Radiology: Cardiothoracic Imaging

Arrhythmic Mitral Valve Prolapse Phenotype: An Unsupervised Machine Learning Analysis Using a Multicenter Cardiac MRI Registry

Journal:	*Radiology: Cardiothoracic Imaging*
Manuscript ID	RYCT-23-0247.R2
Manuscript Type:	Original Research
Manuscript Categorization Terms:	MR-Imaging < 2. MODALITIES/TECHNIQUES, Cardiac < 4. AREAS/SYSTEMS

Arrhythmic Mitral Valve Prolapse Phenotype: An Unsupervised Machine Learning Analysis Using a Multicenter Cardiac MRI Registry

Phenotypic clusters in MVP patients

Authors

Ralph Kwame Akyea, MBChB, PhD[a*]; Stefano Figliozzi, MD[b,c*]; Pedro M. Lopes, MD[d]; Klemens B. Bauer, MD[e;] Sara Moura-Ferreira, MD[f,g]; Lara Tondi, MD[h,i]; Saima Mushtaq, MD[j]; Stefano Censi, MD[k]; Anna Giulia Pavon, MD[l]; Ilaria Bassi, MD[m]; Laura Galian-Gay, MD[n,o]; Arco J. Teske, MD[p]; Federico Biondi, MD[q]; Domenico Filomena, MD[r]; Vasileios Stylianidis, MD[c]; Camilla Torlasco, MD, PhD[s]; Denisa Muraru, MD, PhD[s,t]; Pierre Monney, MD[l,u]; Giuseppina Quattrocchi, MD[m]; Viviana Maestrini, MD, PhD[r]; Luciano Agati, MD, PhD[r]; Lorenzo Monti, MD[b]; Patrizia Pedrotti, MD[m]; Bert Vandenberk, MD, PhD[v]; Angelo Squeri, MD[k]; Massimo Lombardi, MD[h]; António M. Ferreira, MD[d]; Juerg Schwitter, MD[l,u]; Giovanni Donato Aquaro, MD[q]; Gianluca Pontone, MD, PhD[j,w]; Amedeo Chiribiri, MD, PhD[c]; José F. Rodríguez Palomares, MD, PhD[n,o]; Ali Yilmaz, MD, PhD[e]; Daniele Andreini, MD, PhD[x]; Anca-Rezeda Florian, MD, PhD[e]; Marco Francone, MD, PhD[b]; Tim Leiner, MD, PhD[p]; João Abecasis, MD[d]; Luigi Paolo Badano, MD, PhD[s,t]; Jan Bogaert, MD, PhD[v]; Georgios Georgiopoulos, MD, PhD[c,y†]; Pier-Giorgio Masci, MD, PhD[c†]

*contributed equally as first authors; †contributed equally as last authors.

[a] Primary Care Stratified Medicine Research Group, Centre for Academic Primary Care, Lifespan and Population Health Unit, School of Medicine, University of Nottingham, UK;
[b] IRCCS Humanitas Research Hospital, Via Alessandro Manzoni, 56, 20089 Rozzano, Milan, Italy;
[c] School of Biomedical Engineering and Imaging Sciences–Faculty of Life Sciences and Medicine, King's College London, Westminster Bridge Rd, London SE1 7EH, England;
[d] Department of Cardiology, Hospital de Santa Cruz, Centro Hospitalar de Lisboa Ocidental, Carnaxide, Lisbon, Portugal;
[e] Department of Cardiology, University Hospital Muenster, Muenster, Germany;
[f] Department of Cardiology, Hartcentrum, Jessa Hospital, Hasselt, Belgium;
[g] Faculty of Medicine and Life Sciences, Hasselt University, Hasselt, Belgium;
[h] Multimodality Cardiac Imaging Section, IRCSS Policlinico San Donato, San Donato Milanese, Italy;
[i] Department of Radiology, Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Italy;
[j] Department of Perioperative Cardiology and Cardiovascular Imaging, Centro Cardiologico Monzino IRCCS, Milan, Italy;
[k] GVM Care&Research, Maria Cecilia Hospital, Cotignola, Italy;
[l] Center for Cardiac MR, Lausanne University Hospital, CHUV, Lausanne, Switzerland;
[m] Cardiologia-4, Dipartimento Cardio-toraco-vascolare A. De Gasperis, ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy;
[n] Department of Cardiology, Hospital Universitario Vall d'Hebron, Institut de Recerca (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain;
[o] Centro de Investigacion Biomedica en Red, CIBERCV, Madrid, Spain;
[p] Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, the Netherlands;
[q] Fondazione CNR/Regione Toscana G. Monasterio, Pisa, Italy;

