

CLINICAL TRACK

Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis

Bharath Ambale-Venkatesh¹, Xiaoying Yang², Colin O. Wu³, Kiang Liu⁴, W. Gregory Hundley⁵, Robyn McClelland⁶, Antoinette S. Gomes⁷, Aaron R. Folsom⁸, Steven Shea⁹, Eliseo Guallar¹⁰, David A. Bluemke¹¹, João A. C. Lima¹²

¹Department of Radiology, Johns Hopkins University, Baltimore, MD; ²George Washington University, Washington, DC; ³Office of Biostatistics, NHLBI, NIH, Bethesda, MD; ⁴Department of Preventive Medicine, Northwestern University Medical School, Chicago, IL; ⁵Department of Cardiology, Wake Forest University Health Sciences, Winston-Salem, NC; ⁶Department of Biostatistics, University of Washington, Seattle, WA; ⁷Department of Radiology, UCLA School of Medicine, Los Angeles, CA; ⁸Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN; ⁹Departments of Medicine and Epidemiology, Columbia University, New York, NY; ¹⁰Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD; ¹¹Radiology and Imaging Sciences, NIH Clinical Center, Bethesda MD, and; ¹²Department of Medicine, Cardiology and Radiology, Johns Hopkins University, Baltimore, MD.

Running title: CV Event Prediction by Machine Learning: MESA

Subject Terms:

Cardiovascular Disease
Risk Factors
Primary Prevention
Epidemiology
Biomarkers

Address correspondence to:

Dr. João A.C. Lima
Professor of Medicine, Radiology and Epidemiology
Director of Cardiovascular Imaging
Johns Hopkins Hospital
600 N. Wolfe St.
Blalock 524
Baltimore, MD 21287
Tel: 410-614-1284
Fax: 410-614-8222

In July 2016, the average time from submission to first decision for all original research papers submitted to Circulation Research was 12.80 days.

ABSTRACT

Rationale: Machine learning may be useful to characterize cardiovascular risk, predict outcomes and identify biomarkers in population studies.

Objective: To test the ability of random survival forests (RF), a machine learning technique, to predict six cardiovascular outcomes in comparison to standard cardiovascular risk scores.

Methods and Results: We included participants from the Multi-Ethnic Study of Atherosclerosis (MESA). Baseline measurements were used to predict cardiovascular outcomes over 12 years of follow-up. MESA was designed to study progression of subclinical disease to cardiovascular events where participants were initially free of CV disease. All 6814 participants from MESA, aged 45 to 84 years, from 4 ethnicities, and 6 centers across USA were included. 735 variables from imaging and non-invasive tests, questionnaires and biomarker panels were obtained. We used the RF technique to identify the top 20 predictors of each outcome. Imaging, electrocardiography and serum biomarkers featured heavily on the top-20 lists as opposed to traditional CV risk factors. Age was the most important predictor for all-cause mortality. Fasting glucose levels and carotid ultrasonography measures were important predictors of stroke. Coronary artery calcium score was the most important predictor of coronary heart disease and all atherosclerotic cardiovascular disease combined outcomes. Left ventricular structure and function, and cardiac troponin-T were among the top predictors for incident heart failure. Creatinine, age and ankle brachial index were among the top predictors of atrial fibrillation. Tissue necrosis factor- α and interleukin-2 soluble receptors, and N-terminal pro-Brain Natriuretic Peptide levels were important across all outcomes. The RF technique performed better than established risk scores with increased prediction accuracy (decreased Brier score by 10-25%).

Conclusions: Machine learning in conjunction with deep phenotyping improve prediction accuracy in cardiovascular event prediction in an initially asymptomatic population. These methods may lead to greater insights regarding subclinical disease markers without apriori assumptions of causality.

Clinical Trial Registration: Multi-Ethnic Study of Atherosclerosis (MESA) <http://mesa-nhlbi.org/>. ClinicalTrials.gov Identifier: NCT00005487

Keywords:

Random survival forests, cardiovascular events, heart failure, atrial fibrillation, stroke, mortality, coronary heart disease, machine learning, deep phenotyping.

Nonstandard Abbreviations and Acronyms:

ABI	ankle-brachial index
AF	atrial fibrillation
AIC	Akaike Information Criterion
CAC	coronary artery calcium score
CHD	coronary heart disease
Cox-PHM	Cox Proportional Hazard regression model
CVD	cardiovascular disease
ECG	electrocardiography
HF	heart failure
IL2 SR	interleukin-2 soluble receptor
IMT	intima media thickness
IQR	interquartile range
JNC	Joint National Committee
LA	left atrium
LASSO	least absolute shrinkage and selection operator
LV	left ventricle
MESA	Multi-Ethnic Study of Atherosclerosis
MRI	magnetic resonance imaging
NT pro-BNP	N-terminal pro-Brain Natriuretic peptide
RF	Random Survival Forests
RV	right ventricle
SBP	systolic blood pressure
TNF- α SR	tissue necrosis factor- α soluble receptor



INTRODUCTION

Event prediction has been the cornerstone of cardiovascular epidemiology as exemplified by the Framingham study and other prospective studies that function as pillars for much of what comprises current cardiovascular medicine.¹ A fundamental goal of such efforts has been event prediction over relatively long periods of time such as ten years or a lifetime. These efforts have allowed us to characterize sub-clinical disease processes and target key risk factors for modification (e.g. smoking cessation, statin therapy, blood pressure control).² Epidemiological studies used to derive such predictive models frequently contain hundreds or thousands of variables. It is in this context that machine learning methods might be useful as a means to identify the best predictors of outcomes from among millions of phenotypic data points.

The Cox-PHM is often limited for data mining purposes due to correlation between variables, nonlinearity of variables (including potential complex interactions among them), as well as the possibility of over-fitting. On the other hand, machine learning methods, such as Random Survival Forests (RF), use a non-parametric decision tree approach to overcome these issues.³ The purpose of this study was to (a) compare machine learning approaches to Cox-PHM and traditional risk scores for cardiovascular event prediction; and (b) identify the important predictors for each of six cardiovascular clinical outcomes in a large epidemiologic study.

METHODS

The design of the MESA study have been described previously.⁴ Briefly, MESA is a prospective, population-based observational cohort study of 6814 men and women representing four racial/ethnic groups, aged 45-84 years and free of clinical cardiovascular disease at enrollment. As part of the baseline examination (2000-2002), study participants were recruited at six field centers in the United States. Institutional review boards of all field centers approved the study protocol and all participants gave informed consent. Information regarding assessment of markers within MESA has been described previously,⁴ a detailed description is provided in the supplement and a list of markers is shown in Table 1. We included markers from questionnaires, demographics, traditional risk factors, anthropometry, medication use, biochemistry, MRI of the heart and aorta, coronary computed tomography, carotid ultrasound, ECG exams, and ABI.

Outcomes.

All-cause death, stroke, all cardiovascular disease (CVD), coronary heart disease (CHD), atrial fibrillation (AF) and heart failure (HF) events adjudicated as part of MESA were used as end-points (details in Supplement). A telephone interviewer contacted each participant (or representative) every six–nine months to inquire about all interim hospital admissions, outpatient diagnoses, and deaths. Two physicians reviewed all medical records for independent end-point classification and assignment of event dates. Stroke was defined as rapid onset of a documented focal neurologic deficit (vascular causes) lasting 24 hours or until death, or if < 24 hours, when there was a clinically relevant brain lesion. Criteria for CHD included any of myocardial infarction, resuscitated cardiac arrest, definite angina, probable angina followed by revascularization and CHD death. CVD outcomes represented a composite of CVD death, stroke, and CHD. Criteria for incident HF as an end-point included symptomatic HF diagnosed by a physician for a patient receiving medical treatment for HF and 1) pulmonary edema/congestion, and/or 2) dilated ventricle or poor LV function, or evidence of LV diastolic dysfunction. Criteria for incident AF as an end-point required in-hospital AF diagnosis according to ICD9 codes.

Statistical analysis.

Figure 1 shows the statistical analysis procedures followed in this study. Data transformation, indexing and imputation (details in Supplement) was performed as necessary to generate data points to predict six outcomes over the follow-up period. 66·6% of the dataset was randomly selected from the overall group of participants for training; the remaining 33·3% were used for testing/validation. The training dataset was used for model construction using different approaches and optimized to reduce prediction error. These models were then tried on the test dataset to examine model performance and identify the best predictors.

Models tested.

We tested nine different models in our analysis. The first model used the random survival forest (RF) algorithm on all available variables.³ RF is an ensemble tree method for analysis of right-censored data. Details of the RF implementation are provided in the supplement. In short, trees are grown by binary recursive splitting of data. At each split, a candidate variable that maximizes the difference in cumulative hazard between the daughter nodes (and the cut-off that identifies this maximum difference) is chosen. The splitting stops at the terminal nodes when the data at hand can no longer be split such that each terminal node has at least one unique outcome. For each tree, the cumulative hazard rate of a case is determined based on the terminal node that contains it. An ensemble hazard function (and survival probability) is estimated by averaging over all trees in a forest.

The Akaike Information Criterion⁵ for Cox regression with backward stepwise (AIC-BW Cox) elimination and the AIC-Cox with forward stepwise (AIC-FW Cox) regression, as well as the least absolute shrinkage and selection operator for Cox regression (LASSO-Cox), were tested in addition to the

Cox-PHM. LASSO-Cox minimizes the log partial likelihood subject to the sum of the absolute values of the parameters being bounded by a constant. It shrinks coefficients and produces some coefficients that are zero, allowing efficient variable selection.⁶

While RF can be used instead of Cox regression analysis for risk prediction, it can also be used as an efficient variable selection technique. For variable selection using RF, the variables were ranked by the mean of the minimal depth of the maximal subtree (highest point in the tree of a variable) over the entire forest (averaged over 1000 trees). Variables appearing higher on the tree have a higher rank. The top-20 ranked variables were then used again with RF, AIC-FW-Cox, AIC-BW-Cox, LASSO-Cox and Cox-PHM models.

Performance evaluation.

We assessed the performance of each prediction model to discriminate outcomes on the test dataset using Harrell's concordance index (C-index),⁷ and the accuracy of prediction (mean squared distance between the predicted probabilities and actual outcomes) using the Brier score (BS).⁸ Higher C-index and lower BS indicate better prediction performance. C-index and BS for nested models generated using subsets of predictors (chosen based on increasing variable importance) were calculated to assess problems of overfitting. We also compared the results of RF techniques to published risk scores.⁹⁻¹¹ When the models failed to converge, no results were reported and this was considered the worst possible outcome. Data analysis was performed using R software, using publically available libraries for Cox-PHM,¹² Lasso-Cox,⁶ AIC-Cox,⁵ and RF methods.¹³

RESULTS

A total of 6814 participants were included. The average age was 62 years with 53% women, 28% African-American, 38% Caucasian, 12% Chinese-American and 22% Hispanic. 13% of the participants were diabetic, 45% classified as hypertensive based on JNC VI criteria, and 50% were current or former smokers. Over a median of 11.2 years (IQR :10.6 – 11.7), MESA identified 831 all-cause deaths, 710 cardiovascular events (CVD) including 498 CHD events among which 229 were non-fatal myocardial infarctions, 200 strokes, 259 cases of incident HF and 317 incident AF events (Table 2).

Predictors by outcomes.

Table 3 shows the top-20 predictors using RF for each of the outcomes ordered by the minimal depth of maximal subtree. These were the predictors used for the RF-20, Cox-20, Lasso-20, and AIC-20 models. Figure 2 shows Lowess plots of the 12-year survival probability calculated from the RF method over the range of values for the top-5 variables.

Increasing age, perhaps reflecting duration of risk exposure, was the most important predictor of all-cause death. Inflammation and immune response measured by increased interleukin-6, fibrinogen, homocysteine, TNF- α SR and IL2 SR levels; and abnormal hemostasis measured by increased D-Dimer, plasmin-antiplasmin complex and factor VIII levels were among the top 20 markers of all-cause death underlining the role of inflammation and thrombosis as common pathways for chronic diseases leading to death. Similarly, cardiac stress measured by increased NT pro-BNP levels, and myocyte damage by increased cardiac troponin T levels were among the top predictive markers of death reflecting the role of cardiac failure on mortality. In addition to biomarkers that featured so prominently, low and high values of ABI, increased carotid IMT, increased CAC score, aortic dimension and distensibility also showed lower survival probability, indicating the importance of atherosclerosis and vascular abnormalities to mortality in the general population. Importantly also, economic status/income was among the top 20

markers of all-cause death in MESA highlighting the potential power of inequality as a mortality risk factor even when one considers the theoretical distance between such a risk factor and death in the accepted causation chain for most chronic degenerative diseases.

