JACC: CARDIOVASCULAR IMAGING
VOL. ■, NO. ■, 2017

ISSN 1936-878X/$36.00
PUBLISHED BY ELSEVIER
http://dx.doi.org/10.1016/j.jcmg.2016.10.012

Precision Phenotyping in Heart Failure and Pattern Clustering of Ultrasound Data for the Assessment of Diastolic Dysfunction

Alaa Mabrouk Salem Omar, MD, PhD,[a,b] Sukrit Narula, BA,[a] Mohamed Ahmed Abdel Rahman, MD,[c] Gianni Pedrizzetti, PhD,[d] Hala Raslan, MD,[b] Osama Rifaie, MD,[c] Jagat Narula, MD, PhD,[a] Partho P. Sengupta, MD, DM[a]

ABSTRACT

OBJECTIVES The aim of this study was to investigate whether cluster analysis of left atrial and left ventricular (LV) mechanical deformation parameters provide sufficient information for Doppler-independent assessment of LV diastolic function.

BACKGROUND Medical imaging produces substantial phenotyping data, and superior computational analyses could allow automated classification of repetitive patterns into patient groups with similar behavior.

METHODS The authors performed a cluster analysis and developed a model of LV diastolic function from an initial exploratory cohort of 130 patients that was subsequently tested in a prospective cohort of 44 patients undergoing cardiac catheterization. Patients in both study groups had standard echocardiographic examination with Doppler-derived assessment of diastolic function. Both the left ventricle and the left atrium were tracked simultaneously using speckle-tracking echocardiography (STE) for measuring simultaneous changes in left atrial and ventricular volumes, volume rates, longitudinal strains, and strain rates. Patients in the validation group also underwent invasive measurements of pulmonary capillary wedge pressure and LV end diastolic pressure immediately after echocardiography. The similarity between STE and conventional 2-dimensional and Doppler methods of diastolic function was investigated in both the exploratory and validation cohorts.

RESULTS STE demonstrated strong correlations with the conventional indices and independently clustered the patients into 3 groups with conventional measurements verifying increasing severity of diastolic dysfunction and LV filling pressures. A multivariable linear regression model also allowed estimation of E/e′ and pulmonary capillary wedge pressure by STE in the validation cohort.

CONCLUSIONS Tracking deformation of the left-sided cardiac chambers from routine cardiac ultrasound images provides accurate information for Doppler-independent phenotypic characterization of LV diastolic function and noninvasive assessment of LV filling pressures. (J Am Coll Cardiol Img 2017;■:■–■) © 2017 by the American College of Cardiology Foundation.

Heart failure (HF) is a major public health problem in the United States. It is estimated that, by 2030, HF prevalence will grow by 25% and annual costs of care will increase from $21 to $53 billion (1). All efforts must be invested for investigating clinical, laboratory, and imaging data for better phenotypic characterization of HF (2,3) and designing cost-effective strategies for a reliable identification of high-risk populations at early stages. Amongst various cardiac imaging modalities, 2-dimensional (2D) and Doppler echocardiography techniques are most widely used in patients with HF

From the [a]Department of Cardiology, Icahn School of Medicine at Mount Sinai University, New York, New York; [b]Department of Internal Medicine, Medical Division, National Research Centre, Cairo, Egypt; [c]Department of Cardiology, Ain Shams University, Cairo, Egypt; and the [d]Department Engineering and Architecture, University of Trieste, Trieste, Italy. Dr. Pedrizzetti is in a research partnership with Tomtec. Dr. Jagat Narula has received research support as an equipment grant to his institution from Panasonic, Philips, and GE Healthcare; and has received minimal honorarium from GE Healthcare. Dr. Sengupta has served as an advisor to Heart Test Laboratory, Tele Health Robotics; as a consultant for Hitachi-Aloka; and has received research grants from Forest Laboratories and Heart Test Laboratory. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose. A. Jamil Tajik, MD, served as the Guest Editor for this paper.

Manuscript received May 5, 2016; revised manuscript received September 28, 2016, accepted October 3, 2016.

ABBREVIATIONS AND ACRONYMS

AV = single beat simultaneous atrioventricular measurement

AV-S = atrio-ventricular longitudinal strain at peak left ventricular systole

LA = left atrium

LV = left ventricle

SRA = peak longitudinal strain rate during left atrial contraction

SRE = peak longitudinal strain rate during early left ventricular diastole

SRS = peak longitudinal strain rate during left ventricular systole

VRA = peak volume expansion rate during left atrial contraction

VRE = peak volume expansion rate during early left ventricular diastole

VRS = peak volume expansion rate during left ventricular systole

for the assessment of left ventricular (LV) structural and functional abnormalities (4). The newer approaches in tracking natural myocardial markers, or speckles, in 2D cardiac ultrasound images for computing myocardial deformation provide incremental characterization of myocardial functional abnormalities beyond ejection fraction (EF) (5). Recent multicenter studies and global scientific consortia have therefore endorsed standardization and automation of speckle tracking echocardiography (STE) for routine clinical application (6–8).

STE provides large sets of spatial and temporal measurements; therefore, novel big data analytic approaches may be well-suited for STE databases for pattern recognition and superior staging of cardiac muscle dysfunction (9). In this investigation, we hypothesized that the cumulative information obtained from STE-based measurements is similar to that obtained from the conventional 2D echocardiograms and Doppler measurements for characterizing LV diastolic function and LV filling pressures. Therefore, we measured the STE-derived parameters in an exploratory subset of patients with HF for understanding the relationships between STE and conventional variables. In a separate validation group of patients with invasive pressure measurements, we subsequently tested the accuracy of the multivariable models derived from the exploratory set for the assessment of Doppler-independent phenotypic characterization of the LV diastolic dysfunction and noninvasive assessment of LV filling pressures.

METHODS

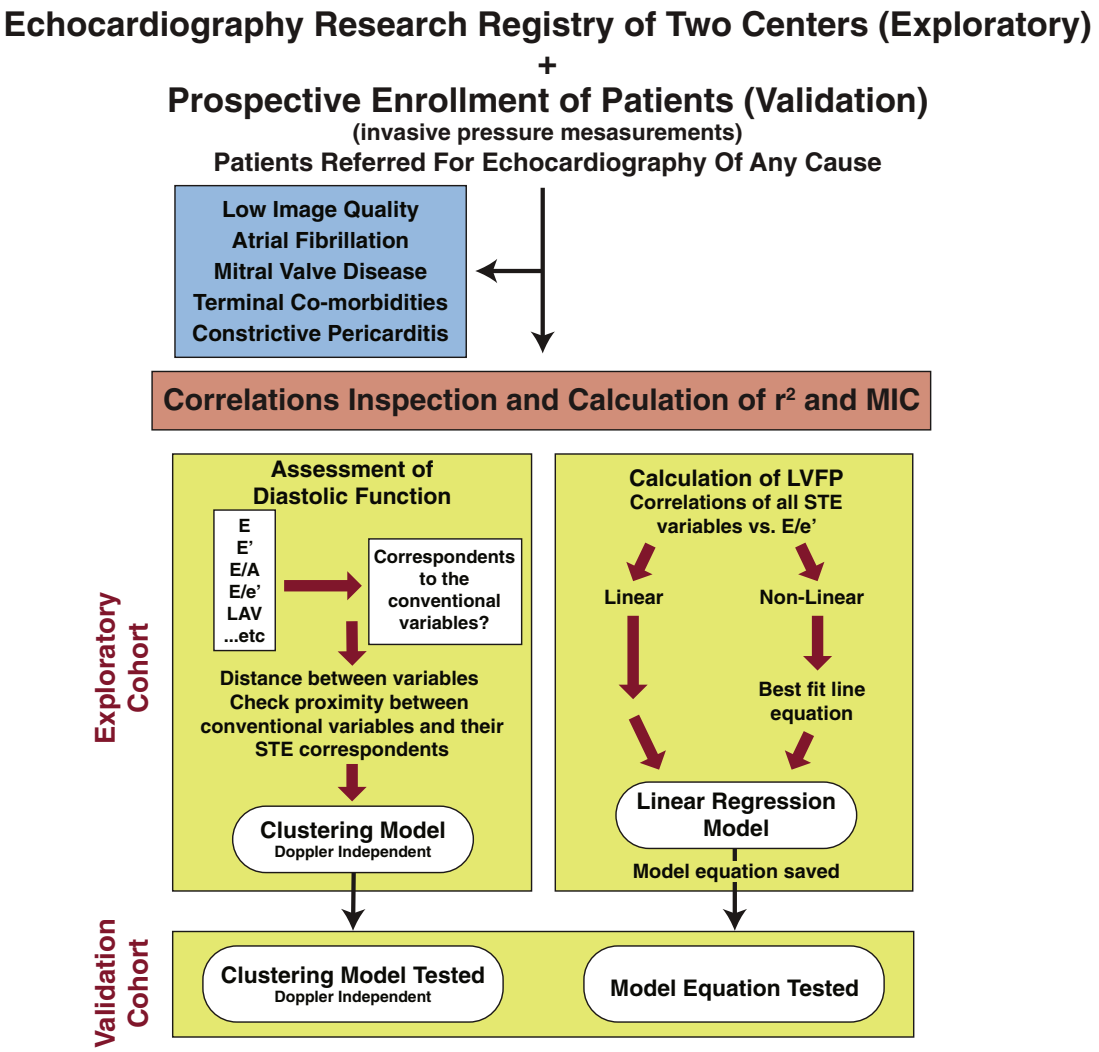
STUDY POPULATION. Patients for exploratory and validation cohorts (**Figure 1**) were recruited from 2 centers; the Ain Shams University Hospital, Cairo, Egypt (CAI) and the Icahn School of Medicine at Mount Sinai, New York, New York (NY). The local ethics committees of both institutions approved the study. A single specialist analyzed echocardiographic studies from both institutions (Mount Sinai Core Laboratory).