[r] Department of Clinical, Internal, Anesthesiology and Cardiovascular Sciences, Sapienza University of Rome, Italy;

[s] Department of Cardiology, Istituto Auxologico Italiano, IRCCS, Milan, Italy;
[t] Department of medicine and surgery, University of Milano-Bicocca, Milan, Italy;
[u] Faculty of Biology and Medicine, University of Lausanne, Until, Lausanne, Switzerland;
[v] Gasthuisberg University Hospital, Leuven, Belgium;
[w] Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy;
[x] Department of Clinical Sciences and Community Health, Cardiovascular Section, University of Milan, Milan, Italy;
[y] Department of Clinical Therapeutics, National and Kapodistrian University of Athens, Athens, Greece.

Ralph Kwame Akyea: ralph.akyea1@nottingham.ac.uk;
Stefano Figliozzi: stefanofigliozzi@hotmail.it;
Pedro M. Lopes: pedro_fagalopes@hotmail.com;
Klemens B. Bauer: klemens.bauer@ukmuenster.de;
Sara Moura-Ferreira: sara.mouraferreira@gmail.com;
Lara Tondi: tondi.lara@gmail.com;
Saima Mushtaq: saima.mushtaq@cardiologicomonzino.it;
Stefano Censi: doc.censi@gmail.com;
Anna Giulia Pavon: annagiulia.pavon@gmail.com;
Ilaria Bassi: ilaria.bassi@ospedaleniguarda.it;
Laura Galian-Gay: laura.galian@vallhebron.cat;
Arco J. Teske: a.j.teske-2@umcutrecht.nl;
Federico Biondi: biondi.federico@yahoo.it;
Domenico Filomena: domenico.filom@gmail.com;
Vasileios Stylianidis: vasileios.stylianidis@kcl.ac.uk
Camilla Torlasco: c.torlasco@auxologico.it;
Denisa Muraru: denisa.muraru@unimib.it;
Pierre Monney: pierre.monney@chuv.ch;
Giuseppina Quattrocchi: giuseppina.quattrocchi@ospedaleniguarda.it;
Viviana Maestrini: vivianamaestrini@gmail.com;
Luciano Agati: luciano.agati@uniroma1.it;
Lorenzo Monti: lorenzo.monti@humanitas.it;
Patrizia Pedrotti: patrizia.pedrotti@ospedaleniguarda.it;
Bert Vandenberk: bert.vandenberk@uzleuven.be;
Angelo Squeri: angelo.squeri@yahoo.it;
Massimo Lombardi: massimo.lombardi@grupposandonato.it;
António M. Ferreira: amcsferreira@chlo.min-saude.pt;
Juerg Schwitter: jurg.schwitter@chuv.ch;
Giovanni Donato Aquaro: aquarogd@gmail.com;
Gianluca Pontone: gianluca.pontone@cardiologicomonzino.it;
Amedeo Chiribiri: amedeo.chiribiri@kcl.ac.uk;
José F. Rodríguez Palomares: jfrodriguezpalomares@gmail.com;
Ali Yilmaz: ali.yilmaz@ukmuenster.de;
Daniele Andreini: daniele.andreini@unimi.it;
Anca-Rezeda Florian: ancaflorian@yahoo.com;
Marco Francone: marco.francone@hunimed.eu;

Tim Leiner: leiner.tim@mayo.edu;
João Abecasis: joaoabecasis@hotmail.com;
Luigi Paolo Badano: luigi.badano@unimib.it;
Jan Bogaert: jan.bogaert@uzleuven.be;
Georgios Georgiopoulos: georgiopoulosgeorgios@gmail.com;
Pier-Giorgio Masci: pier_giorgio.masci@kcl.ac.uk.

Address for Correspondence:
Dr. Pier-Giorgio Masci, MD, PhD
School of Biomedical Engineering and Imaging Sciences, King's College London,
Westminster Bridge Road, London, SE1 7EH
Email: pier_giorgio.masci@kcl.ac.uk; pgmasci@gmail.com
Phone: +44 747 1551780
Fax: +44 747 1551780
Twitter: @masci_pier. Refining the risk of sudden cardiac death by unsupervised machine learning in isolated mitral valve prolapse: #cluster characterized by more extensive mitral valve degeneration, cardiac remodeling and fibrosis is at higher risk. #MVP #SCD #whyCMR #AI

Word count for text: 3,557

Manuscript Type: Original Research.

