Machine Learning–Based Phenogrouping in Mitral Valve Prolapse Identifies Profiles Associated With Myocardial Fibrosis and Cardiovascular Events

Olivier Huttin, Nicolas Girerd, Antoine Jobbe-Duval, Anne-Laure Constant Dit Beaufils, Thomas Senage, Laura Filippetti, Caroline Cueff, Kevin Duarte, Antoine Fraix, Nicolas Piriou, et al.

MACHINE LEARNING-BASED PHENOGROUPING IN MITRAL VALVE PROLAPSE IDENTIFIES PROFILES ASSOCIATED WITH MYOCARDIAL FIBROSIS AND CARDIOVASCULAR EVENTS

Olivier Huttin[a], MD; Nicolas Girerd[b]*, MD, PhD; Antoine Jobbe-Duval[c], MD; Anne-Laure Constant Dit Beaufils[c], MD; Thomas Senage[c,d], MD, PhD; Laura Filippetti[a], MD; Caroline Cueff[c,e], MD; Kevin Duarte[b], PhD; Antoine Fraix[a] MD, PhD, Nicolas Piriou[c], MD; Damien Mandry[a], MD; Nathalie Pace[a] MD; Solena Le Scouarnec[e], PhD; Romain Capoulade[e], PhD; Matthieu Echivard [a] MD; Jean Marc Sellal[a] MD; Marie Marrec[c] Marine Beaumont[f], PhD; Gabriella Hossu[f], PhD; Jean-Noel Trochu[c,e], MD, PhD; Nicolas Sadoul[a] MD, PhD; Pierre-Yves Marie[a], MD, PhD; Charles Guenancia, MD; Jean-Jacques Schott[e], PhD; Jean-Christian Roussel[c,e], MD, PhD; Jean-Michel Serfaty[e], MD, PhD; Christine Selton-Suty[a], MD*; Thierry Le Tourneau[c,e], MD, PhD

a. Service de Cardiologie, Institut Lorrain du Cœur et des Vaisseaux, CHU de Nancy, Nancy, France

b. Université de Lorraine, INSERM, Centre d'Investigations Cliniques-1433, and INSERM U1116; CHRU Nancy; F-CRIN INI-CRCT, (Cardiovascular and Renal Clinical Trialists), Nancy, France.

c. CHU Nantes, Université de Nantes, l'Institut du thorax, CIC 1413, F-44000 Nantes, France;

e. Université de Nantes, CHU de Nantes, CNRS, INSERM, l'Institut du thorax, F-44000 Nantes, France;

f. CIC-IT, U1433, CHRU de Nancy, France ; INSERM U1254, IADI, Université de Lorraine, Nancy, France

Funding

Fondation Cœur et Recherche (TLT, 2013, Paris, France), French Ministry of Health "PHRC-I 2012" (TLT, API12/N/019, Paris, France). T Le Tourneau was supported by an INSERM

Translational Research Grant (2012-2016, Paris, France). The STAMP study (OH, CSS) was supported by a grant from the French Ministry of Health (APJ 2015, n°: 2016-A00954-47). Pr Girerd is supported by the French National Research Agency Fighting Heart Failure (ANR-15-RHU-0004), by the French PIA project Lorraine Université d'Excellence GEENAGE (ANR-15-IDEX-04-LUE) programs, and the Contrat de Plan Etat Région Lorraine and FEDER IT2MP

Disclosures:

Pr Girerd is supported by the French National Research Agency Fighting Heart Failure (ANR-15-RHU-0004), by the French PIA project Lorraine Université d'Excellence GEENAGE (ANR-15-IDEX-04-LUE) programs, and the Contrat de Plan Etat Région Lorraine and FEDER IT2MP and reports honoraria from Novartis, AstraZeneka, Boehringer and Vifor.

Address for correspondence:

Pr. Olivier Huttin

Département de Cardiologie,

CHRU Nancy, Site Brabois,

Rue du Morvan,

54500 Vandoeuvre,

France

Acknowledgments:.

We would like to thank the staff of the CIC-IT and the staff of the MRI service of the "CHRU de Nancy-Brabois", for their daily work in both the service of patients and science. We would especially like to thank Fabienne ANTOINE, Agnès BASILE, Emilien MICARD, and Claire BANASIAK for their support.

ABSTRACT

BACKGROUND

Structural changes and myocardial fibrosis quantification by cardiac imaging have become increasingly important to predict cardiovascular events in patients with mitral valve prolapse (MVP). In this setting, it is likely that an unsupervised approach, using machine learning, may improve their risk assessment.

OBJECTIVES

This study used machine learning to improve the risk assessment of patients with MVP by identifying echocardiographic phenotypes and their respective association with myocardial fibrosis and prognosis.

METHODS

Clusters were constructed using echocardiographic variables in a bicentric cohort of patients with MVP (n=429 patients; 54±15 years) and subsequently investigated for their association with myocardial fibrosis (assessed by CMR imaging) and cardiovascular outcomes.

RESULTS

Mitral regurgitation (MR) was severe in 195 (45%) patients. Four clusters were identified: Cluster 1 comprised no remodeling with mainly mild MR; Cluster 2 was a transitional cluster; Cluster 3 included significant LV and left atrial remodeling with severe MR; Cluster 4 included remodeling with a drop in LV systolic strain. Clusters 3 and 4 featured more myocardial fibrosis than clusters 1 and 2 ($p<0.0001$) and were associated with higher rates of cardiovascular events. Cluster analysis significantly improved diagnostic accuracy over conventional analysis. The decision tree identified severity of MR along with LV systolic strain < 21% and indexed LA volume > 42ml/m² as the 3 most relevant variables to correctly classify participants into one of the echocardiographic profiles.

CONCLUSION

Clustering enabled to identify 4 clusters with distinct echocardiographic LV and LA remodeling profiles associated with myocardial fibrosis and clinical outcomes. Our findings suggest that a simple algorithm based on only 3 key variables (severity of MR, LV systolic strain and indexed LA volume) may help risk stratification and decision-making in patients with MVP.

Clinical Trial: NCT03884426 and NCT02879825

Key words : Mitral valve prolapse , Mitral regurgitation , machine learning , prognosis value , myocardial fibrosis , echocardiography , cardiac magnetic resonance imaging

Abbreviations :

MVP : Mitral valve prolapse MR : mitral regurgitation

LV: left ventricular

LA : left atrial

GLS: Global longitudinal strain

EF: ejection fraction

LGE : late gadolinium enhancement

MAD: Mitral Annular disjunction

PASP : Pulmonary artery systolic pressure

PVC: premature ventricular contraction

CMR: Cardiac Magnetic resonance imaging

CV : Cardio-Vascular

HF : Heart failure

Introduction

Mitral valve prolapse (MVP) is a common echocardiographic finding with an overall benign course (1,2) before progression to primary mitral regurgitation (MR).

The onset of MR-related symptoms is considered a definite indication for mitral valve surgery. On the other hand, indication for surgery in asymptomatic patients is mainly driven by echocardiographic parameters (3). However, this approach assesses each parameter individually and has only been validated on retrospective evolution of patients with MR. Other imaging biomarkers have been reported to improve prediction of cardiovascular outcome and optimize timing of surgery (4,5).

Chronic volume overload leads to progressive LV remodeling with extracellular matrix expansion and myocardial fibrosis. The evaluation of structural change and fibrosis quantification by CMR has become increasingly important in MVP (7). In addition, the degree of remodeling does not parallel MR severity in some patients, and these variant features might be associated with adverse outcomes (8,9), suggesting that the prognostic assessment of MVP could be optimized. .

In this setting, it is likely that an unsupervised approach, using machine learning, may help subdivide patients with heterogeneous into homogeneous phenotypes However, no study to date has taken advantage of the capacities of artificial intelligence to construct imaging phenotypes relevant to MVP pathophysiology.

Hence, we aimed this to identify echocardiographic phenotypes in MVP and investigate their association with myocardial fibrosis and cardiovascular events. We hypothesized that 1) cluster analysis - while removing *a priori* considerations - can identify relevant subgroups based on echocardiographic parameters; and 2) such echocardiographic phenotypes are associated with different levels of fibrosis and long-term cardiovascular events.

Methods

The present study gathered 429 MVP patients with trace to severe MR enrolled either in a MVP genetic study (NCT03884426) in Nantes (n=277) or in the MVP STAMP study (NCT02879825) in Nancy (n=152), France. Patients were enrolled between June 2010 and January 2020 (Flowchart, **Supplemental Figure 1**). All patients underwent comprehensive echocardiography and CMR to assess left and right ventricular volumes and ejection fraction (EF), as well as myocardial replacement fibrosis by late gadolinium enhancement (LGE). Patients with previous cardiac surgery, myocardial infarction or symptoms of coronary artery disease, with non-ischemic cardiomyopathy, congenital cardiac disease other than isolated bicuspid aortic valve, or with more than mild aortic regurgitation or stenosis were not included in the study. Patients with coronary artery disease diagnosed by angio-coronarography during the pre-operative work-up without myocardial ischemia during the ischemic test were included. An additional cohort of 133 MVP patients was considered as a replication group for validation. Both MVP genetic and MVP STAMP studies are in compliance with the Declaration of Helsinki and were approved by the local ethics committees. Written informed consent was obtained from all participants.