Exploratory group. A convenience sample was developed from data retrieved from 2 centers (CAI and NY). The CAI cohort was obtained prospectively and included 108 consecutive patients with HF symptoms referred between June 2013 and March 2014 to a single operator (Dr. Abdel Rahman) who performed all of the echocardiograms. Patients were excluded if they had poor echocardiographic images, inadequate visualization of LV and left atrial (LA) biplane views, inadequate data for assessing LV diastolic function and filling pressures, systemic comorbidities (e.g., malignancies, terminal hepatic failure, end-stage chronic renal disease on dialysis), more than a mild degree of valve disease, and pericardial diseases. As such, 25 patients were excluded from subsequent study analyses because of significant mitral regurgitation (19 patients) and insufficient echocardiographic quality (6 patients). We further enriched this sample with 79 retrospectively identified patients from NY with HF symptoms who had undergone echocardiograms in the same period with a single physician (Dr. Sengupta). After applying the same exclusion criteria as used in CAI, 47 patients were selected, with 32 excluded because of significant mitral valve disease (22 patients) and insufficient image quality (10 patients). In summary, the exploratory group comprised 130 patients with HF symptoms (83 patients from CAI, and 47 patients from NY).

Validation group. We prospectively identified 44 consecutive patients with HF symptoms who were undergoing left and right heart catheterizations. The exclusion criteria used in the exploratory cohort were also observed for the validation group. Echocardiographic examinations were performed by an investigator blinded to the exploratory group analyses (Dr. Omar) and were acquired using the same standardized protocol. Echocardiographic examinations were performed just before right heart and left heart catheterization studies. Pulmonary capillary wedge pressure (PCWP) and left ventricular end-diastolic pressure (LVEDP), were measured by an investigator (Dr. Rifaie) blinded to echocardiographic data (4).

ECHOCARDIOGRAPHIC EXAMINATION. Two-dimensional echocardiography. All echocardiographic studies were performed with a commercially available echocardiography system equipped with a 2.5-MHz multifrequency phased array transducer (Vivid 7 or E9, GE-Vingmed, Horton, Norway). Digital, routine gray-scale 2D loops from apical 2- and 4-chamber views with 3 consecutive beats were obtained with both the left ventricle and left atrium clearly and completely visualized. LV end-diastolic volume, end-systolic volume, and EF were calculated using the biplane Simpson method of discs and left atrial maximum volume (LAVmax) and minimum volume (LAVmin) were calculated using the biplane area-length method. All measurements were made in ≥3 consecutive cardiac cycles, and average values were used for the final analyses.