Acknowledgments: We are grateful to Dr. Silvia Pica, MD, for her relenteless and unforgotten contribution to developing the present international cohort of patients.

Data sharing statement: Data generated or analyzed during the study are available from the corresponding author by request.

Disclosures:

Ralph Kwame Akyea: none.
Stefano Figliozzi: none.
Pedro M. Lopes: none.
Klemens B. Bauer: none.
Sara Moura-Ferreira: none.
Lara Tondi: none.
Saima Mushtaq: none.
Stefano Censi: none.
Anna Giulia Pavon: none.
Ilaria Bassi: none.
Laura Galian-Gay: none.
Arco J. Teske: none.
Federico Biondi: none.
Domenico Filomena: none.
Camilla Torlasco: none.
Denisa Muraru: none.
Pierre Monney: none.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Giuseppina Quattrocchi: none.
Viviana Maestrini: none.
Luciano Agati: none.
Lorenzo Monti: none.
Patrizia Pedrotti: none.
Bert Vandenberk: none.
Angelo Squeri: none.
Massimo Lombardi: none.
António M. Ferreira: none.
Juerg Schwitter: none.
Giovanni Donato Aquaro: none.
Gianluca Pontone: none.
Amedeo Chiribiri: none.
José F. Rodríguez Palomares: none.
Ali Yilmaz: none.
Daniele Andreini: none.
Anca-Rezeda Florian: none.
Marco Francone: none.
Tim Leiner: none.
João Abecasis: none.
Luigi Paolo Badano: none.
Jan Bogaert: none.
Georgios Georgiopoulos: none.
Pier-Giorgio Masci: none.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Arrhythmic Mitral Valve Prolapse Phenotype: An Unsupervised Machine Learning Analysis Using a Multicenter Cardiac MRI Registry

Article Type: Original Research.

Summary Statement

- In patients with mitral valve prolapse, cardiac MRI parameters pinpointing degenerative changes of the mitral apparatus, left and right chamber remodeling and myocardial fibrosis identified a phenotype at increased arrhythmic risk.

Key Points

- Unsupervised machine learning was applied to 474 patients with mitral valve prolapse without moderate-severe mitral regurgitation or left ventricular dysfunction undergoing cardiac MRI and clinical follow-up evaluating an arrhythmic endpoint (i.e., unexplained syncope, sustained ventricular tachycardia or sudden cardiac death).
- Among the two phenotypic clusters identified, cluster-2 patients (n=199/474, 42%) had more severe mitral valve degeneration, left and right heart chamber remodeling, and myocardial fibrosis than those in cluster-1; demographic and clinical features had negligible contributions in differentiating the clusters.
- Cluster-2 patients showed a higher risk of developing the arrhythmic endpoint (HR: 3.91, 95% CI: 1.22–12.47) over a median follow-up of 3.3 years.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Abbreviations

LGE: Late Gadolinium Enhancement;
LV: Left Ventricle;
MAD: Mitral Annulus Disjunction;
MVP: Mitral Valve Prolapse;
RV: Right Ventricle:
SCD: Sudden Cardiac Death;
VT: Ventricular Tachycardia.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Abstract

Purpose: To use unsupervised machine learning to identify phenotypic clusters with increased risk of arrhythmic mitral valve prolapse (MVP).

Materials and Methods: This retrospective study included patients with MVP without significant mitral regurgitation or left ventricular (LV) dysfunction undergoing late gadolinium enhancement (LGE) cardiac MRI between October 2007 and June 2020 in 15 European tertiary centers. The study endpoint was a composite of sustained ventricular tachycardia, (aborted)-sudden cardiac death, or unexplained syncope. Unsupervised data-driven hierarchical k-mean algorithm was utilized to identify phenotypic clusters. The association between clusters and the study endpoint was assessed by Cox proportional hazards model.

