significantly increased the risk for incident heart failure, especially combinations including diabetes and MI. This suggests that individuals with these conditions, singly or in combination, should be prioritized in efforts to predict and prevent heart failure incidence.

Key words: Incident heart disease; Risk Factors; Diabetes; Myocardial Infarction; Prediction; Survival Analysis; Prevention

Introduction

Heart failure is an important health care issue facing the United States because of its high prevalence, morbidity and mortality, and cost of care.[1] Heart failure currently affects over 5 million people in the United States.[2] More than half of those who develop heart failure die within just five years of diagnosis.[3] Moreover, as of 2011, the cost of heart failure was greater than 30 billion dollars per year.[4] Although hospitalizations for heart failure may have slightly decreased,[5] an increase in both heart failure incidence and cost of care is expected as a result of the improvements in health care that extend survival times of heart failure patients and the ongoing aging of the population.[4]

There are several clinical conditions and demographic factors that contribute to heart failure.[6] These clinical risk factors include those that can be modified by patients, like hypertension, hyperlipidemia, diabetes, obesity, smoking, lack of physical activity, and ischemic heart disease (IHD), and those that cannot be modified like age, gender, and family history of premature cardiovascular disease. Of the clinical risk factors, hypertension and IHD account for approximately three fourths of all cases.[7] Most of the work on prediction of heart failure is based on certain specific cohort studies, with the most prominent being the Framingham Heart Study (computed 4-year probability of heart failure) and the Health Aging and Body Composition (ABC) study (predicted 5-year probability of heart failure), both of which have been used to develop prediction scores for incident heart failure.[8,9] Unfortunately, these prediction scores only have moderate discriminative value for identifying those persons likely to develop new heart failure. The bigger problem with such short-term risk prediction tools is that the risk is dramatically lower in young persons and it preferentially identifies elderly people as being at high risk. In addition to such prediction scores, other researchers have used circulating biomarkers[10] and cardiac imaging methods[11] to detect individuals with early heart failure for whom appropriate interventions can be implemented to prevent heart failure progression. Research indicates that, as opposed to being a population-wide screening strategy for preventing heart failure, a combination of prediction score, biomarkers and appropriate imaging may be a more efficient strategy for individuals already identified at being at higher than normal risk.[12]

Population projections estimate that the prevalence of heart failure in 2030 will increase by more than 25% from 2013 levels.[2] It is crucial to predict those likely to develop heart failure and implement appropriate preventive interventions because 30-40% of patients typically die

within 1 year of their initial HF event[13,14] emphasizing the critical need to prevent HF. Even in a well-studied condition like IHD where improved treatment has enhanced survival, there is increased likelihood of residual myocardial dysfunction in survivors with up to 53% developing HF. No recent rigorous RCT's of HF prevention have been conducted. The older RCTs of heart failure prevention include BP-lowering trials and the Study of Left Ventricular Dysfunction (SOLVD) Prevention trial.[15] However, it is very difficult to do primary prevention trials of heart failure and most of the evidence on heart failure prevention is extrapolated from observational studies. Prevention is recommended primarily by optimizing management of HF risk factors, particularly reducing obesity and BP lowering in high-risk individuals, but many patients may have several risk factors in combination. When individuals have several comorbidities, the management is not straightforward, treatment guidelines are not clear and patients have difficulty adhering to treatment, leading to suboptimal risk factor management.

When applying a prevention strategy for HF, it is recommended to take a lifetime perspective on HF risk rather than use risk estimates derived from short-term prediction models since those will predominantly identify elderly persons as being at high risk.[16] The American College of Cardiology/American Heart Association Guideline for the Management of Heart Failure[17,18] provides an efficient framework, where they recommend a staging classification (A to D) for HF, with Stage A and Stage B comprising asymptomatic individuals eligible for primary prevention of heart failure. Stage A individuals are at high risk for HF but without structural heart disease while stage B individuals are at even higher risk for HF with structural heart disease. Stage A includes persons who have hypertension, diabetes, obesity, metabolic syndrome, or atherosclerotic heart disease. Stage B includes those with previous MI and structural heart disease. This staging system recognizes that HF has established risk factors and develops in a progressive manner. To prevent progression to HF, intense effort is needed with the best time to intervene being before Stage A (primordial prevention), or at Stage A or Stage B (primary prevention).[15,19]

While the risk factors that contribute to the development of heart failure are well established, i.e. the conditions comprising Stage A or Stage B, little is known about their relative magnitude, how their effects are transmitted in combination, and the best way to predict those likely to develop heart failure. These topics are important because if we can predict those most likely to develop heart failure, personalized care and proactive monitoring could be implemented to prevent heart failure. [17,18] Further, if these predictions are based on the readily available heart failure risk factors, they can be identified easily from clinical databases. In this study, we

evaluate the predictive ability of easily identifiable clinical risk factors to identify those persons most at risk for developing heart failure, determine if multiple clinical risk factors in the same individual are related to greater risk for developing heart failure, and elucidate the combination of clinical conditions that provides the best prediction of those likely to develop heart failure. This population-based cohort study has potential to provide important insights on heart failure prevention in the US.