Echocardiography

Echocardiograms were carried out in the 2 centers by experienced investigators according to current recommendations (10-12) using a commercially available ultrasound system (Vivid 7, 9 or E95, GE Medical Systems, Milwaukee, WI, USA) and were all reviewed on an ECHOPAC workstation (version 110.1.0, GE Healthcare). All chamber volumes and mass measurements and myocardial deformation (Global LV systolic strain (%)) were performed according to current guidelines and indexed to body surface area (BSA)(6)

In addition, mitral valve (MV) apparatus was comprehensively assessed as previously described (13). Mitral regurgitation (MR) was graded according to recent guidelines (11) by at least 2 methods (PISA, volumetric method, continuity equation) and classified as trace-mild

(regurgitant volume ≤ 15 mL), moderate (16-59 mL) or severe (≥ 60 mL).

Cardiovascular magnetic resonance (CMR) imaging

CMR image acquisition was performed by either 1.5 or 3 Tesla systems (Siemens Medical Solutions, Erlangen, Germany) in 2, 3 and 4 chambers in both longitudinal and small axis planes using an ECG-gated cine steady state free precession sequence (SSFP) for each exam. Fifteen minutes after intravenous administration of gadolinium, delayed contrast-enhanced images were acquired using inversion-recovery fast gradient-echo pulse sequences into long axis and short axis views. Left and right ventricular volumes and EF, MR volume and grade, and the presence and location of LGE (considered if ≥ one segment) were collected. Post-processing analysis was performed using SyngoVIA software (version VB40B, Siemens Healthcare GmbH, Erlangen, Germany).

Follow-up

Clinical investigation focused on MVP-related symptoms (chest pain, palpitations, lightheadedness or syncope) and functional status. Presence of premature ventricular contraction (PVC) was based on 24-hour ECG Holter recordings.

Patient follow-up was conducted by clinical examination, medical reports, and phone contact with patients, relatives or general practitioner. The primary endpoint was a composite of cardiac death, heart failure, new-onset atrial fibrillation, and arterial embolism, whether occurring before or after MV surgery

Machine learning approach: cluster analysis and decision tree construction.

The Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME) guidelines (12) checklist is shown in Supplemental Table 7.

For the machine learning approach, a model-based clustering using latent class model (LCM) was performed based on several available echocardiographic data, e.g. LV mass index, LV end-diastolic volume index, LA volume index (LAVi), E/A ratio, pulmonary artery systolic pressure (PASP), MR severity, mitral annulus disjunction (**Figure 1, supplemental**

Figure 2). To simplify the classification algorithm, a decision tree was then performed to predict the clusters obtained from LCM clustering (Figure 4). More details about model-based clustering using LCM approach and decision tree are provided in Supplemental Material.

Statistical analysis

Categorical variables are described as frequencies (percentages) and continuous variables as means ± standard deviation and median (25^{th} and 75^{th} percentiles). Comparisons across echocardiographic phenotypes were carried out using the Kruskal-Wallis test for continuous variables and χ^2 or Fisher's exact test for categorical variables.

Association of the clusters with baseline myocardial fibrosis evaluated by CMR (LGE) was assessed using the logistic regression model and with CV outcomes using a Cox model. CV event-free survival curves for clusters were plotted using the Kaplan-Meier method and compared with the log-rank test. In addition, C-Statistic, integrated discrimination improvement (IDI), and the net reclassification improvement (NRI) were performed to assess the additional value of cluster stratification in predicting CV outcomes as compared to a baseline model including parameters previously reported to be independently associated with CV outcomes in MR: age, sex, LV systolic diameter (LVESD) > 40 mm, LAVi > 60mL/m² and PASP > 60mmHg.

In the validation cohort, clusters were built from the decision tree and the association of clusters with endpoints was assessed and compared to a LASSO cox analysis including the same variables used in the cluster analysis (see Supplemental Material for more details on LASSO Cox analysis).

All statistical analyses were performed using R version 3.4.0 (R Development Core Team, Vienna, Austria). A p value < 0.05 was considered statistically significant.

Results

Study population characteristics

Baseline characteristics of the 429 MVP patients (54±15 years, 57% male) are

summarized in Table 1. Twenty-five percent of the participants had hypertension, 3% had diabetes mellitus and 9% were NYHA class III-IV. MR was severe in 195 (45.5%) patients. Mean LVEF was 65.5±7.9 % with a mean GLS 21.5± 3.1%.

The prevalence of LGE+ was 29% in the myocardial wall and 7% in papillary muscles. Focal fibrosis was mainly located in the midwall layer and occasionally extended to the subendocardial layer.

Clusters derived from echocardiographic data

Our cluster analysis identified 4 echocardiographic phenotypes groups (**Table 2 and Figure 1**):

- Cluster 1 (N=131) mainly included patients with mild MR without LA or LV remodeling, preserved LVEF and normal LV systolic strain. This cluster encompassed the youngest participants with less cardiovascular risk factors and a higher proportion of patients with mitral annular disjunction (MAD). This cluster is labeled in the subsequent text as the **"Mild MR"** cluster.
- Cluster 2 (N=104) had echocardiographic characteristics that were in-between cluster 1 and clusters 3/4. This cluster is an intermediate group with moderate MR, preserved LV function but mild LV and LA remodeling. This cluster is labeled in the subsequent text as the **"Moderate MR with LA remodeling"** cluster.
- Cluster 3 (N=120) predominantly included males with severe MR. These patients had the highest (absolute) LV systolic strain and preserved LVEF. This cluster had increased LV mass, LV volumes and LA volumes, elevated PASP but normal RV function. This cluster is labeled in the subsequent text as the **"Severe MR with LV/LA remodeling**" cluster.
- Cluster 4 (N=74) included patients of similar mean age to cluster 3, who had the highest E wave and E/e' ratio. MR was severe and frequently due to flail leaflet (84.3%), usually without MAD (16.2 %). LV and LA remodeling were similar to cluster 3 but with lower LV longitudinal systolic strain values and cardiac output and higher PASP. This cluster is

labeled in the subsequent text as the **"Severe MR with LV longitudinal systolic dysfunction"** cluster.

Association of echocardiographic phenotypes with fibrosis and clinical outcomes

1/ Myocardial fibrosis

When considering cluster 1 as reference, clusters 2, 3 and 4 were associated with higher rates of myocardial fibrosis [crude odds ratio (OR) (95%CI)=3.55 (1.87 - 6.74) p=0.0001 for cluster 3, and 5.25 (2.61 - 10.56), p<0.0001 for cluster 4] (**Supplemental Table 2**). Nearly half of the patients in cluster 4 had myocardial fibrosis (n = 30 (44.1%) p<0.0001). After adjustment for age and sex, cluster 4 still exhibited a 4-fold higher rate of myocardial fibrosis [adjusted-OR (95%CI)=4.24 (1.86 - 9.63), p=0.0006] compared to cluster 1.

2/ Cardiovascular outcome

During a mean follow-up of 2.1 ± 1.9 years, CV events (n=92) included 8 cardiovascular deaths (4 related to VT, 3 to HF, 1 to IE), 49 worsening heart failure status (among which 4 had concomitant SV arrhythmia, one IE, one stroke and one VT), 26 de novo atrial fibrillation/SVT (4 of which the precipited HF, one in the setting of IE and one concomitant to VT) and 9 arterial embolisms/strokes (2 in the context of SVT, 2 in the context of IE). Regarding surgery, 43 occurred before and 49 after operation. When considering cluster 1 as reference, clusters 3 (severe MR with LV/LA remodeling) and 4 (severe MR with LV longitudinal systolic dysfunction) were significantly associated with higher rates of CV events [crude hazard ratio (HR) (95%CI)=5.34 (2.25 - 12.69), p=0.0002 for cluster 3 and =7.88 (3.27 - 18.96), p<0.0001 for cluster 4] (table 3 and **Supplemental Table 2**). This association remained highly significant after adjustment for age and sex [crude-HR (95%CI)=5.17 (2.07 - 12.93), p=0.0004 in cluster 3 and =6.86 (2.56 - 18.37), p=0.0001 in cluster 4].(**Figure 2**)

When adding clusters to the baseline model, no clinical or echocardiographic parameters remained significantly associated with outcome, whereas clusters 3 and 4 were

strongly associated with CV outcome [crude-OR for cluster 4 (95%CI)= 13.55 (2.36 - 77.70) p = 0.003). Adding clusters significantly improved the accuracy for outcome prediction on top of baseline parameters (NRI 0.384 [0.127 to 0.641], p = 0.003; AUROC 0.74 vs. 0.70, p=0.01) (**Figure 3**). When excluding patients with atrial fibrillation, results remained unchanged and clusters were still significantly associated with outcome (sensitivity analysis, data not shown)

The sensitivity analysis uncensored to but adjusted on MV surgery considered as binary time-dependent variable provided an HR similar to the previous analysis. In addition, the association of clusters with outcome remained significant when considering only pre-surgery events (**Supplemental Table 3).**

Decision tree for identifying echocardiographic phenotypes

The decision tree identified severity of MR with LV systolic strain (threshold value 21%) and indexed LA volume (threshold value 42ml/m²) as the three most relevant variables to correctly classify participants into one of the four echocardiographic profiles (**Figure 4**). When excluding MR grade as a candidate variable for the construction of the decision tree, the decision algorithm was slightly modified with regurgitant volume considered as a relevant variable to discriminate the clusters. The other 2 cardinal variables in the algorithm remained LAVi and GLS (**Supplemental Figure 3**).