Increased fasting glucose levels were the most important risk factor for stroke, while high blood pressure, a known stroke risk, and age also featured in the top-20 list. Measures of atherosclerosis (with carotid stenosis being most important) and inflammation were also top-5 predictors of stroke. Expectedly, a composite of atherosclerosis measures (low and high ABI, increased carotid IMT, decreased aortic distensibility) were among the most important predictors of CHD which represents a subset of CVD events, with CAC being by far the most important, reflecting the specific influence of coronary atherosclerosis. Increased LV regional wall thickness (myocardial hypertrophy), decreased ejection fraction (myocardial function) and increased aortic cross-sectional area (aortic dilatation), as well as biomarkers of abnormal hemostasis, inflammation, myocardial stress and damage featured among the other top predictors of CHD. Importantly, ECG LV hypertrophy as well as major Q-wave and repolarization abnormalities were markers of CHD and CVD events in MESA. Additionally, among traditional risk factors, pack years of smoking and pulse pressure were among the top predictors for CHD, while systolic blood pressure and age were among the top predictors for CVD.

For incident HF as the endpoint, cardiac chamber stress (increased LV volume, and increased NT-proBNP levels), and decreased LV function from MRI were the most important markers. LV hypertrophy on ECG, a lengthened QT interval indicating increased risk for tachyarrhythmias, increased creatinine levels, increased vascular stiffness, atherosclerosis as measured by CAC and ABI, and inflammation were also among the top predictors for HF. Increased pulse pressure and increased waist-to-hip ratio were also among the top risk factors for incident HF reflecting the role of obesity and hypertension on incident HF development.

For incident AF as the endpoint, inflammation, higher levels of creatinine, atherosclerosis (CAC and ABI), and repolarization abnormalities were the most important markers. Decreased LA function, and increased age and pulse pressure were also among the top risk factors for AF development.

Predictors across outcomes.

In general, variables from imaging markers, ABI, and serum biomarkers were of intermediate-to-high prediction importance while questionnaires and medication exposures were of lower importance. Components of ECG related to ST segments were of intermediate importance while other ECG indices ranged from low to intermediate prediction importance. As illustrated in Online Figure III, just the first five-six variables listed by the RF algorithm produced C-indices greater than 0.75 for CVD, CHD, incident HF and AF prediction reflecting the importance of NT-proBNP (HF and AF) and CAC (CHD and CVD). Prediction of all-cause mortality and stroke with a C-index greater than 0.75 required a larger group of variables.

Comparison of prediction models.

Table 4 shows the C-index and the Brier score (BS) for the eight tested models using the test datasets. The standard Cox, Lasso-Cox and AIC-Cox methods failed to converge when all the 735 variables were included, and hence BS and C-index could not be calculated. As shown in Online Figure III, using the nested models, less than 20 variables were necessary to obtain a stable and high C-index for the RF method. Addition of more variables into the model beyond 30 resulted in minimal improvement of the C-index, if any. Figure 3 shows variable importance measured using the minimum depth of the maximal subtree, for each of the 735 variables used in analysis. Lower values correspond to greater prediction importance.

For all outcomes of interest, the RF model with all 735 covariates showed a very high C-index and low BS. The RF-20 model was comparable and even outperformed the RF model with all covariates in some cases. Both the RF models outperformed the AIC-FWCoX across all endpoints with higher C-index and lower BS. The use of RF for variable selection with top 20 (RF-20) covariates and subsequent application of Lasso-Cox and AIC-BWCoX resulted in fewer variables selected into the final models for most of the outcomes. The C-indices from these standard Cox, Lasso-Cox and AIC-BWCoX models were high and the BS low in general, and very similar to that of the RF-20 model. Figures 4a-4f show the C-index values over time for all the models (models that did not converge are not shown). In general, the C-index values were higher for prediction of short-term as compared to long-term events.

The models from machine learning that included biomarkers and measures of sub-clinical disease were, as expected, better than the AHA/ASCVD (C-index: 0.73, BS: 0.11) and the Framingham (C-index: 0.73, BS: 0.089) risk scores for incident CVD prediction (see Table 4). The performance of the RF-20 model for incident CHD prediction was better than the Framingham CHD risk score (C-index: 0.69, BS: 0.072) with a higher C-index and lower BS.

When comparing the population-specific risk scores, the RSF-20 model from machine learning that included biomarkers and measures of sub-clinical disease was better than the MESA CHD Risk score (C-index: 0.79, BS: 0.074) for incident CHD prediction¹⁴ (see Table 4). The performance of the RF-20 model for incident CHF prediction was better than the MESA-HF¹⁵ risk score (C-index: 0.80, BS: 0.038) with a higher C-index and lower BS.

DISCUSSION

The results of this study suggest that machine learning methods are well-suited for meaningful risk prediction in extensively phenotyped large-scale epidemiological studies. The RF based method of risk prediction provided better event prediction over standard risk scores. RF based methods of variable selection followed by Cox regression methods also allowed for improved prediction of outcomes, without problems of overfitting and non-convergence while accounting for nonlinearities. The results also suggest the importance of deep phenotyping using subclinical markers defined by imaging, electrocardiographic and blood biochemistry, as revealed by their prominent presence on the lists of top-20 predictors, for cardiovascular disease event prediction.

This work is unique by demonstrating patterns of predictors that vary for specific disease outcomes. While age, inflammation, cardiac stress and vascular disease dominate the prediction of death in the MESA study, impaired glucose metabolism and hypertension lead in the prediction of stroke and sub-clinical atherosclerosis markers occupy center stage in forecasting overall cardiovascular events be they limited to the heart (CHD) or involving the systemic circulation. For incident heart failure and atrial fibrillation, a combination of markers reflecting increased cardiac chamber stress coupled with electric dysregulation are at the forefront of potential outcome determinants.

Another important pattern of findings from this investigation was the underrepresentation of certain traditional CV risk factors such as gender, race/ethnicity, and therapy exposure (medication use) among the top predictors of disease outcome. Important exceptions to such trend were the presence of socioeconomic status as one of the top predictors of death and the role of hypertension as a top predictor of stroke, CVD, CHD as well as incident HF and AF. The lower than expected representation of traditional risk factors may stem from the fact that because fundamental to the genesis, maintenance and progression of CV diseases, they are intrinsic components of other phenotypes, particularly sub-clinical phenotypes that are more distal to disease initiation but closer to adverse outcomes. Even though some of

these risk factors did not feature in the top 20, they remain crucial to medical practice, particularly, disease prevention.

Machine learning and deep phenotyping.

The application and use of machine learning tools for CVD are still controversial.¹⁰ This is so, even as there has been an increasing use of imaging tests, ECG exams and lab tests in recent years.¹⁶ In most cases, even when multiple markers are acquired, not all are used for diagnosis.¹⁷ For example, regional function measures from imaging, large portions of biomarker panels or ECG signals are frequently ignored by many clinicians. As we move into the age of precision medicine, understanding the utility of phenotypic data and methods to analyze already acquired information is of paramount importance. Machine learning methods, and RF particularly, have been used before for CVD risk prediction.¹⁸⁻²⁷ In this study, we were able to use deeply phenotyped data to predict outcomes in a population study that accounts for time to event.²⁸ The added advantage is that these methods can be extended and refined, regularly, with new data. They also account for non-linearity in relationships (Figure 2), for example, both high and low values of ABI (as previously shown) were predictive of incidence of CVD.

This work also confirms the influence of certain markers and risk factors on CV events. From this analysis, the importance of markers, heretofore underestimated, such as TNF- α SR and IL2 SR, come to fore with machine learning. In this regard, machine learning opens the possibility of discovering new relationships that are not hypothesis driven and without prior assumptions. Identifying effective disease markers and discovering unknown mechanisms may be of benefit for effective screening strategies, and suggest specific targets for risk reduction. Yet another advantage of this technique is the ability to recognize the best predictors within a domain (questionnaires, imaging, etc), as well as their importance with respect to predictors from other domains. This approach to biomarker identification may be of particular benefit in intermediate risk groups where underlying subclinical risk is not apparent in traditional CV risk factors.

Methodological considerations.

In this study, We used the minimum depth of the maximal subtree as the main measure of variable importance because of prior research that showed inherent advantages to using this over permutation testing²⁹. While there are other methods to do the same, the change in Gini index, for example, they may not utilize survival data and hence may not be the best method in the case of survival analysis as ours. However, we provide the top-10 variables from both the Gini index (12-year cut-off) and permutation testing in the Supplement.

While deep phenotyping might help in biomarker discovery, it is seen from Online Figure III, that far fewer than the measured variables are necessary for obtaining a high C-index. To this end, RF methods may help in identifying important variables. We have shown the top-20 variables, as well as the C-index and BS using just the top-20 models. It is plausible that the optimal number of variables is less than 20. However, formal methods need to be developed with consideration to cost, appropriateness, ease of access, and reproducibility of measurements for a more judicious approach to variable selection for event prediction.

MESA, designed to study progression of subclinical disease to manifest symptoms and outcomes, was performed in a middle-aged population free of CV disease at baseline. Therefore, results may not generalize to other study populations. We did not include genetic data; the identification of the phenome-genome interaction and assessment of their combined prediction ability may potentially improve our findings.³⁰ While phenomorphing (longitudinal covariate data) and risk prediction is of interest, it is out of

the scope of this study. While this study identifies top predictors as a method of “biomarker discovery”, further work is required including validation in other populations, as the training and test datasets were drawn from within the MESA study population.

Conclusion.

In an extensively phenotyped population free of CV disease at baseline, using random forests, we show efficient cardiovascular risk prediction for specific outcomes including death, stroke, CV events, incident heart failure and atrial fibrillation. Inflammation, subclinical atherosclerosis, myocardial damage, and cardiac chamber stress were among the most important predictors across all outcomes. We provide a framework for “big data” applications to obtain meaningful risk prediction, biomarker identification, and generate data-driven hypotheses.

ACKNOWLEDGEMENTS

The information contained herein (for the MESA Columbia Field Center) was derived in part from data provided by the Bureau of Vital Statistics, New York City Department of Health and Mental Hygiene. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. The MESA protocol, including information about the populations from which recruitment occurred, detailed exclusion criteria, investigators, and other information, is available at www.mesa-nhlbi.org. A full list of participating MESA investigators and institutions can also be found.

SOURCES OF FUNDING

This research was supported by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute and by grants UL1-TR-000040 and UL1-TR-001079 from NCRR.

DISCLOSURES

The authors have nothing to disclose related to the manuscript.

REFERENCES

1. Lloyd-Jones DM. Cardiovascular Risk Prediction: Basic Concepts, Current Status, and Future Directions. *Circulation*. 2010;121:1768-1777.
2. Wong ND. Epidemiological studies of CHD and the evolution of preventive cardiology. *Nat Rev Cardiol*. 2014;11:276-289.
3. Gorodeski EZ, Ishwaran H, Kogalur UB, Blackstone EH, Hsieh E, Zhang Z-m, Vitolins MZ, Manson JE, Curb JD and Martin LW. Use of Hundreds of Electrocardiographic Biomarkers for Prediction of Mortality in Postmenopausal Women The Women's Health Initiative. *Circulation: Cardiovascular Quality and Outcomes*. 2011:CIRCOUTCOMES. 110.959023.
4. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M and Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156:871-81.
5. Akaike H. Likelihood of a model and information criteria. *Journal of econometrics*. 1981;16:3-14.
6. Tibshirani R. The lasso method for variable selection in the Cox model. *Statistics in medicine*. 1997;16:385-395.