Methods

Study design and study sample

We analyzed data from 14,407 participants in the National Health and Nutritional Examination Survey (NHANES) Epidemiologic Follow-up Study (NHEFS). NHEFS is a national longitudinal study designed to investigate the relationships between clinical, nutritional, and behavioral factors assessed in the original NHANES I study and their subsequent morbidity, mortality, and hospital utilization. The sample design of NHEFS was a multistage, stratified, probability sample of the US civilian non-institutionalized adult population. Some groups were over-sampled including the elderly, women of childbearing age, minorities, and low-income subjects. NHEFS is the cohort study of NHANES I participants ages 25 – 74 years in 1971 to 1975, which was the baseline evaluation. Data was also collected at the follow-up visits in 1982 – 1984, 1986, 1987, and 1992.[20-23]

Measurements

The baseline examination for participants included an interview by trained professionals who collected demographic information, socioeconomic data, and a detailed medical history. Additionally, a standardized physical examination and blood draw were administered to participants to collect biometric data. The first follow-up took place from 1982-1984,[20] including all members of the NHEFS cohort, and consisted of personal interviews with subjects or proxies, as well as measurement of pulse rate, weight, and blood pressure. At this point, hospital and nursing home records were reviewed and death certificates adjudicated. The second follow-up was conducted during 1986 and included all cohort members ages 55-74 at baseline who were not known to be deceased as of the first follow-up.[23] The data collection

consisted of a 30-minute computer-assisted telephone interview as well as collecting hospital and nursing home records, and death certificates. The third follow-up took place in 1987[22] and the fourth in 1992.[21] During both of these follow-ups, a 30-minute computer-assisted telephone interview was conducted with all surviving cohort members. Hospital and nursing home records and death certificates were also collected.

Dependent and Independent Variables

This project focused on identifying a few variables as important risk factors for heart failure. We used variable-importance measurements obtained from classification tree analysis (a.k.a. Classification and Regression Tree; CART)[24] to determine the top predictor variables for heart failure incidence, which were: diabetes, hypertension, myocardial infarction (MI), and obesity. A self-report of previous physician-diagnosed diabetes was used to define diabetes status. Hypertension was defined by mean systolic blood pressure exceeding 140 mm Hg, or mean diastolic blood pressure exceeding 90 mm Hg, or self-reported usage of anti-hypertensive medication. Prior MI was defined according to self-report of physician diagnosis of a previous MI. Body mass index (BMI) was defined as weight in kilograms divided by the square of the height in meters (kg/m^2). Obesity was defined as having a BMI ≥ 30. Additional demographic and lifestyle factors such as age, sex, race, educational attainment, physical activity, smoking, and alcohol consumption were also considered as covariates.

The main dependent variable studied in this project was heart failure diagnosis. We determined heart failure incidence from ICD-9 codes from hospitalizations and where heart failure was the cause of mortality, from the National Death Index. To create the final analytic dataset, we excluded individuals who had a diagnosis of heart failure prior to baseline.

Statistical Analysis

We used variable-importance measures from the CART analysis to identify important risk factors of heart failure incidence among a large pool of candidate variables including: hypertension, MI, obesity, total cholesterol, current smoking status, activity level, and alcohol use. CART provides a flexible and model-free method to evaluate the data as an exploratory analysis. We selected variables *a priori* based on current understanding of risk factors related

to heart failure incidence, and whether they are readily available in clinical documentation. Variable importance of each predictor, an indication of which predictors are most useful for predicting the response variable using the CART analysis, was ranked based on the contribution that predictor made to the model. The Youden Index (J), a measure of diagnostic effectiveness, calculated as J = [sensitivity + specificity] – 1 generates values ranging from 0 to 1. There is a known trade-off between sensitivity and specificity, and the Youden index provides an overall measure to summarize the performance of the test. A maximum Youden index value implies a cut-point that optimizes a model's differentiating ability when equal weight is given to sensitivity and specificity. We used a cut-point for the Youden Index of 0.49 based on the data. As part of implementing the CART model, we randomly selected 7,000 records (approximately 50%) to train the model and used the remaining records to test, and select the best loss parameter. After identifying the top predictors for heart failure incidence, we examined descriptive statistics by presence of the heart failure risk factor of interest.

Using the product limit method, the underlying statistical method used to generate cumulative incidence curves for survival analysis, we created HF incidence curves to depict the HF incidence rates and to assess differences in heart failure incidence between groups using the log-rank test. We also assessed follow-up time, number of heart failure events, and cumulative incidence rate. Cox proportional hazards models were used to evaluate the effect of each risk factor alone or in combination with other risk factors on heart failure incidence while controlling for age, sex, race, and education.[25] A full model evaluated the effect of all combinations of diabetes, hypertension, MI, and obesity on heart failure incidence while controlling for age, sex, race, and education.