Results: Four-hundred-seventy-four patients (mean age, 47 ± [SD] 16 years; 244 female, 230 male) Two phenotypic clusters were identified. Cluster-2 patients (n=199/474, 42%) had more severe mitral valve degeneration (i.e., bi-leaflet MVP and leaflet displacement), left and right heart chamber remodeling, and myocardial fibrosis by LGE cardiac MRI than those in cluster-1. Demographic and clinical features (i.e., symptoms, arrhythmias at Holter monitoring) had negligible contribution in differentiating the two clusters. Compared with cluster-1, the risk of developing the study endpoint over a median follow-up of 39 months was significantly higher in cluster-2 patients (hazard ratio: 3.79, 95% CI:1.19–12.12, p=0.024) after adjustment for LGE extent.

Conclusion: Among patients with MVP without significant mitral regurgitation or LV dysfunction, unsupervised machine learning enabled the identification of two phenotypic clusters with distinct arrhythmic outcomes based primarily on cardiac MRI features. These results encourage the use of in-depth imaging-based phenotyping for implementing arrhythmic risk prediction in MVP.

Keywords: Cardiac MRI; Mitral Valve Prolapse; Cluster Analysis; Ventricular Arrhythmia; Sudden Cardiac Death; Unsupervised Machine Learning.

Introduction

Mitral valve prolapse (MVP) is the most common valvular disease, with a prevalence of 2-3% in the general population and an overall good prognosis in the absence of significant mitral regurgitation or left ventricular (LV) dysfunction.[1] This scenario is challenged by growing evidence suggesting that a subset of patients with MVP is exposed to sustained ventricular tachycardia (VT) and sudden cardiac death (SCD) despite the absence of moderate-to-severe mitral regurgitation or LV dysfunction.[2-6] This entity is referred to as '*arrhythmic MVP*' and represents a conundrum for physicians given the difficulty in estimating the arrhythmic risk. Recent cross-sectional studies identified some clinical and imaging features associated with the arrhythmic phenotype,[7] including mitral annulus disjunction (MAD), prolapse severity, and myocardial fibrosis.[2-8] However, the relative importance of each of these components in compounding the arrhythmic risk remains uncertain. For instance, our group has recently disputed the role of MAD as an arrhythmic maker highlighting, in contrast, the importance of myocardial fibrosis as identified by late gadolinium enhancement (LGE) cardiac MRI.[3]

The multidimensionality and complexity of demographic, clinical and cardiac MRI features hinder the in-depth characterization of MVP and, thereby, a meaningful identification of the pro-arrhythmic features. One successful way to circumvent this limitation is the use of novel machine learning techniques, such as unsupervised hypothesis-free machine learning data-driven analysis. This approach has been successfully used in identifying new phenotypes in complex and heterogeneous diseases such as chronic heart failure,[9] dilated cardiomyopathy,[10] and Parkinson disease.[11]

In this study, we leveraged unsupervised machine learning to interrogate a multidimensional and complex dataset from an international multicenter registry of patients with MVP studied by cardiac MRI to identify novel phenotypic features associated with the arrhythmic outcome on longitudinal analysis.

Materials and Methods

Study design

This study originates from an international multicenter longitudinal retrospective registry including patients with MVP without significant mitral regurgitation or LV dysfunction studied by cardiac MRI.[3] The ethics committee of the 15 included centers approved the study in agreement with the ethics approval issued at the leading center (King's College London; research ethics committee no. 15/NS/0030). All patients provided written informed consent.

Study sample

Patient selection and inclusion in the registry have been previously described elsewhere.[3] Briefly, patients were included if they met the following criteria: i) aged ≥18 years; ii) MVP was present at cardiac MRI; iii) clinical information and continuous electrocardiogram (ECG) monitoring were available within 3 months from cardiac MRI; iv) LGE imaging was carried out. Exclusion criteria were as follows: i) cardiomyopathy; ii) LV ejection fraction <40%; iii) ischemic heart-disease; iv) congenital heart-disease; v) inflammatory heart disease; vi) mitral regurgitation grade ≥ moderate (as per transthoracic echocardiography or mitral regurgitation fraction >20% at cardiac MRI); vii) participation in competitive sport. Patients with LV ejection fraction <55% but ≥40% were included in the study given that mildly reduced systolic function can be associated with MVP in the absence of significant mitral regurgitation.[12,13]