7. Harrell FE, Califf RM, Pryor DB, Lee KL and Rosati RA. Evaluating the yield of medical tests. *Jama*. 1982;247:2543-2546.
8. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly weather review*. 1950;78:1-3.
9. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM and Kannel WB. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation*. 2008;117:743-753.
10. Goff DC, Lloyd-Jones DM, Bennett G, O'Donnell C, Coady S and Robinson J. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *J Am Coll Cardiol*. 2014.
11. Lloyd-Jones DM, Wilson PW, Larson MG, Beiser A, Leip EP, D'Agostino RB and Levy D. Framingham risk score and prediction of lifetime risk for coronary heart disease. *The American journal of cardiology*. 2004;94:20-24.
12. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34:187-220.
13. Ishwaran H, Kogalur UB, Blackstone EH and Lauer MS. Random survival forests. *The Annals of Applied Statistics*. 2008;841-860.
14. McClelland RL, Jorgensen NW, Budoff M, Blaha MJ, Post WS, Kronmal RA, Bild DE, Shea S, Liu K, Watson KE, Folsom AR, Khera A, Ayers C, Mahabadi A-A, Lehmann N, Jöckel K-H, Moebus S, Carr JJ, Erbel R and Burke GL. 10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Study). *Journal of the American College of Cardiology*. 2015;66:1643-1653.
15. Chahal H, Bluemke DA, Wu CO, McClelland R, Liu K, Shea SJ, Burke G, Balfour P, Herrington D and Shi P. Heart failure risk prediction in the Multi-Ethnic Study of Atherosclerosis. *Heart*. 2015;101:58-64.
16. Andrus BW and Welch HG. Medicare services provided by cardiologists in the United States: 1999–2008. *Circulation: Cardiovascular Quality and Outcomes*. 2012;5:31-36.
17. Lanktree MB, Hassell RG, Lahiry P and Hegele RA. Phenomics: expanding the role of clinical evaluation in genomic studies. *Journal of Investigative Medicine*. 2010;58:700-706.
18. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132:1920-1930.
19. Ishwaran H, Blackstone EH, Pothier CE and Lauer MS. Relative risk forests for exercise heart rate recovery as a predictor of mortality. *Journal of the American Statistical Association*. 2004;99:591-600.
20. Inuzuka R, Diller G-P, Borgia F, Benson L, Tay EL, Alonso-Gonzalez R, Silva M, Charalambides M, Swan L and Dimopoulos K. Comprehensive use of cardiopulmonary exercise testing identifies adults with congenital heart disease at increased mortality risk in the medium term. *Circulation*. 2012;125:250-259.
21. Hsich E, Gorodeski EZ, Blackstone EH, Ishwaran H and Lauer MS. Identifying important risk factors for survival in patient with systolic heart failure using random survival forests. *Circulation: Cardiovascular Quality and Outcomes*. 2011;4:39-45.
22. Sitar-tăut A, Zdrenghea D, Pop D and Sitar-tăut D. Using machine learning algorithms in cardiovascular disease risk evaluation. *Age*. 2009;1:4.
23. Colombet I, Ruelland A, Chatellier G, Gueyffier F, Degoulet P and Jaulent M. Models to predict cardiovascular risk: comparison of CART, multilayer perceptron and logistic regression. *Proceedings of the AMIA Symposium*. 2000:156.
24. Park G-M, Han S, Kim SH, Jo M-W, Her SH, Lee JB, Lee MS, Kim HC, Ahn J-M and Lee S-W. Model for assessing cardiovascular risk in a Korean population. *Circulation: Cardiovascular Quality and Outcomes*. 2014;7:944-951.
25. Shardell MD, Alley DE, Hicks GE, El-Kamary SS, Miller RR, Semba RD and Ferrucci L. Low-serum carotenoid concentrations and carotenoid interactions predict mortality in US adults: the Third National Health and Nutrition Examination Survey. *Nutrition research*. 2011;31:178-189.

26. Rizza S, Copetti M, Rossi C, Cianfarani M, Zucchelli M, Luzi A, Pecchioli C, Porzio O, Di Cola G and Urbani A. Metabolomics signature improves the prediction of cardiovascular events in elderly subjects. *Atherosclerosis*. 2014;232:260-264.
27. Rebholz CM, Grams ME, Matsushita K, Inker LA, Foster MC, Levey AS, Selvin E and Coresh J. Change in Multiple Filtration Markers and Subsequent Risk of Cardiovascular Disease and Mortality. *Clinical Journal of the American Society of Nephrology*. 2015:CJN. 10101014.
28. Yoo I, Alafaireet P, Marinov M, Pena-Hernandez K, Gopidi R, Chang J-F and Hua L. Data mining in healthcare and biomedicine: a survey of the literature. *Journal of medical systems*. 2012;36:2431-2448.
29. Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ and Lauer MS. High-Dimensional Variable Selection for Survival Data. *Journal of the American Statistical Association*. 2010;105:205-217.
30. Benjamin I, Brown N, Burke G, Correa A, Houser SR, Jones DW, Loscalzo J, Vasan RS and Whitman GR. American Heart Association Cardiovascular Genome-Phenome Study Foundational Basis and Program. *Circulation*. 2015;131:100-112.



Circulation Research

ONLINE FIRST

FIGURE LEGENDS

Figure 1. A flowchart describing the general framework of the study. Models were built using the training dataset, and the test dataset was used for computing the C-index and the Brier Score shown in Table 4.

Figure 2. Plots showing Lowess curves (for continuous variables) and box plots (for categorical variables) of the survival probability vs variable values for the top-5 predictors for each of the outcomes at 12 years. The y-axis represents survival probability calculated from the RF-20 algorithm (range: 0 to 1). The x-axis spans the range (or categories) of the variable of interest. **Abbreviations:** NT pro-BNP = N-terminal pro-Brain Natriuretic peptide, TNF- α SR = tissue necrosis factor- α soluble receptor, IL2 SR = interleukin-2 soluble receptor, CAC = coronary artery calcium score, LVESV = left ventricle end-systolic volume. **Units for each variable:** NT pro-BNP – pg/ml, TNF- α SR – pg/ml, IL2 SR – pg/ml, CAC – Agatston's units, cardiac troponin T – ng/ml, ABI – ratio, age – years, fasting glucose – mg/dl.

Figure 3. Plots showing the variable importance for each of the 735 variables used in analysis. The color of the dots represents the category or type of measurement. The legend on the right provides the phenotype category ordered from left-to-right on the individual plots. The variable importance is measured using the minimum depth of the maximal subtree, with lower values representing greater importance of corresponding variable. **Abbreviations:** NT pro-BNP = N-terminal pro-Brain Natriuretic peptide, TNF- α SR = tissue necrosis factor- α soluble receptor, IL2 SR = interleukin-2 soluble receptor, CAC = coronary artery calcium score, ABI = ankle-brachial index, IMT = intima media thickness, SBP = systolic blood pressure.

Figure 4. The concordance index for each of the models tested over time. The full models (models with all 735 variables) did not converge for the LASSO-Cox, AIC-Cox and the Cox PHM models, and hence are not shown here. The prediction ability of conventional risk scores for heart failure (MESA HF risk score), cardiovascular disease (AHA/ASCVD risk score) and coronary heart disease (Framingham CHD risk score) are also shown (yellow curve). In general, the C-index for all variables decreased over time.

ONLINE FIRST

NOVELTY AND SIGNIFICANCE

What Is Known?

- Machine learning techniques, such as the random survival techniques, may be an effective statistical methodology for handling biomedical data of increased volume, velocity, and variety, under the curse of dimensionality.
- These methods do not require a priori assumptions regarding causality and may thus be suitable for defining the role of novel biomarkers in cardiovascular disease prediction.

What New Information Does This Article Contribute?

- Machine learning methods are better suited for meaningful risk prediction in extensively phenotyped large-scale epidemiological studies than regular Cox proportional Hazards models or risk scores.
- Random survival forests may be an effective machine learning strategy for incident cardiovascular event prediction and risk stratification in large populations with large phenotypic datasets.

There is a lack of studies using machine-learning techniques with deep phenotyping (multiple evaluations of different aspects of a specific disease process) for cardiovascular event prediction. We examined the ability of combining deep phenotyping with machine learning for cardiovascular event prediction in the Multi-Ethnic Study of Atherosclerosis (MESA). The random survival forests based method of risk prediction yielded an entirely unexpected perspective on event prediction of specific outcomes such as death, stroke, cardiovascular events, incident heart failure and atrial fibrillation, with superior predictive power and improved accuracy than established risk scores. The results also suggest the importance of subclinical disease markers determined by imaging, electrocardiography and blood tests, as revealed by their prominent presence on the lists of the top-20 phenotyping predictors for the selected outcomes. This strategy could yield insights about specific use of variables for specific event prediction and guiding strategies to prevent cardiovascular disease outcomes. Potentially, these techniques could be applied retrospectively to analyze large phenotyping datasets for identifying disease mechanisms, and as a means of hypothesis generation, without prior assumptions.

TABLES

Table 1. A list of the markers that were used for prediction in this study.

Traditional Risk Factors, Demographics, Anthropometry, Site

Age, gender, race, body mass index, body surface area, waist-to-hip ratio, systolic blood pressure, diastolic blood pressure, pulse pressure, diabetes, smoking status, pack years, high density lipoprotein cholesterol, low density lipoprotein cholesterol, total cholesterol, triglycerides, heart rate, creatinine, site, waist circumference, Hip circumference, fasting glucose

Medication use

All hypertension, Angiotensin Converting Enzyme, Angiotensin-II Receptor Blockers, lipid-control, statins, beta-blockers, calcium channel blockers

Atherosclerotic markers – computed tomography, carotid ultrasonography

Coronary artery calcium score, ankle-brachial index, common and internal carotid artery intima media thickness, maximum carotid stenosis

Questionnaire

Family history of heart attacks, Alcohol use, Number of drinks per week, emphysema, asthma, arthritis, Cancer, liver disease, education level, economic status/income, exercise metabolic equivalents

Magnetic Resonance Imaging (MRI) markers

Left ventricular (LV) mass, LV End-diastolic volume, LV End-systolic volume, LV Ejection fraction, LV mass-volume ratio, LV stroke volume, LV sphericity index at end-diastole and end-systole, LV cardiac output, LV end-diastolic wall thickness, LV end-systolic wall thickness, ascending aortic distensibility, descending aortic distensibility, pulse wave velocity, maximum ascending aortic area, maximum descending aortic area, aortic arch distance, Maximum left atrial (LA) volume, Minimum LA volume, Maximum LA strain, Total LA ejection fraction, Passive LA ejection fraction, Active LA ejection fraction, Right ventricular (RV) mass, RV End-diastolic volume, RV End-systolic volume, RV Ejection fraction, RV stroke volume.

Lab Biomarkers

Interleukin-2 soluble receptor, Plasmin-Antiplasmin Complex, D-dimer, Factor viii, N-Terminal pro-Brain Natriuretic Peptide, cardiac troponin T, C-reactive protein, Interleukin-6, fibrinogen, homocysteine, Tissue necrosis factor-a soluble receptor

Electrocardiographic (ECG) main

PR duration, QRS duration, QT duration, P-axis, QRS axis, T-axis, Minnesota codes, ECG-LV hypertrophy by cornell voltage and novacode, heart rate variability short-term and overall components, Cornell voltage

ECG all

P, P', Q, R, R', S, S', T and T' wave duration, amplitude, area, and intrinsicoid; Middle and End of ST segment amplitudes; Amplitude at the point of 60 msec from J-point; STJ amplitude; total QRS area, balance, deflection balance, intrinsicoid; for each of the leads (AVL, AVR, AVF, I, II, III, V1, V2, V3, V4, V5, V6).

Table 2. General characteristics of the MESA sample at baseline, 2000-2002.

Variable	Value
Age (in years)	62·15 (10·23)
Gender (% female)	52·85
Race	
% African-American	27·78
% Caucasian	38·48
% Chinese-American	11·78
% Hispanic	21·95
Body Mass index (kg/m ²)	28·34 (5·48)
Diabetes	
% Impaired Fasting Glucose	13·83
% Treated	10·01
% Untreated	2·64
Systolic Blood Pressure (mm Hg)	126·59 (21·48)
Use of Hypertension Medication (%)	37·23
Heart Rate (bpm)	63·13 (9·66)
Smoking status	
% current	13·06
% former	36·62
Total Cholesterol (mg/dL)	194·16 (35·73)
HDL cholesterol (mg/dL)	50·96 (14·83)
Lipid Medication use (%)	16·15
Heart Failure, n (%)	259 (3·8)
All Cardiovascular Disease, n (%)	710 (10·4)
Coronary Heart Disease, n (%)	498 (7·3)
Atrial Fibrillation, n (%)	317 (4·7)
All-cause Death, n (%)	831 (12·2)
Stroke, n (%)	200 (2·9)



Table 3. The top-20 ranked variables by the variable importance from the random survival forest method for each of the outcomes of interest. The relative variable importance (RVI) of each variable can be assessed using the normalized minimal depth of the maximal subtree (which can be seen in Figure 3). The normalized RVI values vary from 0 (most important) to 1 (least important).