Validation in the replication cohort

In the replication cohort, Clusters built according to the decision tree remained significantly associated with CV outcomes: Considering cluster 1 as reference, cluster 3/4 were significantly associated with CV outcomes with HR within the same range than the one observed in the derivation cohort [crude-HR (95%CI)= 5.21 (1.23 - 22.04), p; 0.025]. When we used Cox models build in the derivation to predict the risk of CV events in the replication cohort, the cluster based c-index values were higher as compared to pooled LASSO Cox model **(Supplemental Tables 4, 5 and 6)**.

Discussion

In the present analysis, in an unbiased/without *a priori* approach, we were able to identify 4 MVP echocardiographic homogeneous phenotypes despite the heterogeneity of patients with MVP using cluster analysis. The two clusters with the highest risks of CV events exhibited the most abnormal echocardiographic profiles with significant LV and LA remodeling without (cluster 3) or with (cluster 4) LV longitudinal dysfunction. These 2 clusters had also a higher burden of myocardial fibrosis. Our findings suggest that a simple algorithm based on "classical" variables such as MR severity, LA remodeling and LV strain identifies distinct individual profiles with different risks of MVP-related outcomes.

MVP echocardiographic phenotypes and outcomes: limits of established echocardiographic parameters

Appropriate timing for surgery and risk stratification in patients with severe primary MR still remains a challenge. Among established risk factors, LV remodeling, LA dilatation and elevated pulmonary pressure significantly differed across the 4 clusters, suggesting that these variables were indeed key features of heterogeneity in our MVP population. However, despite the fact that these echocardiographic variables were pivotal for clustering, each of these variables taken individually has previously been reported not to be sufficiently specific to clearly identify high-risk patients (15–17). This was also the case in our baseline model where only LA volume was associated with outcomes. When adding clusters to the baseline model, the clinical and echocardiographic parameters were no longer associated, whereas cluster association with CV outcome was almost threefold higher. Such an approach already proved its interest in asymptomatic middle-aged individuals and allowed to build a decision tree to predict overt heart failure among a multitude of highly correlated variables. (18).

Interestingly, LA and LV remodeling mainly discriminated clusters 1 to 3. However, focusing only on the LV and LA remodeling pattern does not allow discriminating between clusters 3 and 4 (**Figure 1**). Cluster 4, which exhibited significantly worse CV outcomes,

differed primarily by a greater impairment in systolic function. However, CV events were recorded in as high as 20%bof patients in cluster 2, although they were relatively young and, had initially moderate MR without LV remodeling. These result highlight the need of a better characterization and possibly targeted follow-up of MVP patients.

Thus, among this large cohort of patients with MVP, our findings show that while LA and LV dimensions are able to differentiate homogeneous groups of MVP patients, the sole use of these variables is not sufficient at more advanced stages of the disease process.

The key importance of global systolic function

The present study shows that systolic function, assessed by GLS, is as an important factor of risk stratification in MVP. Indeed, the main difference between clusters 3 and 4 was the decrease in longitudinal strain in cluster 4 with a cut-off value of 21% as defined by the derived decision tree (**Figure 4**). LVEF was not sufficiently sensitive to stratify the risk. Only a few studies have shown the association of GLS with outcome after surgery (19,20). In our cohort, the cluster 4 phenotype was associated with the highest proportion of CV events, which emphasizes the key importance of assessing systolic function in risk-stratification of MVP patients, consistently with previous reports (19,21). Of note, the GLS cut-off value (21%) was similar to the one proposed in studies based only on a prognostic approach (4). This threshold is higher than the lower limit of normal, which indicates that strain values that are usually perceived as normal in MR are actually already associated with increased CV outcomes. So, values lower than 21%should be considered as abnormal and associated with worse prognosis and myocardial fibrosis in MVP with MR (22). The prognostic data on GLS are in line with and support the clinical relevance of the stratification proposed in the current study. Furthermore, our study specifically showed a biphasic evolution of GLS according to cluster severity with an increase in GLS values from cluster 1 to cluster 3 and a decrease in cluster 4. In the late stage of MVP (cluster 4), longitudinal dysfunction develops despite a preserved LVEF. This reinforces the value of GLS as a more sensitive and accurate

measurement of LV function (5) than LVEF. Indeed, current guidelines underscore its potential incremental value over LVEF for risk stratification in patients with severe primary MR (23).

The results of the present study thus suggest that a systematic analysis of myocardial deformation in MVP would appear mandatory in routine practice to help identify patients with severe MR-induced ventricular remodeling and subsequent myocardial scarring (17,24,25), in whom the use of CMR for detection of myocardial fibrosis may be relevant. Furthermore, from a clinical standpoint, our results suggest intervening prior to the decrease in strain given the very high risk observed in cluster 4.

Need of stratification of patients with less than moderate MR

There is growing evidence that moderate MR and/or disproportionate LV remodeling according to MR severity is associated with adverse outcomes (8,9). Echocardiography remains the first-line imaging modality for MR assessment, especially in the setting of moderate MR (3).

The degree of MR severity in the present analysis increased from cluster 1 to 4. Our results moreover identified a pattern of patients without severe MR but associated with an intermediate prognosis (cluster 2, 3-fold higher risk of cardiovascular outcome compared to cluster 1). This cluster can be considered as the transitional state in which LV and LA remodeling are in progression. The discrimination of this transitional state from a low-risk situation (cluster 1) is obtained solely using LA volume (cut-off of 42 ml/m²) in our decision tree. This concept of identifying a subgroup of patients with moderate MR severity but poor prognosis has previously been studied (26). Accordingly, Mantovani et al. highlighted the value of LA diameter for differentiating patients with moderate MR and a higher risk of CV events, which is consistent with the algorithm proposed herein, placing emphasis on LA volume.

Patients with moderate primary MR have been demonstrated to have a significantly

higher risk of CV events comparatively to patients with no or mild MR (27)(28)(29). Patients with less than moderate MR due to MVP have already been shown to potentially exhibit early LV and LA remodeling, which neither predicts MR progression nor mortality. In case of discordances between LV/LA remodeling and MR severity, female sex, together with premature ventricular contractions were considered as potential confounding factors (27,30). It was recently speculated that this latter discordance could be part of the phenotypic expression of a genetically mediated process, but also related to MV leaflet interaction or to asymptomatic atrial fibrillation (29).

The difference in outcome among the various grades of MR in the present study is beyond the scope of this investigation. A recent report using cardiac MRI enabled addressing this concept of disproportionate remodeling in patients with non-severe MR (9). Our current findings confirm the role of a MVP outcome effect independently of the effect exerted by regurgitant volume on the left heart chambers. Fibrosis in the setting of significant MR may rather be associated with MR-induced ventricular remodelling, neuro-humoral activation and/or induced by recurring stretch exerted by the prolapsing leaflet on the myocardium (15,20,21). Further studies with longitudinal follow-up will enhance our understanding of myocardial fibrosis and LV remodeling genesis in patients without significant MR.

Progressive myocardial fibrosis continuum and clinical perspective

Cardiac magnetic resonance imaging is considered as the gold standard for assessing cardiac remodeling and is further able to detect myocardial fibrosis scars, a classical consequence of maladaptive ventricular remodeling (31). In our recent companion publication, LGE was observed in 25-30% of patients with MVP (32) and appeared as a landmark of MVP-CM, being mainly localized in papillary muscles and in the LV inferolateral basal wall. LGE is attributed to chronic maladaptive LV remodeling and abnormal constraints exerted on the myocardial wall and papillary muscle. Previous reports suggest that regional replacement fibrosis is clinically relevant in patients with MVP to

further identify patients who are at high risk of cardiac events (32). Accordingly, CMR-based LGE diagnosis of myocardial fibrosis gradually increased from 13% in cluster 1 to 44% in cluster 4 in the present study. Indeed, the increasing proportion of myocardial fibrosis concomitant with the parallel increasing severity of the clusters provides the pathophysiological scaffold to our stratification based on a non-*a priori* approach.

Data-driven echocardiographic phenotypes may assist clinicians in stratifying the level of risk of MVP patients, with increasing levels of cardiovascular remodeling and cardiac fibrosis and subsequent higher risk of CV events. The good prognostic performance of this approach suggests its potential helpfulness in determining the follow-up strategy and possibly the therapeutic management to be applied in these patients. Importantly, the value of the clustering approach was here superior to the LASSO approach as its prediction performance was higher in the validation cohort.

For instance, rapid advances in new percutaneous techniques of mitral valve repair, raised the issue of optimal timing of intervention, specifically for patients with moderate or asymptomatic severe MR. However, whether the use of this decision tree to guide MVP management can ultimately improve clinical outcome will need to be tested in clinical trials **(Figure 5)**.

Study Limitations

While our clustering approach enabled overcoming key limitations in clinical practice in assessing MR degree and symptomatic status, our study nevertheless featured several limitations. The algorithm integrated GLS measurements with potential bias among vendors (35). Regarding CMR, MR quantification using CMR could have helped to further characterize MVP and better phenotype MVP. Such MR quantification was however not available in the current analysis. On the same manner, T1 mapping allowing a quantitative assessment of the interstitial component of cardiac fibrosis could have been interesting to better characterize myocardial but it was not performed systematically in our patients. Lastly,

we did not systematically assess diffuse LV myocardial interstitial fibrosis in our cohort of patients. Larger multicenter cohorts would provide more robust validation of our phenotyping algorithm.

CONCLUSION

In this large series of patients with MVP and various degrees of MR, our constructed clusters were able to predict the presence of myocardial fibrosis and clinical outcome. The four identified echocardiographic phenotypes confirmed that the interplay of GLS and remodeling are the main predictors of the risk of CV events in MVP, irrespective of the level of MR. Prognostic implications of cluster-guided therapeutic strategy require further evaluation. Notwithstanding, this simple algorithm may be of particular interest in the identification of high-risk MVP patients with moderate MR or severe asymptomatic MR who require close follow-up.

