



Published in final edited form as:

J Am Soc Echocardiogr. 2020 March ; 33(3): 322–331.e1. doi:10.1016/j.echo.2019.09.010.

Association of Machine Learning-Derived Pheno-groupings of Echocardiographic Variables with Heart Failure in Stable Coronary Artery Disease: The Heart and Soul Study.

Rakesh K. Mishra^{a,b}, Geoffrey Tison^a, Qizhi Fang^a, Rebecca Scherzer^b, Mary A. Whooley^b, Nelson B. Schiller^{a,b}

^aDivision of Cardiology, Department of Medicine, UCSF

^bDepartment of Medicine, San Francisco VAMC.

Abstract

BACKGROUND: Many individual echocardiographic variables have been associated with heart failure in patients with stable coronary artery disease (CAD), but their combined utility for prediction has not been well studied.

METHODS: Unsupervised model-based cluster analysis was performed blinded to the study outcome in 1000 patients with stable CAD on 15 transthoracic echocardiographic (TTE) variables. We evaluated associations of cluster membership with HF hospitalization using Cox proportional hazards regression analysis.

RESULTS: The echo-derived clusters partitioned subjects into 4 pheno-groupings: phenogroup 1 (n=85) had the highest levels, phenogroups 2 (n=314) and 3 (n=205) displayed intermediate levels and phenogroup 4 (n=396) had the lowest levels of cardiopulmonary structural and functional abnormalities. Over 7.1±3.2 years of follow-up, there were 198 HF hospitalizations. After multivariable adjustment for traditional cardiovascular risk factors, phenogroup 1 was associated with a nearly 5-fold increased risk (HR 4.8; 95%CI: 2.4–9.5), phenogroup 2 was associated with a nearly 3-fold increased risk (HR 2.7; 95%CI: 1.4–5.0), and phenogroup 3 was associated with a nearly 2-fold increased risk (HR 1.9; 95%CI: 1.0–3.8) of HF hospitalization, relative to phenogroup 4.

CONCLUSIONS: TTE variables can be used to classify stable CAD patients into separate pheno-groupings that differentiate cardiopulmonary structural and functional abnormalities, and can predict HF hospitalization, independent of traditional cardiovascular risk factors.

Corresponding Author: Rakesh Mishra, MD, San Francisco VA Medical Center, 4150 Clement St. (111C), San Francisco, CA 94122, rakesh.mishra@ucsf.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Transthoracic echocardiography (TTE) provides multiple measurements, such as left ventricular ejection fraction (LVEF), left ventricular mass index (LVMI), left ventricular outflow tract velocity-time integral (VTI_{LVOT}) and severity of mitral regurgitation (MR) and diastolic dysfunction (DD), that predict adverse outcomes in coronary artery disease (CAD) (1–5). We previously published a TTE heart failure index (HFI) based on 5 measurements (LVMI, left atrial volume index (LAVI), MR, VTI_{LVOT} and DD) that predicted heart failure (HF) hospitalizations in a large cohort of participants with stable, ambulatory CAD(6). These 5 measurements were selected from 15 echocardiographic variables and weighted according to their prognostic value using Cox proportional hazard models.

However, additive risk scores such as this do not capture the potentially complex, joint associations of variables that may interact and move together or separately. An alternative method to incorporate information from potentially discordant variables is unsupervised cluster analysis, an agnostic multivariable method that segregates similar cases without the constraint of an *a priori* diagnostic system (7). This type of analysis has been successfully applied to various areas in cardiology, including phenotypic clustering of DD parameters, pheno-grouping heart failure patients to identify responders to cardiac resynchronization therapy and identifying cardiopulmonary structural and functional phenotypes in patients with HIV (7–9). Therefore, we sought to (i) use an unsupervised, model-based clustering method to generate pheno-groupings that are prognostically meaningful, using all 15 TTE variables; (ii) compare the prognostic value of these pheno-groups against the HFI to predict HF hospitalizations.

Methods

The Heart and Soul Study is a prospective cohort study evaluating the impact of psychosocial factors on cardiovascular outcomes. The study methods have been previously described in detail(10). Participants were enrolled between 2000 and 2002 from 2 Veterans' Affairs hospitals, an academic medical center, and 9 public health clinics in the San Francisco area. All participants had CAD defined by either a history of myocardial infarction, angiographic evidence of $\geq 50\%$ stenosis in a coronary vessel, evidence of inducible ischemia by treadmill electrocardiography or nuclear perfusion stress imaging, or a history of coronary revascularization. Patients were excluded if they were unable to walk >1 block, had a history of acute coronary syndrome within the prior 6 months, or intended to move out of the local area within 3 years. A total of 1,024 study participants provided informed consent and completed baseline echocardiographic and laboratory testing, including 549 (54%) with a history of myocardial infarction, 237 (23%) with a history of revascularization but not myocardial infarction, and 238 (23%) with a diagnosis of coronary disease that was documented by their physician (on the basis of a positive angiogram or treadmill test in $>98\%$ of cases). Of the 1,024 participants, 1,000 participants with all the requisite echocardiographic measurements were included in this analysis. The institutional review boards at the University of California San Francisco, the San Francisco Veterans Affairs Medical Center, the Veterans Affairs Palo Alto Health Care System, and the Community Health Network of San Francisco approved this protocol. All participants

provided written informed consent. The investigation was performed in accordance with the Declaration of Helsinki.

Echocardiographic measurements

A complete resting 2-dimensional echocardiogram with an Acuson Sequoia ultrasound system (Siemens Medical Solutions, Mountain View, California) with a 3.5-MHz transducer and Doppler ultrasound examination was performed in all patients. Standard 2-dimensional parasternal short-axis and apical 2- and 4-chamber views during quiet respiration or held expiration were obtained. Two highly experienced sonographers made all sonographic measurements, and a single cardiologist reader (N.B.S.), who was blinded to clinical and laboratory information, evaluated, confirmed, and—when needed—corrected each measurement.

The fifteen candidate echocardiographic variables chosen *a priori* by the investigators were the same as the variables chosen to create the HFI(6):

1. left ventricular end-systolic volume index (LVESVI)
2. left ventricular end-diastolic volume index (LVEDVI)
3. left ventricular ejection fraction (LVEF)
4. left atrial volume index (LAVI)
5. right atrial volume index
6. left ventricular mass index (LVMI)
7. pulmonary artery peak systolic pressure (PASP)
8. right ventricular outflow tract velocity-time integral (VTI_{RVOT})
9. left ventricular outflow tract velocity-time integral (VTI_{LVOT})
10. aortic valve area
11. right atrial pressure
12. diastolic dysfunction (DD)
13. mitral regurgitation (MR) severity
14. tricuspid regurgitation severity
15. resting wall motion score index

Standard apical 2- and 4-chamber views were obtained. LV end-systolic and end-diastolic volumes were obtained by planimetry with the biplane method of discs as described (11). The LVEF was calculated as (end-diastolic volume - end-systolic volume)/end-diastolic volume.