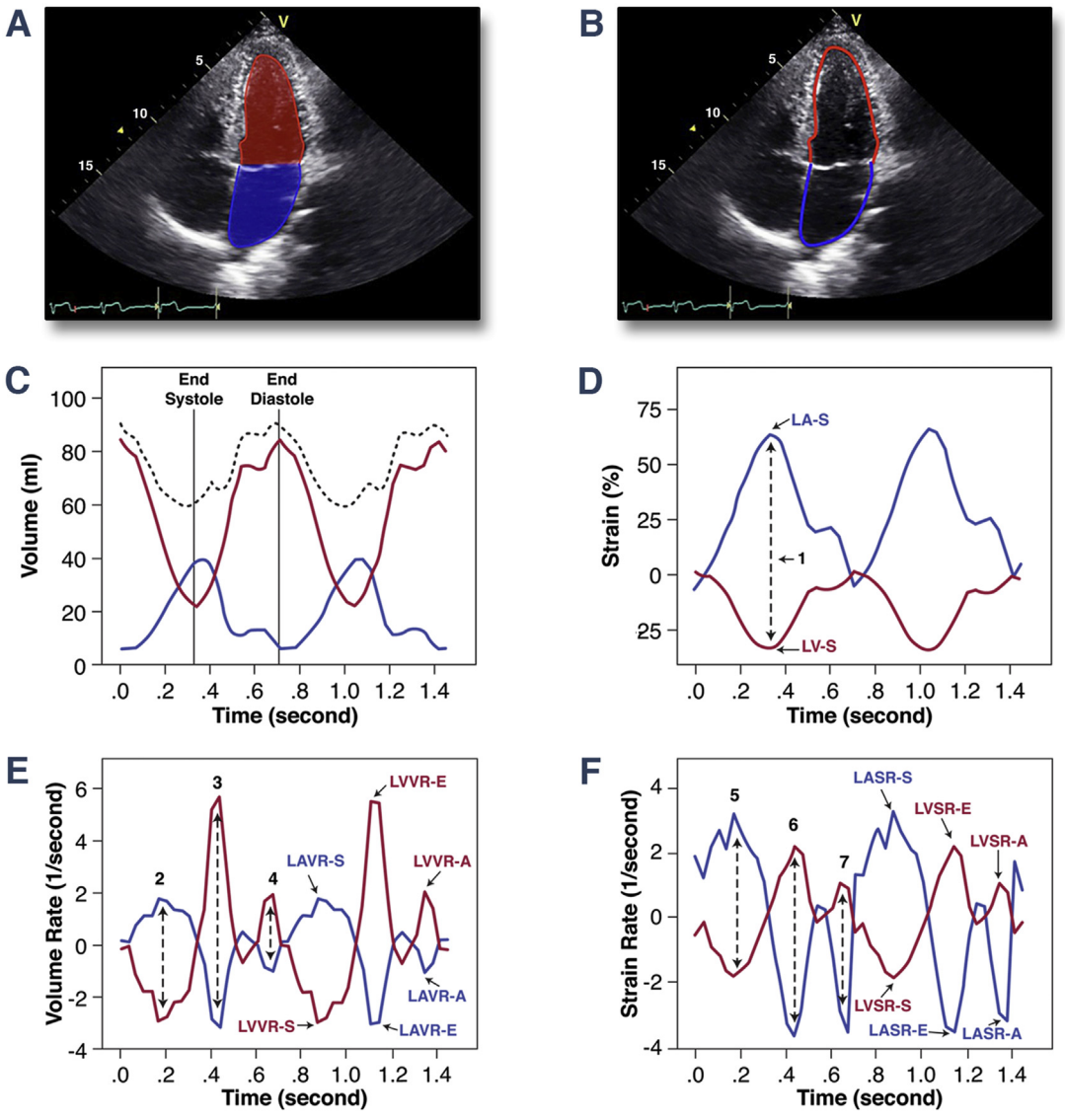
FIGURE 1 Study Protocol and Workflow



Patients for exploratory and validation cohorts were recruited from 2 centers. All patients had a comprehensive echocardiographic examination, with special emphasis on traditional echocardiographic predictors of left ventricular (LV) diastolic function and left ventricular filling pressures (LVFP), in addition to speckle tracking echocardiography (STE) calculation of average atrioventricular longitudinal wall motion (strain and strain rate) and volume expansion (volume and volume rates) parameters. Patients in the validation cohort also had invasive assessment of LVFP. Correlations and similarity assessments between the STE and conventional parameters were first tested in exploratory groups. Subsequently, the extracted data were used to construct 2 multivariable models: a clustering model for the assessment of the severity of LV diastolic dysfunction and a linear regression model for the calculation of LVFP. Both models were then saved and tested in the validation group for both noninvasive and invasive assessment of diastolic dysfunction. E = pulsed Doppler derived mitral flow early diastolic velocity; E' = tissue Doppler derived mitral annular early diastolic velocity; E/A = Doppler derived mitral flow early to late diastolic velocity ratio; E/e' = ratio of Doppler derived mitral flow early diastolic velocity to tissue Doppler derived mitral annular early diastolic velocity; LAV = left atrial maximum volume; MIC = maximal information criteria; r^2 = squared Pearson correlation coefficient.

Pulsed-wave Doppler examination. The pulsed-wave Doppler-derived transmitral velocity and tissue Doppler-derived mitral annular velocities were obtained from the apical 4-chamber view. The early diastolic wave velocity (E), late diastolic atrial contraction wave velocity (A), and the E-wave deceleration time were measured using pulsed-wave Doppler recording. Spectral pulsed-wave tissue Doppler-derived early and late diastolic velocities (e′ and a′) were averaged from the septal and lateral

FIGURE 2 Speckle Tracking-Derived Atrioventricular Measurements



(A) STE of volume expansion. **(B)** STE of wall deformation. **(C)** Volumes plotted against time during a cardiac cycle for the left atrium **(blue)**, left ventricle **(red)**, and total left heart volume in black dotted line. **(D)** LA-S **(blue)** and LV-S **(red)**. Atrioventricular strain was calculated as one-half the sum of the instantaneous maximal absolute values of LAS and LVS **(black dotted arrow 1)**. **(E)** Volume rates plotted against time during a cardiac cycle for the left atrium **(blue)** and left ventricle **(red)**. Atrioventricular volume rate during systole **(black dotted arrow 2)**, during early diastole **(black dotted arrow 3)**, and during late diastole **(black dotted arrow 4)** were calculated as one-half the sum of the instantaneous maximal absolute corresponding values of the left atrium and left ventricle. **(F)** Strain rate plotted against time during a cardiac cycle for the left atrium **(blue)** and left ventricle **(red)**. Atrio-ventricular strain rate during systole **(black dotted arrow 5)**, during early diastole **(black dotted arrow 6)**, and during late diastole **(black dotted arrow 7)** were calculated as one-half the sum of the instantaneous maximal absolute corresponding values of the left atrium and left ventricle. LA-S = left atrial longitudinal strain; LV-S = left ventricular longitudinal strain; LASR-A = peak left atrial strain rate during left atrial contraction; LASR-E = early diastolic Peak left atrial strain rate; LASR-S = peak left atrial strain rate during left ventricular systole; LAVR-A = peak left atrial volume expansion rate during left atrial contraction; LAVR-E = early diastolic Peak left atrial volume expansion rate; LAVR-S = peak left atrial volume expansion rate during left ventricular systole; LVSR-A = peak left ventricular strain rate during left atrial contraction; LVSR-E = early diastolic Peak left ventricular strain rate; LVSR-S = peak left ventricular strain rate during left ventricular systole; LVVR-A = peak left ventricular volume expansion rate during left atrial contraction; LVVR-E = early diastolic Peak left ventricular volume expansion rate; LVVR-S = peak left ventricular volume expansion rate during left ventricular systole; other abbreviations as in Figure 1.

mitral annular positions. The averaged E/e′ ratio was calculated as a Doppler echocardiographic estimate of the left ventricular filling pressure (LVFP).

STE. Two-dimensional cardiac performance analysis software (2D CPA, Tomtec Imaging Systems, Unterschleissheim, Germany) was used for a simultaneous frame-by-frame movement assessment of the stable patterns of LV and LA speckles in apical 4- and 2-chamber views. The endocardial borders of both the LA and LV were traced at the end-diastolic frame, identified as 1 frame before mitral valve closure at end-diastole. The software allowed delineating endpoints at the annulus for both LV and the LA tracking lines, with a specific marker point that indicated the position of the annulus. LA and LV speckle tracking was then performed during the cardiac cycle; the instantaneous changes in volume, volume rates, longitudinal strain, and strain rate were obtained from both apical views and averaged (Figure 2).

Single-beat speckle tracking–derived volume and strain measurements. STE-derived volume curves were used for defining LV end-diastolic volume and end-systolic volume, LAVmax, and LAVmin, and total left heart volume during ventricular systole (TLVs) and diastole (TLVd) (Figure 2C). From the strain curves, simultaneous peak LV systolic strain and peak LA strain during LV systole were also measured (Figure 2D). Atrioventricular strain (AV-S) was calculated as the average of the magnitude of global left atrial strain (LAS) and left ventricular strain (LVS) [AV-S = (LAS + LVS)/2], where both LV systolic strain and LA strain during LV systole are taken as positive values. From the volume rate and strain rate curves, simultaneous diastolic volume rates at early and late diastole and strain rates at early and late diastole and in peak ventricular systole of the left ventricle and of the left atrium were measured (Figures 2E and 2F). Finally, AV volume rate at early and late diastole as well as strain rates during early and late diastole and peak ventricular systole (VR-E_{AV}, VR-A_{AV}, SR-E_{AV}, SR-A_{AV}, SR-S_{AV}, respectively) were calculated by averaging the respective LV and LA absolute values.

CARDIAC CATHETERIZATION STUDIES. Invasive hemodynamic data were available for the validation cohort only. The echocardiographic examination was performed just before the patients were wheeled in for right and left cardiac catheterization; invasive pressures were measured through a fluid-filled balloon-tipped catheter (10). Fluoroscopically verified mean PCWP and LVEDP were obtained at end-expiration with the 0-level set at the midaxillary line and represent the average of 5 cardiac cycles (11). Significant elevation of LVFP was defined as PCWP >18 mm Hg (12,13).

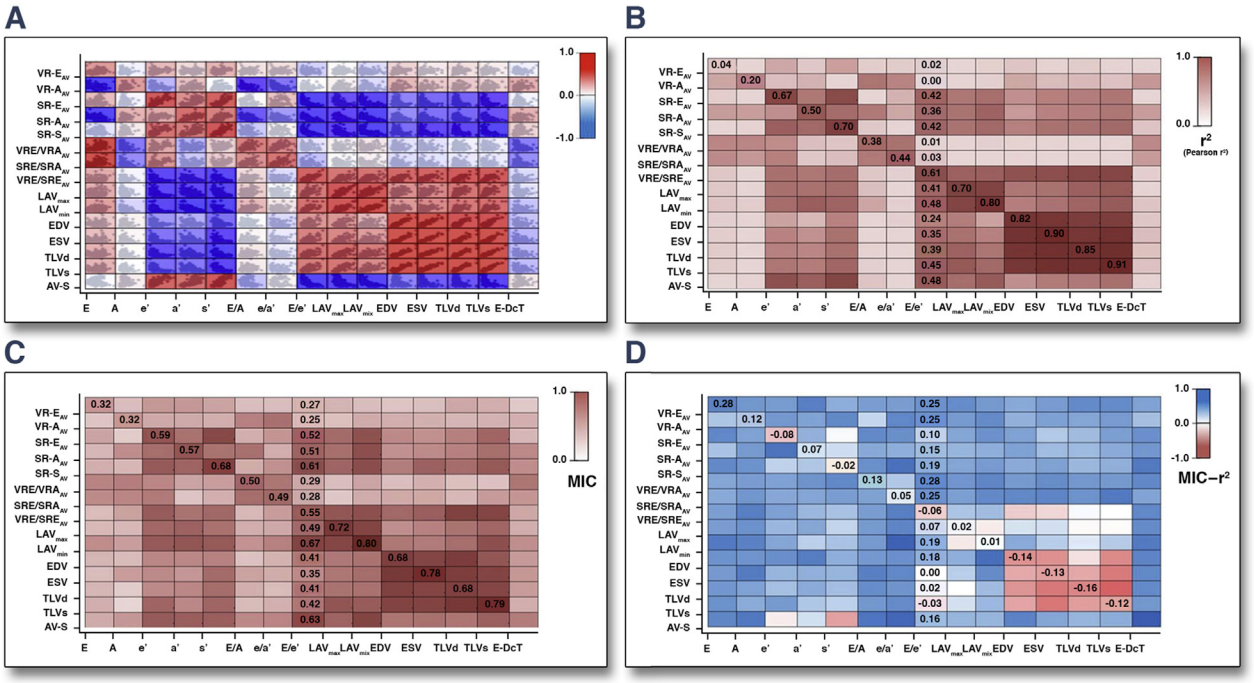
TABLE 1 Demographic, Clinical, and Echocardiographic Variables for Both Study Groups