Rank	Death	RVI	Stroke	RVI
1	Age	0.00	Fasting glucose	0.00
2	Tissue Necrosis Factor- α soluble receptor	0.07	Interleukin-2 soluble receptor	0.09
3	Interleukin-2 soluble receptor	0.09	Maximum carotid stenosis	0.11
4	N-Terminal pro-Brain Natriuretic Peptide	0.16	Tissue Necrosis Factor- α soluble receptor	0.13
5	Ankle-Brachial Index	0.21	N-Terminal pro-Brain Natriuretic Peptide	0.16
6	Coronary Artery Calcium score	0.25	Internal carotid intima media thickness	0.18
7	Common carotid intima media thickness	0.26	Systolic blood pressure	0.24
8	Internal carotid intima media thickness	0.32	Pulse pressure	0.28
9	Descending aortic distensibility	0.33	Descending aortic distensibility	0.32
10	Plasmin-Antiplasmin Complex	0.35	Ankle-Brachial Index	0.32
11	Cardiac Troponin-T	0.37	Coronary Artery Calcium score	0.32
12	D-dimer	0.37	R Amplitude in Lead V2	0.32
13	Maximum ascending aortic area	0.38	R Amplitude in Lead V6	0.35
14	Ascending aortic distensibility	0.39	Minnesota code 1 score: F lead group	0.35
15	Homocysteine	0.39	Ascending aortic distensibility	0.37
16	Thoracic aorta arch length	0.41	Age	0.38
17	R Amplitude in Lead V	0.41	Cardiac output	0.39
18	Interleukin-6	0.41	JT Duration	0.40
19	Economic status/income	0.42	LV mass-volume ratio	0.40
20	Maximum descending aortic area	0.42	End-diastolic septal anterior wall thickness	0.41

Rank	Coronary Heart Disease	RVI	All CVD	RVI
1	Coronary Artery Calcium score	0.00	Coronary Artery Calcium score	0.00
2	Tissue Necrosis Factor- α soluble receptor	0.28	Tissue Necrosis Factor- α soluble receptor	0.24
3	Cardiac Troponin-T	0.31	N-Terminal pro-Brain Natriuretic Peptide	0.25
4	N-Terminal pro-Brain Natriuretic Peptide	0.35	Interleukin-2 soluble receptor	0.28
5	Minnesota code 1 score: F lead group	0.36	Cardiac Troponin-T	0.35
6	Ankle-Brachial Index	0.37	Ankle-Brachial Index	0.40
7	Common carotid intima media thickness	0.44	Common carotid intima media thickness	0.41
8	Interleukin-2 soluble receptor	0.48	Pulse pressure	0.41
9	Pack years of smoking	0.50	Maximum ascending aortic area	0.42
10	Internal carotid intima media thickness	0.50	Internal carotid intima media thickness	0.42

11	Factor VIII	0·50	Age	0·42
12	End-systolic mid-ventricular septal wall thickness	0·52	R Amplitude in Lead V4	0·44
13	Maximum descending aortic area	0·53	Systolic blood pressure	0·44
14	End-systolic mid-ventricular antero-septal wall thickness	0·54	Factor VIII	0·46
15	Maximum ascending aortic area	0·54	Ascending aortic distensibility	0·47
16	S Amplitude in Lead AVR	0·55	Waist-to-hip ratio	0·47
17	End-diastolic basal septal wall thickness	0·55	Minnesota code 1 score: F lead group	0·48
18	Left ventricular ejection fraction	0·56	End-diastolic basal septal wall thickness	0·48
19	Pulse pressure	0·56	Plasmin-Antiplasmin Complex	0·49
20	Descending aortic distensibility	0·56	End-diastolic basal inferior wall thickness	0·49

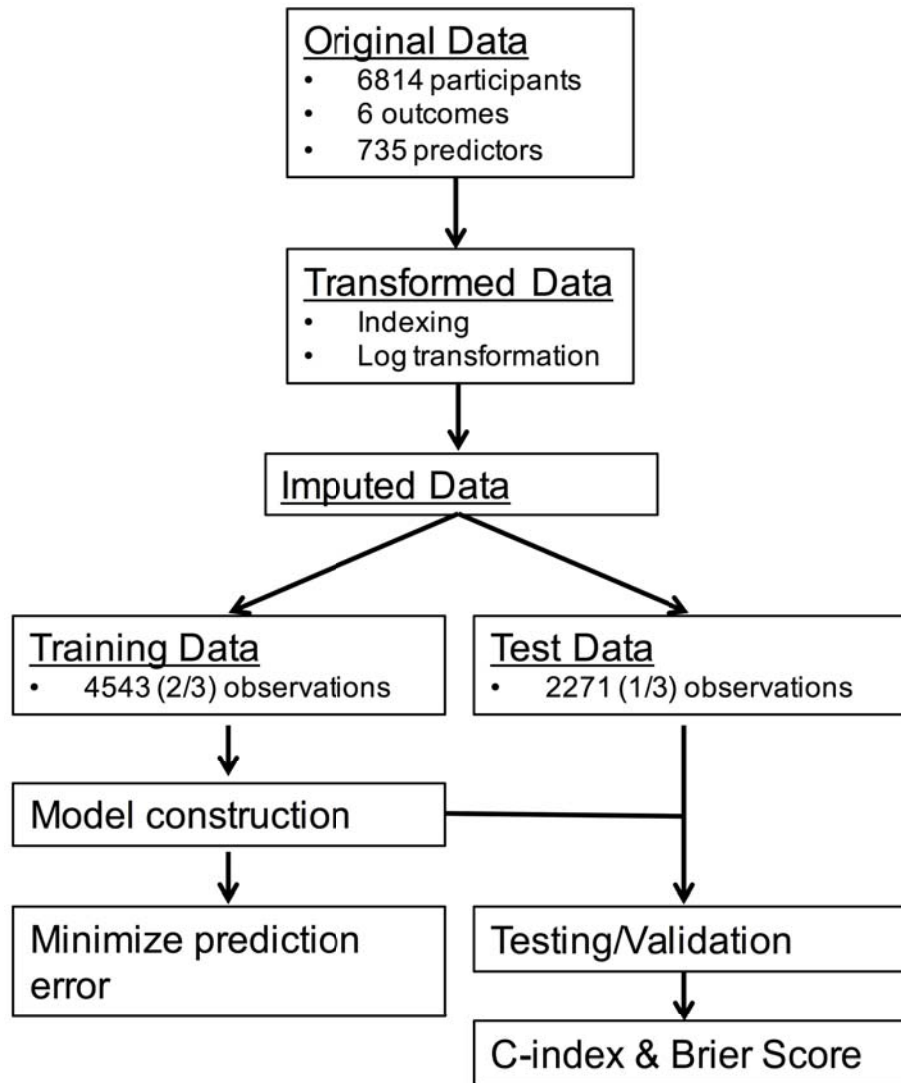
Rank	Heart Failure	RVI	Atrial Fibrillation	RVI
1	N-Terminal pro-Brain Natriuretic Peptide	0·00	N-Terminal pro-Brain Natriuretic Peptide	0·00
2	Tissue Necrosis Factor- α soluble receptor	0·07	Coronary Artery Calcium score	0·23
3	Coronary Artery Calcium score	0·07	Age	0·27
4	End-systolic left ventricular volume	0·10	Creatinine	0·31
5	Cardiac Troponin-T	0·20	Ankle-Brachial Index	0·36
6	End-diastolic left ventricular volume	0·21	Interleukin-2 soluble receptor	0·39
7	Left ventricular ejection fraction	0·24	Tissue Necrosis Factor- α soluble receptor	0·41
8	QTC INTERVAL	0·32	Common carotid intima media thickness	0·45
9	QT Index	0·34	R Amplitude in Lead V4	0·53
10	Interleukin-2 soluble receptor	0·36	STJ Amplitude in Lead V5	0·53
11	Waist-to-hip ratio	0·38	Internal carotid intima media thickness	0·55
12	Ankle-Brachial Index	0·42	Pulse pressure	0·55
13	PR INTERVAL	0·45	Estimate of overall heart rate variability	0·56
14	Creatinine	0·45	End-systolic basal lateral wall thickness	0·56
15	Pulse pressure	0·46	End-systolic mid-ventricular anterior wall thickness	0·56
16	End-diastolic left ventricular mass	0·47	Heart Rate	0·57
17	Estimate of overall heart rate variability	0·51	QRS AXIS (degrees)	0·57
18	T Amplitude in Lead V1	0·51	Cardiac Troponin-T	0·57
19	Minnesota code 1 score: V lead group	0·51	Total left atrial ejection fraction	0·57
20	Minnesota code 1 score: F lead group	0·52	Pack-years of smoking	0·58

Table 4. The number of variables and the performance (Concordance-index and Brier score) for each of the models tested as well as for the risk scores at the end of follow-up.

	DTH	STRK	CHD	CVD	HF	AF
Number of variables						
<i>RSF with all covariates</i>	735	735	735	735	735	735
<i>RSF with top 20 covariates</i>	20	20	20	20	20	20
<i>AIC-Cox with Forward Selection</i>	13	9	5	6	5	6
<i>Cox with top 20 RSF covariates</i>	20	20	20	20	20	20
<i>Lasso Cox with top 20 RSF covariates</i>	19	17	19	19	10	15
<i>AIC Cox Backward Selection with top 20 RSF covariates</i>	16	12	13	13	11	12
Concordance-index @ 12 years						
<i>RSF with all covariates</i>	0.86	0.77	0.81	0.81	0.84	0.82
<i>RSF with top 20 covariates</i>	0.84	0.75	0.80	0.80	0.84	0.75
<i>AIC-Cox with Forward Selection</i>	0.78	0.70	0.74	0.74	0.78	0.79
<i>Cox with top 20 RSF covariates</i>	0.80	0.66	0.75	0.76	0.81	0.78
<i>Lasso Cox with top 20 RSF covariates</i>	0.80	0.67	0.75	0.76	0.82	0.78
<i>AIC Cox Backward Selection with top 20 RSF covariates</i>	0.80	0.68	0.75	0.76	0.80	0.78
Brier Score @ 12 years						
<i>RSF with all covariates</i>	0.083	0.031	0.067	0.083	0.035	0.039
<i>RSF with top 20 covariates</i>	0.076	0.030	0.065	0.079	0.033	0.038
<i>AIC-Cox with Forward Selection</i>	0.088	0.032	0.069	0.087	0.035	0.045
<i>Cox with top 20 RSF covariates</i>	0.086	0.031	0.070	0.087	0.035	0.037
<i>Lasso Cox with top 20 RSF covariates</i>	0.086	0.031	0.070	0.087	0.033	0.038
<i>AIC Cox Backward Selection with top 20 RSF covariates</i>	0.086	0.031	0.069	0.087	0.035	0.038

HF: heart failure, CVD: all cardiovascular disease, CHD: coronary heart disease, AF: atrial fibrillation, DTH: death, STRK: stroke, RSF: random survival forest, LASSO: least absolute shrinkage and selection operator, AIC: Akaike Information Criteria.

FIGURE 1



Models

1. Random forest with all predictors
2. Cox model with all predictors
3. LASSO-Cox with all predictors
4. AIC-Cox Backward Stepwise Regression with all predictors
5. RF with top 20 RF predictors
6. Cox model with top 20 RF predictors
7. LASSO-Cox with top 20 RF predictors
8. AIC-Cox Backward Stepwise Regression with top 20 RF predictors
9. AIC-Cox Forward Regression with all predictors