Cardiac MRI Protocol and Analysis

The cardiac MRI protocol and image analysis have been previously described.[3] All patient scans were carried out on a 1.5-Tesla system using dedicated cardiac software, phased-array surface receiver coil, and electrocardiogram triggering. Protocol and sequence parameters were previously described.[3] Briefly, ventricular volumes, mass, and function as well as atrial areas

were analyzed according to the current Society of Cardiovascular Magnetic Resonance recommendations.[14] MVP was defined as ≥2.0-mm displacement of the mitral valve leaflet into the left atrium on the cine 3-chamber image at end-systole.[15] MAD was defined as an anatomic variant of the posterior mitral annulus resulting in a separation (≥2.0 mm) between the left atrial wall/mitral-valve junction and LV inferolateral wall on the cine 3-chamber image at end-systole.[4,5] On post-contrast images, LGE was deemed present in the LV walls or papillary muscles if at least one of the following conditions was fulfilled: i) LGE visible in two orthogonal views; ii) LGE visible on the same image orientation after swapping phase-frequency direction.[14,16] When present in the LV walls, LGE extent was quantified as myocardium with signal intensity >5 standard deviations than normal myocardium.[4] The signal intensity of the normal (nulled) myocardium was measured by manually drawing a region of interest in the non-enhanced myocardium devoid of artifacts. LGE was expressed as a percentage of LV mass (%LV).

Variables used for cluster analyses

generating phenotypic clusters based on demographic, clinical, and Cardiac MRI features at baseline. There were 32 demographic, clinical, and cardiac MRI variables for analyses **(Supplementary Table 1)**. History of malignant ventricular arrhythmias (i.e., sustained VT, ventricular fibrillation or aborted SCD) at baseline was excluded given their well-established adverse prognostic impact[17] necessitating implantable cardioverter defibrillator for secondary prevention of SCD.[7] To improve the cluster analysis and mitigate collinearity, 5 highly correlated variables were excluded based on clinical judgement/importance **(Supplementary Figure 1** shows the correlation**)**. The remaining 27 variables were used for the cluster analysis **(Supplementary Table 1)**. Finally, a sensitivity analysis was carried out by excluding patients presenting with malignant ventricular arrhythmias (n=18) at baseline and including the baseline

burden of ventricular arrhythmias at continuous ECG monitoring (ventricular ectopic beats≤10,000/24 hours versus ventricular ectopic beats>10,000/24 hours and/or at least one episode of non-sustained ventricular tachycardia)[3] in the cluster analysis.

Cluster generation

A combination of approaches was used to determine the optimal number of clusters (Elbow method, Gap statistics, and using the NbClust package in R version 4.2.1), **(Supplementary Figure 2)**. The NbClust package uses 30 different clustering indices to determine the optimal number of clusters based on the highest frequency of selection from all 30 indices.[18] To identify phenotypic groups of patients (i.e., clusters) with similar clinical and cardiac MRI characteristics, a combined k-means and hierarchical agglomerative approach, called hierarchical k-means clustering, was used.[19] This hierarchical k-means process allows for the k-means–based approach to speed up the traditional k-means algorithm in both training and query phases, which allows for a much larger number of centroids to be used and in turn leads to much better learning.[19] In the process, k is selected as the branching factor, which defines the number of clusters at each level of the clustering hierarchy. To ensure the robustness of the clusters identified, 1,000 initializations (i.e.,, random starting points) were carried out. The screen plot for the principal component analysis dimensions was generated **(Supplementary Figure 3)**. A gradient boosting model was applied, using the h2o package (http://www.h2o.ai), to identify as well as rank the variables that predict each of the identified phenotypic clusters. SHapley Additive exPlanations (SHAP) were used to assess the discriminative influence of the variables for each of the identified clusters.[20]

Outcome measures

The study endpoint was a composite of sustained VT, (aborted)-SCD, or unexplained syncope at follow-up. The clinical follow-up started at the cardiac MRI date and lasted until the common closing date of June 2020 (minimum and maximum intervals were 6 months and 156 months, respectively). Patients with non-cardiac death were censored at the event date. Events were adjudicated by two experienced cardiologists (20 and 23 years of experience) who were blinded to cardiac MRI results but had full access to clinical records and contacted the treating physicians whenever needed. A consensus was reached between the two cardiologists in case of disagreement.

Statistical analysis

For each cluster, descriptive characteristics are provided, reporting proportion (%) for categorical variables and mean (standard deviation) or median (interquartile range) for continuous variables. Kruskal-Wallis and χ-squared tests were used to compare across clusters for continuous and categorical data, respectively. Incidence rates per 1,000 person-years with 95% confidence intervals (CI) were provided. Kaplan-Meier curves were used to plot the event-free survival of the two phenotypic clusters identified; log-rank test was used to compare the difference between the event-free survivals between the two clusters. Multivariable Cox proportional hazards analysis was performed to determine the association between the clusters and risk of the study endpoint. The hazard ratios (HR) with 95% CI were reported. The proportional hazards assumption was assessed using Schoenfeld residuals. All statistical analyses were performed using Stata SE version 17 (StataCorp LP) and R version 4.2.1. An alpha level of 0.05 was used. All tests were two-tailed.