Left and right atrial volumes were obtained at end-ventricular systole by manual planimetry with the biplane method of discs for the left atrium and single plane method of discs for the right atrium, as previously described and validated (11). All chamber volumes were subsequently indexed to body surface area. LV mass was calculated with a truncated

ellipsoid equation and indexed to body surface area as previously described and validated (11, 12). The tricuspid regurgitation jet was visualized with color flow mapping, and continuous wave Doppler was used to capture the flow signal from measurement of peak tricuspid regurgitant velocity. The peak tricuspid regurgitant velocity for the current study was the highest measurement obtainable by Doppler imaging among the parasternal, apical, and subcostal views. The right ventricular systolic pressure was estimated with the modified Bernoulli equation ($p = 4v^2$) and added to the estimated right atrial pressure to obtain the pulmonary artery systolic pressure (13).

The VTI_{RVOT} was obtained by placing a pulsed wave Doppler sample volume in the proximal right ventricular outflow tract at the level of the pulmonic valve, in the parasternal short-axis view and tracing the outer boundaries of the spectral Doppler signal to obtain the VTI. The sample volume was placed such that the opening valve Doppler signal was greater than or equal to the closing signal. The VTI_{LVOT} was obtained by placing a pulsed wave Doppler sample volume in the left ventricular outflow tract (LVOT) immediately proximal to the aortic valve in the anteriorly angulated apical 4-chamber view and tracing the outer boundaries of the peak spectral Doppler signal to obtain the VTI_{LVOT} . Proper location in the LVOT was confirmed by visualization of the aortic valve closure signal (14). The LVOT diameter was measured at the level of the aortic annulus from the parasternal long-axis view in mid-systole. The aortic valve area was derived from the velocity-time integrals of the aortic valve and the LVOT tract with the continuity equation (15).

Diastolic dysfunction was defined as the presence of at least 1 of the following: impaired relaxation defined as a ratio of peak mitral early diastolic to atrial contraction velocity (E/A) of < 0.75 with systolic dominant pulmonary vein flow; pseudonormal defined as $0.75 < E/A < 1.5$ with diastolic dominant pulmonary vein flow; restrictive filling defined as an $E/A \geq 1.5$ with diastolic dominant pulmonary vein flow (16). Diastolic dysfunction was only determined if both pulmonary vein flow and E/A were both recorded. Pulmonary vein flow was recorded in 1,011 patients, and E/A was recorded in 971 patients.

The severity of mitral and tricuspid regurgitation was determined semi-quantitatively according to American Society of Echocardiography guidelines (17). The right atrial pressure was obtained by inspection of the inferior vena cava during respiration as previously described (18).

Regional LV function was assessed with a standard 16-segment model (11). Segmental scores were assigned as follows: normal or hyperkinesis = 1; hypokinesis = 2; akinesis = 3; dyskinesis = 4; and aneurysmal = 5. The wall motion score index (WMSI) was derived as the sum of all scores divided by the number of segments visualized.

Construction of clusters

The goal of the clustering procedure was to simplify the data from 15 distinct echocardiographic measurements to partition subjects into a small number of pheno-groups based on the totality of echocardiographic information. These groupings were made based solely on the aggregate echocardiographic data and did not utilize clinical characteristics or

subsequent outcomes. We first examined unadjusted Spearman correlations between markers. We standardized each TTE measure (to a mean of zero and variance of one) so that variables with larger variances would not have more effect on the clusters. We used finite Gaussian mixture modeling with the EM algorithm to perform unsupervised clustering of TTE variables, using the Bayesian Information Criterion (BIC) to determine the number of clusters(19). Groups of participants having similar patterns of echocardiographic measurements can be identified as “pheno-groups”. We handled missing TTE variables using multiple imputation, generating 10 imputed datasets to ensure high relative efficiency. For each participant, cluster membership was defined as the consensus assignment across the 10 imputed datasets(20). Our clustering procedure identified 4 distinct subgroups of participants, which we call phenogroups, based on the solution yielding the optimal BIC.

Management of Missing Data

Rates of missing data were low (<3%) for all parameters with the exception of LV diastolic function (6.2% missing) and PASP (19% missing; supplemental table 1). We handled missing TTE variables using multiple imputation, generating 10 imputed datasets to ensure high relative efficiency. For each participant, cluster membership was defined as the consensus assignment across the 10 imputed datasets(20) as shown in supplemental table 2. Overall agreement across imputed datasets was moderate to good ($\kappa = 0.61$).

Transthoracic Echocardiographic Heart Failure Index

The derivation of the transthoracic echocardiographic heart failure index (HFI) has been previously described(6). Briefly, the association of the aforementioned 15 TTE measurements with subsequent HF hospitalization was evaluated using Cox proportional hazard models. The 5 TTE measurements that independently predicted HF - left ventricular mass index (LVMI), left atrial volume index (LAVI), mitral regurgitation (MR), left ventricular outflow tract velocity-time integral (VTI_{LVOT}), diastolic dysfunction (DD) - were combined into an index. Variables were defined as normal or abnormal on the basis of dichotomous cutoffs determined from the American Society of Echocardiography. Abnormal variables in each measurement were assigned points, ranging from 0 to 8, based on the strength of association with HF: 3 points for LVMI, 2 points for DD, 1 point for VTI_{LVOT} , MR, and LAVI. Furthermore, the HFI was categorized as follows: low <3 points, medium 3–4 points, high 5–6 points, very high 7–8 points.

Candidate Co-variates

Each participant completed a detailed questionnaire that included age, gender, race or ethnicity, medical history (including history of HF), level of physical activity and current smoking. Study personnel recorded all current medications and measured height and weight. Total, low-density and high-density lipoprotein cholesterol, creatinine, C-reactive protein (CRP), high-sensitivity troponin T (hs-TnT) and N-terminal pro-brain natriuretic peptide (NT-proBNP) were measured from fasting serum samples. Estimated glomerular filtration rate (eGFR) was calculated using Modification of Diet in Renal Disease equation (21). Covariates used in multivariable models were selected if there was a statistically significant difference among the clusters ($p < 0.05$).

Outcomes

To ascertain the primary outcome of HF hospitalization, we conducted annual telephone follow-up interviews and questioned participants or their proxies regarding recent emergency room visits and hospital stays. Medical records, death certificates, and coroner's reports were retrieved. Participants were censored at point of the first HF admission, when lost to follow-up, or upon death. Admission for HF was defined as a clinical syndrome with a minimum 1-night hospital stay and involving at least 2 of the following: paroxysmal nocturnal dyspnea, orthopnea, elevated jugular venous pressure, pulmonary rales, a third heart sound, cardiomegaly on chest radiography, or pulmonary edema on chest radiography. These clinical signs and symptoms must have represented a clear change from the normal clinical state of the patient and must have been accompanied by either failing cardiac output as determined by peripheral hypoperfusion (in the absence of other causes such as sepsis or dehydration) or peripheral or pulmonary edema treated with intravenous diuretics, inotropes, or vasodilators.