CLINICAL PERSPECTIVES

- Clustering enabled to identify clusters with distinct echocardiographic remodeling profiles associated with myocardial fibrosis and clinical outcomes.
- Our findings suggest that a simple algorithm based on only 3 key variables (severity of mitral regurgitation, left ventricular systolic strain and indexed left atrial remodeling) may help risk stratification in patients with MVP.
- This strategy may be of particular interest for decision-making in MVP patients with moderate MR or severe asymptomatic MR

TRANSLATIONAL OUTLOOK

This study shows that machine learning using echocardiographic data can be used to decipher the phenotypes of patients in a given disease area. Herein, we show that machine learning can 1) identify homogenous phenotypes, 2) provide a simple algorithm to predict the membership to these phenotypes. It also shows that this phenotyping is strongly associated with outcome. In MVP, this machine learning derived phenotyping performs better than and improve prognostic accuracy on top of usual risk-stratifying approaches. Our study consequently appears as a promising example of the potential of machine-learning fueled by imaging to better categorize CV disease and provide ready-to-use algorithm that could be used in routine clinical practice.

References

1. Nishimura RA, McGoon MD, Shub C, Miller FA, Ilstrup DM, Tajik AJ. Echocardiographically documented mitral-valve prolapse. Long-term follow-up of 237 patients. N Engl J Med. 1985;313(21):1305–9.
2. Düren DR, Becker AE, Dunning AJ. Long-term follow-up of idiopathic mitral valve prolapse in 300 patients: a prospective study. J Am Coll Cardiol. 1988;11(1):42–7.
3. Ra N, Cm O, Ro B, Ba C, Jp E, La F, et al. 2017 AHA/ACC Focused Update of the 2014 AHA/ACC Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines [Internet]. Journal of the American College of Cardiology. 2017 [cited 2020 Dec 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/28315732/
4. Prognostic implications of global LV dysfunction: a systematic review and meta-analysis of global longitudinal strain and ejection fraction - PubMed [Internet]. [cited 2020 Dec 14]. Available from: https://pubmed.ncbi.nlm.nih.gov/24860005/
5. Penicka M, Vecera J, Mirica DC, Kotrc M, Kockova R, Van Camp G. Prognostic Implications of Magnetic Resonance-Derived Quantification in Asymptomatic Patients With Organic Mitral Regurgitation: Comparison With Doppler Echocardiography-Derived Integrative Approach. Circulation. 2018 27;137(13):1349–60.
6. Prediction of left ventricular ejection fraction 6 months after surgical correction of organic mitral regurgitation: the value of exercise echocardiography and deformation imaging - PubMed [Internet]. [cited 2020 Dec 14]. Available from: https://pubmed.ncbi.nlm.nih.gov/22504944/
7. D K, F N, Rj K, Ro B, Ma K, J X, et al. Myocardial Fibrosis in Patients With Primary Mitral Regurgitation With and Without Prolapse [Internet]. Journal of the American College of Cardiology. 2018 [cited 2020 Dec 13]. Available from:

https://pubmed.ncbi.nlm.nih.gov/30115220/

8. Lt Y, Sw A, Z L, G B, R M, M T, et al. Mitral Valve Prolapse Patients with Less than Moderate Mitral Regurgitation Exhibit Early Cardiac Chamber Remodeling [Internet]. Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography. 2020 [cited 2020 Dec 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/32222479/
9. El-Tallawi KC, Kitkungvan D, Xu J, Cristini V, Yang EY, Quinones MA, et al. Resolving the Disproportionate Left Ventricular Enlargement in Mitral Valve Prolapse Due to Barlow Disease: Insights From Cardiovascular Magnetic Resonance. JACC Cardiovasc Imaging [Internet]. 2020 Oct 29 [cited 2020 Dec 13]; Available from: http://www.sciencedirect.com/science/article/pii/S1936878X20308251
10. S K, Y L, T K, S Y, Ic H, Jb P, et al. Unsupervised Cluster Analysis of Patients With Aortic Stenosis Reveals Distinct Population With Different Phenotypes and Outcomes [Internet]. Circulation. Cardiovascular imaging. 2020 [cited 2020 Dec 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/32418453/
11. Zoghbi WA, Adams D, Bonow RO, Enriquez-Sarano M, Foster E, Grayburn PA, et al. Recommendations for Noninvasive Evaluation of Native Valvular Regurgitation. J Am Soc Echocardiogr. 2017;30(4):303–71.
12. Sengupta PP, Shrestha S, Berthon B, Messas E, Donal E, Tison GH, et al. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. JACC Cardiovasc Imaging. 2020 Sep;13(9):2017–35.
13. Marbac M, Sedki M. VarSelLCM: an R/C++ package for variable selection in model-based clustering of mixed-data with missing values. Bioinforma Oxf Engl. 2019 Apr 1;35(7):1255–7.
14. CRAN - Package rpart [Internet]. [cited 2022 May 21]. Available from: https://cran.r-

project.org/web/packages/rpart/index.html

15. Hong-TaoYuan, Yang M, Zhong L, Lee YH, Vaidya VR, Asirvatham SJ, et al. Ventricular premature contraction associated with mitral valve prolapse. Int J Cardiol. 2016;221:1144–9.
16. Chesler E, King RA, Edwards JE. The myxomatous mitral valve and sudden death. Circulation. 1983;67(3):632–9.
17. Vohra J, Sathe S, Warren R, Tatoulis J, Hunt D. Malignant ventricular arrhythmias in patients with mitral valve prolapse and mild mitral regurgitation. Pacing Clin Electrophysiol PACE. 1993;16(3):387–93.
18. Kobayashi M, Huttin O, Magnusson M, Ferreira JP, Bozec E, Huby AC, et al. Machine Learning-Derived Echocardiographic Phenotypes Predict Heart Failure Incidence in Asymptomatic Individuals. JACC Cardiovasc Imaging. 2021 Sep 8;S1936-878X(21)00556-8.
19. Alashi A, Mentias A, Abdallah A, Feng K, Gillinov AM, Rodriguez LL, et al. Incremental Prognostic Utility of Left Ventricular Global Longitudinal Strain in Asymptomatic Patients With Significant Chronic Aortic Regurgitation and Preserved Left Ventricular Ejection Fraction. JACC Cardiovasc Imaging. 2018;11(5):673–82.
20. Mentias A, Naji P, Gillinov AM, Rodriguez LL, Reed G, Mihaljevic T, et al. Strain Echocardiography and Functional Capacity in Asymptomatic Primary Mitral Regurgitation With Preserved Ejection Fraction. J Am Coll Cardiol. 2016 01;68(18):1974–86.
21. Hiemstra YL, Tomsic A, van Wijngaarden SE, Palmen M, Klautz RJM, Bax JJ, et al. Prognostic Value of Global Longitudinal Strain and Etiology After Surgery for Primary Mitral Regurgitation. JACC Cardiovasc Imaging. 2020;13(2 Pt 2):577–85.
22. STretch and Myocardial Characterization in Arrythmogenic Mitral Valve Prolapse-2 - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 May 21]. Available from:

https://clinicaltrials.gov/ct2/show/NCT04852731

23. Joint Task Force on the Management of Valvular Heart Disease of the European Society of Cardiology (ESC), European Association for Cardio-Thoracic Surgery (EACTS), Vahanian A, Alfieri O, Andreotti F, Antunes MJ, et al. Guidelines on the management of valvular heart disease (version 2012). Eur Heart J. 2012 Oct;33(19):2451–96.
24. Kligfield P, Levy D, Devereux RB, Savage DD. Arrhythmias and sudden death in mitral valve prolapse. Am Heart J. 1987;113(5):1298–307.
25. Kitkungvan D, Nabi F, Kim RJ, Bonow RO, Khan MA, Xu J, et al. Myocardial Fibrosis in Patients With Primary Mitral Regurgitation With and Without Prolapse. J Am Coll Cardiol. 2018 Aug;72(8):823–34.
26. Mantovani F, Clavel MA, Michelena HI, Suri RM, Schaff HV, Enriquez-Sarano M. Comprehensive Imaging in Women With Organic Mitral Regurgitation: Implications for Clinical Outcome. JACC Cardiovasc Imaging. 2016 Apr;9(4):388–96.
27. Enriquez-Sarano M, Avierinos JF, Messika-Zeitoun D, Detaint D, Capps M, Nkomo V, et al. Quantitative determinants of the outcome of asymptomatic mitral regurgitation. N Engl J Med. 2005 Mar 3;352(9):875–83.
28. Hayashi H, Abe Y, Morita Y, Yamaji Y, Nakane E, Haruna Y, et al. Prognostic significance of moderate primary mitral regurgitation and concomitant paroxysmal atrial fibrillation. J Cardiol. 2020;75(3):309–14.
29. Yang LT, Ahn SW, Li Z, Benfari G, Mankad R, Takeuchi M, et al. Mitral Valve Prolapse Patients with Less than Moderate Mitral Regurgitation Exhibit Early Cardiac Chamber Remodeling. J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr. 2020 Jul;33(7):815-825.e2.
30. Yiginer O, Keser N, Ozmen N, Tokatli A, Kardesoglu E, Isilak Z, et al. Classic mitral valve prolapse causes enlargement in left ventricle even in the absence of significant mitral regurgitation. Echocardiogr Mt Kisco N. 2012 Feb;29(2):123–9.