Results

Study sample characteristics

Four hundred and seventy-four patients with isolated MVP were included in the study. The mean age was 47 ± 16 years, with 244/474 (51.5%) female patients and 230/474 (48.5%) male patients. Two phenotypic clusters were identified. The dendrogram and principal component analysis dimensions for the identified clusters are shown in **Figure 1**.

Phenotypic clusters

Table 1 summarizes the two clusters' main demographic, clinical, and cardiac MRI characteristics. Patients in the two clusters showed similar age and sex distribution. Cluster-2 included patients with a higher prevalence of bi-leaflet prolapse and MAD, more pronounced anterior or posterior leaflet displacement, and MAD longitudinal extent as compared with patients in cluster-1. The accentuated structural and functional mitral valve derangements in cluster-2 patients were associated with a higher prevalence and extent of LGE compared with cluster-1 patients. Finally, patients in cluster-2 had larger biventricular and atrial dimensions than cluster-1.

Grading the importance of variables for clusters

Variables contributed to the model's prediction of clusters with different magnitudes (feature importance) and directions (sign). The contributions are accounted for by Shapley values (**Figure 2)**. Using a gradient boosting model, bi-leaflet MVP, posterior and anterior leaflets displacement, and left and right end-diastolic volumes and LGE extent were identified as the most important variables for predicting phenotypic clusters.

Phenotypic clusters and clinical outcomes

During a median follow-up of 39 months (6 – 156 months), 18 patients experienced the study endpoint. The overall incidence rate was 12.0 per 1.000 person-years (95% CI: 7.6 – 19.1). Cluster-1 had a lower incidence rate (6.4 per 1000 person-years; 95% CI: 2.9 – 14.2) compared with cluster-2 (21.5 per 1,000 person-years; 95% CI: 12.2 – 37.9), yielding an HR of 5.30 (95% CI: 1.79-15.74). Given the prognostic importance of LGE in the prior study,[3] we adjusted the Cox proportional hazards regression analysis by LGE extent. Patients in cluster-2 had a significantly higher risk of the study endpoint than those in cluster-1 after adjusting for LGE extent (HR: 3.79; 95% CI: 1.19 – 12.12; p=0.024) **(Table 2, Figure 3**).

Sensitivity Analysis

In the sensitivity analysis sample (n=456), two phenotypic clusters were identified mirroring the results of the core analysis including the whole study sample. Ventricular ectopic beats burden and/or non-sustained ventricular tachycardia by continuous ECG monitoring at baseline were similar between the two clusters **(Supplementary Table 2)**. Of importance, ECG monitoring had a negligible contribution in the identification of clusters based on Shapley analysis **(Supplementary Figure 4).** Cox proportional hazards analysis confirmed that cluster-2 patients had higher likelihood of experiencing the study endpoint during follow-up than those in the cluster-1 after adjusting for LGE extent (HR: 4.03; 95% CI:1.06-15.32; p=0.041) (**Supplementary Table 3, Supplementary Figure 5**).

Discussion

LV dysfunction and moderate-to-severe mitral regurgitation are well-established risk factors for adverse cardiovascular events in patients with MVP.[1,21-23] However, only 1 out of 5 cases of SCD-related with MVP occurs in patients with severe mitral regurgitation [2], rendering the characterization of arrhythmic MVP phenotype elusive. In this multicenter registry including

474 patients with isolated MVP without significant LV dysfunction or mitral regurgitation, we utilized data-driven unsupervised machine learning to identify meaningful non-apriori features underpinning the arrhythmic MVP phenotype. This enabled us to identify two phenotypic clusters based on demographic, clinical, and cardiac MRI variables. Cluster-2 patients had more severe MVP, as epitomized by more pronounced atrial leaflet displacement and higher prevalence of bi-leaflet prolapse, larger ventricles and atria as well as higher prevalence and extent of myocardial fibrosis as compared with patients in cluster-1. Of importance, cluster-2 patients had a four times greater likelihood of experiencing (aborted)-SCD, sustained VT, or unexplained syncope during a median follow-up of more than three years (**Figure 4**).