	Exploratory Group (n = 130)	Validation Group (n = 44)
Age, yrs	53.6 ± 16.4	57.7 ± 7.9
Sex, male/female	94 (72)/36 (28)	29 (66)/15 (34)
Heart rate, beats/min	77.3 ± 13.5	78.5 ± 14.9
NYHA functional class		
I/II/III/IV	35 (28)/61 (46)/31 (22)/3 (3)	3 (7)/28 (64)/12 (27)/1 (2)*
NYHA functional class >II	34 (26)	13 (30)
Mean blood pressure, mm Hg	94.5 ± 19	103.8 ± 22*
Risk factors		
Diabetes	38 (29)	17 (39)
Hypertension	56 (43)	27 (61)*
Smoking	38 (29)	10 (23)
Hyperlipidemia	23 (18)	4 (9)
>2 risk factors	44 (34)	20 (45)
Type of presentation		
Dilated cardiomyopathy	58 (45)	20 (45)
Ischemic heart disease	23 (18)	24 (55)
Restrictive cardiomyopathy	11 (8)	–
Hypertension	14 (11)	–
Other	24 (18)	–
Treatment		
Beta-blockers	62 (48)	25 (57)
Renin angiotensin-aldosterone blockers	44 (34)	17 (39)
Spironolactone	23 (18)	3 (7)*
Furosemide	46 (35)	15 (34)
Digoxin	15 (12)	1 (2)
Statins	47 (36)	16 (36)
Nitrates	22 (17)	24 (55)*
Aspirin	50 (38)	29 (66)*
Clopidogrel	11 (8)	13 (30)*
Calcium channel blockers	10 (7)	5 (11)
Conventional variables		
LAVmax, ml	66 ± 25.8	62.4 ± 22.3
LAVmin, ml	30.7 ± 21.6	24.9 ± 15.1
ESV, ml	70.2 ± 58	54.1 ± 29
EF, %	53.1 ± 16	55.3 ± 12.6
EF >50%/EF <50%	73/57	25/19
E, cm/s	82.4 ± 22.4	76.8 ± 18.9
A, cm/s	68.6 ± 29.8	77.1 ± 26.2
E-DcT, ms	180.6 ± 74.7	192 ± 66.4
e′, cm/s	7.7 ± 4	7.1 ± 2
a′, cm/s	7.8 ± 3.1	8.6 ± 2.7
E/A	1.41 ± 0.65	1.15 ± 0.67*
E/e′	13.1 ± 6.6	11.5 ± 4.1

Nominal data are expressed as n (%) and continuous data are expressed as mean ± SD. *p < 0.05.

A = mitral flow late diastolic velocity; a′ = tissue Doppler derived mitral annular late diastolic velocity; e′ = tissue Doppler derived mitral annular early diastolic velocity; EDV = end-diastolic volume; EF = ejection fraction; E = mitral flow early diastolic velocity; E-DcT = mitral E-wave deceleration time; ESV = end-systolic volume; LAVmax = maximal left atrial volume in milliliters; LAVmin = minimal left atrial volume; NYHA = New York Heart Association; s′ = tissue Doppler-derived mitral annular ejection systolic velocity; TLVs = total left heart volume during ventricular systole.

FIGURE 3 Heat Maps of Correlations Between STE and Conventional Variables in the Exploratory Group