Two blinded physician adjudicators reviewed each event, and if there was agreement, the outcome classification was binding. If there was disagreement, a third blinded physician adjudicator reviewed the event and determined the outcome classification.

Statistical analyses

We compared the demographic, anthropometric, laboratory and echocardiographic variables and co-morbidities and medication use across phenogroups using chi-square and ANOVA tests for categorical and continuous variables, respectively. To study the association of the phenogroups with HF hospitalization, we generated Kaplan-Meier curves and used unadjusted and multivariable-adjusted Cox proportional hazards models. To compare the prognostic value of the addition of phenogroup membership vs. the HFI to a reference model, we estimated discrimination using Harrell's c statistic for concordance and, in addition, calculated the integrated discrimination index (IDI) (22). A 2-sided *P* value of <0.05 was considered statistically significant. All statistical methods were implemented in StataIC 13 (StataCorp LP, College Station, Texas), R package of mclust for cluster memberships, and survIDINRI for IDI.

Results

Construction of Echocardiographic Pheno-Groupings

We identified four distinct echocardiographic phenogroups using an unsupervised (*i.e.* agnostic of outcomes) automated clustering method. We started with generating Spearman correlations among the 15 echocardiographic variables. In unadjusted analysis, we found moderately strong correlations ($r \geq 0.5$) among LVEDVI, LVESVI, LVEF and WMSI and between VTI_{LVOT} and VTI_{RVOT} (Figure 1); all other correlations were moderate ($r < 0.5$) to weak ($r < 0.3$). The strongest correlation was between LVEDVI and LVESI ($r = 0.89$), as expected.

On average, persons in phenogroup 4 had smaller LVEDVI and LVESVI and normal LVMI and WMSI (Table 1). Phenogroups 1–3 had lower VTI_{LVOT} and VTI_{RVOT} and phenogroup 1

had borderline LVEF. Thirty seven percent and 50% of participants had greater than mild MR in phenogroups 1 and 2, respectively, and 34% and 60% of participants had greater than mild TR in phenogroups 1 and 2, respectively; no participant in phenogroup 3 or 4 had greater than mild MR or TR. Though every phenogroup had a mean calculated aortic valve area of $>2\text{cm}^2$, it increased stepwise from phenogroup 1 to 4. No participants in phenogroup 4 had pseudonormal or restrictive DD while 34% of participants in phenogroup 1 did. RA pressure and PASP were also elevated in phenogroup 1. LAVI and RAVI were increased in phenogroups 1 and 2. Consistent with the stratification of the phenogroups by LVMI, LAVI, MR, VTI_{LVOT} and DD, there was an incremental decrease in the HFI from phenogroup 1 (4.4 ± 2.1) to 4 (2.2 ± 1.7).

Cohort Characteristics

Overall, the cohort was of older age, predominantly male and mostly white (Table 1). There was a high prevalence of hypertension (71%). As is to be expected in a CAD cohort, a majority of participants had a history of myocardial infarction (54%) and coronary revascularization (59%). A sizeable proportion also had a history of CHF (18%).

Phenogroup 4 was younger on average ($p<0.0001$) and had the greatest proportion of women ($p=0.002$; Table 1). Phenogroups 1 and 3 had a greater proportion of diabetes ($p=0.014$) and history of CHF ($p<0.0001$). The percentage of participants with myocardial infarction ($p<0.0001$) and coronary revascularization ($p<0.0001$) rose incrementally from phenogroup 4 through 1. eGFR decreased while serum creatinine increased from phenogroup 4 through 1 ($p<0.0001$). Phenogroups 3 and 4 had higher total cholesterol ($p=0.017$) and LDL (0.01) than phenogroups 1 and 2. Serum NT-proBNP levels ($p=0.0001$) increased stepwise from phenogroup 4 through phenogroup 1 while phenogroup 1 had the highest hs-TnT levels ($p<0.0001$). There was no significant difference in β -blocker ($p=0.19$) or statin ($p=0.24$) use among the phenogroups, but aspirin ($p=0.026$) use was lowest and ACE-inhibitor use was highest in phenogroup 4 ($p<0.0001$).

Association of Echocardiographic Pheno-Groupings with Heart Failure

During a mean follow-up of 7.1 ± 3.2 years, the rate of HF hospitalizations was highest in phenogroup 1 (86 per 1000 person-years, 95%CI 78–94) and lowest in phenogroup 4 (10 per 1000 person-years, 95%CI 9–11), with phenogroups 2 and 3 showing intermediate event rates (Figure 2a). Similarly, in unadjusted analyses, phenogroup 1 had a greater than 8-fold risk (HR 8.4; 95%CI 4.1 – 17.1), phenogroup 2 had a 4-fold risk (HR 4.3; 95% CI 1.9 – 9.6), and phenogroup 3 had a 3-fold risk of HF hospitalization (HR 3.0; 95% CI 1.6 – 5.3), compared to phenogroup 4 (Table 2). After multivariable adjustment for traditional CVD risk factors, including demographics, comorbidities, laboratory values and medications, the elevated risk of HF hospitalization among those in phenogroups 1 and 2 persisted (Table 2). In the Kaplan-Meier analysis, there was a stepwise decrease in survival from phenogroup 4 to phenogroup 1 (Figure 2b).

Performance of Pheno-Groupings versus the Heart Failure Index in Predicting Heart Failure

The integrated discrimination index (IDI) for HF hospitalization was significant when adding either the heart failure index (HFI; 6.4%, 95% CI 1.5 – 11.9%; $p=0.03$) or the phenogroups (5.1%, 95% CI 2.0 – 8.9%; $p=0.01$; Table 3) to a reference model that included demographic variables, comorbidities including a history of MI, HF, angina and revascularization, laboratory values including NT-proBNP and hs-TnT and medications. The C-statistic was the same for HFI and phenogroup membership added to a reference model (0.84; 95% CI 0.81 – 0.87).

For each category of the HFI (low, medium, high and very high), phenogroup 1 had highest event rates for HF hospitalization (Figure 3). For the high and low HFI, the event rates decreased stepwise from phenogroup 3 to 1. The medium HFI did not exhibit such a trend.

Discussion

In this cohort of 1,000 participants with stable CAD, we found that automated model-based unsupervised clustering based on 15 echocardiographic measures identifies four distinct pheno-groupings, which successfully partition subjects into categories of risk of HF hospitalization. Our group had previously used multivariable Cox proportional hazard models to select 5 of these 15 echocardiographic measures to create a heart failure index (HFI) that independently predicted HF hospitalizations in the same cohort(6). We found that the machine learning-derived phenogroups performed as well as the HFI in predicting HF hospitalization.

Machine learning algorithms can discover hidden patterns in complex and heterogeneous data. In addition, unsupervised learning, such as cluster analysis, analyzes the intrinsic structure within data and can evaluate complex nonlinear interaction among variables, without the *a priori* consideration of outcomes(7). Unsupervised machine learning algorithms have been used to analyze left ventricular function to characterize heart failure with preserved ejection fraction(23), discover phenotypic clustering of left ventricular diastolic function (7), classify prognostic categories using exercise echocardiography in heart failure with preserved ejection fraction(24) and pheno-group patients with heart failure to identify response to cardiac resynchronization therapy(8). Ours is the first study to apply an unsupervised machine learning algorithm to TTE measurements among patients with stable CAD, a population known to be at elevated risk of HF, to successfully derive clusters of risk for HF hospitalization.