Our findings indicate that the sole presence of MVP is unlikely to harbinger an untoward prognosis unless associated with advanced degenerative processes of mitral apparatus coupled with chambers dilatation and myocardial fibrosis. Patients with MVP and Cluster-1 characteristics, including single-leaflet prolapse of limited entity, no or limited heart chambers' dilation, and no or limited myocardial fibrosis by LGE, showed an extremely low incidence of adverse outcomes.

The severity of MVP, underpinned as bi-leaflet prolapse or displacement of the mitral leaflets, was the most important feature in cluster-2. The mechanical stress in the papillary muscles and adjacent walls brings about electrophysiological derangements encompassing a decrease of action potential duration and stretch-mediated early after depolarization[24,25], and favors, in parallel, the development of myocardial fibrosis.[26] In turn, myocardial fibrosis may act as substrate for re-entry ventricular arrhythmias [27], and unsurprisingly LGE was a key feature in the cluster analysis, with higher prevalence and greater extent in cluster-2 than cluster-1 patients. This result is in keeping with our and other groups' findings highlighting the association between myocardial fibrosis and adverse clinical outcome.[3,4] However, given the low incidence of clinical outcomes and relatively high prevalence of myocardial fibrosis by

LGE at baseline, this feature alone is unlikely to identify patients at high risk of SCD. By demonstrating an independent prognostic value of cluster analysis over LGE in patients with MVP, our data encourage a more holistic and granular approach integrating the myocardial fibrosis with other morpho-functional parameters such as MVP severity and chamber dilatation to better identify MVP patients at highetned risk of SCD. In our study, the presence of MAD had a negligible impact in cluster discrimination. This finding aligns with recent evidence from our and other groups which did not find any prognostic value of MAD in patients with MVP or healthy individuals harboring this condition[3,28] and diverge from earlier studies where MAD was associated with worse outcome.[5] This discripancy likely reflects differences in study samples and method to asses and define MAD.[5,28,29,30] It is worth noting that the longitudinal extent of MAD, which was greater in Cluster-2, contributed to cluster differentiation. This result is consonant with previous studies showing an association between MAD longitudinal length greater than 8.5 mm[31] or 10 mm[32] and ventricular arrhythmias. This evidence holds pathophysiological plausibility, given that a higher degree of prolapse and/or MAD concurs to mechanical tension on the papillary muscles and adjacent myocardium, prompting electrophysiocal derangements and myocardial fibrosis.[26]

Finally, cluster-2 patients had larger ventricles and atria than cluster-1 patients. In patients with MVP but less than moderate mitral regurgitation, LV and atrial dilatation have been reported in prior studies.[12,13] It remains uncertain whether these abnormalities result from a genetically mediated process[33] or from a volume overload secondary to the 'third chamber' effect, which refers to the formation of a functional 'chamber' underlying the ventricular side of the prolapsing mitral leaflets.[12] In our study, the prevalence of bi-leaflet MVP and leaflet displacement magnitude were associated with LV and atrial dilatation. We also found that right ventricular (RV) dilatation was a main feature differentiating cluster-2 from cluster-1 patients. Large studies integrating cardiac MRI and genome-wide-association analysis may help to

clarify the underpinnings of left and right chamber remodeling and dysfunction in patients with MVP.

Unlike cardiac MRI-based features, demographic and clinical features had a low or negligible impact in cluster analysis, thus supporting the use of advanced imaging-based phenotyping for MVP. This finding was also supported by the sensitivity analysis which showed a negligible contribution of ECG monitoring in the identification of the phenotypic clusters. Of note, RV dilatation and myocardial fibrosis were two key features in cluster discrimination, and both can be accurately detected and quantified by cardiac MRI. We acknowledge that cardiac MRI cannot be routinely performed in unselected patients with MVP given its limited availability and relatively high costs. However, one may argue that this imaging modality may be implemented in a subset of patients harboring some 'red-flag'[34] features at transthoracic echocardiography, such as bi-leaflet prolapse and/or severe prolapsing leaflets. Dedicated large prospective longitudinal cohort studies incorporating transthoracic echocardiography, cardiac MRI, continuous cardiac rhythm monitoring, and health economics are warranted to delineate the most cost-effective strategy in patients with MVP without hemodynamically significant mitral regurgitation and/or severe LV dysfunction.