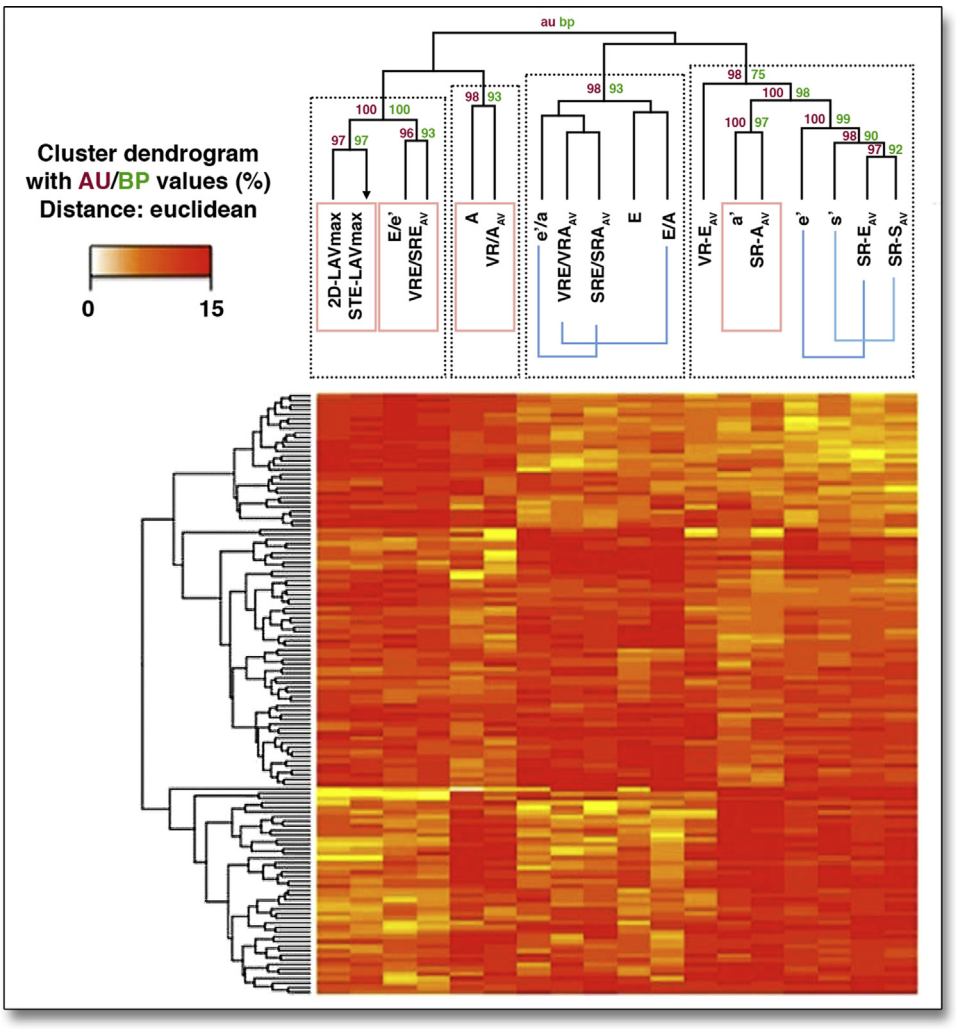


(A) Heat maps overlapping the dots plots of correlations for illustrative purposes. Direct correlations are shown in shades of red (the darker the color, the more it approves correlation coefficient of 1.0), and inverse correlations are sown in shades of blue (the darker the color, the more it approves correlation coefficient of -1.0). **(B)** Squared correlation coefficient (r^2), **(C)** MIC, and, **(D)** the difference between both (MIC − r^2) are shown for the STE that best represent correlates with the conventional variables as well as between all STE variables and E/e′, as a surrogate for LVFP. Correlations were first visually inspected as in **(A)** and linear correlations were expressed as Pearson's correlation equation. Nonlinear correlations were confirmed by the value MIC − r^2 and were expressed as the equation of the best line of fit. AV-S = atrioventricular strain; EDV = end-diastolic volume; ESV = end-systolic function; LAVmax = left atrial maximum volume; LAVmin = left atrial minimum volume; LVFP = left ventricular filling pressure; MIC = maximal information criteria; SR-A_{AV} = peak atrioventricular strain rate during atrial contraction; SR-E_{AV} = early diastolic peak atrioventricular strain rate; SRE/SRA_{AV} = ratio between atrioventricular strain rate at early diastole and during atrial contraction; TLVd = total left heart volume during ventricular diastole; TLVs = total left heart volume during ventricular systole; VR-A_{AV} = peak atrioventricular volume expansion rate at left atrial contraction; VR-E_{AV} = early diastolic peak atrioventricular volume expansion rate; VRE/SRE_{AV} = ratio between atrioventricular volume expansion rate and strain rate at early diastole; VRE-VRA_{AV} = ratio between atrioventricular volume expansion rate at early diastole and during atrial contraction.

STATISTICAL ANALYSIS. Categorical variables were expressed as number (%) and were compared using the chi-square test. Continuous variables were expressed as mean ± SD, and were compared using the independent sample Student *t* test. Correlations between conventional versus STE parameters were tested using linear regression, expressed as Pearson correlation coefficient (r) and by exploring the maximal information criteria (MIC). The difference between the absolute values of MIC and the squared Pearson correlation coefficient (r^2) (i.e., MIC − $r^2 \geq 0.1$), was used as a marker of nonlinear correlation, confirmed subsequently by visual inspection of the correlation plots (14). Correlations were then used to construct 2 multivariable models comprising STE variables: 1) a clustering model for the assessment of the severity of diastolic dysfunction; and 2) a linear regression model to compute the LVFP. Details of the statistical methods used to construct the models can be found in the Online Appendix (Online Tables 1 and 2). Intraobserver, interobserver, and beat to beat variability were also checked and were expressed as mean ± SD of the differences and interclass correlation coefficients (Online Tables 3 and 4).

All analyses were performed with commercially available software (SPSS version 21.0, SPSS, Inc., Chicago, Illinois, and R: A language and environment for statistical computing, version 3.0.1, R Foundation for Statistical Computing, Vienna, Austria). A p value of <0.05 was considered statistically significant. The detailed statistical methods are presented in the following section.

FIGURE 4 Clustering Dendrograms for Conventional Variables and Their STE Correspondents



The correlation plots in Figure 1 suggested STE correspondents of conventional variables. Clustering dendrograms using STE and conventional variables together. The dissimilarity matrix is given as a heat map of Euclidean distance **(red)**. The AU **(red numbers)** and BP **(green numbers)** were calculated. AU values are shown only for leaflets that had an AU >95% (considered statistically significant). Significant proximity of variables in the clustering leaflets were decided using 2D-LAVmax, E/e′, A-wave velocity, and a′ velocity were shown to be in perfect proximity with their STE counterparts STE-LAVmax, VR-E/SR-E$_{AV}$, VR-A$_{AV}$, and SR-A$_{AV}$, respectively (AU = 97%, 96%, 98%, and 100%, respectively). The conventional parameters e′/a′ and E/A were also in significant proximity to their STE counterparts SR-E/SR-A$_{AV}$ and VR-E/VR-A$_{AV}$, respectively (AU = 98%) and also between e′ and s′ and their STE counterparts SR-E$_{AV}$ and SR-S$_{AV}$, respectively (AU = 100%). AU = approximately unbiased probability; 2D = 2-dimensional; VR-E = rate of volume expansion at early diastole; other abbreviations as in Figures 1 and 3.

RESULTS

The demographic, clinical, and echocardiographic data for patients from the exploratory and validation groups are summarized in Table 1. Both groups were similar in age, sex distribution, and New York Heart Association (NYHA) functional class. All risk factors were also similar in both groups except hypertension, which was more prevalent in the validation cohort (p = 0.028). Both groups had a similar degree of LV remodeling and severity of diastolic function.

EXTRACTION OF STE CORRESPONDENTS OF CONVENTIONAL VARIABLES. Correlations between STE characteristics and conventional measurements. Figure 3 shows the correlations between the conventional and Doppler-based

TABLE 2 Description of Properties of the Doppler Independent Clusters

	Cluster 1	Cluster 2	Cluster 3
Exploratory (n = 130)			
Number	34 (26)	51 (39)	45 (35)
Age, yrs	37.3 ± 16.0	58.1 ± 13.0	61.2 ± 10.0*†
Males	23 (68)	34 (67)	37 (82)
Diabetic	7 (21)	12 (24)	19 (42)
Hypertensive	8 (24)	25 (49)	23 (51)*
NYHA functional class >2	0 (0)	9 (18)	25 (56)*
EF, %	63.5 ± 10.0	52.8 ± 16.0	44.9 ± 15.5*†‡
EF <50%	2 (6)	8 (16)	24 (53)*
E, cm/s	90.2 ± 13.2	70.4 ± 22.3	90.2 ± 22.4*†‡
A, cm/s	63.0 ± 19.9	81.1 ± 26.9	58.7 ± 34.4*†‡
E-DcT	170.2 ± 56.0	216.3 ± 76.0	148.5 ± 68.6*†‡
e′, cm/s	13.2 ± 2.2	6.5 ± 2.1	5.0 ± 1.5*†‡
a′, cm/s	9.8 ± 2.4	9.1 ± 2.3	4.7 ± 1.8*‡
s′, cm/s	9.6 ± 1.7	6.7 ± 1.4	4.8 ± 1.5*†‡
E/A	1.53 ± 0.42	0.97 ± 0.49	1.4 ± 0.65*†‡
e′/a′	1.42 ± 0.46	0.76 ± 0.34	1.25 ± 0.65*†‡
E/e′	7.1 ± 1.8	11.6 ± 4.0	19.2 ± 6.3*†‡
LAVmax, ml	45.7 ± 10	58.9 ± 15.2	66.0 ± 25.0*†‡
Validation (n = 44)			
Number	7 (16)	29 (66)	8 (18)
Age, yrs	53.4 ± 10.0	58.9 ± 7.6	57.4 ± 6.8
Males	6 (86)	19 (66)	4 (50)
Diabetic	5 (71)	10 (34)	2 (25)
Hypertensive	5 (71)	19 (56)	3 (38)
NYHA functional class >2	2 (29)	9 (31)	2 (25)
EF, %	57.3 ± 11.5	58.5 ± 10.1	42.2 ± 10.9*†‡
EF <50%	3 (43)	9 (31)	7 (88)*
E, cm/s	83.1 ± 20.3	69.6 ± 13.2	97.1 ± 20.8*‡
A, cm/s	62.5 ± 19.6	84.3 ± 23.1	63.6 ± 33.4
E-DcT	144.6 ± 46.0	212.9 ± 59.4	152.0 ± 72.0*†‡
e′, cm/s	9.0 ± 2.3	7.1 ± 1.7	5.5 ± 1.0*†‡
a′, cm/s	8.6 ± 1.4	9.6 ± 2.3	5.0 ± 1.7*‡
s′, cm/s	7.5 ± 1.4	7.3 ± 1.5	4.9 ± 1.25*‡
E/A	1.6 ± 1.0	0.87 ± 0.27	1.8 ± 0.74*†‡
e′/a′	1.1 ± 0.46	0.78 ± 0.23	1.2 ± 0.35*†‡
E/e′	9.4 ± 1.5	10.4 ± 3.2	17.7 ± 1.8*‡
LAVmax, ml	49.0 ± 14.0	58.3 ± 20.0	88.8 ± 15.7*‡
PCWP, mm Hg	14.6 ± 4.1	16.1 ± 6.4	27.3 ± 5.4*‡
LVEDP, mm Hg	18.4 ± 5.6	20.0 ± 5.5	36.0 ± 6.6*‡

Values are n (%) or mean ± SD. *Overall $p < 0.05$. †$p < 0.05$ between clusters 1 and 2. ‡$p < 0.05$ between clusters 2 and 3. Cluster 1 represents a Doppler-independent group of patients with low left ventricular filling pressures as identified by the conventional variables, cluster 2 represents a Doppler-independent group of patients with mildly elevated left ventricular filling pressures as identified by the conventional variables, and cluster 3 represents a Doppler-independent group of patients with significantly elevated left ventricular filling pressures as identified by the conventional variables.