Finally, to assess the predictive accuracy of the final Cox model to predict incident heart failure, we calculated Harrell and Uno concordance statistics (C-statistics) that allowed us to evaluate discrimination of the risk prediction models.[26,27] These C-statistics are a generalization from methods for dichotomous outcomes that are available for logistic models. These methods to evaluate the discriminative ability of the Cox model characterizes the ability of the model to accurately classify participants by the actual occurrence of new heart failure events. All analyses incorporated the complex sample design by including the strata and cluster variables to provide population estimates. We conducted statistical analyses using SAS 9.4 and R. All reported p-values are two-sided.

Results

The CART model, with cost parameters tuned by maximizing the Youden Index, showed that hypertension, diabetes, obesity, and MI were the four most important risk factors to predict heart failure incidence. Table 1 summarizes the baseline socio-demographic characteristics of the study population by these risk factor groups.

Figure 1 displays the age- and sex-standardized cumulative incidence rate of heart failure by risk factor group. All risk factor groups had a greater cumulative incidence rate of heart failure when compared to the no risk factor group. The hypertension and obesity groups displayed similar curves while the MI and diabetes groups had similar curves and had the highest cumulative incidence rates.

Table 2 depicts the age and sex-adjusted effect of a single risk factor alone or each risk factor with additional risk factors on heart failure incidence. It shows the number of participants in each group, the total years of follow-up for that group of participants, the number of incident heart failure events, and the age- and sex-adjusted heart failure incidence rate per 1000 person-years. The incidence rate of the no risk factor group was the smallest with a rate of 3.2/1000 person-years. Each risk factor alone had a higher incidence rate than that of the no risk factor group, with the diabetes (14.9/1000 person-years) and MI (5.6/1000 person-years) groups having notably higher incidence rates. Participants with risk factors in addition to a single risk factor all had higher incidence rates than the participants with the single risk factor alone. This was particularly marked for diabetes and MI; diabetes with additional risk factors raised the incidence rate to 18.9/1000 person-years (from 14.9/1000 person-years) while in MI additional risk factors raised the incidence rate to 14.9/1000 person-years (from 5.6/1000 person-years).

Table 3 depicts the unadjusted and adjusted effect of all risk factor combinations on incident heart failure. It shows the number of participants in each group, the total years of follow-up for that group of participants, the number of incident heart failure events, and the age- and sex-adjusted heart failure incidence rate per 1000 person-years. Again, the lowest incidence rate was seen among participants with no risk factors. The highest incidence rate was seen in the group with diabetes and MI at a rate of 108.4/1000 person-years. Other combinations that included diabetes and MI had particularly high incidence rates though, in general, the rates were higher in those with additional risk factors when compared to groups with a single risk factor. Table 3 also depicts the age, sex, race, and education adjusted hazard ratios for each combination of the risk factors, all compared to the no risk factor reference

group. Like the incident rate results, the highest hazard ratios were among participants with diabetes and MI.

To further evaluate the effect of risk factors in addition to the primary risk factor, we created a 3-level variable (no risk factor, the single risk factor alone, and the single risk factor with additional risk factors) and evaluated the relationship between this variable and time to HF incidence in separate models by risk factor while controlling for age, sex, race, and education (Figure 2). We display the results as multivariate-adjusted hazard ratios, 95% confidence intervals, and the associated p values for each risk factor alone and each risk factor with additional risk factors (the no risk factor group was the referent). We also reported the number of events included in each model. For all models, participants with additional risk factors had higher hazard ratios than those with the single risk factor alone. This trend was most apparent for diabetes and MI where the hazard ratios rise from 1.8 (95% CI 0.8, 3.9) for diabetes alone to 6.0 (95% CI 4.0, 8.9) for diabetes with additional risk factors; and 2.4 (95% CI 1.2, 5.0) for MI alone to 5.4 (95% CI 3.5, 8.3) for MI with additional risk factors.

Finally, to evaluate the ability of the model to predict future new heart failure events, we calculated C-statistics for the final Cox-models. A higher C-statistic indicates that the model has higher discriminating ability between subjects who develop heart failure and those who do not. The Harrell's C-statistic was calculated as 0.841 (SE 0.008; 95% CI 0.825, 0.857). Harrell's C-statistic discards pairs that cannot be compared because of censoring. Uno's C-statistic models the distribution of the censoring variable to weight the uncensored variables while estimating the C-statistic, i.e. the estimates are independent of the censoring variable. The corresponding Uno's C-statistic was 0.797 (SE 0.138, 95% CI 0.527, 1.000).

Discussion

We evaluated population-based US national data to evaluate the relationship between readily available clinical factors and subsequent occurrence of new heart failure. All four of the individual risk factors were related to high risk for incident heart failure, with diabetes and MI being particularly potent. For all four conditions, having additional risk factors was related to a much higher risk, especially for combinations including diabetes and MI together.

The Framingham Heart Study provided evidence regarding the major clinical risk factors for HF,[6] with IHD and hypertension being responsible for three fourths of all cases. More

recently there has been recognition of the importance of diabetes in increasing the risk for HF.[28] Our paper quantifies the relative importance of the different risk factors in the US population and identifies diabetes and a previous MI as particularly strong risks for heart failure. Previous work has reported on this[29,30] and its potential structural and mechanistic reasons.[31,32] The results of this study also indicate that individuals who have the combination of both diabetes and MI (with or without other risk factors) are at a dramatically increased risk for incident heart failure.