FIGURE 2

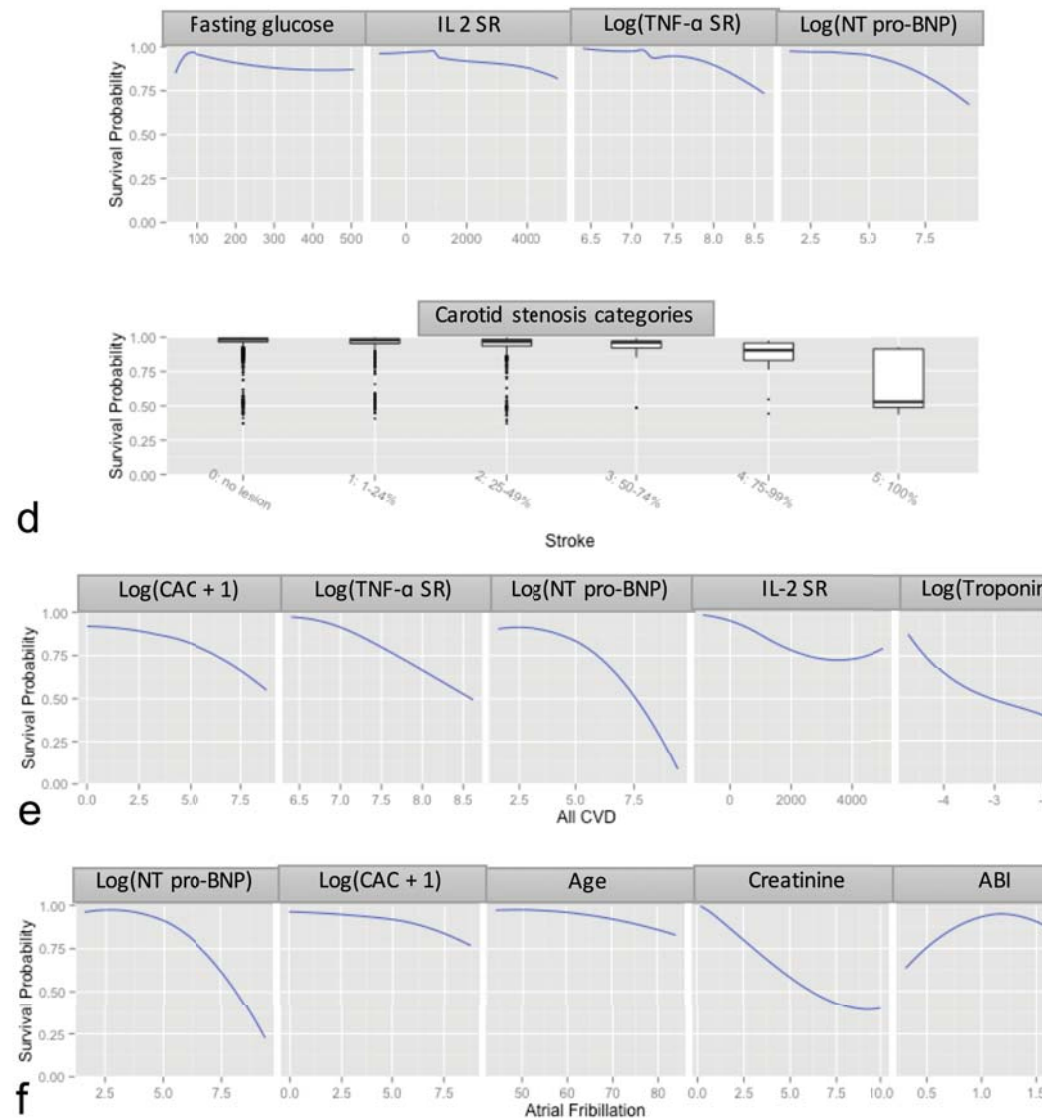
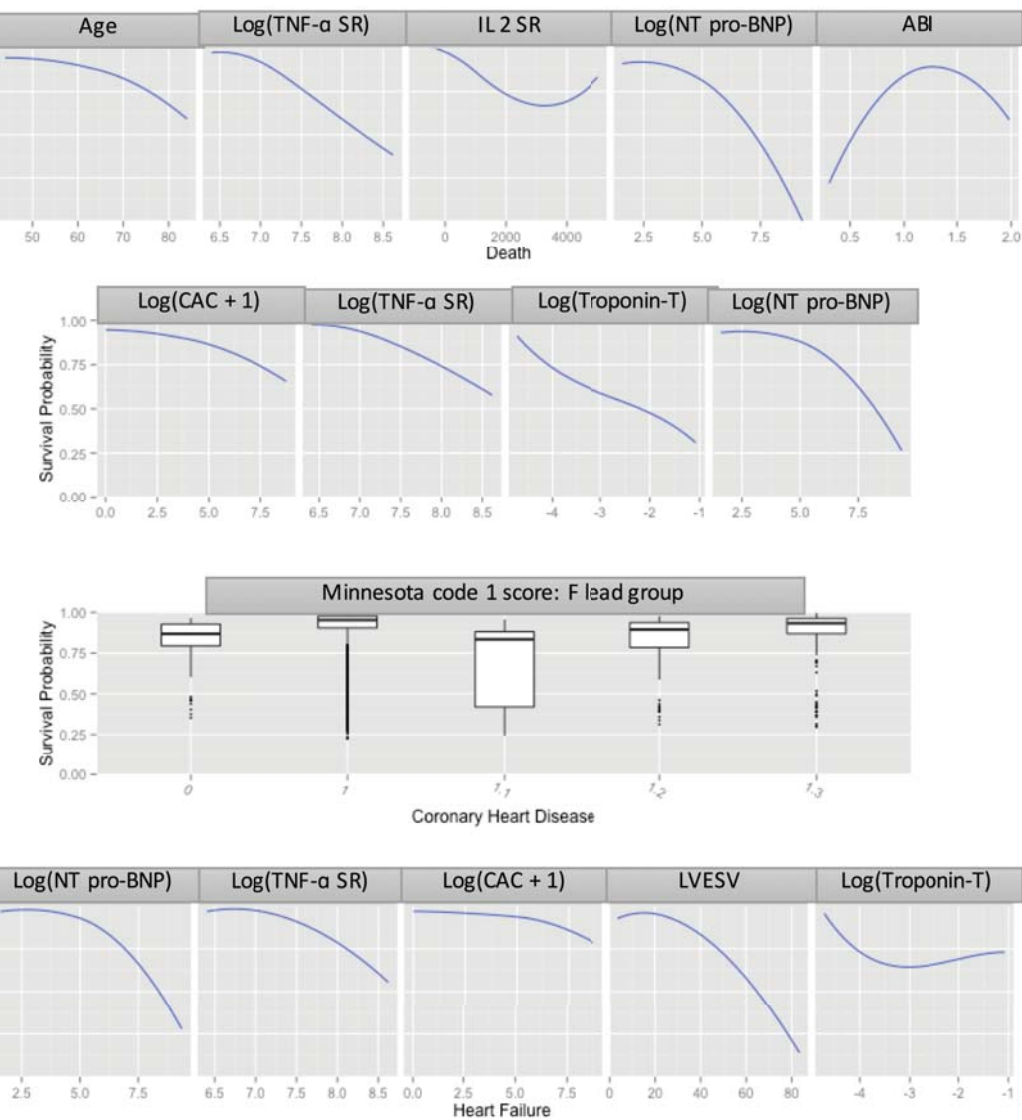


FIGURE 3

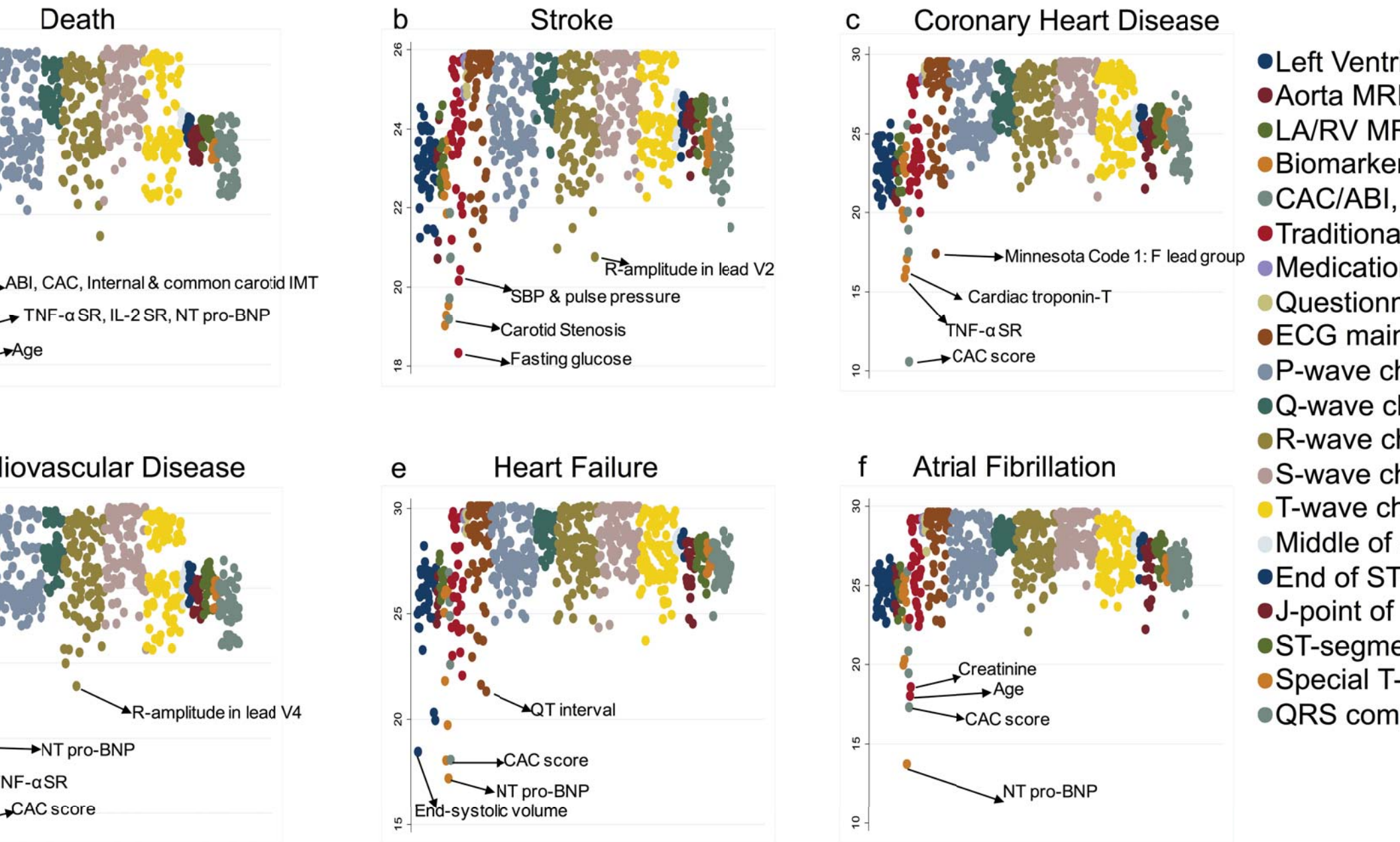
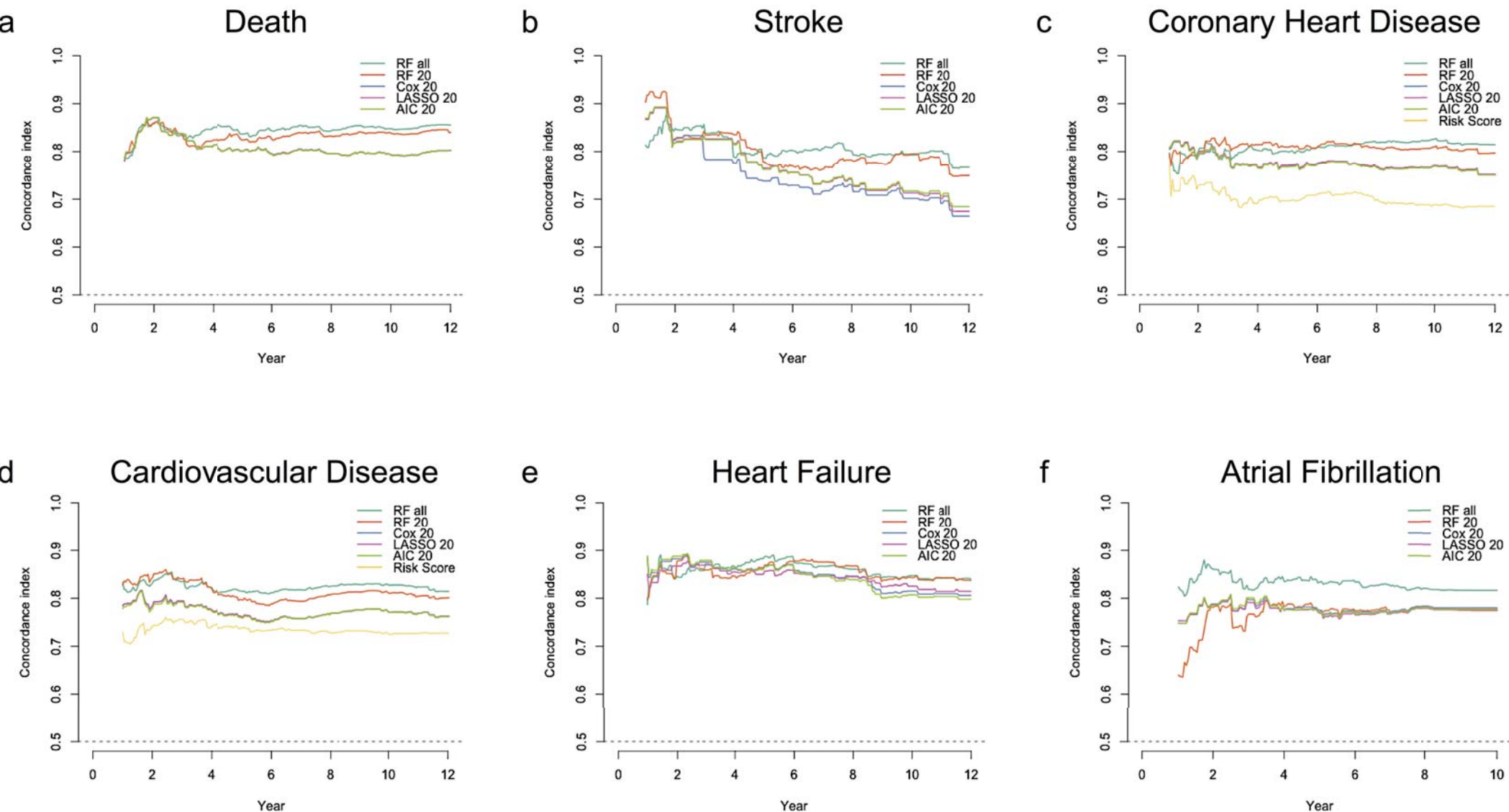


FIGURE 4



Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis

Bharath Ambale-Venkatesh, Xiaoying yang, Colin O Wu, Kiang Liu, W G Hundley, Robyn L McClelland, Antoinette S Gomes, Aaron R Folsom, Steven Shea, Eliseo Guallar, David A Bluemke and João A Lima

Circ Res. published online August 9, 2017;

Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2017 American Heart Association, Inc. All rights reserved.

Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://circres.ahajournals.org/content/early/2017/08/09/CIRCRESAHA.117.311312>

Data Supplement (unedited) at:

<http://circres.ahajournals.org/content/suppl/2017/08/09/CIRCRESAHA.117.311312.DC1>

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation Research* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:
<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Circulation Research* is online at:
<http://circres.ahajournals.org/subscriptions/>

Title

Online only Supplemental Material for: “Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis”

eMethods

Assessment of predictors

Information regarding examinations within MESA has been described previously.¹ Questionnaires were used to ascertain each participant's demographic, smoking status, alcohol consumption, medical conditions, access to medical care, family history of CVD, and use of medications. Physical activity was measured by using a questionnaire adapted from the Cross-Cultural Activity Participation Study. During the examination, anthropometric quantities were measured. Resting blood pressure was measured three times in the seated position using a sphygmomanometer and the average of the last two measurements were used in analysis. Chest computed tomography (CT) performed using either multidetector row CT or electron beam CT measured Agatston score, a well-established measure of coronary artery calcium score (CAC) to evaluate the CAC score per patient.² Cardiac magnetic resonance imaging (MRI) was performed using scanners with 1.5-T magnets, and with phased-array surface coils and electrocardiogram gating. Imaging consisted of cine images of the whole heart for size and function assessment of the left atrium^{3,4} and ventricle,^{5,6} phase contrast flow images of measure aortic structure⁷ and stiffness⁸ and tagged MRI of the left ventricle for regional deformation assessment.^{9,10} For carotid ultrasonography, images of the right and left common and internal carotid arteries were captured, using high-resolution B-mode ultrasound, and the bulb and distensibility of the distal common carotid artery were obtained.¹¹ To obtain the ankle-brachial index, blood pressure was measured with a Doppler probe in the bilateral brachial, dorsalis pedis, and posterior tibial arteries.¹² For electrocardiography, three 12-lead recordings were obtained using a Marquette MAC-PC instrument, Minnesota coding was used to classify abnormalities.¹³⁻¹⁹ Blood was drawn from participants, and aliquots prepared for central analysis. Measurements were performed to assess lipids and lipoproteins,²⁰ systemic inflammation and hemostasis,²¹ N-terminal pro-B-type natriuretic peptide and cardiac troponin-T,²² and fibrinolysis,^{23,24} and insulin resistance.²⁵ A random urine sample was collected, with one aliquot being analyzed centrally for creatinine. Online Table I provides a list of the variables used.

Event Adjudication

Events adjudicated as incident heart failure (HF), atrial fibrillation (AF), Dementia, coronary heart disease (CHD), all cardiovascular disease (CVD) and death (or all-cause mortality) as part of the MESA study were used as end-points. In addition to MESA follow-up examinations every two years, a telephone interviewer contacted each participant (or representative) every six–nine months to inquire about all interim hospital admissions, cardiovascular outpatient diagnoses, and deaths. Two physicians reviewed all records for independent end-point classification and assignment of event dates.

Criteria for CHD included any of – myocardial infarction (MI), resuscitated cardiac arrest (RCA), definite angina, probable angina (if followed by revascularization) and CHD death. Reviewers classified MI as definite, probable, or absent, based primarily on combinations of symptoms, ECG, and cardiac biomarker levels. In most cases, definite or probable MI required either abnormal cardiac biomarkers (two times upper limits of normal) regardless of pain or ECG findings; evolving Q waves regardless of pain or biomarker findings; or a combination of chest pain, and ST-T evolution or new LBBB, and biomarker levels one-two times upper limits of normal. Reviewers classified RCA when a patient successfully recovered from a full cardiac arrest through cardiopulmonary resuscitation (including cardioversion). Definite or probable angina required symptoms of typical chest pain or atypical symptoms. Probable angina required, in addition to symptoms, a physician diagnosis of angina and medical treatment for it. Definite angina required one or more additional criteria, including CABG surgery or other revascularization procedure; 70% or greater obstruction on coronary angiography; or evidence of ischemia by stress tests or by resting ECG. Coronary revascularization or, a physician diagnosis of angina, or CHD, in the absence of symptoms, was not considered to be angina. Fatal CHD was classified as definite, possible, or absent. Definite fatal CHD required a documented MI within the previous 28 days, chest pain within the 72 hours before death, or a history of CHD, and required the absence of a known non-atherosclerotic or non-cardiac cause of death. If the definite fatal CHD criteria were not met, possible fatal CHD could be assigned with an underlying cause of death consistent with fatal CHD and required the absence of a known non-atherosclerotic or non-cardiac cause of death.

CVD was considered a composite of MI, RCA, definite or probable angina, stroke, stroke death, CHD death, atherosclerotic death and CVD death. Stroke was classified as present or absent and consisted of rapid onset of a documented focal neurologic deficit lasting 24 hours or until death, or if < 24 hours, there was a clinically relevant lesion on brain imaging. Patients with focal neurologic deficits secondary to brain trauma, tumor, infection, or other non-

vascular cause were excluded. Cause of death was assigned for potential CVD deaths through committee review as part of MESA.

Reviewers classified HF as definite, probable, or absent. Definite or probable HF required heart failure symptoms, such as shortness of breath or edema, as asymptomatic disease is not a MESA endpoint. In addition to symptoms, probable HF required HF diagnosed by a physician and patient receiving medical treatment for HF. Definite HF required one or more other criteria, such as pulmonary edema/congestion by chest X-ray; dilated ventricle or poor LV function by echocardiography or ventriculography; or evidence of left ventricular diastolic dysfunction. We considered participants not meeting any criteria, including just a physician diagnosis of HF without any other evidence, as having no HF.

AF cases were detected using hospital discharge International Classification of Diseases, Ninth Revision (ICD9) diagnosis codes for AF or atrial flutter (427.31 or 427.32). MESA ascertained hospital discharge ICD-9 codes and Centers for Medicare and Medicaid Services (CMS) inpatient hospital claims. AF events during a hospital stay with coronary artery bypass surgery or valve replacement surgery were not counted as incident events.

Death was based on all-cause mortality while an additional endpoint – event-free survival was also examined. Event-free survival was classified as all participants with none of the seven outcomes examined during the follow-up period.

Statistical Analysis

Figure 1 shows the statistical analysis procedures followed in this study. Following data imputation, sufficient data for all predictors and endpoints was available in 6814 participants leading to a total of 5,199,082 data points to predict six outcomes over 10.3 median years of follow-up. As a training data set, 66.6% of the dataset was randomly selected from the overall group of participants; the remaining 33.3% were used as the validation dataset. The training dataset was used for model construction using the different approaches and optimized to reduce prediction error and maximize prediction ability. These models were then tested on the testing dataset to examine model performance and identify the best predictors. Data assembly was performed with STATA version 13.1 (Texas, USA). Analyses were performed using R software (www.r-project.org), using publically available libraries for Cox proportional hazards model (PHM),^{26, 27} Lasso-Cox,²⁸⁻³⁰ AIC-Cox,^{31, 32} and RF methods.^{33, 34} The adaptive tree imputation method³⁵ was used for imputation of missing data.

Survival Random Forest

A binary tree is a decision tool that uses a binary tree-like graph or model of decisions and their possible consequences. It is a flowchart like structure where each node represents a decision (based on a selected variable) and the two branches of the node represent the outcome of the test. Each branch could lead to two leaf nodes and to further subtrees based on the classification. It is one of the most popular techniques used in data mining or machine learning.

The survival random forests (RSF) method, introduced by Ishwaran et al,³³ is an ensemble tree method for analysis of right-censored data. A high-level description of the algorithm can be stated as: (a) drawing bootstrap samples from the data, with each bootstrap sample excluding one-third of the data called out-of-bag data. This is similar to k-fold cross-validation to limit over-fitting; (b) growing a binary survival tree for each bootstrap sample by recursive splitting of tree nodes starting at the root node, which is the top of the tree comprising all data; (c) At each node of the tree including the root node, a subset of all the variables in the dataset are randomly selected - this allows each variable to assert its' importance to event prediction; (d) each node is then split based on the candidate variable (from the subset) that maximizes the survival difference between daughter nodes – this gives us the maximum cumulative hazard function at that node; (e) the tree is grown to full size such that each terminal node, that is the most extreme nodes in a fully grown tree, has at least one unique outcome – this is both an efficient and sufficient stopping criterion; (f) calculating a cumulative hazard rate function for each terminal node within a tree - all cases within the node in a tree will therefore have the same cumulative hazard rate; (g) averaging over all trees to obtain an ensemble hazard function; and (h) using the out-of-bag sample to obtain an out-of-bag prediction error. An illustration is provided in Online Figure I.

While RSF can be used instead of Cox regression analysis for prediction, it can also be used as an efficient variable selection technique either on its own or in conjunction with other methods. For variable selection using RSF, the variables are ranked by the mean of the minimal depth of the maximal subtree over the entire forest. A subtree using a particular variable is defined as a maximal subtree (of that variable) if there is no other subtree (of the full tree) closer to the root node using the variable of interest. The minimal depth of a node is defined as the distance of the node from the root node.

In general, variables appearing higher on the tree, closer to the root node, have a higher rank (and hence are more important) as compared to the variables that appear first near the terminal nodes.

Data transformations, indexing, and imputation

While transformation of variables is not necessary for the RF method, it is necessary for other regression methods to work efficiently. Therefore, we used logarithm transformation for variables as necessary based on the visual histogram (NT pro-BNP, TNF α SR, Troponin-T, C-reactive protein, interleukin-6, homocysteine, aortic distensibility, pulse wave velocity, and Agatston's calcium score). In addition, to allow for easier interpretation (in keeping with prior literature), certain MRI variables were indexed to body surface area – LV mass, volumes, aortic dimensions, and left atrial volumes.

The adaptive tree imputation method introduced by Ishwaran et al was used for imputation of missing data. This algorithm works by adaptively imputing missing data even as the tree is grown by drawing randomly from the set of non-missing in-bag data within the working node. At the end of the imputation stage, each missing value is replaced by the corresponding average imputed values (or the majority vote for categorical variable) from the entire forest. For the other models, such as the AIC-Cox, LASSO-Cox and Cox PHM models, the imputed dataset with all the missing values replaced using the random forest imputation procedure detailed above was used. All variables available in less than 40% of the population were excluded from analysis. Online Figure II shows the numbers available distribution.

Models tested

We tested eight different models in our analysis in addition to the null model (all regression coefficients = zero). The first model used the random survival forest (RF) algorithm on all available variables.³⁵ The RF method is an ensemble tree method for analysis of right-censored data. Using the training dataset, 1000 binary survival trees were grown to form an ensemble RF with each tree built on a randomly selected bootstrap sample, which, on average, excluded 37% of the data. From the bootstrap sample, each tree is grown by recursive binary splitting of data, among a subset of 28 randomly selected variables (out of the 735 total). Each split is made by determining a split point of a single variable that maximizes the cumulative hazard of the two resulting data subset, i.e., the daughter nodes. The splitting stops when the data at hand can no longer be split, i.e., reaching the terminal nodes. For each tree, the cumulative hazard rate of a case is determined based on the terminal node that contains it. An ensemble hazard function (and the survival probability) is then estimated by averaging over all trees. The AIC-Cox using the Akaike Information Criterion (AIC) for Cox regression with backward stepwise elimination as well as the LASSO-Cox, a method for variable selection and shrinkage in Cox's proportional hazards model, were also tested in addition to the Cox-PHM.

While RF can be used instead of Cox regression analysis for risk prediction, it can also be used as an efficient variable selection technique either on its own or in conjunction with other methods. For variable selection using RF, the variables were ranked by the mean of the minimal depth of the maximal subtree (highest point in the tree of a variable) over the entire forest (averaged over all 1000 trees). In general, variables appearing higher on the tree, closer to the root node, had a higher rank. The top-20 ranked variables were then used again with RF, AIC-Cox, LASSO-Cox and regularized Cox PHM models.