PCWP = pulmonary capillary wedge pressure; LVEDP = left ventricular end diastolic pressure; other abbreviations as in Table 1.

measurements versus STE-based indices. The distribution of the STE-based measurements showed several correlations with the conventional parameters (Figure 3A, Online Table 1), as shown next.

The best correlations for strain rate-based parameters found between SR-E_{AV} with e′ ($r^2 = 0.67$, MIC = 0.59; $p < 0.001$), SR-A_{AV} with a′ ($r^2 = 0.50$, MIC = 0.57; both $p < 0.001$), and SR-S_{AV} with s′ ($r = 0.7$, MIC = 0.68; both $p < 0.001$). From the values, it can be concluded that all correlations were linear (all MIC – $r^2 < 0.1$) (Figures 2B to 2D, Online Table 2).

The best volume rate-based parameters correlations were found between VR-E_{AV} with E wave velocity ($r^2 = 0.04$, $p = 0.03$; and MIC 0.32; $p < 0.001$) and VR-A_{AV} with A wave velocity ($r^2 = 0.2$, MIC = 0.32; both $p < 0.001$). Both correlations were of nonlinear nature (both MIC – $r^2 > 0.1$) (Figures 2B to 2D).

The best correlations for ratio-based parameters were found between VR-E/VR-A_{AV} with the ratio E/A ($r^2 = 0.38$, MIC = 0.5; both $p < 0.001$), SR-E/SR-E_{AV} with the ratio e′/a′ ($r^2 = 0.44$, MIC = 0.49; both $p < 0.001$), and VR-E/SR-E_{AV} with the ratio E/e′ ($r^2 = 0.61$, MIC = 0.55; both $p < 0.001$). As such, the correlation with E/A was found to be nonlinear (MIC – $r^2 > 0.1$) and the correlations with e′/a′ and E/e′ were linear (MIC – $r^2 < 0.1$) (Figures 2B to 2D).

Strong correlations were found between the volume-based parameters 2D and STE-derived LV (EDV: $r^2 = 0.82$, MIC = 0.68, both $p < 0.001$; ESV: $r^2 = 0.9$, MIC = 0.78, both $p < 0.001$) and LA volumes (LAVmax: $r^2 = 0.7$, MIC = 0.72, both $p < 0.001$; LAVmin: $r^2 = 0.8$, MIC = 0.8, both $p < 0.001$). Moreover, 2D and STE-derived total left heart volumes (TLVd and TLVs), also correlated strongly (TLVd: $r^2 = 0.85$, MIC = 0.68, both $p < 0.001$; TLVs: $r^2 = 0.91$, MIC = 0.79, both $p < 0.001$). All correlations were of a linear nature, as suggested by MIC – $r^2 < 0.1$ (Figures 2B to 2D).

Variable clustering and similarity assessment. Next, a clustering model was constructed using the traditionally used parameters E, A, e′, and a′ velocities and LAVmax, in addition to the ratios E/A, e′/a′, and E/e′, in addition to their correlation-derived corresponding STE variables (Figure 4). The model identifies the closest coupled parameters and segregates them into groups according to the distance between the coupled parameters. In this model, all STE variables were perfectly coupled or showed significant proximity to their conventional counterparts within the clustering dendrograms (approximately unbiased probability >95%), suggesting a high statistical level of overlap between the STE-derived variables and the conventional volume and Doppler variables (log likelihood ratio: -1,413.0, BIC: -4,287.5; distance: minimum 1.2, median 5.522, mean 5.671, maximum 13.380) (Figure 4).

As such, it can be inferred from the correlations that correspondents to the conventionally used parameters

can be found within the STE data, and that these parameters, according to the similarity analysis, might be functioning similarly to the conventional variables. As a result, the extracted STE variables were used as correspondents of the conventional parameters in deriving the Doppler independent models.

DEVELOPMENT OF DOPPLER-INDEPENDENT CLUSTERING MODEL FOR ASSESSING DIASTOLIC DYSFUNCTION SEVERITY. A 2-step clustering model was constructed using all available STE parameters. On this basis, the model divided the patients into 3 groups (log likelihood ratio: -451, BIC: -1,939.3) (Table 2). These clusters showed progressive worsening of diastolic functions and LVFP, as suggested by the conventional variables E/A, e′, LAVmax, and E/e′ (Table 2), and were accompanied with progressively increasing age, higher prevalence of hypertension, increasing NYHA functional class, and worsening EF.

DEVELOPMENT OF MULTIVARIABLE LINEAR REGRESSION MODEL FOR CALCULATION OF FILLING PRESSURES. Correlations between all STE variables and E/e′ (as a surrogate of LVFP) were rechecked. Visual inspection identified that correlations were of a linear nature for TLVs, TLVd, and VR-E/SR-E_{AV} ($r = 0.67, 0.62, 0.78$, respectively; all $p < 0.001$) and were nonlinear for AV-S, SR-E_{AV}, SR-A_{AV}, and SR-S_{AV} (best fit $r = 0.73, 0.77, 0.65, 0.7$; all $p < 0.001$). A multivariate regression model suggested that AV-S, SR-E_{AV}, and VR-E/SR-E_{AV} were the best independent predictors of E/e′ (beta = 0.28, 0.187, 0.448; $p = 0.01$, 0.032, <0.001; model-adjusted $r^2 = 0.684$, $p < 0.001$) (Figures 5A and 5B), whereas other variables lost statistical significance. The model equation using these 3 variables was saved and tested in the validation cohort (Online Appendix).

APPLICATION OF THE FINDINGS OF THE EXPLORATORY GROUP TO THE VALIDATION GROUP. Clustering model. The Doppler-independent 2-step clustering model, when applied to the validation group, reproduced the same features of clustering as seen in the exploratory cohort. The clinical and functional characteristics of the exploratory cohorts were also retained in the validation cohort. Although patient's age, risk factors, and NYHA functional class were not statistically different between clusters, systolic and diastolic function parameters were significantly different between the groups.

Moreover, to overcome the differences between the exploratory and validation cohorts, a 2-step cluster analysis initiated from a 1:1 propensity score-derived exploratory group was undertaken. The originally noted differences between the clusters persisted in the validation cohort (Online Appendix, Online Table 5).