The HFI was created initially using a group of 15 historically and intuitively promising echocardiographic measurements. From these, after eliminating the measurements that were statistically redundant in linear models, 5 measurements were selected to create the HFI: LVMI, LAVI, MR, VTI_{LVOT} , and DD. Some candidate variables that have a rich history as predictors of poor cardiovascular outcomes, mortality and HF such as LVEDVI, LVESVI and PASP were not selected to form the HFI either because they were either strongly inter-correlated or were not independent predictors (25–27). In contrast to traditional methods of risk stratification, based on selection of predictors based on their association with outcomes,

in the current study, participants were separated into four distinct pheno-groupings based upon an unsupervised machine-learning algorithm, which considered all 15 TTE variables.

These pheno-groupings exhibited distinct baseline demographic characteristics, comorbidities and laboratory features, demonstrating that they represent distinct phenotypes of participants with stable CAD. Furthermore, the pheno-groupings demonstrated varying risk of HF hospitalization from low risk (phenogroup 4) to high (phenogroup 1) risk groups. Of relevance to subsequent HF hospitalization, phenogroup 1 had the highest rates of history of HF, MI and revascularization, lowest eGFR and the highest levels of CRP, hs-cTnT and NT-proBNP (28). These pheno-groupings also partitioned the TTE measurements into meaningful groups; the high risk (phenogroup 1) group had the highest LVEDVI, LVESVI, LVMI and WMSI and the lowest LVEF. Moreover, the HFI increased from the lowest risk (phenogroup 4) to the highest risk (phenogroup 1) group. In summary, the four pheno-groupings generated by a machine learning algorithm have clinical, laboratory and echocardiographic values (*i.e.* history of HF, higher values of hs-TnT and NT-proBNP, lower eGFR, higher LVEDVI, LVESVI, LVMI and WMSI and lower LVEF), which are consistent with informal risk assessment for HF performed by physicians treating patients with stable CAD in the clinic.

Pheno-groupings were prognostically meaningful in predicting HF hospitalization. As expected, the high risk (phenogroup 1), with all the high-risk demographic, comorbid, laboratory and echocardiographic features, had the highest risk of HF hospitalization. Many of the features of the high risk pheno-grouping such as elevated LAESVI, PASP, LVESVI and LVMI, LV diastolic dysfunction and reduced LVOT and RVOT VTI have been previously shown to be associated with adverse outcomes in this and other cohorts(1–3, 16, 26, 27). While individual TTE measures showed statistically significant associations with adverse outcomes, the full panel was not simultaneously predictive, making it difficult to reconcile the findings of multiple measures. The unsupervised clustering method used in the current study were likely able to identify the complex associations between the TTE variables and, thereby, successfully distinguish pheno-groupings with different HF risk profiles.

As evaluated by the IDI analysis and Harrell's c-statistic, the pheno-groupings of TTE variables performed equally as well as the HFI for the prediction of HF hospitalizations. This indicates that an unsupervised machine learning algorithm that is agnostic of outcomes can partition the cohort into prognostically meaningful categories of risk as well as standard regression methods that rely on the inclusion of outcomes in creating their models of risk. This is the first study to directly compare unsupervised machine learning clustering of TTE variables with a score derived from linear methods, dependent on outcomes, applied to the same variables to evaluate risk of HF hospitalization among participants with stable CAD.

Study Limitations

The primary outcome is HF hospital stay, which might be prone to error despite rigorous adjudication, depending on clinical assessment, coding, and accuracy of chart review. In particular, the clinical assessment of physical findings like the determination of jugular venous pressure are fraught with significant variability in its evaluation and, therefore, are

prone to error(29). In addition, TTE measurements were limited to those available at the time of the initial echocardiographic examination. Therefore, we were not able to incorporate newer techniques that are widely available now such as tissue Doppler. In addition, we no longer have access to the original echocardiograms to perform speckle tracking strain analysis.

The objective of our study was to provide a proof of principle of the prognostic value of using unsupervised clustering, agnostic of outcomes, to identify clinically relevant phenogroups. Therefore, the specific phenogroups we derive are not intended to generalize to other cohorts, but instead we demonstrate that a similar approach can be taken to derive phenogroups in other cohorts. An additional limitation is that missing TTE data were handled using multiple imputation, resulting in some uncertainty in the phenogroup assignments. However, we found moderate to good agreement in cluster assignment across the imputed datasets, despite this missing data.

The distribution of participants across the phenogroups is unequal and this may result in under-or overestimation of the point estimates. Most of the participants enrolled in the study had preserved LVEF (> 50%) and, therefore, our conclusions may not apply to participants with heart failure with reduced LVEF. Furthermore, we cannot comment on how this clustering method would apply to individual patients without stable CAD.

Conclusions

Automated unsupervised clustering of echocardiographic measures successfully partitioned a population with stable CAD into pheno-groupings of risk for HF hospitalization. This machine learning algorithm performed as well as the HFI, derived previously in the same cohort using traditional scalar methods. This suggests the possibility that automated machine learning algorithms could be applied to echocardiographic measurements to derive pheno-groupings by being incorporated into echocardiographic platforms or the electronic medical record.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Ristow B, Ali S, Whooley MA, Schiller NB. Usefulness of left atrial volume index to predict heart failure hospitalization and mortality in ambulatory patients with coronary heart disease and comparison to left ventricular ejection fraction (from the Heart and Soul Study). *Am J Cardiol.* 2008;102(1):70–6. [PubMed: 18572038]
2. Turakhia MP, Schiller NB, Whooley MA. Prognostic significance of increased left ventricular mass index to mortality and sudden death in patients with stable coronary heart disease (from the Heart and Soul Study). *Am J Cardiol.* 2008;102(9):1131–5. [PubMed: 18940278]
3. Ristow B, Na B, Ali S, Whooley MA, Schiller NB. Left ventricular outflow tract and pulmonary artery stroke distances independently predict heart failure hospitalization and mortality: the Heart and Soul Study. *J Am Soc Echocardiogr.* 2011;24(5):565–72. [PubMed: 21324645]
4. Ren X, Na B, Ristow B, Whooley MA, Schiller NB. Usefulness of diastolic dominant pulmonary vein flow to predict hospitalization for heart failure and mortality in ambulatory patients with