From these results, one clinical implication is to support an aggressive approach to preventing heart failure in persons with these risk factors; in particular, those with diabetes or MI. Individuals with diabetes and a previous MI (with or without other risk factors) are at particularly high risk for developing heart failure, should be identified as being at high risk and prioritized in efforts to predict and prevent heart failure. It has been shown that individuals who have good control of their HF risk factors have significantly lower risk for HF.[33,34] While this study did not address ways in which these patients could be prioritized, proactive monitoring could be implemented including more frequent serum natriuretic peptides and imaging (e.g. echocardiograms).

This project may have certain potential limitations. We used self-reported data for diabetes and MI, which may have caused misclassification. However, it is not likely to have been differential. Further, because of the time that this data was collected, diabetes and MI diagnoses only captured the more severe cases leading any bias to shift the results towards the null. There could be errors in diagnosing heart failure as the cause of death, though it is unlikely that any misclassification bias would be differential. Also, the lack of information regarding renal function and other vascular risk factors meant that we could not control for them in the analysis. However, we have refrained from controlling for mediators that may lie in the pathway of the outcome. The small number of events for certain specific risk factor groups (e.g. diabetes and MI) make their estimate appear less accurate. However, the totality of the evidence regarding combinations that include MI and diabetes is strong. Finally, the 20-year time period over which the data was collected is both a limitation, since there were changes in diagnostic criteria and treatment options during this time, and strength, due to the lengthy period of follow-up.

The strengths of this data set stem from it being a national sample, which included a diverse population with a wide spectrum of ages, different races, and SES classes from different parts of the country. The study utilized standardized data collection over many consecutive years with strict attention to quality control implying valid measurements. Data were available

for many consecutive years and this long follow-up meant that there were more outcomes to provide more power to the analysis. These strengths make the NHEFS sample a strong cohort from which to analyze relationships that would be generalizable to future populations.

Future studies should prospectively evaluate the relationship between clinical comorbidities, new biochemical measures, and genetic markers on heart failure incidence. More importantly, it is critical to test interventions targeting modifiable risk factors that are key to treating these chronic conditions. It is also important to evaluate health system approaches to consistently intervene on heart failure risk factors and assess early signs of ventricular dysfunction since even modest changes can improve quality of life and longevity in such patients.[35]

In summary, having diabetes, MI, hypertension or obesity significantly increased the risk for incident heart failure, especially the combination of diabetes and MI. This suggests that individuals with this constellation of conditions should be prioritized in efforts to predict and prevent heart failure incidence.

Institutional Review Board (IRB) Approval

This was an analysis of publicly available data. Informed consent was obtained directly from all participants as part of NHANES I. Since we were using publicly available de-identified data, we obtained exempt review from the New York University School of Medicine IRB.

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation.* 2017.
2. Heidenreich PA, Trogdon JG, Khavjou OA, et al. Forecasting the Future of Cardiovascular Disease in the United States: A Policy Statement From the American Heart Association. *Circulation.* 2011.
3. Barker WH, Mullooly JP, Getchell W. Changing incidence and survival for heart failure in a well-defined older population, 1970-1974 and 1990-1994. *Circulation.* 2006;113(6):799-805.
4. Heidenreich PA, Albert NM, Allen LA, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circulation Heart failure.* 2013;6(3):606-619.
5. Chen J, Normand SL, Wang Y, Krumholz HM. National and regional trends in heart failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008. *Jama.* 2011;306(15):1669-1678.
6. Kenchaiah S, Narula J, Vasan RS. Risk factors for heart failure. *The Medical clinics of North America.* 2004;88(5):1145-1172.
7. Velagaleti RS, Vasan RS. Heart failure in the twenty-first century: is it a coronary artery disease or hypertension problem? *Cardiology clinics.* 2007;25(4):487-495; v.
8. Kannel WB, D'Agostino RB, Silbershatz H, Belanger AJ, Wilson PW, Levy D. Profile for estimating risk of heart failure. *Archives of internal medicine.* 1999;159(11):1197-1204.
9. Butler J, Kalogeropoulos A, Georgiopoulou V, et al. Incident heart failure prediction in the elderly: the health ABC heart failure score. *Circulation Heart failure.* 2008;1(2):125-133.
10. Florea VG, Anand IS. Biomarkers. *Heart failure clinics.* 2012;8(2):207-224.
11. Patel MR, White RD, Abbara S, et al. 2013 ACCF/ACR/ASE/ASNC/SCCT/SCMR appropriate utilization of cardiovascular imaging in heart failure: a joint report of the American College of Radiology Appropriateness Criteria Committee and the American College of Cardiology Foundation Appropriate Use Criteria Task Force. *Journal of the American College of Cardiology.* 2013;61(21):2207-2231.
12. Redfield MM. Strategies to screen for stage B as a heart failure prevention intervention. *Heart failure clinics.* 2012;8(2):285-296.
13. Cowie MR, Wood DA, Coats AJ, et al. Survival of patients with a new diagnosis of heart failure: a population based study. *Heart.* 2000;83(5):505-510.
14. Roger VL, Weston SA, Redfield MM, et al. Trends in heart failure incidence and survival in a community-based population. *Jama.* 2004;292(3):344-350.
15. Schocken DD, Benjamin EJ, Fonarow GC, et al. Prevention of heart failure: a scientific statement from the American Heart Association Councils on Epidemiology and Prevention, Clinical Cardiology, Cardiovascular Nursing, and High Blood Pressure Research; Quality of Care and Outcomes Research Interdisciplinary Working Group; and Functional Genomics and Translational Biology Interdisciplinary Working Group. *Circulation.* 2008;117(19):2544-2565.
16. Yusuf S, Pitt B. A lifetime of prevention: the case of heart failure. *Circulation.* 2002;106(24):2997-2998.
17. Yancy CW, Jessup M, Bozkurt B, 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. *Circulation.* 2017;136(6):e137-e161.
18. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology