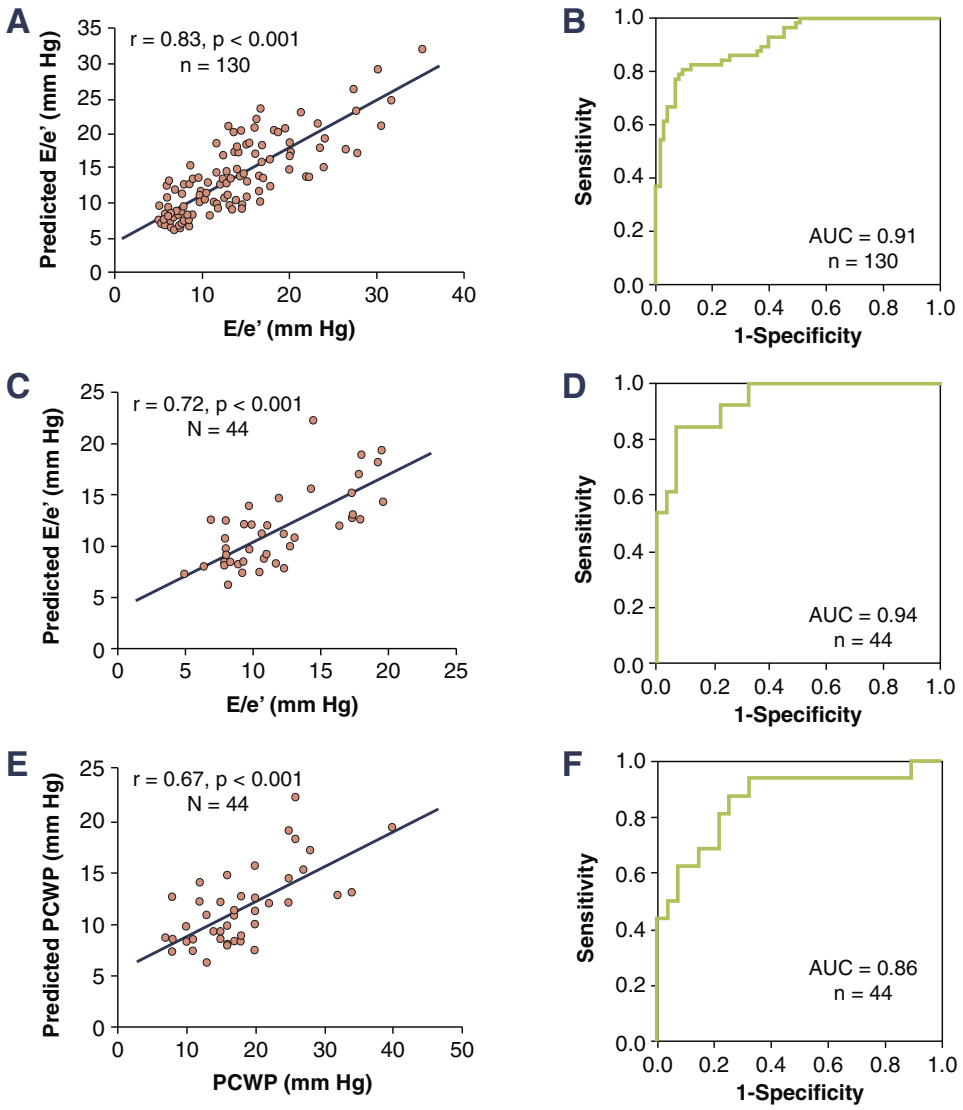
Linear regression model. Using the multivariate linear regression model equation saved from the exploratory cohort, the predicted E/e′ was found to correlate significantly with E/e′ ($r = 0.72$, $p < 0.001$) (Figure 5C). Receiver operator characteristic curves suggested areas under the curve for detection of E/e′ >13 of 0.94 (95% confidence interval: 0.872 to 0.999) (Figure 5D).

VALIDATING THE CLUSTERING AND THE LINEAR REGRESSION MODELS IN THE ASSESSMENT OF INVASIVELY MEASURED HEMODYNAMICS. Interestingly, in the 2-step clustering model produced in the validation cohort, the values of PCWP and LVEDP corresponded to the severity of diastolic function (Table 2). Of more importance, the multivariate linear regression equation was able to predict PCWP ($r = 0.67$, $p < 0.001$) (Figure 5E). Receiver operator characteristic curves suggested areas under the curve for detection of PCWP >18 of 0.86 (95% confidence interval: 0.738 to 0.985) (Figure 5F).

DISCUSSION

The key findings of this study are: 1) a high statistical level of overlap was seen between STE-derived data and conventional echocardiographic methods of diastolic function assessment; 2) clustering of the patients based on STE data into 3 different groups that corresponded to worsening severity of diastolic dysfunction grades as verified by the conventional parameters; and 3) building a linear multivariable model from Doppler-independent STE data demonstrated a good diagnostic accuracy in predicting LV filling pressures. These findings suggest that the information content of STE variables corresponds to that derived from 2D and Doppler-based analysis and can provide an independent assessment of diastolic function and LV filling pressures. There is growing shift toward using data-driven analytics for precision medicine (15). Use of these techniques in the field of cardiac imaging is enabling automated interpretation of echocardiography images using machine-learning approaches (8,16). The data presented in this study thus represent the first steps in computer-driven unsupervised classification of recurring patterns in speckle tracking-derived data that will potentially open up new opportunities for fully automated assessment of diastolic function in clinical practice.

FIGURE 5 Multivariate Linear Regression Model for Correlations in the Exploratory and Validation Cohorts



(A) Dot plot showing the correlation of the output of the equation of a multivariate model constructed of SR-E_{AV}, AV-S, VR-E/SR-EAV (predicted E/e′), and the Doppler surrogate of LVFP E/e′ in the exploratory group. **(B)** Receiver operator characteristic curve of the ability of the model (predicted E/e′) to predict E/e′ >13 in the exploratory group. **(C)** Dot plot showing the correlation of the output of the same multivariate model equation (predicted E/e′) and the Doppler surrogate of LVFP E/e′ in the validation group. **(D)** Receiver operator characteristic curve of the ability of the model (predicted E/e′) to predict of E/e′ > 13 in the validation group. **(E)** Dot plot showing the correlation of the output of the same multivariate model equation and the invasively measured PCWP in the validation group. **(F)** Receiver operator characteristic curve of the ability of the model (predicted PCWP) to predict of PCWP >18 mm Hg in the validation group.
AUC = area under the curve; PCWP = pulmonary capillary wedge pressure; other abbreviations as in Figures 1 to 4.

CONVENTIONAL ASSESSMENT OF DIASTOLIC FUNCTION: STRENGTHS AND PITFALLS. The assessment of LV diastolic functions is essential in the course of management of patients with heart failure (17). Methods used have evolved over the decades from clinical evaluation to invasive pressure measurements and noninvasive advanced imaging, with echocardiography commonly used at the bedside (17). However, no single echocardiographic parameter has been shown to adequately address diastolic function, and conventional

approaches incorporate several 2D, Doppler, and tissue Doppler variables in decision tree-based algorithms. This occurs mainly because the development of diastolic dysfunction involves complex interactions, which include defects in active and passive diastolic properties occurring separately or in combination. Moreover, LV diastolic function is closely coupled with atrial function; thus, the transmitral flow parameters of the left ventricle have to be interpreted in light of LV relaxation, LA function, and loading parameters, with each one affecting LV filling pressures. Because of the high dimensionality and complexity of variables that affect LV diastolic function and filling pressures, a multiparametric approach is necessary during echocardiographic assessment (4,18).

Although currently recommended algorithms continue to be successfully implemented in clinical practice, they require a high level of training and expertise to be effectively used (4). The validity of the consensus-driven classification algorithm versus data-driven clustering tree classification could be an important consideration for further streamlining of the existing guidelines. However, the use of a multistep multiparametric approach using different 2D and pulsed Doppler measurements requires considerable laboratory effort for standardization (19). In addition, these parameters are compiled from different cardiac cycles at different cardiac locations and are thus susceptible to time and hemodynamic load-related measurement variances (19-21). In contrast, the instantaneous integration of atrial and ventricular function and geometry from a single heart beat may be useful to overcome variability related to heart rate and respiratory load-related changes. Several STE parameters such as SR-E have been previously investigated as surrogate variables for measuring LVFP or to assess of LV diastolic function (22-24). Such STE-based assessments are attractive, alternative ways of assessing cardiac function because multivariable assessment of cardiac function can be performed using just gray scale-based cardiac ultrasound motion data with multiple spatial assessments that can be integrated for more than 1 cardiac chamber.

ROLE OF STE-DERIVED LARGE DATA ANALYTIC PLATFORMS IN PRECISION MEDICINE. The field of big data analytics has operationalized precision medicine as an approach to establishing clinical phenotypic characterization of different diseases while taking into account genetic and environmental variability (15). Phenomapping approaches using unbiased cluster analysis have recently been proposed for meaningful categorization of patients with HF (2). As shown in our study, STE is capable of generating useful data from a single echocardiographic loop and can, therefore, improve imaging-based cardiovascular phenotypic characterization. Moreover, recent standardization efforts (6,7) have revealed that STE measurements are more reproducible than conventional 2D and Doppler indices. The emergence of automated STE approaches may further improve efficiency and reduce interobserver and intraobserver variability (8). In the near future, the ability to extract large-scale information from echocardiography database will enable the emergence of machine learning models capable of capturing data for automated analyses, so that after images are uploaded, information will be automatically extracted and analyzed for providing decisions in real time. This may be helpful in increasing diagnostic throughput and efficiency in the face of the growing burden of cardiovascular disease in the community and the existing work shortage in the field (8,25,26).