- coronary heart disease (from the Heart and Soul Study). *Am J Cardiol.* 2009;103(4):482–5. [PubMed: 19195506]
5. Grigioni F, Detaint D, Avierinos JF, Scott C, Tajik J, Enriquez-Sarano M. Contribution of ischemic mitral regurgitation to congestive heart failure after myocardial infarction. *J Am Coll Cardiol.* 2005;45(2):260–7. [PubMed: 15653025]
 6. Stevens SM, Farzaneh-Far R, Na B, Whooley MA, Schiller NB. Development of an echocardiographic risk-stratification index to predict heart failure in patients with stable coronary artery disease: the Heart and Soul study. *JACC Cardiovasc Imaging.* 2009;2(1):11–20. [PubMed: 19356527]
 7. Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic Clustering of Left Ventricular Diastolic Function Parameters: Patterns and Prognostic Relevance. *JACC Cardiovasc Imaging.* 2018.
 8. Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, et al. Machine learning-based pheno-grouping in heart failure to identify responders to cardiac resynchronization therapy. *Eur J Heart Fail.* 2019;21(1):74–85. [PubMed: 30328654]
 9. Scherzer R, Shah SJ, Secemsky E, Butler J, Grunfeld C, Shlipak MG, et al. Association of Biomarker Clusters With Cardiac Phenotypes and Mortality in Patients With HIV Infection. *Circ Heart Fail.* 2018;11(4):e004312. [PubMed: 29615435]
 10. Ruo B, Rumsfeld JS, Hlatky MA, Liu H, Browner WS, Whooley MA. Depressive symptoms and health-related quality of life: the Heart and Soul Study. *JAMA.* 2003;290(2):215–21. [PubMed: 12851276]
 11. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr.* 2015;28(1):1–39 e14. [PubMed: 25559473]
 12. Schiller NB, Skioldebrand CG, Schiller EJ, Mavroudis CC, Silverman NH, Rahimtoola SH, et al. Canine left ventricular mass estimation by two-dimensional echocardiography. *Circulation.* 1983;68(1):210–6. [PubMed: 6851047]
 13. Yock PG, Popp RL. Noninvasive estimation of right ventricular systolic pressure by Doppler ultrasound in patients with tricuspid regurgitation. *Circulation.* 1984;70(4):657–62. [PubMed: 6478568]
 14. Goldman JH, Schiller NB, Lim DC, Redberg RF, Foster E. Usefulness of stroke distance by echocardiography as a surrogate marker of cardiac output that is independent of gender and size in a normal population. *Am J Cardiol.* 2001;87(4):499–502, A8. [PubMed: 11179548]
 15. Otto CM, Pearlman AS, Comess KA, Reamer RP, Janko CL, Huntsman LL. Determination of the stenotic aortic valve area in adults using Doppler echocardiography. *J Am Coll Cardiol.* 1986;7(3):509–17. [PubMed: 3950230]
 16. Ren X, Ristow B, Na B, Ali S, Schiller NB, Whooley MA. Prevalence and prognosis of asymptomatic left ventricular diastolic dysfunction in ambulatory patients with coronary heart disease. *Am J Cardiol.* 2007;99(12):1643–7. [PubMed: 17560867]
 17. Zoghbi WA, Adams D, Bonow RO, Enriquez-Sarano M, Foster E, Grayburn PA, et al. Recommendations for Noninvasive Evaluation of Native Valvular Regurgitation: A Report from the American Society of Echocardiography Developed in Collaboration with the Society for Cardiovascular Magnetic Resonance. *J Am Soc Echocardiogr.* 2017;30(4):303–71. [PubMed: 28314623]
 18. Kircher BJ, Himelman RB, Schiller NB. Noninvasive estimation of right atrial pressure from the inspiratory collapse of the inferior vena cava. *Am J Cardiol.* 1990;66(4):493–6. [PubMed: 2386120]
 19. Yeung KY, Fraley C, Murua A, Raftery AE, Ruzzo WL. Model-based clustering and data transformations for gene expression data. *Bioinformatics.* 2001;17(10):977–87. [PubMed: 11673243]
 20. Basagana X, Barrera-Gomez J, Benet M, Anto JM, Garcia-Aymerich J. A framework for multiple imputation in cluster analysis. *Am J Epidemiol.* 2013;177(7):718–25. [PubMed: 23445902]

21. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med.* 2006;145(4):247–54. [PubMed: 16908915]
22. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008;27(2):157–72; discussion 207–12. [PubMed: 17569110]
23. Sanchez-Martinez S, Duchateau N, Erdei T, Kunszt G, Aakhus S, Degiovanni A, et al. Machine Learning Analysis of Left Ventricular Function to Characterize Heart Failure With Preserved Ejection Fraction. *Circ Cardiovasc Imaging.* 2018;11(4):e007138. [PubMed: 29661795]
24. Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala W. Contribution of Cardiovascular Reserve to Prognostic Categories of Heart Failure With Preserved Ejection Fraction: A Classification Based on Machine Learning. *J Am Soc Echocardiogr.* 2019;32(5):604–15. [PubMed: 30718020]
25. Vasan RS, Larson MG, Benjamin EJ, Evans JC, Levy D. Left ventricular dilatation and the risk of congestive heart failure in people without myocardial infarction. *N Engl J Med.* 1997;336(19):1350–5. [PubMed: 9134875]
26. McManus DD, Shah SJ, Fabi MR, Rosen A, Whooley MA, Schiller NB. Prognostic value of left ventricular end-systolic volume index as a predictor of heart failure hospitalization in stable coronary artery disease: data from the Heart and Soul Study. *J Am Soc Echocardiogr.* 2009;22(2):190–7. [PubMed: 19084372]
27. Ristow B, Ali S, Ren X, Whooley MA, Schiller NB. Elevated pulmonary artery pressure by Doppler echocardiography predicts hospitalization for heart failure and mortality in ambulatory stable coronary artery disease: the Heart and Soul Study. *J Am Coll Cardiol.* 2007;49(1):43–9. [PubMed: 17207721]
28. Bansal N, Hyre Anderson A, Yang W, Christenson RH, deFilippi CR, Deo R, et al. High-sensitivity troponin T and N-terminal pro-B-type natriuretic peptide (NT-proBNP) and risk of incident heart failure in patients with CKD: the Chronic Renal Insufficiency Cohort (CRIC) Study. *J Am Soc Nephrol.* 2015;26(4):946–56. [PubMed: 25278510]
29. Bredthardt T, Moreno-Weidmann Z, Uthoff H, Sabti Z, Aeppli S, Puelacher C, et al. How accurate is clinical assessment of neck veins in the estimation of central venous pressure in acute heart failure? Insights from a prospective study. *Eur J Heart Fail.* 2018;20(7):1160–2. [PubMed: 29314487]