Foundation/American Heart Association Task Force on practice guidelines. *Circulation.* 2013;128(16):1810-1852.
19. Young JB, Narula J. Prevention Should Take Center Stage. *Cardiology clinics.* 2007;25:xi–xiii.
20. Cohen BB, Barbano HE, Cox CS, et al. Plan and operation of the NHANES I Epidemiologic Followup Study: 1982- 84. *Vital Health Stat 1.* 1987(22):1-142.
21. Cox CS, Mussolino ME, Rothwell ST, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1992. *Vital Health Stat 1.* 1997(35):1-231.
22. Cox CS, Rothwell ST, Madans JH, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1987. *Vital Health Stat 1.* 1992(27):1-190.
23. Finucane FF, Freid VM, Madans JH, et al. Plan and operation of the NHANES I Epidemiologic Followup Study, 1986. *Vital Health Stat 1.* 1990(25):1-154.
24. Friedman J, Hastie T, Tibshirani R. *The elements of statistical learning.* Vol 1. New York, NY, USA: Springer; 2001.
25. Cox DR. Regression Models with Life Tables. *Journal of the Royal Statistical Society, Series B.* 1972;74:187-220.
26. Harrell FE. The PHGLM Procedure. *SUGI Supplemental Library Guide, Version 5 Edition*. Cary, NC:: SAS Institute Inc.; 1986.
27. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statistics in medicine.* 2011;30(10):1105-1117.
28. Horwich TB, Fonarow GC. Glucose, obesity, metabolic syndrome, and diabetes relevance to incidence of heart failure. *Journal of the American College of Cardiology.* 2010;55(4):283-293.
29. Patel N, Chen O, Donahue C, et al. Impact of diabetes on heart failure incidence in adults with ischemic heart disease. *Journal of diabetes and its complications.* 2017.
30. He J, Ogden LG, Bazzano LA, Vupputuri S, Loria C, Whelton PK. Risk factors for congestive heart failure in US men and women: NHANES I epidemiologic follow-up study. *Archives of internal medicine.* 2001;161(7):996-1002.
31. Donnino R, Patel S, Nguyen AH, et al. Comparison of quantity of left ventricular scarring and remodeling by magnetic resonance imaging in patients with versus without diabetes mellitus and with coronary artery disease. *The American journal of cardiology.* 2011;107(11):1575-1578.
32. Spahillari A, Talegawkar S, Correa A, et al. Ideal Cardiovascular Health, Cardiovascular Remodeling, and Heart Failure in Blacks: The Jackson Heart Study. *Circulation Heart failure.* 2017;10(2).
33. Folsom AR, Yamagishi K, Hozawa A, Chambless LE, Atherosclerosis Risk in Communities Study I. Absolute and attributable risks of heart failure incidence in relation to optimal risk factors. *Circulation Heart failure.* 2009;2(1):11-17.
34. Munoz MA, Real J, Del Val JL, et al. Impact of the sustained control of cardiovascular risk factors on first episode heart failure: The relevant role of primary care. *The European journal of general practice.* 2015;21(4):224-230.
35. Avery CL, Loehr LR, Baggett C, et al. The population burden of heart failure attributable to modifiable risk factors: the ARIC (Atherosclerosis Risk in Communities) study. *Journal of the American College of Cardiology.* 2012;60(17):1640-1646.