Performance Evaluation

We assessed the performance of each prediction model to discriminate outcomes at different event times using Harrell's concordance index (C-index),³⁶⁻³⁸ and the accuracy of the predictions (mean squared distance between the predicted and actual probabilities) using the Brier score (BS).^{39, 40} Performance measures (C-index and BS) for nested models (of the overall models) using subsets of all the predictors (added based on increasing variable importance) were also calculated to assess problems of overfitting. We also compared the results of RF techniques to other known risk scores.⁴¹⁻⁴³

References

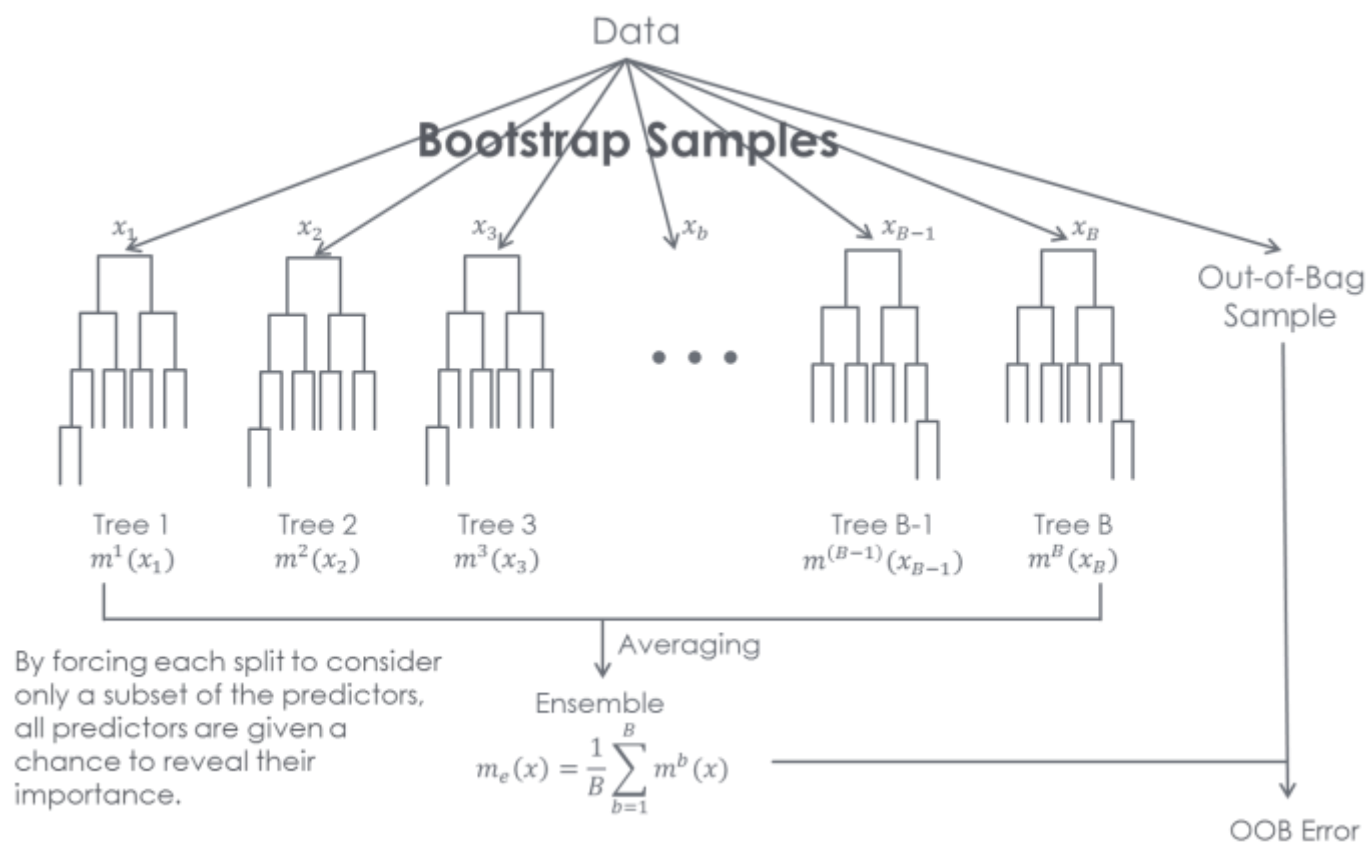
1. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M and Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol*. 2002;156:871-81.
2. Carr JJ, Nelson JC, Wong ND, McNitt-Gray M, Arad Y, Jacobs DR, Sidney S, Bild DE, Williams OD, Detrano RC, Jacobs Jr. DR, Sidney S, Bild DE, Williams OD and Detrano RC. Calcified Coronary Artery Plaque Measurement with Cardiac CT in Population-based Studies: Standardized Protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) Study1. *Radiology*. 2005;234:35-43.

3. Habibi M, Chahal H, Opdahl A, Gjesdal O, Helle-Valle TM, Heckbert SR, McClelland R, Wu C, Shea S, Hundley G, Bluemke DA and Lima JA. Association of CMR-Measured LA Function With Heart Failure Development: Results From the MESA Study. *JACC Cardiovasc Imaging*. 2014;7:570-579.
4. Zareian M, Habibi M, Venkatesh BA, Opdahl A, Chamera EH, Wu C, Zemrak F, Bluemke D and Lima JA. Left atrial structure and functional quantitation using cardiac magnetic resonance: comparison of manual delineation vs. multimodality tissue tracking based semi-automated methods. *Journal of Cardiovascular Magnetic Resonance*. 2014;16:P348.
5. Natori S, Lai S, Finn JP, Gomes AS, Hundley WG, Jerosch-Herold M, Pearson G, Sinha S, Arai A and Lima JA. Cardiovascular function in multi-ethnic study of atherosclerosis: normal values by age, sex, and ethnicity. *American Journal of Roentgenology*. 2006;186:S357-S365.
6. Cheng S, Fernandes VRS, Bluemke DA, McClelland RL, Kronmal RA and Lima JAC. Age-related left ventricular remodeling and associated risk for cardiovascular outcomes: the Multi-Ethnic Study of Atherosclerosis. *Circulation: Cardiovascular Imaging*. 2009;2:191-198.
7. Malayeri AA, Natori S, Bahrami H, Bertoni AG, Kronmal R, Lima JAC and Bluemke DA. Relation of aortic wall thickness and distensibility to cardiovascular risk factors (from the Multi-Ethnic Study of Atherosclerosis [MESA]). *Am J Cardiol*. 2008;102:491-496.
8. Redheuil A, Yu WC, Wu CO, Mousseaux E, de Cesare A, Yan R, Kachenoura N, Bluemke D and Lima JA. Reduced ascending aortic strain and distensibility: earliest manifestations of vascular aging in humans. *Hypertension*. 2010;55:319-26.
9. Ambale-Venkatesh B, Armstrong AC, Liu CY, Donekal S, Yoneyama K, Wu CO, Gomes AS, Hundley GW, Bluemke DA and Lima JA. Diastolic function assessed from tagged MRI predicts heart failure and atrial fibrillation over an 8-year follow-up period: The multi-ethnic study of atherosclerosis. *European Heart Journal Cardiovascular Imaging*. 2014;15:442-449.
10. Venkatesh BA, Donekal S, Yoneyama K, Wu C, Fernandes VR, Rosen BD, Shehata ML, McClelland R, Bluemke DA and Lima JA. Regional myocardial functional patterns: Quantitative tagged magnetic resonance imaging in an adult population free of cardiovascular risk factors: The multi-ethnic study of atherosclerosis (MESA). *J Magn Reson Imaging*. 2014;15:24749.
11. Polak JF, Szklo M, Kronmal RA, Burke GL, Shea S, Zavodni AE and O'Leary DH. The value of carotid artery plaque and intima-media thickness for incident cardiovascular disease: the multi-ethnic study of atherosclerosis. *Journal of the American Heart Association*. 2013;2:e000087.
12. Criqui MH, McClelland RL, McDermott MM, Allison MA, Blumenthal RS, Aboyans V, Ix JH, Burke GL, Liu K and Shea S. The ankle-brachial index and incident cardiovascular events in the MESA (Multi-Ethnic Study of Atherosclerosis). *Journal of the American College of Cardiology*. 2010;56:1506-1512.
13. Prineas RJ, Crow RS and Zhang Z-M. *The Minnesota code manual of electrocardiographic findings*: Springer Science & Business Media; 2009.
14. Beinart R, Zhang Y, Lima JA, Bluemke DA, Soliman EZ, Heckbert SR, Post WS, Guallar E and Nazarian S. The QT interval is associated with incident cardiovascular events: the MESA study. *Journal of the American College of Cardiology*. 2014;64:2111-2119.
15. Chrispin J, Jain A, Soliman EZ, Guallar E, Alonso A, Heckbert SR, Bluemke DA, Lima JA and Nazarian S. Association of electrocardiographic and imaging surrogates of left ventricular hypertrophy with incident atrial fibrillation: MESA (Multi-Ethnic Study of Atherosclerosis). *Journal of the American College of Cardiology*. 2014;63:2007-2013.
16. Ebong IA, Bertoni AG, Soliman EZ, Guo M, Sibley CT, Chen Y-DI, Rotter JI, Chen Y-C and Goff Jr DC. Electrocardiographic abnormalities associated with the metabolic syndrome and its components: the multi-ethnic study of atherosclerosis. *Metabolic syndrome and related disorders*. 2012;10:92-97.
17. Ilkhanoff L, Liu K, Ning H, Nazarian S, Bluemke DA, Soliman EZ and Lloyd-Jones DM. Association of QRS duration with left ventricular structure and function and risk of heart failure in middle-aged and older adults: the Multi-Ethnic Study of Atherosclerosis (MESA). *European journal of heart failure*. 2012;14:1285-1292.
18. Jain A, Tandri H, Dalal D, Chahal H, Soliman EZ, Prineas RJ, Folsom AR, Lima JA and Bluemke DA. Diagnostic and prognostic utility of electrocardiography for left ventricular hypertrophy defined by magnetic resonance imaging in relationship to ethnicity: the Multi-Ethnic Study of Atherosclerosis (MESA). *American heart journal*. 2010;159:652-658.
19. Li Y, Dawood FZ, Chen H, Jain A, Walsh JA, Alonso A, Lloyd-Jones DM and Soliman EZ. Minor isolated Q waves and cardiovascular events in the MESA study. *The American journal of medicine*. 2013;126:450. e9-450. e16.
20. Ariyo AA, Thach C and Tracy R. Lp(a) lipoprotein, vascular disease, and mortality in the elderly. *N Engl J Med*. 2003;349:2108-2115.