STUDY LIMITATIONS. This study used novel clustering approaches for the analysis of large-scale STE data for characterizing diastolic function grades and LV filling pressures. However, STE measurements in our study were obtained from biplane views for simultaneously enabling chamber quantification using the biplane Simpson method. This allowed extraction of functional and geometric measurements that correspond to conventional 2D and Doppler-based functional assessments. The 3-chamber view of the left ventricle and left atrium was not included; therefore, the data may not be truly representative of global LV and LA mechanics. Second, only longitudinal velocity, strain, and strain rate parameters were used in the STE database because longitudinal LV mechanical parameters are better standardized and currently more reproducible. The incremental value of radial and circumferential strain parameters and LV twist mechanics was not tested in the present investigation. Third, the overall sample size was small, especially for subgroup analysis for patients with preserved or reduced EF and the prognostic information of clustered groups was not evaluated in the current study. The sample size for exploratory and validation cohort were in a 70:30 ratio, which is concordant with designs in previous studies (27,28). Finally, the correlation of the model used for calculation of PCWP in the validation cohort was only moderate, which might be related to the small sample size and the narrow range of pressures included. However, the area under the curve for prediction of PCWP >18 mm Hg was 0.86. Further studies should

include more samples to validate if such findings can be sufficient for clinical decision-making.

CONCLUSIONS

STE offers large-scale data with a high level of information overlap with the existing 2D and Doppler-based indices of diastolic function. Cluster patterns of STE-based data may be useful for phenotypic characterization of LV diastolic functions in patients with diastolic dysfunction.

ACKNOWLEDGMENTS The authors thank Dr. Shameer Khader and Dr. Joel Dudley for their advise on statistical analysis and Dr. Rashid Ahmed for his efforts in preparing this manuscript.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. Partho P. Sengupta, Icahn School of Medicine at Mount Sinai, Mount Sinai Medical Center, One Gustave L. Levy Place, P.O. Box 1030, New York, New York 10029. E-mail: partho.sengupta@mountsinai.org.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: The currently recommended algorithms for echocardiographic assessment of left ventricular diastolic function and filling pressures are complex, requiring multistep 2D, Doppler and tissue Doppler acquisitions. Alternately, the same information could be automatically assembled from left atrial and ventricular speckle tracking echocardiography database which contains substantial phenotypic information for applying modern clustering and classification algorithms.

TRANSLATIONAL OUTLOOK: Application of clustering and classification algorithm to speckle tracking derived database would be useful for development of machine learning algorithms that can allow automated diastolic function assessment. This would be potentially useful for standardized echocardiographic evaluations and improving the quality of interpretation.

REFERENCES

1. Heidenreich PA, Albert NM, Allen LA, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. Circ Heart Fail 2013;6:606–19.

2. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation 2015;131:269–79.

3. Francis GS, Cogswell R, Thenappan T. The heterogeneity of heart failure: will enhanced phenotyping be necessary for future clinical trial success? J Am Coll Cardiol 2014;64:1775–6.

4. Nagueh SF, Smiseth OA, Appleton CP, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr 2016;29:277–314.

5. Lang RM, Badano LP, Mor-Avi V, 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;28:1–39.

6. Yang H, Marwick TH, Fukuda N, et al. Improvement in strain concordance between two major vendors after the strain standardization initiative. J Am Soc Echocardiogr 2015;28:642–8.

7. Voigt JU, Pedrizzetti G, Lysyansky P, et al. Definitions for a common standard for 2D speckle tracking echocardiography: consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. J Am Soc Echocardiogr 2015;28:183–93.

8. Knackstedt C, Bekkers SCAM, Schummers G, et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: the FAST-EFs Multicenter Study. J Am Coll Cardiol 2015;66:1456–66.

9. Narula J. Are we up to speed? From big data to rich insights in CV imaging for a hyperconnected world. J Am Coll Cardiol Img 2013;6:1222–4.

10. Swan HJ, Ganz W, Forrester J, Marcus H, Diamond G, Chonette D. Catheterization of the heart in man with use of a flow-directed balloon-tipped catheter. N Engl J Med 1970;283:447–51.

11. Nagueh SF, Middleton KJ, Kopelen HA, Zoghbi WA, Quinones MA. Doppler tissue imaging: a noninvasive technique for evaluation of left ventricular relaxation and estimation of filling pressures. J Am Coll Cardiol 1997;30:1527–33.

12. Forrester JS, Diamond G, Chatterjee K, Swan HJ. Medical therapy of acute myocardial infarction by application of hemodynamic subsets (first of two parts). N Engl J Med 1976;295:1356–62.

13. Nohria A, Tsang SW, Fang JC, et al. Clinical assessment identifies hemodynamic profiles that predict outcomes in patients admitted with heart failure. J Am Coll Cardiol 2003;41:1797–804.

14. Reshef DN, Reshef YA, Finucane HK, et al. Detecting novel associations in large data sets. Science 2011;334:1518–24.

15. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015;372:793–5.

16. Sengupta PP, Huang YM, Bansal M, et al. Cognitive machine-learning algorithm for cardiac imaging: a pilot study for differentiating constrictive pericarditis from restrictive cardiomyopathy. Circ Cardiovasc Imaging 2016;9:e004330.

17. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013;62:e147–239.

18. Mottram PM, Marwick TH. Assessment of diastolic function: what the general cardiologist needs to know. Heart 2005;91:681–95.

19. Park JH, Marwick TH. Use and limitations of E/e' to assess left ventricular filling pressure by echocardiography. J Cardiovasc Ultrasound 2011;19:169–73.

20. Bhella PS, Pacini EL, Prasad A, et al. Echocardiographic indices do not reliably track changes in left-sided filling pressure in healthy subjects or patients with heart failure with preserved ejection fraction. Circ Cardiovasc Imaging 2011;4:482–9.

21. Mullens W, Borowski AG, Curtin RJ, Thomas JD, Tang WH. Tissue Doppler imaging in the estimation of intracardiac filling pressure in decompensated patients with advanced systolic heart failure. Circulation 2009;119:62–70.

22. Chen S, Yuan J, Qiao S, Duan F, Zhang J, Wang H. Evaluation of left ventricular diastolic function by global strain rate imaging in patients with obstructive hypertrophic cardiomyopathy: a simultaneous speckle tracking echocardiography and cardiac catheterization study. Echocardiography 2014;31:615–22.

23. Kuwaki H, Takeuchi M, Chien-Chia Wu V, et al. Redefining diastolic dysfunction grading: combination of E/A ≤0.75 and deceleration time >140 ms and E/epsilon' ≥10. J Am Coll Cardiol Img 2014;7:749–58.

24. Wang J, Khoury DS, Thohan V, Torre-Amione G, Nagueh SF. Global diastolic strain rate for the assessment of left ventricular relaxation and filling pressures. Circulation 2007;115:1376–83.

25. Bonow RO, Smith SC Jr. Cardiovascular manpower: the looming crisis. Circulation 2004; 109:817–20.

26. Rodgers GP, Conti JB, Feinstein JA, et al. ACC 2009 survey results and recommendations: addressing the cardiology workforce crisis. A report of the ACC board of trustees workforce task force. J Am Coll Cardiol 2009;54:1195–208.

27. Nahato KB, Harichandran KN, Arputharaj K. Knowledge mining from clinical datasets using rough sets and backpropagation neural network. Comput Math Methods Med 2015; 2015:460189.

28. Krasteva V, Jekova I, Leber R, Schmid R, Abacherli R. Superiority of classification tree versus cluster, fuzzy and discriminant models in a heartbeat classification system. PLoS One 2015; 10:e0140123.

KEY WORDS big-data analytics, diastolic dysfunction, left ventricular filling pressures, speckle-tracking echocardiography

APPENDIX For supplemental text and tables, please see the online version of this article.