	AVA	LVEF	VTI LVOT	VTI RVOT	LVMl	WMSI	LVESVI	LVEDVI	DD	RAP	TRS	PASP	MRS	RAVI	LAVI
LAVI	-0.11	-0.22	0.22	0.12	0.36	0.17	0.40	0.43	0.07	0.11	0.15	0.20	0.22	0.36	1.00
RAVI	-0.11	-0.13	0.09	0.07	0.12	0.06	0.19	0.18	0.06	0.17	0.17	0.16	0.14	1.00	0.36
MRS	-0.10	-0.22	-0.07	-0.06	0.18	0.29	0.28	0.28	0.11	0.03	0.27	0.16	1.00	0.14	0.22
PASP	-0.13	0.07	0.18	0.08	0.12	0.02	0.00	0.06	0.12	0.22	0.33	1.00	0.16	0.16	0.20
TRS	-0.04	-0.05	-0.01	-0.08	0.02	0.02	0.10	0.13	0.15	0.09	1.00	0.33	0.27	0.17	0.15
RAP	-0.08	-0.08	0.01	0.05	0.03	0.05	0.06	0.05	0.16	1.00	0.09	0.22	0.03	0.17	0.11
DD	-0.09	-0.18	-0.15	-0.11	0.15	0.17	0.13	0.09	1.00	0.16	0.15	0.12	0.11	0.06	0.07
LVEDVI	-0.06	-0.45	-0.19	-0.03	0.47	0.51	0.89	1.00	0.09	0.05	0.13	0.06	0.28	0.18	0.43
LVESVI	-0.07	-0.75	-0.30	-0.10	0.48	0.57	1.00	0.89	0.13	0.06	0.10	0.00	0.28	0.19	0.40
WMSI	-0.15	-0.53	-0.32	-0.09	0.38	1.00	0.57	0.51	0.17	0.05	0.02	0.02	0.29	0.06	0.17
LVMl	-0.11	-0.36	-0.03	0.01	1.00	0.38	0.48	0.47	0.15	0.03	0.02	0.12	0.18	0.12	0.36
VTI RVOT	-0.17	0.17	0.50	1.00	0.01	-0.09	-0.10	-0.03	-0.11	0.05	-0.08	0.08	-0.06	0.07	0.12
VTI LVOT	-0.02	0.36	1.00	0.50	-0.03	-0.32	-0.30	-0.19	-0.15	0.01	-0.01	0.18	-0.07	0.09	0.22
LVEF	0.09	1.00	0.36	0.17	-0.36	-0.53	-0.75	-0.45	-0.18	-0.08	-0.05	0.07	-0.22	-0.13	-0.22
AVA	1.00	0.09	-0.02	-0.17	-0.11	-0.15	-0.07	-0.06	-0.09	-0.08	-0.04	-0.13	-0.10	-0.11	-0.11

Figure 1.

Heatmap showing Spearman correlation coefficients between TTE variables.

Abbreviations: TTE = transthoracic echocardiography; LAVI = left atrial end-systolic volume index; RAVI = right atrial end-systolic volume index; MRS = severity of mitral regurgitation; PASP= Pulmonary artery systolic pressure; TRS = severity of tricuspid regurgitation; RAP = right atrial pressure; DD = diastolic dysfunction; LVEDVI = left ventricular end-diastolic volume index; LVESVI = left ventricular end-systolic volume index; WMSI = wall motion score index; LVMl = left ventricular mass index; VTI RVOT= right ventricular outflow tract velocity-time integral; VTI LVOT = left ventricular outflow tract velocity-time integral; LVEF = left ventricular ejection fraction; AVA = aortic valve area.

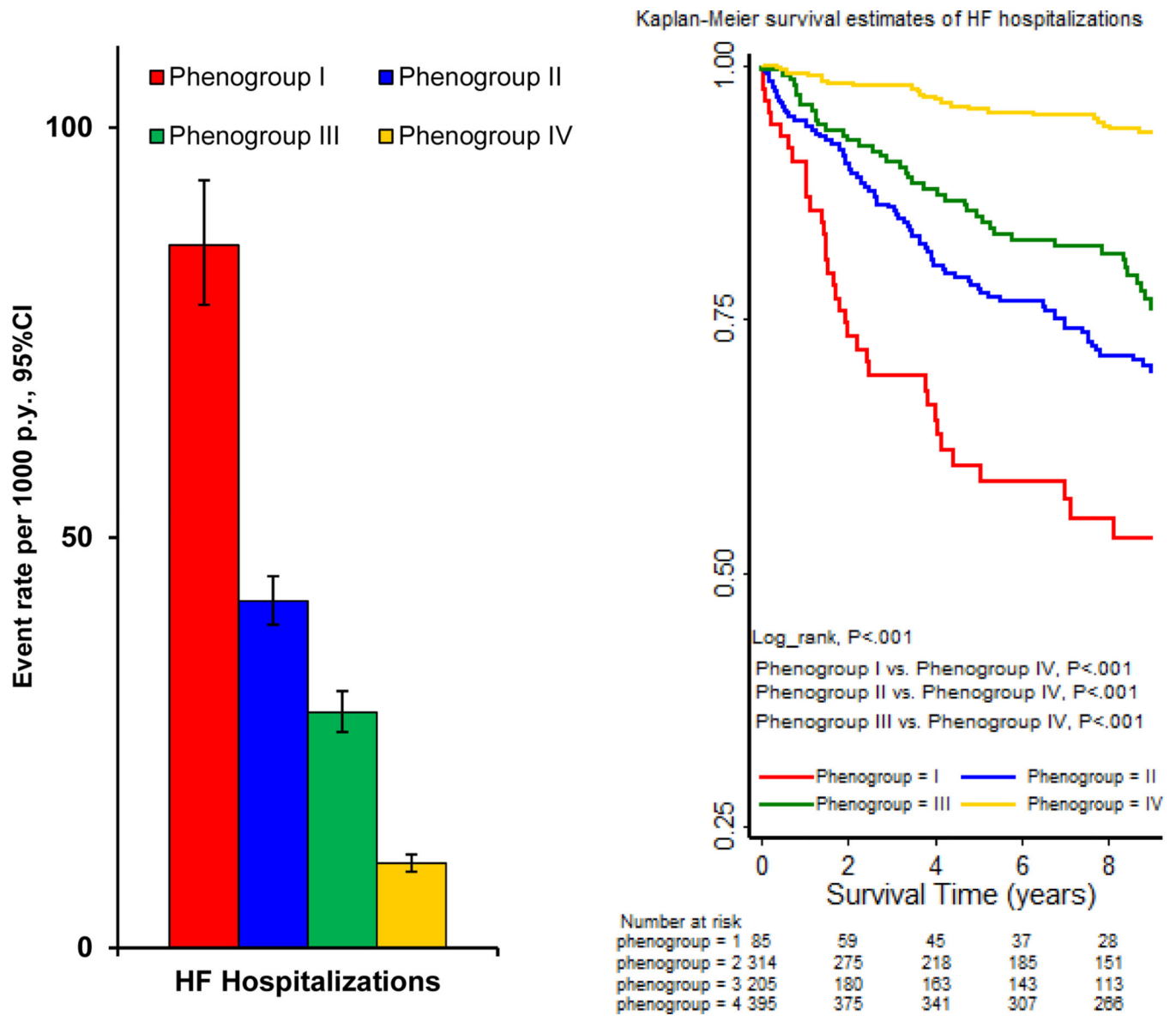


Figure 2. (A) Clinical outcome event rates, stratified by TTE-derived phenotype. py, Person-years. (B) Kaplan-Meier curves of HF hospitalizations by phenogroupings. Abbreviation. TTE = transthoracic echocardiography; HF = heart failure; py = person-years.