Table 1: Baseline characteristics of NHANES I participants by risk factors*					
	No Risk Factor (NRF)	**Diabetes (DM)**	**Myocardial Infarction (MI)**	**Hypertension (HTN)**	**Obesity (OBS)**
Unweighted sample size (weighted)	4801 (38,651,629)	709 (4,349,254)	504 (3,365,231)	5205 (35,555,730)	2731 (17,820,516)
Age (SE)	41.5 (0.29)	54.9 (0.81)	59.4 (0.56)	52.6 (0.25)	49.7 (0.39)
Sex (SE)					
% Male	46.8 (0.88)	51.2 (2.76)	70.1 (3.10)	54.6 (0.96)	35.9 (1.72)
% Female	53.2 (0.88)	48.8 (2.76)	29.9 (3.10)	45.4 (0.96)	64.1 (1.72)
Race (SE)					
% White	90.8 (0.62)	85.0 (1.91)	93.1 (1.41)	86.5 (1.25)	85.3 (1.59)
% Non-White	9.2 (0.62)	15.0 (1.91)	6.9 (1.41)	13.5 (1.25)	14.7 (1.59)
Education (SE)					
% < High School	0.44 (0.12)	1.7 (0.57)	0.6 (0.37)	1.3 (0.28)	1.6 (0.48)
% High School	73.3 (1.22)	86.3 (2.41)	84.7 (2.18)	80.9 (1.19)	83.1 (1.57)
% Some College	15.8 (0.89)	9.5 (2.19)	10.3 (2.01)	11.0 (0.85)	10.3 (1.20)
% ≥ College	10.5 (0.75)	2.5 (0.84)	4.4 (1.25)	6.8 (0.66)	5.1 (0.81)
Mean BMI (SE)	23.4 (0.07)	27.1 (0.32)	26.6 (0.29)	27.3 (0.10)	31.5 (0.21)
% Smoking	38.2 (1.44)	32.4 (2.89)	23.4 (3.45)	29.6 (1.21)	24.7 (1.45)
Mean Systolic Blood Pressure	117.7 (0.26)	144.5 (1.52)	142.7 (1.27)	151.8 (0.53)	143.1 (0.84)

% High Cholesterol	23.9 (1.05)	40.2 (2.92)	44.7 (3.05)	39.2 (0.95)	40.2 (1.29)
% Occasional/ Regular Alcohol Consumption	75.8 (1.54)	50.6 (3.29)	57.4 (3.87)	66.4 (1.79)	70.0 (2.26)
Physical Activity					
% Inactive	8.5 (0.62)	16.2 (2.31)	16.3 (2.52)	10.8 (0.71)	12.3 (1.07)
% Moderate	40.8 (1.02)	49.7 (2.86)	53.6 (3.77)	45.7 (1.23)	47.3 (1.54)
% Very Active	50.7 (0.99)	34.1 (3.27)	30.1 (3.22)	43.5 (1.39)	40.4 (1.87)

*These analyses are not mutually exclusive regarding each risk factor. Standard error is shown for each mean and proportion.

Table 2. Incidence and Incidence Rate of Heart Failure (by each risk factor alone and by each risk factor with additional risk factors, as well as participants without any risk factors) *

	No RFs	DM Alone	DM & other RFs	MI Alone	MI & other RFs	HTN Alone	HTN & other RFs	OBS Alone	OBS & other RFs
Number of Participants	4801	152	707	124	379	2912	5198	826	2731
Person Years of Follow Up	70166	1959	7557	1230	4747	38275	66476	12114	36822
Number of Heart Failure Events	170	23	182	23	141	387	854	51	405
HF Incidence Rate Per 1000 Person Years**	3.2	14.9	18.9	5.6	14.9	5.4	10.0	5.0	9.0

*The participants in each of these analyses are not mutually exclusive regarding each risk factor.

** Age- and sex-adjusted incidence rates.

RF Risk Factor; DM Diabetes Mellitus; MI Myocardial Infarction; HTN Hypertension; OBS Obesity

Table 3: Incidence Rate and Hazard Ratios of Heart Failure for all combinations of the four risk factors