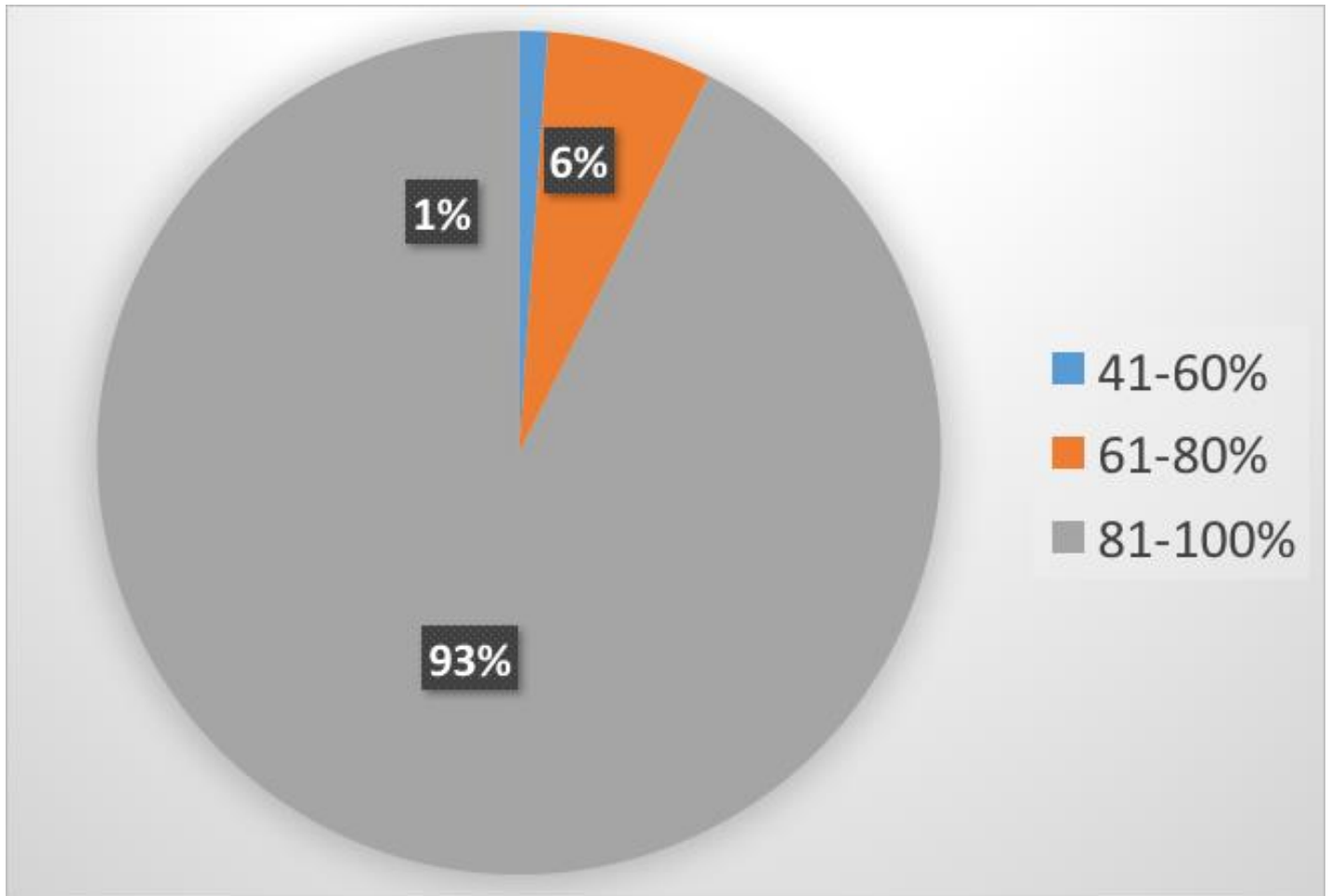
21. Arnett DK, McClelland RL, Bank A, Bluemke DA, Cushman M, Szalai AJ, Jain N, Gomes AS, Heckbert SR and Hundley WG. Biomarkers of inflammation and hemostasis associated with left ventricular mass: The Multiethnic Study of Atherosclerosis (MESA). *International journal of molecular epidemiology and genetics*. 2011;2:391.
22. Daniels LB, Clopton P, Sanchez OA, Bahrami H, Lima JA, Tracy RP, Siscovick D, Bertoni AG, Greenland P and Cushman M. Serial measurement of N-terminal pro-B-type natriuretic peptide and cardiac troponin T for cardiovascular disease risk assessment in the Multi-Ethnic Study of Atherosclerosis (MESA). *American Heart Journal*. 2015;170:1170-1183.
23. Smiles AM, Macy EM, Sakkinen PA, Bovill EG, Mann KG and Tracy RP. Biological and analytical variability in plasma prothrombin concentration: implications for use in epidemiology. *Blood Coag Fibrinolysis*. 1998;9:525-531.
24. Lutsey P, Cushman M, Steffen L, Green D, Barr R, Herrington D, Ouyang P and Folsom A. Plasma hemostatic factors and endothelial markers in four racial/ethnic groups: the MESA study. *Journal of Thrombosis and Haemostasis*. 2006;4:2629-2635.
25. Bertoni AG, Wong ND, Shea S, Ma S, Liu K, Preethi S, Jacobs DR, Wu C, Saad MF and Szklo M. Insulin Resistance, Metabolic Syndrome, and Subclinical Atherosclerosis The Multi-Ethnic Study of Atherosclerosis (MESA). *Diabetes Care*. 2007;30:2951-2956.
26. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34:187-220.
27. Therneau T and Lumley T. A Package for Survival Analysis in S. version 2.38. Available online at <http://cran.r-project.org/web/packages/snow/index.html>. 2015.
28. Tibshirani R. The lasso method for variable selection in the Cox model. *Statistics in medicine*. 1997;16:385-395.
29. Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*. 2010;33:1.
30. Simon N, Friedman J, Hastie T and Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *Journal of statistical software*. 2011;39:1-13.
31. Akaike H. Likelihood of a model and information criteria. *Journal of econometrics*. 1981;16:3-14.
32. Venables WN and Ripley BD. *Modern applied statistics with S-PLUS*: Springer Science & Business Media; 2013.
33. Ishwaran H, Kogalur UB, Blackstone EH and Lauer MS. Random survival forests. *The Annals of Applied Statistics*. 2008:841-860.
34. Ishwaran H and Kogalur U. Random Forests for Survival, Regression and Classification (RF-SRC), R package version 1.6. URL <http://cran.r-project.org/package=randomForestSRC>. 2014.
35. Gorodeski EZ, Ishwaran H, Kogalur UB, Blackstone EH, Hsieh E, Zhang Z-m, Vitolins MZ, Manson JE, Curb JD and Martin LW. Use of Hundreds of Electrocardiographic Biomarkers for Prediction of Mortality in Postmenopausal Women The Women's Health Initiative. *Circulation: Cardiovascular Quality and Outcomes*. 2011:CIRCOUTCOMES.110.959023.
36. Harrell FE, Califf RM, Pryor DB, Lee KL and Rosati RA. Evaluating the yield of medical tests. *Jama*. 1982;247:2543-2546.
37. Uno H, Cai T, Pencina MJ, D'Agostino RB and Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statistics in medicine*. 2011;30:1105-1117.
38. Gerds TA, Kattan MW, Schumacher M and Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Statistics in Medicine*. 2013;32:2173-2184.
39. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly weather review*. 1950;78:1-3.
40. Gerds TA and Schumacher M. Consistent Estimation of the Expected Brier Score in General Survival Models with Right-Censored Event Times. *Biometrical Journal*. 2006;48:1029-1040.
41. Lloyd-Jones DM, Wilson PW, Larson MG, Beiser A, Leip EP, D'Agostino RB and Levy D. Framingham risk score and prediction of lifetime risk for coronary heart disease. *The American journal of cardiology*. 2004;94:20-24.
42. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM and Kannel WB. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. *Circulation*. 2008;117:743-753.
43. Chahal H, Bluemke DA, Wu CO, McClelland R, Liu K, Shea SJ, Burke G, Balfour P, Herrington D and Shi P. Heart failure risk prediction in the Multi-Ethnic Study of Atherosclerosis. *Heart*. 2015;101:58-64.

Online Figures and Tables

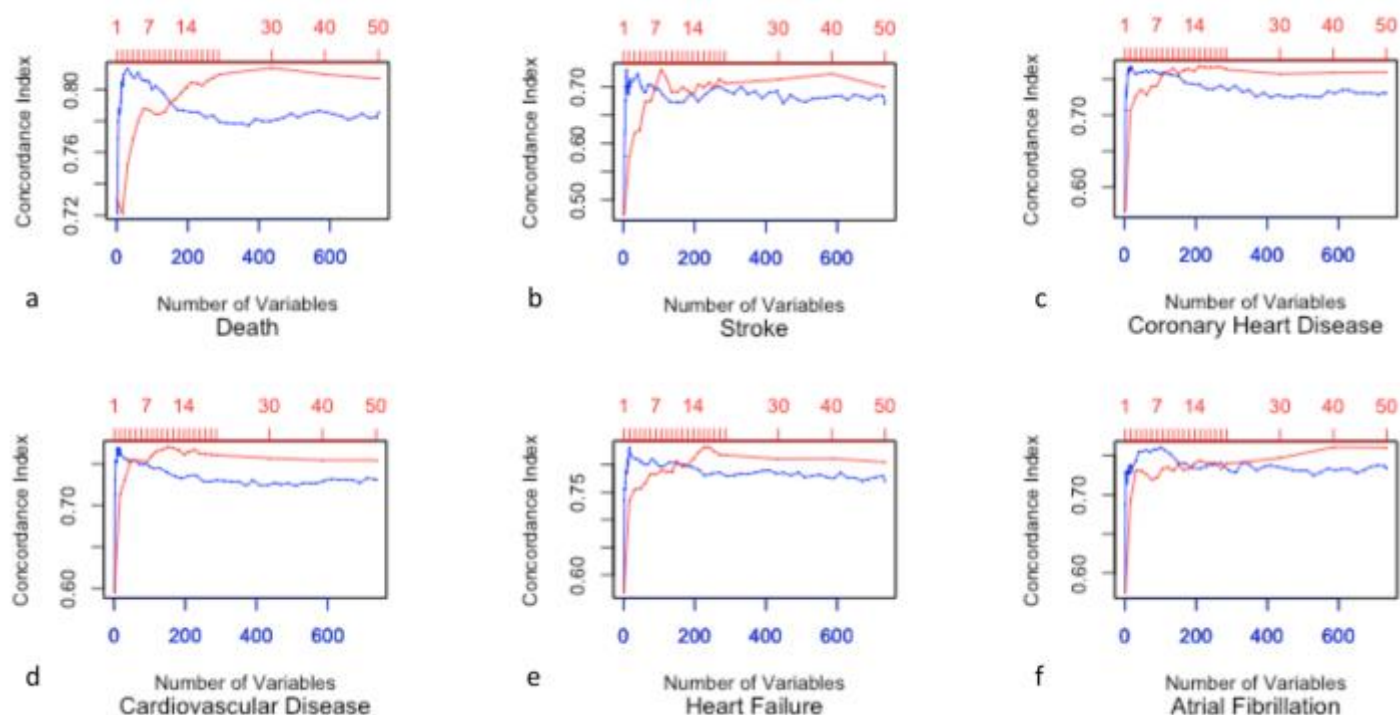
Online Figure I. An illustration of the random survival forest built from a collection of trees.



Online Figure II. The pie chart shows the distribution of available data. The legends indicate the number of participants the variables were available in. The pie chart shows that 93% of the variables were available in over 80% of the population, while only 1% of the variables were available in less than 20% of the population. In all, only 1% of the variables considered were excluded from the analysis.



Online Figure III. Nested RF models showing the change in concordance index with the addition of each of the 735 variables (in blue) by order of increasing rank (variable importance). The red curve indicates the same but only for the top-50 ranked variables on a magnified scale.



Online Table I: The top-10 ranked variables by the variable importance from the random survival forest method for each of the outcomes of interest. The variable importance of each variable was assessed here using permutation testing.

	CHF	CVDA	CHDA	AF	DTH	STRK
1	NT pro BNP	calcium score	calcium score	NT pro BNP	Age	Interleukin-2 soluble receptor
2	LV end-systolic volume	NT pro BNP	TNf a soluble receptor	Age	TNf a soluble receptor	calcium score
3	cardiac troponin T	Interleukin-2 soluble receptor	NT pro BNP	Interleukin-2 soluble receptor	Interleukin-2 soluble receptor	systolic blood pressure
4	TNf a soluble receptor	TNf a soluble receptor	Interleukin-2 soluble receptor	TNf a soluble receptor	NT pro BNP	LV anteroseptal basal end-diastolic wall thickness
5	calcium score	cardiac troponin T	cardiac troponin T	calcium score	calcium score	pulse pressure
6	Age	Age	ABI	creatinine	common carotid IMT	Fasting glucose
7	QT Index	internal carotid CMT	internal carotid CMT	STJ Amplitude in Lead V5 (uV)	ABI	Maximum carotid stenosis
8	LV end-diastolic volume	pulse pressure	maximum ascending aortic area	common carotid IMT	LV anteroseptal basal end-diastolic wall thickness	LV inferior basal end-diastolic wall thickness
9	QTC INTERVAL (msec)	maximum ascending aortic area	Age	internal carotid CMT	LV mass-volume ratio	LV mass-volume ratio
10	Interleukin-2 soluble receptor	ABI	maximum descending aortic area	homocysteine	internal carotid CMT	end-diastolic wall thickness (odbsc1)

Online Table II: The top-10 ranked variables by the variable importance from the random survival forest method for each of the outcomes of interest. The variable importance of each variable was assessed here using the change in Gini Index.

	CHF	CVDA	CHDA	AF	DTH	STRK
1	LV end-systolic volume	calcium score	calcium score	NT pro BNP	Age	Interleukin-2 soluble receptor
2	NT-pro BNP	NT pro BNP	TNf a soluble receptor	Age	TNf a soluble receptor	T Wave Area (T + T'), Lead AVF
3	Calcium score	TNf-a soluble receptor	ABI	calcium score	Interleukin-2 soluble receptor	Fasting glucose
4	TNf-a soluble receptor	Interleukin-2 soluble receptor	NT pro BNP	creatinine	NT pro BNP	Maximum carotid stenosis
5	LV end-diastolic volume	internal carotid intima media thickness	common carotid intima media thickness	Interleukin-2 soluble receptor	calcium score	internal carotid CMT
6	cardiac troponin-T	cardiac troponin T	cardiac troponin T	common carotid IMT	ABI	R Amplitude in Lead V2 (uV)
7	QTC INTERVAL (msec)	Age	internal carotid intima media thickness	TNf a soluble receptor	common carotid IMT	LV cardiac output
8	LV Ejection fraction	common carotid IMT	LV septal mid-ventricular end-systolic wall thickness	STJ Amplitude in Lead V5	homocysteine	P Wave Area, Lead V2
9	QT Index	maximum ascending aortic area	LV septal basal end-diastolic wall thickness (odbsc1)	ABI	internal carotid intima media thickness	Total QRS Area, Lead V1
10	Interleukin-2 soluble receptor	ABI	Interleukin-2 soluble receptor	R Amplitude in Lead V4 (uV)	maximum ascending aortic area	systolic blood pressure