31. Pradella S, Grazzini G, Brandani M, Calistri L, Nardi C, Mori F, et al. Cardiac magnetic resonance in patients with mitral valve prolapse: Focus on late gadolinium enhancement and T1 mapping. Eur Radiol. 2019 Mar;29(3):1546–54.
32. Constant Dit Beaufils AL, Huttin O, Jobbe-Duval A, Senage T, Filippetti L, Piriou N, et al. Replacement Myocardial Fibrosis in Patients with Mitral Valve Prolapse: Relation to Mitral Regurgitation, Ventricular Remodeling and Arrhythmia. Circulation. 2021 Mar 12;
33. Kitkungvan D, Yang EY, El Tallawi KC, Nagueh SF, Nabi F, Khan MA, et al. Extracellular Volume in Primary Mitral Regurgitation. JACC Cardiovasc Imaging. 2021 Jun;14(6):1146–60.
34. Weinsaft JW, Kim J. Beyond the Mitral Valve: Myocardial Fibrosis for Therapeutic Decision-Making and Prognostication of Degenerative Mitral Regurgitation. JACC Cardiovasc Imaging. 2021 Oct 7;S1936-878X(21)00630-6.
35. Ünlü S, Mirea O, Duchenne J, Pagourelias ED, Bézy S, Thomas JD, et al. Comparison of Feasibility, Accuracy, and Reproducibility of Layer-Specific Global Longitudinal Strain Measurements Among Five Different Vendors: A Report from the EACVI-ASE Strain Standardization Task Force. J Am Soc Echocardiogr Off Publ Am Soc Echocardiogr. 2018 Mar;31(3):374-380.e1.

Figure 1: **Echocardiographic phenotypes according to cluster.** Cluster analysis: Radar charts of echocardiographic phenotypes illustrating the differences in cardiac parameters across the 4 clusters. Values are standardized values expressed as z-scores (SD) from mean values, and both radar charts are constructed without max and min values of each variable. Abbreviations are provided in **Table 1**.

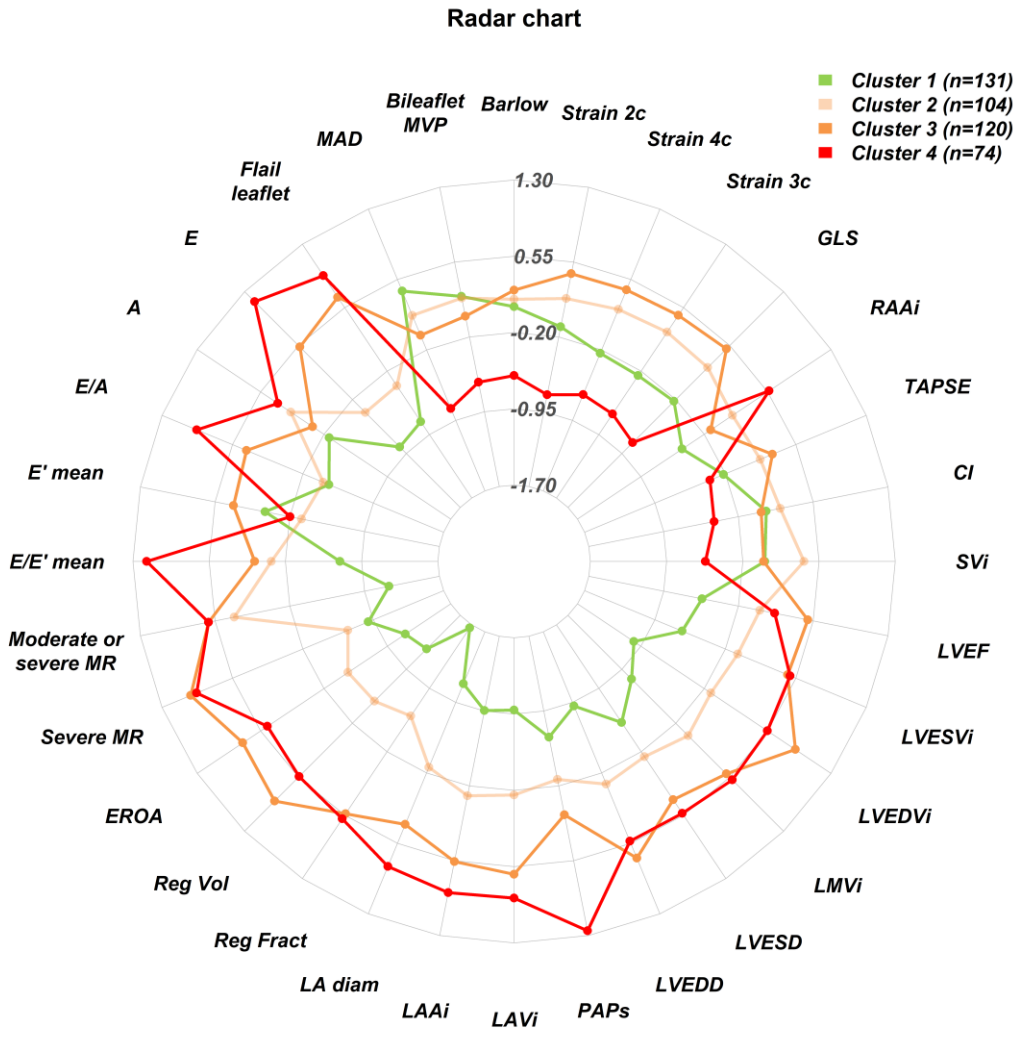


Figure 2: Incidence of cardiovascular events according to echocardiographic cluster.

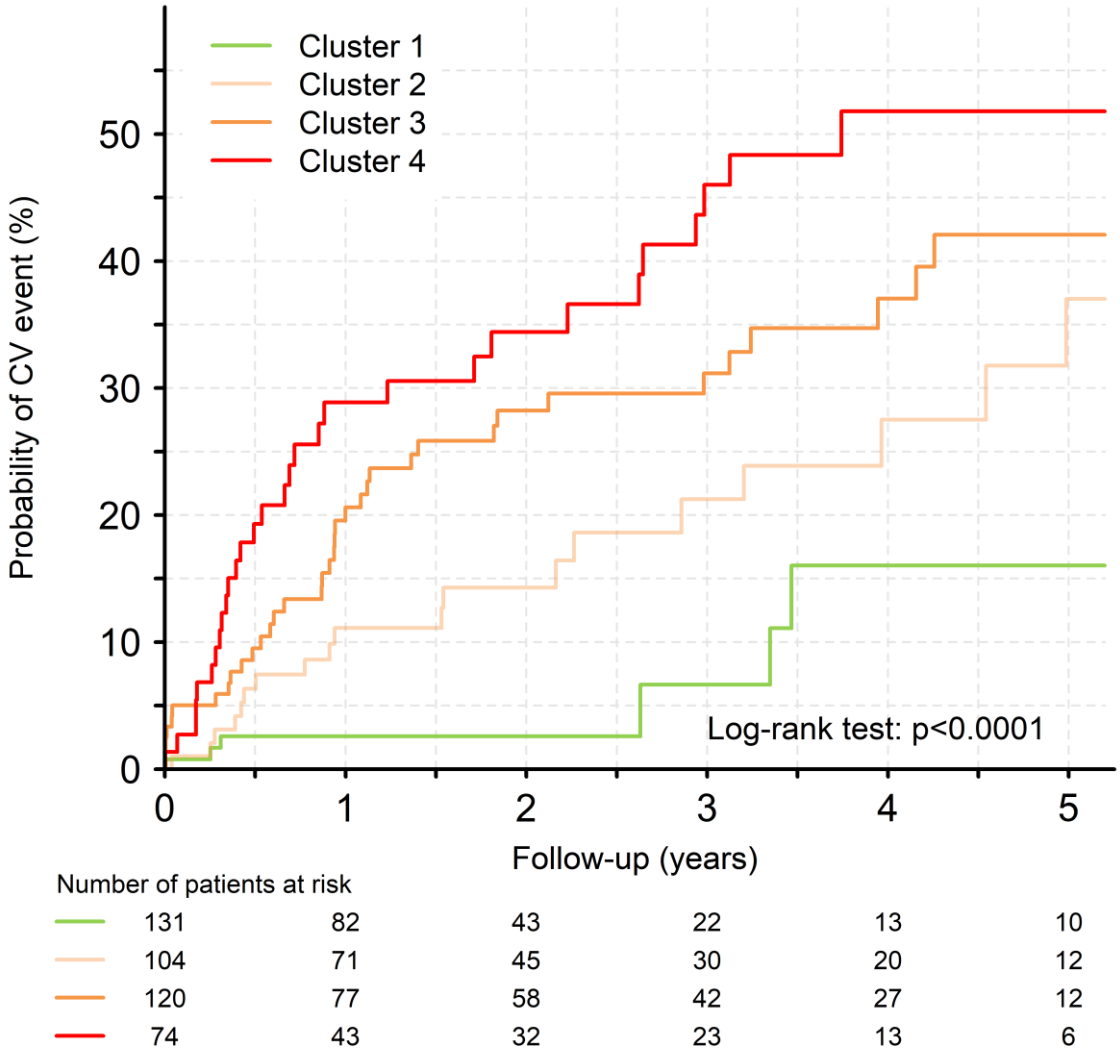
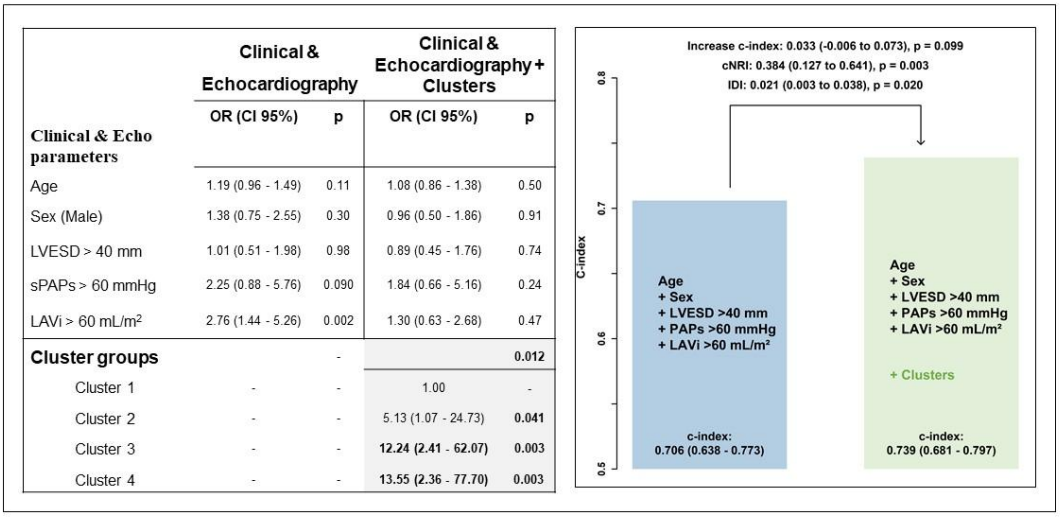
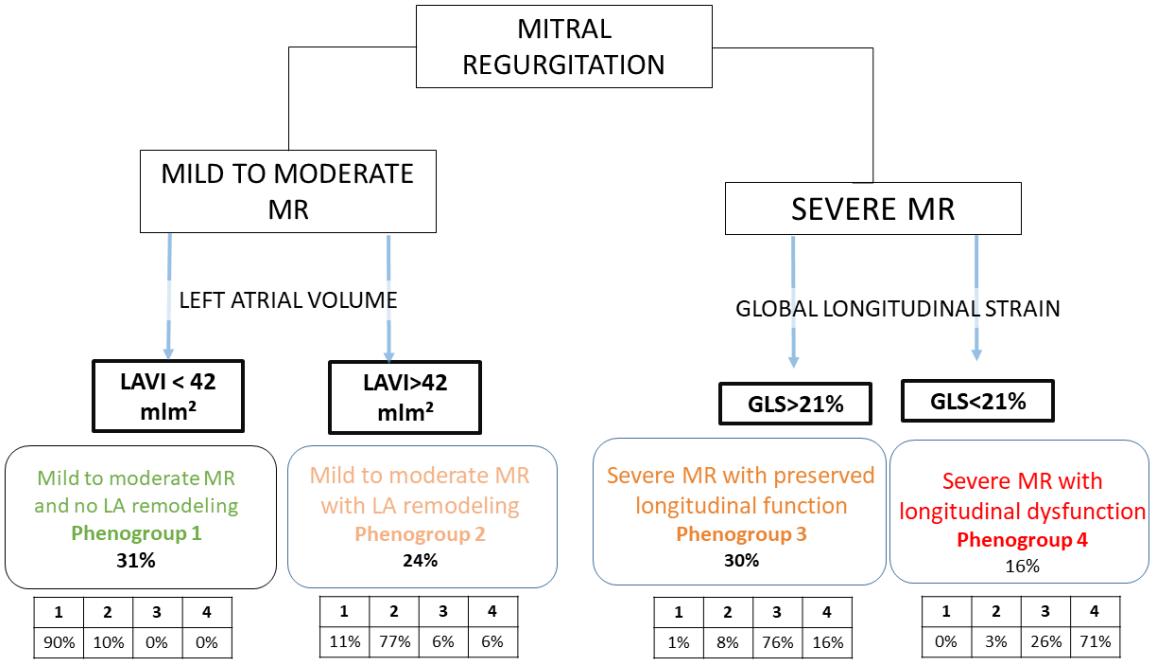