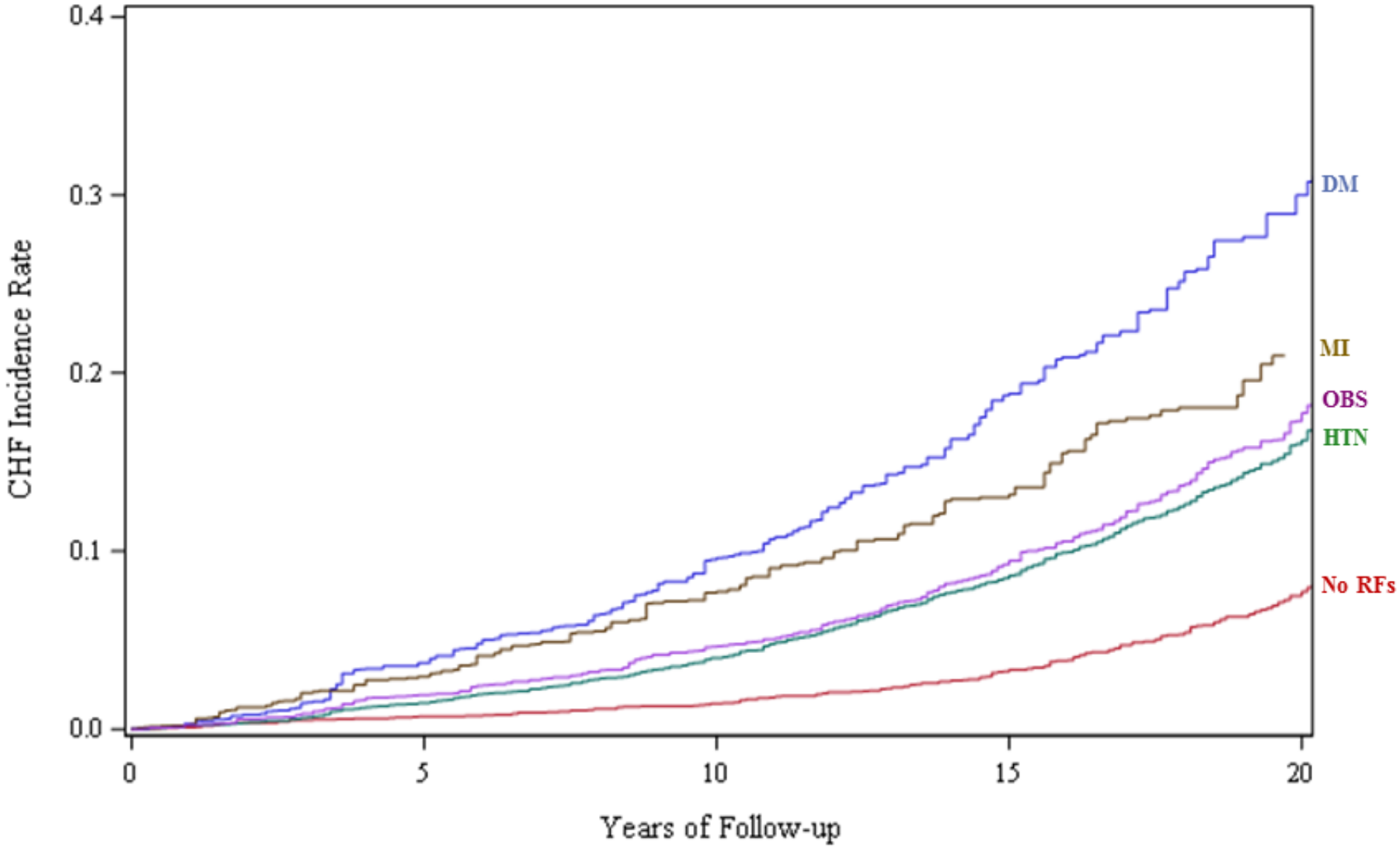
	No RFs	DM Alone	MI Alone	HTN Alone	OBS Alone	DM & MI	DM & HTN	DM & OBS	MI & HTN	MI & OBS	HTN & OBS	DM, MI & HTN	DM, MI & OBS	DM, HTN & OBS	MI, HTN & OBS	DM, MI, HTN & OBS
Participants	4801	152	124	2912	826	11	235	54	189	22	1522	38	5	188	90	24
Total Years of Follow Up	70166	1959	1230	38275	12114	68	2277	660	1781	213	20713	342	33	2008	871	210
Heart Failure Events #	170	23	23	387	51	4	58	7	49	5	237	19	1	64	34	6
Incidence Rate/1000 Person Years**	3.2	14.9	5.6	5.4	5.0	108.4	19.0	12.2	9.0	10.1	7.8	19.5	8.3	21.1	15.5	28.5
Hazard Ratio	1 (Ref)	1.8	2.4	1.8	1.6	33.9	4.6	1.3	4.3	7.6	2.8	9.2	8.9	7.5	4.8	8.1
95% CI Lower Upper		0.8 3.9	1.2 4.9	1.3 2.4	1.0 2.6	6.7 170.7	3.0 7.0	0.4 4.2	2.7 6.8	2.9 19.5	2.1 3.8	3.8 22.2	1.1 69.5	4.3 12.9	2.3 10.3	2.3 28.1
p-value*		0.16	0.019	0.0006	0.066	<0.0001	<0.0001	0.64	<0.0001	<0.0001	<0.0001	<0.0001	0.038	<0.0001	<0.0001	0.0011

RF Risk Factor; DM Diabetes Mellitus; MI Myocardial Infarction; HTN Hypertension; OBS Obesity. *Cox models controlled for age, sex, race, and education. The p-values are from Cox models (No RF is the Referent). ** Age- and sex-adjusted incidence rates.