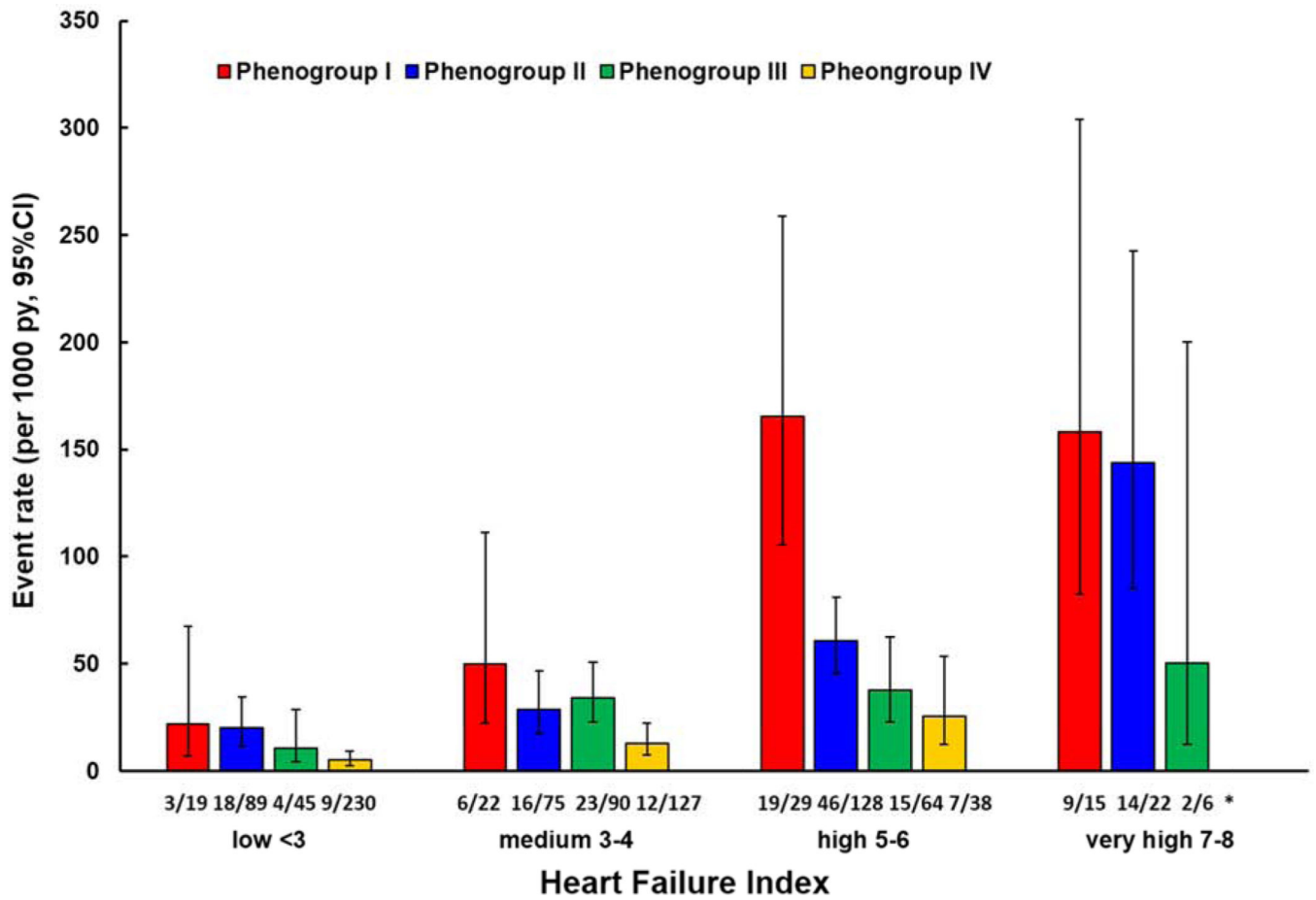


Figure 3.
 HF hospitalizations Stratified by Pheno-Grouping in Each Heart Failure Index Category.
 *number of patients with heart failure hospitalizations in each heart failure index category, stratified by phenogroup.
 Abbreviations: py = person-years.

Table 1.

Summary of baseline characteristics, stratified by TTE-derived pheno-groupings.

	Overall N=1000	Phenogroup I 85 (9%)	Phenogroup II 314 (31%)	Phenogroup III 205 (21%)	Phenogroup IV 396 (40%)	p value
Age (y)	67±11	67±11	71±11	68±10	63±10	<0.0001
Male	816 (82%)	67 (79%)	265 (84%)	181 (88%)	303 (77%)	0.002
Race (non white)	401 (40%)	30(35%)	116(37%)	80 (39%)	175 (44%)	0.17
BMI (kg/m²)	28±5	29±7	27±5	29±5	29±5	<0.0001
Medical history and Behavior risk factors						
Hypertension	704 (71%)	61 (72%)	214 (68%)	149 (73%)	280 (71%)	0.70
Diabetes	260 (26%)	26 (31%)	63 (20%)	65 (32%)	106 (27%)	0.014
CHF	178 (18%)	22 (26%)	65 (21%)	46 (23%)	45 (11%)	<0.0001
Stroke	143 (14%)	11 (13%)	47 (15%)	33 (16%)	52 (13%)	0.72
Myocardial infarction	537 (54%)	53 (63%)	187 (60%)	124 (61%)	173 (44%)	<0.0001
Revascularization	587 (59%)	56 (67%)	209 (67%)	130 (64%)	192 (49%)	<0.0001
Physically inactive	363 (36%)	38 (45%)	116 (37%)	67 (33%)	142 (36%)	0.31
Smoking	196 (20%)	23 (27%)	42 (13%)	42 (21%)	89 (23%)	0.005
Laboratory values						
eGFR (ml/min/1.73m²)	71±22	62±24	66±21	69±22	77±22	<0.0001
LDL (mg/dL)	104±34	99±35	100±33	108±33	107±34	<0.01
HDL (mg/dL)	46±14	46±15	46±14	46±13	45±14	0.72
Total Cholesterol (mg/dL)	178±42	173±47	172±42	181±39	181±42	0.017
NT-proBNP (pg/mL)	175 (73 453)	478 (184 1103)	296 (136 751)	204 (99 465)	94 (41 200)	<0.0001
hs-cTroponin T (pg/mL)	5.7 (0 13.4)	9.9 (4.1 20.3)	6.7 (0.2 16.2)	8.3 (0.2 15.3)	8.3 (0.2 15.3)	<0.0001
CRP (ug/mL)	3.4 (1.6 8.9)	3.7 (1.8 12.5)	3.5 (1.6 9.0)	3.6 (1.8 9.5)	3.2 (1.5 8.5)	0.28
Medications						
Medications Beta blocker	578 (59%)	48 (58%)	198 (63%)	113 (55%)	219 (56%)	0.19
ACE-inhibitor	512 (52%)	59 (72%)	163 (52%)	111 (54%)	179 (46%)	<0.0001
Statin	640 (65%)	52 (63%)	215 (69%)	134 (65%)	239 (62%)	0.24
Aspirin	723 (73%)	49 (59%)	233 (75%)	153 (75%)	288 (74%)	0.026
TTE measurements						
LAVI	33±12	39±13	38±15	31±11	29±8	<0.0001
RAVI	23±10	30±13	26±12	21±9	21±8	<0.0001
MR						
None to trace	812 (81%)	54 (64%)	157 (50%)	205 (100%)	396 (100%)	
Mild	179 (18%)	27 (32%)	152 (48%)	0	0	<0.0001
Moderate or severe	9 (1%)	4 (5%)	5 (2%)	0	0	
PASP	32±9	41±15	34±10	31±7	29±6	<0.0001
TR						
None to trace	784 (78%)	56 (66%)	127 (40%)	205 (100%)	396 (100%)	
Mild	206 (21%)	24 (28%)	182 (58%)	0	0	<0.0001