Figure 3. Additional value of cluster to predict CV outcomes : Additional value of cluster implementation on top of a baseline guideline model in predicting cardiovascular outcomes at two years.



Clinical & Echo parameters	Clinical & Echocardiography		Clinical & Echocardiography + Clusters	
	OR (CI 95%)	p	OR (CI 95%)	p
Age	1.19 (0.96 - 1.49)	0.11	1.08 (0.86 - 1.38)	0.50
Sex (Male)	1.38 (0.75 - 2.55)	0.30	0.96 (0.50 - 1.86)	0.91
LVESD > 40 mm	1.01 (0.51 - 1.98)	0.98	0.89 (0.45 - 1.76)	0.74
sPAPs > 60 mmHg	2.25 (0.88 - 5.76)	0.090	1.84 (0.66 - 5.16)	0.24
LAVi > 60 mL/m²	2.76 (1.44 - 5.26)	0.002	1.30 (0.63 - 2.68)	0.47
Cluster groups		-		**0.012**
Cluster 1	-	-	1.00	-
Cluster 2	-	-	5.13 (1.07 - 24.73)	**0.041**
Cluster 3	-	-	**12.24 (2.41 - 62.07)**	**0.003**
Cluster 4	-	-	**13.55 (2.36 - 77.70)**	**0.003**

Figure 4: **Echo phenotype based algorithm** : Decision trees depicting the identification of echocardiographic phenotypes using entry echocardiographic data (n = 429 patients).



1	2	3	4
90%	10%	0%	0%

1	2	3	4
11%	77%	6%	6%

1	2	3	4
1%	8%	76%	16%

1	2	3	4
0%	3%	26%	71%

Figure 5: Central Illustration : Framework of the machine-learning echocardiographic phenotypes in mitral valve prolapse.

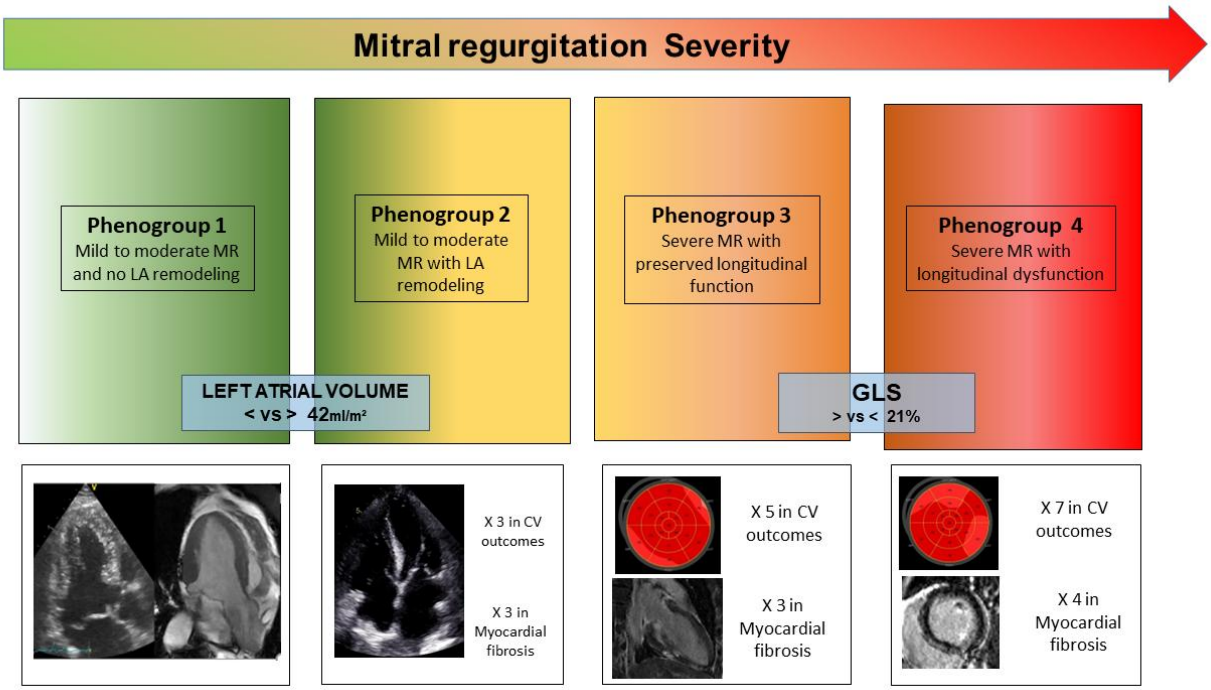


Table 1: Baseline characteristics of the study population according to echocardiographic phenotypes (n=429 patients).