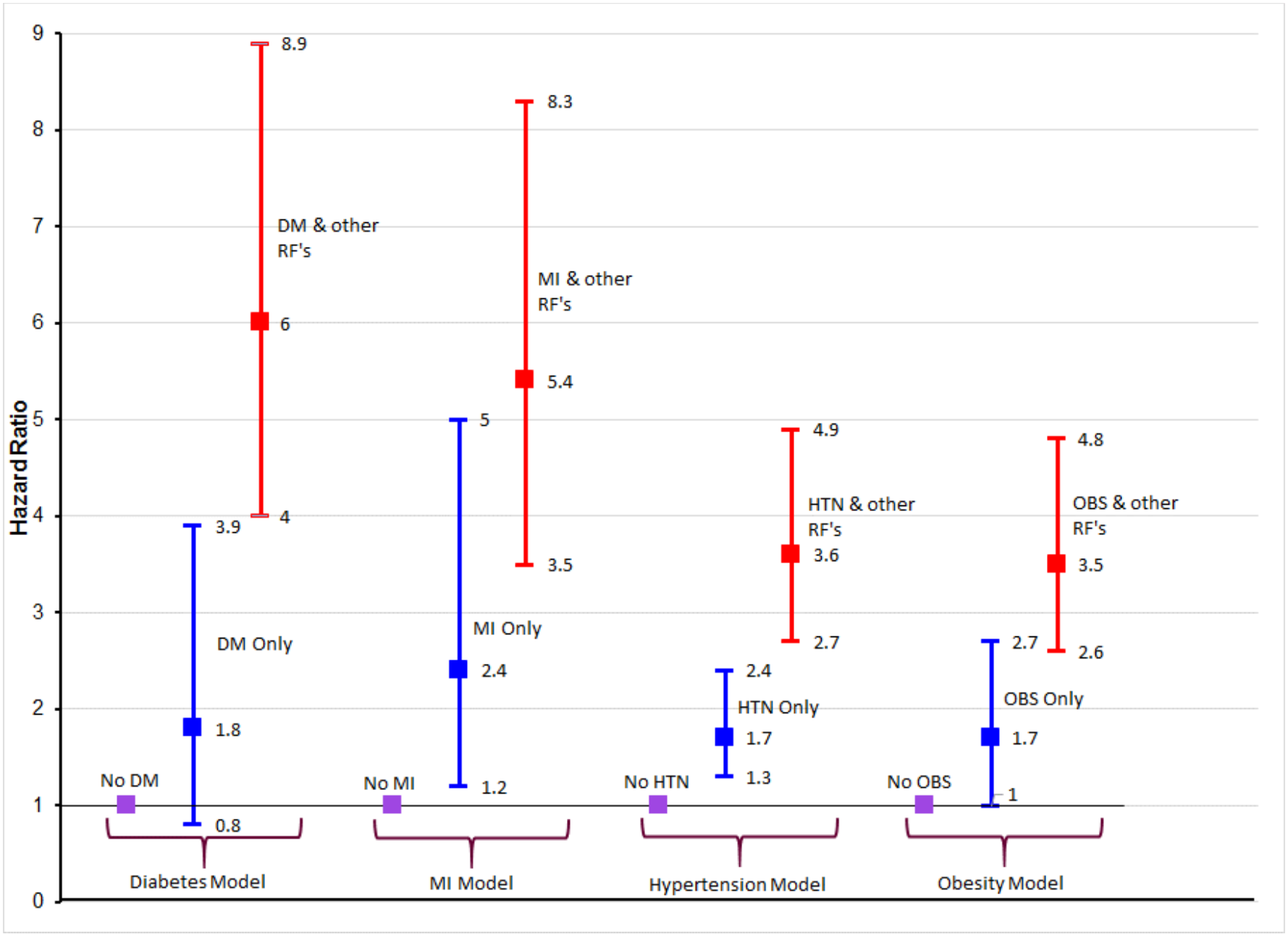
Figure 1: Age- and sex-standardized cumulative incidence rates of heart failure by risk factor group. All risk factor groups had a greater cumulative incidence rate of heart failure as compared to the no risk factor group. Age- and sex-standardized 20-year cumulative incidence rates by group: DM 29.9%; MI 21.0%; OBS 17.7%; HTN 16.2%; No RFs 7.7%. The participants in each of these analyses are not mutually exclusive regarding each risk factor. (RF Risk Factor; DM Diabetes Mellitus; MI Myocardial Infarction; HTN Hypertension; OBS Obesity).

Figure 2: Multivariable Hazard Ratios for Incident Heart Failure (age, sex, race, and education adjusted). Results displayed as multivariate-adjusted hazard ratios and 95% confidence intervals for each risk factor alone and each risk factor with additional risk factors (the no risk factor group was the referent category). The no risk factor group does not include the risk factor of interest but may have other risk factors. For instance, in the hypertension model, participants without hypertension serve as the referent but this group can have individuals with diabetes, obesity or MI. Each model includes the reference (absence of the specific chronic condition), the specific chronic condition alone, and the chronic condition along with other risk factors. For all models, participants with additional risk factors had higher hazard ratios than those with the single risk factor alone. This was most apparent for diabetes and MI. The number of HF events for each group are as follows: DM only 170; DM & other RFs 182; MI only 23; MI & other RFs 141; HTN only 387; HTN & other RFs 854; OBS only 51; OBS & other RFs 405.

Figure 1


0.4
0.3
0.2
0.1
0.0
CHF Incidence Rate
0
5
10
15
20
Years of Follow-up
DM
MI
OBS
HTN
No RFs

Figu


Hazard Ratio
9
8
7
6
5
4
3
2
1
0
8.9
DM & other RF's
6
4
3.9
DM Only
1.8
0.8
No DM
Diabetes Model
8.3
MI & other RF's
5.4
3.5
5
MI Only
2.4
1.2
No MI
MI Model
4.9
HTN & other RF's
3.6
2.7
2.4
HTN Only
1.7
1.3
No HTN
Hypertension Model
4.8
OBS & other RF's
3.5
2.6
2.7
OBS Only
1.7
1
No OBS
Obesity Model