	Overall	Phenogroup I	Phenogroup II	Phenogroup III	Phenogroup IV	p value
	N=1000	85 (9%)	314 (31%)	205 (21%)	396 (40%)	
Moderate or severe	10 (1%)	5 (6%)	5 (2%)	0	0	
Right atrial pressure	5.3±1.4	8.6±3.7	5±0	5±0	5±0	<0.0001
Diastolic Function						
Normal	604 (64%)	36 (51%)	164 (58%)	54 (27%)	350 (90%)	
Impaired relaxation	224 (24%)	11 (15%)	69 (24%)	107 (54%)	37 (10%)	<0.0001
Pseudonormal or restrictive	110 (12%)	24 (34%)	50 (18%)	36 (18%)	0	
LVEDVI	52±18	65±32	57±17	54±19	44±10	<0.0001
LVESVI	22±16	40±39	24±12	22±13	15±4	<0.0001
LVMI	100±34	119±78	102±27	109±32	89±17	<0.0001
Wall motion score Index	1.2±0.4	1.4±0.6	1.3±0.5	1.1±0.2	1.0±0.1	<0.0001
VTI_{RVOT}	0.19±0.04	0.20±0.05	0.19±0.04	0.19±0.04	0.21±0.04	<0.0001
VTI_{LVOT}	0.22±0.05	0.22±0.09	0.21±0.05	0.21±0.04	0.23±0.04	<0.0001
LVEF	62±10	54±16	59±10	61±10	66±5	<0.0001
Aortic valve area	2.5±0.6	2.3±0.7	2.4±0.7	2.5±0.7	2.6±0.5	0.004
Heart Failure Index	3.2±2.0	4.4±2.1	3.9±2.0	3.7±1.6	2.2±1.7	<0.0001

Data are summarized as N (%) for categorical variables. Continuous variables are summarized as mean ± SD or median (IQR).

Abbreviations: TTE = transthoracic echocardiography; BMI = body mass index; eGFR = estimated glomerular filtration rate; LDL = low density lipoprotein; HDL = high density lipoprotein; NT-proBNP = N-terminal pro-brain natriuretic peptide; hs-cTroponinT = high sensitivity cardiac troponin T; CRP = C-reactive protein; ACE = angiotensin converting enzyme.

Data are summarized as N (%) for categorical variables; Continuous variables are summarized as mean ± SD or median (IQR).

Abbreviations: LAVI = left atrial end-systolic volume index; RAVI = right atrial end-systolic volume index; MR = mitral regurgitation; PASP= Pulmonary artery systolic pressure; TR = tricuspid regurgitation; LVEDVI = left ventricular end-diastolic volume index; LVESVI = left ventricular end-systolic volume index; LVMI = left ventricular mass index; VTI_{LVOT} = left ventricular outflow tract velocity–time integral; VTI_{RVOT}= right ventricular outflow tract velocity– time integral; LVEF = left ventricular ejection fraction.

Table 2.

Associations of TTE-derived pheno-groupings with HF hospitalizations

Outcome	Phenogroup I HR (95%CI)	Phenogroup II HR (95%CI)	Phenogroup III HR (95%CI)	Phenogroup IV (reference)
HF hospitalizations				
Event rate (per 1000, 95%CI)	85.7 (78.4 93.6)	42.3 (39.4 45.3)	28.7 (26.3 31.3)	10.3 (9.3 11.4)
Unadjusted HR (95% CI)	8.4 (4.1, 17.1), p<0.001	4.3 (1.9, 9.6), p=0.001	3.0 (1.6, 5.3), p<0.001	Reference
Model 1: demographics	7.7 (4.0, 14.6), p<0.001	3.6 (1.6, 7.9), p=0.003	2.6 (1.3, 5.4), p=0.011	Reference
Model 2: + comorbidities	6.5 (3.2, 13.2), p<0.001	3.5 (1.7, 6.9), p=0.001	2.2 (1.2, 4.1), p=0.012	Reference
Model 3: + laboratory values	5.0 (2.6, 9.7), p<0.001	2.9 (1.5, 5.3), p=0.001	2.0 (1.0, 3.9), p=0.045	Reference
Model 4: + medications	4.8 (2.4, 9.5), p<0.001	2.7 (1.4, 5.0), p=0.002	1.9 (1.0, 3.8), p=0.056	Reference

Hazard ratios with 95% confidence intervals (CI) from Cox proportional hazards models. Multivariable adjusted models control sequentially for demographics (age, sex, race, body mass index), comorbidities (current smoking, diabetes, history of heart failure, history of myocardial infarction, history of angina or revascularization), laboratory values (total cholesterol, LDL, eGFR, NT-proBNP, hs-cTroponin T), and medication use (antihypertensive therapy and aspirin).

Abbreviations. CI = confidence interval; TTE = transthoracic echocardiography; HF = heart failure; HR = hazard ratio.

Table 3.

Comparison of model fit for TTE-derived Pheno-groupings versus Heart Failure Index score for HF Hospitalization

Outcome	Harrell's C-statistic (95% CI)	Integrated Discrimination Index (95%CI)
HF hospitalization		
0. Reference model*	0.81 (0.78–0.84)	reference
1. Reference + HFI	0.82 (0.79–0.85), p=0.053 [†]	+6.4% (1.5%, 12.5%), p=0.033 [†]
2. Reference + Phenogroup	0.83 (0.80–0.85), p=0.025 [‡]	+5.1% (2.0%, 9.3%), p=0.007 [‡]
1 vs. 2. Ref. + HFI vs. Ref. + Phenogroup	p=0.69	-1.2% (-6.3% 4.0%), p=0.69

* Reference model includes demographics (age, sex, race, body mass index), comorbidities (current smoking, diabetes, history of heart failure, history of myocardial infarction, history of angina or revascularization), laboratory values (total cholesterol, LDL, eGFR, NT-proBNP, hscTroponin T), and medication use (antihypertensive therapy and aspirin)

[†] p-value for model 1 vs. 0;

[‡] p-value for model 2 vs. 0

Abbreviations. TTE = transthoracic echocardiography; CI = confidence interval.