Variables	Mean ± SD / n (%)	Median (Q1 - Q3)
Clinical characteristics		
Age at echocardiography (years)	53.8 ± 15.4	55.0 (43.0 - 65.0)
Male	245 (57.1 %)	
Heart rate (bpm)	69.0 ± 11.3	68.8 (61.8 - 76.0)
Systolic BP (mmHg)	130.8 ± 17.2	130.0 (120.0 - 141.0)
BMI (kg/m²)	23.4 ± 3.7	23.1 (20.7 - 25.5)
Tobacco use	99 (24.2 %)	
Hypertension	108 (25.2 %)	
Diabetes	14 (3.3 %)	
Dyslipidemia	68 (16.6 %)	
History of AF	56 (13.1 %)	
Symptoms		
Chest pain	42 (15.6 %)	
PVC	118 (43.2 %)	
Lypothymia/syncope	7 (5.8 %)	
NYHA class III-IV	39 (9.1 %)	
Events		
CV event	92 (21.4 %)	
Time to follow-up for CV event (years)	2.1 ± 1.9	1.6 (0.6 - 3.2)
Myocardial fibrosis (LGE ≥ one segments)	119 (28.8 %)	
Echocardiographic parameters		
Septal wall thickness at end-diastole (mm)	9.3 ± 2.3	9.1 (7.8 - 10.8)
LV end-diastolic dimension (mm)	55.4 ± 7.4	55.9 (50.0 - 60.1)
LV end-systolic dimension (mm)	35.6 ± 5.6	35.2 (31.6 - 39.0)
LV mass index (g/m²)	107.2 ± 33.9	104.0 (82.8 - 126.2)
Relative wall thickness	0.32 ± 0.09	0.31 (0.26 - 0.37)
LVEDV index (mL/m²)	85.8 ± 28.7	87.3 (61.4 - 106.6)
LVESV index (mL/m²)	28.9 ± 10.7	27.7 (20.8 - 35.0)
LVEF (%)	65.5 ± 7.9	66.3 (60.2 - 70.6)
Cardiac index (L/min/m²)	2.7 ± 0.7	2.6 (2.3 - 3.0)
GLS (%)	21.5 ± 3.1	21.6 (19.6 - 23.5)
RV diameter (mm)	30.0 ± 4.9	30.0 (26.8 - 32.5)
E wave (cm/s)	98.9 ± 35.4	95.0 (71.3 - 123.0)
E/A	1.63 ± 0.80	1.47 (1.12 - 1.91)
Deceleration Time (ms)	189.1 ± 59.1	184.4 (149.3 - 215.3)
E' mean (cm/s)	11.1 ± 3.2	11.0 (9.0 - 13.0)
E/E' lateral	9.3 ± 4.6	8.2 (6.0 - 11.2)
E/E' septal	10.8 ± 7.6	9.3 (6.9 - 12.8)
E/E' mean	9.6 ± 4.3	8.8 (6.4 - 11.5)
AI grade (2 classes)		

Variables	Mean ± SD / n (%)	Median (Q1 - Q3)
No trace	297 (78.8 %)	
Trace	80 (21.2 %)	
MR grade (3 classes)		
Trace-mild	118 (27.5 %)	
Moderate	116 (27.0 %)	
Severe	195 (45.5 %)	
EROA (cm²)	0.33 ± 0.20	0.33 (0.16 - 0.44)
Regurgitant volume (mL)	49.2 ± 30.8	53.0 (21.0 - 70.0)
Regurgitant fraction (%)	41.1 ± 17.2	45.1 (30.7 - 52.4)
LA diameter (mm)	43.1 ± 7.6	43.1 (38.0 - 48.8)
LA area (cm²)	26.4 ± 8.9	26.0 (19.9 - 31.9)
LAV index (mL/m²)	58.6 ± 28.0	55.3 (38.5 - 75.4)
TAPSE (mm)	24.2 ± 4.9	24.0 (20.8 - 27.7)
RA area (cm²)	18.3 ± 8.3	17.0 (13.8 - 20.8)
PAPs (mmHg)	33.4 ± 13.8	30.0 (23.6 - 37.4)
Barlow	273 (63.8 %)	
Anterior leaflet MVP	26 (6.1 %)	
Posterior leaflet MVP	179 (41.9 %)	
Bileaflet MVP	210 (49.2 %)	
Anterior leaflet length (mm)	26.7 ± 3.9	26.1 (24.1 - 29.0)
Posterior leaflet length (mm)	18.7 ± 4.0	18.9 (16.0 - 21.0)
Mitral annulus diameter (mm)	38.3 ± 5.7	38.1 (34.6 - 42.0)
Mitral annulus disjunction	220 (57.0 %)	
Flail leaflet	157 (39.0 %)	
CMR parameters		
LVEDV index (mL/m²)	105.4 ± 28.4	103.0 (84.2 - 123.2)
LVESV index (mL/m²)	41.6 ± 14.7	39.7 (31.6 - 48.7)
LVEF (%)	60.8 ± 6.9	60.4 (56.2 - 65.0)
Myocardial mass index (g/m²)	68.7 ± 17.5	68.8 (56.3 - 80.0)
RVEDV index (mL/m²)	81.9 ± 22.3	80.7 (68.0 - 94.0)
RVEDV index (mL/m²)	39.0 ± 14.3	37.1 (29.3 - 46.7)
RVEF (%)	52.9 ± 8.1	53.0 (47.8 - 58.0)
Regurgitant volume (mL)	41.8 ± 28.4	37.0 (20.1 - 60.0)
Regurgitant fraction (%)	33.4 ± 17.0	33.8 (22.3 - 45.8)

N: number of non-missing data; SD: standard deviation, Q1: 1st quartile; Q3: 3rd quartile.

AF, Atrial Fibrillation ; BP, blood pressure; BMI, body mass index; NYHA, New York Heart Association; Ant: anterior, EDD: end-diastolic diameter, ESD: end-systolic diameter, EROA: effective regurgitant orifice area, GLS: global longitudinal strain, IVS: inter-ventricular septum, LAVol: left atrium volume, LV EF: left ventricular ejection fraction, LA, Left atrium; PAP, Pulmonary artery Pressure; MAD: mitral annulus disjunction, MR: mitral regurgitation, MVP: mitral valve prolapse, PVC, Premature ventricular contraction; PASP: pulmonary artery systolic pressure, Post.: posterior, PW: posterior wall, Reg Vol: regurgitant volume, RV: right ventricle; RV Right ventricle; RA right atrium

Table 2: Baseline characteristics according to echocardiographic phenotypes and identified clusters.

Variables	Cluster 1 (n=131)	Cluster 2 (n=104)	Cluster 3 (n=120)	Cluster 4 (n=74)	p-value
Clinical characteristics					
Age at echocardiography (years)	44.0 (31.0 - 54.0)	56.0 (45.0 - 62.5)	56.0 (46.0 - 65.0)	68.0 (59.0 - 77.0)	<0.0001
Male	42 (32.1 %)	50 (48.1 %)	101 (84.2 %)	52 (70.3 %)	<0.0001
Heart rate (bpm)	69.4 (62.0 - 76.3)	66.0 (60.0 - 73.0)	67.0 (60.7 - 77.0)	70.0 (62.0 - 80.0)	0.085
Systolic BP (mmHg)	124.0 (111.0 - 136.0)	130.0 (120.0 - 146.0)	130.0 (123.0 - 141.0)	133.5 (122.0 - 141.0)	0.0005
BMI (kg/m²)	21.7 (19.8 - 24.3)	23.1 (20.7 - 25.3)	23.9 (21.4 - 25.7)	24.7 (22.2 - 27.7)	<0.0001
Tobacco use	18 (14.8 %)	23 (23.5 %)	30 (25.6 %)	28 (38.9 %)	0.002
Hypertension	11 (8.5 %)	23 (22.1 %)	34 (28.3 %)	40 (54.1 %)	<0.0001
Diabetes	4 (3.1 %)	2 (1.9 %)	2 (1.7 %)	6 (8.1 %)	0.11
Dyslipidemia	10 (8.2 %)	11 (11.2 %)	21 (17.9 %)	26 (36.1 %)	<0.0001
History of AF	3 (2.3 %)	12 (11.7 %)	17 (14.3 %)	24 (32.9 %)	<0.0001
Symptoms					
Chest pain	25 (23.4 %)	12 (20.0 %)	4 (6.7 %)	1 (2.4 %)	0.0009
PVC	61 (56.0 %)	24 (40.7 %)	23 (34.3 %)	10 (26.3 %)	0.003
Pre-syncope/syncope	2 (13.3 %)	4 (13.8 %)	1 (2.2 %)	0 (0.0 %)	0.027
NYHA class III-IV	2 (1.5 %)	9 (8.7 %)	6 (5.0 %)	22 (29.7 %)	<0.0001
Echocardiographic parameters					
Septal wall thickness at end-diastole (mm)	8.1 (7.0 - 9.7)	9.3 (7.3 - 11.0)	9.4 (8.4 - 10.6)	9.9 (8.3 - 11.5)	<0.0001
LV end-diastolic dimension (mm)	48.6 (45.1 - 52.6)	55.0 (51.0 - 58.4)	60.0 (57.4 - 63.8)	59.0 (54.0 - 64.8)	<0.0001
LV end-systolic dimension (mm)	32.3 (29.8 - 35.0)	34.4 (31.3 - 38.3)	38.1 (35.4 - 40.0)	36.9 (33.2 - 43.8)	<0.0001
LV mass index (g/m²)	78.4 (65.9 - 93.1)	98.0 (84.4 - 126.1)	119.1 (105.1 - 142.3)	123.7 (106.5 - 143.1)	<0.0001
Relative wall thickness	0.34 (0.28 - 0.39)	0.31 (0.27 - 0.39)	0.29 (0.25 - 0.33)	0.31 (0.26 - 0.39)	0.0004
LVEDV index (mL/m²)	53.5 (45.6 - 64.5)	84.6 (70.1 - 96.0)	110.1 (97.7 - 123.5)	102.7 (89.3 - 118.0)	<0.0001
LVESV index (mL/m²)	21.1 (17.2 - 25.5)	27.7 (21.9 - 33.7)	32.9 (28.2 - 39.8)	30.7 (23.7 - 42.3)	<0.0001
LVEF (%)	60.8 (56.7 - 66.1)	66.1 (60.8 - 70.3)	69.1 (66.2 - 72.5)	67.4 (60.4 - 74.3)	<0.0001

Variables	Cluster 1 (n=131)	Cluster 2 (n=104)	Cluster 3 (n=120)	Cluster 4 (n=74)	p-value
Cardiac index (L/min/m²)	2.7 (2.3 - 3.0)	2.7 (2.4 - 3.2)	2.6 (2.3 - 3.1)	2.3 (1.9 - 2.9)	<0.0001
GLS (%)	21.0 (19.1 - 22.6)	22.0 (20.2 - 24.2)	22.8 (21.6 - 24.8)	19.2 (17.2 - 21.1)	<0.0001
RV diameter (mm)	25.1 (21.9 - 28.2)	29.9 (27.0 - 32.3)	30.1 (27.0 - 32.3)	31.9 (28.5 - 35.1)	<0.0001
E wave (cm/s)	69.0 (56.0 - 79.0)	86.0 (68.0 - 100.0)	116.0 (101.0 - 136.0)	140.0 (120.0 - 158.0)	<0.0001
E/A	1.20 (0.89 - 1.49)	1.29 (1.05 - 1.49)	1.84 (1.57 - 2.28)	2.03 (1.48 - 2.92)	<0.0001
Deceleration time (ms)	174.5 (149.7 - 213.0)	192.4 (158.8 - 239.8)	192.6 (162.5 - 218.2)	164.2 (128.0 - 195.6)	0.0003
E' lateral (cm/s)	11.7 (9.0 - 15.5)	11.0 (8.0 - 13.0)	12.0 (10.0 - 15.0)	11.5 (9.0 - 15.0)	0.006
E/E' lateral	5.5 (4.7 - 7.0)	7.7 (6.2 - 10.4)	9.3 (7.6 - 11.5)	12.6 (9.8 - 16.2)	<0.0001
E/E' septal	6.8 (5.6 - 8.2)	8.9 (6.8 - 11.6)	10.1 (8.7 - 12.6)	15.3 (13.2 - 19.4)	<0.0001
E/E' mean	6.1 (5.1 - 7.1)	8.0 (6.5 - 10.7)	9.7 (8.6 - 11.5)	13.8 (10.7 - 17.5)	<0.0001
AI grade (2 classes)					0.0008
No trace	97 (90.7 %)	68 (73.9 %)	87 (78.4 %)	45 (67.2 %)	
Trace	10 (9.3 %)	24 (26.1 %)	24 (21.6 %)	22 (32.8 %)	
MR grade (3 classes)					<0.0001
Trace-mild	106 (80.9 %)	12 (11.5 %)	0 (0.0 %)	0 (0.0 %)	
Moderate	24 (18.3 %)	80 (76.9 %)	6 (5.0 %)	6 (8.1 %)	
Severe	1 (0.8 %)	12 (11.5 %)	114 (95.0 %)	68 (91.9 %)	
EROA (cm²)	0.07 (0.04 - 0.13)	0.22 (0.14 - 0.31)	0.43 (0.38 - 0.55)	0.42 (0.32 - 0.50)	<0.0001
Regurgitant volume (mL)	9.0 (5.3 - 14.8)	32.0 (21.0 - 47.0)	72.0 (64.0 - 81.0)	65.0 (55.0 - 79.0)	<0.0001
Regurgitant fraction (%)	10.5 (6.2 - 18.1)	32.7 (20.4 - 41.1)	50.3 (43.8 - 56.1)	51.9 (45.8 - 56.8)	<0.0001
LA diameter (mm)	34.3 (30.8 - 38.1)	41.1 (38.0 - 44.0)	46.0 (42.1 - 50.1)	49.9 (45.7 - 51.8)	<0.0001
LA area (cm²)	17.0 (15.0 - 20.2)	23.8 (22.0 - 28.2)	31.7 (28.0 - 36.9)	33.8 (29.0 - 38.1)	<0.0001
LAV index (mL/m²)	30.6 (24.3 - 37.3)	52.2 (44.3 - 61.0)	72.4 (60.3 - 86.3)	80.2 (64.3 - 91.0)	<0.0001
TAPSE (mm)	22.8 (20.8 - 25.4)	25.0 (20.8 - 28.4)	25.5 (22.3 - 28.9)	22.2 (19.1 - 25.0)	<0.0001
RA area (cm²)	13.6 (11.5 - 16.6)	16.8 (14.9 - 20.4)	17.6 (15.0 - 20.9)	21.0 (15.9 - 24.8)	<0.0001
RA area index (cm²/m²)	8.1 (6.8 - 9.5)	9.6 (7.8 - 11.3)	9.4 (8.3 - 11.0)	11.3 (8.9 - 13.7)	<0.0001
PAPs (mmHg)	22.1 (20.0 - 27.0)	28.0 (24.1 - 33.6)	33.0 (29.0 - 39.0)	49.2 (35.0 - 64.0)	<0.0001
Barlow	87 (66.4 %)	72 (69.9 %)	89 (74.2 %)	25 (33.8 %)	<0.0001

Variables	Cluster 1 (n=131)	Cluster 2 (n=104)	Cluster 3 (n=120)	Cluster 4 (n=74)	p-value
Anterior leaflet MVP	12 (9.2 %)	3 (2.9 %)	3 (2.5 %)	8 (11.0 %)	0.019
Posterior leaflet MVP	42 (32.1 %)	33 (31.7 %)	53 (44.5 %)	51 (69.9 %)	<0.0001
Bileaflet MVP	78 (59.5 %)	61 (58.7 %)	59 (49.6 %)	12 (16.4 %)	<0.0001
Anterior leaflet length (mm)	24.0 (22.0 - 26.0)	26.4 (23.4 - 28.7)	28.0 (24.9 - 31.0)	26.0 (24.7 - 28.6)	<0.0001
Posterior leaflet length (mm)	13.9 (12.1 - 16.2)	18.1 (16.0 - 20.9)	20.0 (17.5 - 22.6)	18.8 (16.6 - 21.0)	<0.0001
Mitral annulus diameter (mm)	33.9 (31.0 - 37.1)	38.3 (35.0 - 42.4)	41.0 (38.0 - 44.7)	39.8 (36.2 - 43.0)	<0.0001
Mitral annulus disjunction	82 (78.1 %)	64 (65.3 %)	63 (54.8 %)	11 (16.2 %)	<0.0001
Flail leaflet	0 (0.0 %)	19 (20.7 %)	79 (71.8 %)	59 (84.3 %)	<0.0001
CMR parameters					
LVEDV index (mL/m²)	85.0 (75.0 - 96.0)	100.9 (83.0 - 112.0)	122.0 (107.0 - 138.1)	121.4 (97.7 - 141.6)	<0.0001
LVEDV index (mL/m²)	36.1 (28.7 - 41.0)	38.5 (30.6 - 45.6)	46.7 (37.7 - 55.2)	43.6 (31.9 - 60.6)	<0.0001
LVEF (%)	58.6 (55.0 - 62.7)	60.6 (56.9 - 64.8)	62.5 (57.2 - 66.6)	62.1 (54.3 - 68.0)	0.001
Myocardial mass index (g/m²)	56.0 (48.0 - 61.0)	67.9 (55.0 - 74.0)	76.0 (69.0 - 87.2)	79.8 (71.4 - 92.0)	<0.0001
RVEDV index (mL/m²)	72.0 (63.3 - 82.0)	82.1 (70.9 - 93.0)	88.7 (76.8 - 101.6)	80.8 (66.6 - 104.9)	<0.0001
RVEDV index (mL/m²)	31.5 (24.3 - 36.8)	37.6 (29.2 - 44.1)	42.9 (36.5 - 52.3)	42.3 (31.5 - 53.2)	<0.0001
RVEF (%)	56.0 (52.0 - 60.0)	54.0 (49.5 - 58.4)	50.0 (46.4 - 53.6)	49.5 (44.3 - 55.6)	<0.0001
Regurgitant volume (mL)	18.0 (8.0 - 27.0)	32.0 (20.8 - 43.9)	60.0 (45.0 - 73.0)	63.1 (45.5 - 85.0)	<0.0001
Regurgitant fraction (%)	20.9 (10.3 - 28.0)	29.4 (23.8 - 37.9)	44.0 (36.2 - 50.0)	48.1 (37.5 - 56.2)	<0.0001

Values are reported as median (Q1 – 3) for continuous variables or frequency (percentage) for categorical variables.

*p-value from Kruskal-Wallis test for continuous variables and chi-square test or Fisher's exact test for categorical variables.

See Table 1 for abbreviations.

Table 3: Proportion of cardiovascular events in each cluster

	Overall (n=429)	**Cluster 1 (n=131)**	**Cluster 2 (n=104)**	**Cluster 3 (n=120)**	**Cluster 4 (n=74)**
CV event	92 (21.4%)	6 (4.6%)	20 (19.2%)	36 (30.0%)	30 (40.5%)
Type of CV event					
CV death	8 (8.7%)	1 (16.7%)	1 (5.0%)	2 (5.6%)	4 (13.3%)
Worsening HF	49 (53.3%)	1 (16.7%)	9 (45.0%)	23 (63.9%)	16 (53.3%)
De novo atrial fibrillation	26 (28.3%)	2 (33.3%)	8 (40.0%)	7 (19.4%)	9 (30.0%)
Systemic embolism (stroke, TIA…)	9 (9.8%)	1 (16.7%)	1 (5.0%)	3 (8.3%)	4 (13.3%)
Sustained VT / VF	7 (7.6%)	1 (16.7%)	2 (10.0%)	3 (8.3%)	1 (3.3%)