Our study had several limitations. First, we included patients undergoing clinically indicated cardiac MRI at tertiary centers and thus cannot exclude selection bias. The large sample size, including consecutive patients with isolated MVP with no co-existent cardiopathy, comorbidities, significant LV dysfunction or mitral regurgitation, and no restrictions on symptoms' presentations, together with the use of an unbiased analysis approach with artificial intelligence, increases the robustness of the study findings compared with previous investigations.[4,5] More extensive studies, including those with patients from non-tertiary centers, remain necessary to confirm our data. With this regard, our research methods are based

on routine and conventional cardiac MRI protocols and post-processing; therefore, they are easily reproducible and potentially have immediate clinical applicability. Second, we were not able to investigate some imaging- and non-imaging-based parameters previously proposed as markers of ventricular arrhythmias in MVP, such as mitral leaflet thickness.[8] Moreover, we were not able to incorporate promising cardiac MRI parameters in MVP such as T1-mapping[35] or global longitudinal strain[36] given that they were not implemented at the time of cardiac MRI examinations for most of the patients. Moreover, we were not able to include ventricular repolarization abnormalities and sites of origin of ventricular ectopic beats in our analysis given that 12-lead ECG recording at the study outset was not available for most of the patients.[4,37] Along this line, continuous ECG monitoring data did not include some features with potential prognostic implications such as rapid non-sustained ventricular tachycardia (>180 bpm) or polymorphic ventricular ectopic beats.[7] Finally, cluster analysis may be limited by the sample size of the dataset and number of clinical features used to determine cluster association.

In conclusion, we identified two distinct phenotypic clusters in patients with MVP without hemodynamically significant mitral regurgitation or LV dysfunction. Cluster-2 patients had more extensive mitral valve degenerative abnormalities, left and right heart chamber remodeling, and myocardial fibrosis than cluster-1 patients, resulting in a nearly 4-fold increased risk of developing (aborted)-SCD, sustained VT or unexplained syncope over more than three years follow-up. By contrast, demographic and clinical features had negligible contribution in differentiating the two clusters, ultimately supporting the role of in-depth phenotyping by advanced cardiovascular imaging for arrhythmic risk stratification.

References

1. Freed LA, Levy D, Levine RA, et al. Prevalence and clinical outcome of mitral-valve prolapse. N Engl J Med 1999;341:1–7. DOI: 10.1056/NEJM199907013410101

2. Han HC, Ha FJ, Teh AW, et al. Mitral valve prolapse and sudden cardiac death: a systematic review. J Am Heart Assoc 2018;7(23):e010584. DOI: 10.1161/JAHA.118.010584

3. Figliozzi S, Georgiopoulos G, Lopes PM, et al. Myocardial Fibrosis at Cardiac MRI Helps Predict Adverse Clinical Outcome in Patients with Mitral Valve Prolapse. Radiology 2023;306(1):112-121. doi: 10.1148/radiol.220454

4. Basso C, Perazzolo-Marra M, Rizzo S, et al. Arrhythmic mitral valve prolapse and sudden cardiac death. Circulation 2015;132(7):556–566. doi: 10.1161/CIRCULATIONAHA.115.016291

5. Dejgaard LA, Skjølsvik ET, Lie ØH, et al. The Mitral Annulus Disjunction Arrhythmic Syndrome. J Am Coll Cardiol 2018;72(14):1600-1609. DOI: 10.1016/j.jacc.2018.07.070

6. Essayagh B, Sabbag A, Antoine C, et al. Presentation and Outcome of Arrhythmic Mitral Valve Prolapse. J Am Coll Cardiol 2020;76(6):637-649. doi: 10.1016/j.jacc.2020.06.029

7. Sabbag A, Essayagh B, Ramírez Barrera JD, et al. EHRA expert consensus statement on arrhythmic mitral valve prolapse and mitral annular disjunction complex in collaboration with the ESC Council on valvular heart disease and the European Association of Cardiovascular Imaging endorsed by the Heart Rhythm Society, by the Asia Pacific Heart Rhythm Society, and by the Latin American Heart Rhythm Society. Europace 2022;24(12):1981-2003. doi: 10.1093/europace/euac125

8. Han Y, Peters DC, Salton CJ, et al. Cardiovascular magnetic resonance characterization of mitral valve prolapse. J Am Coll Cardiol Img 2008;1(3):294-303. DOI: 10.1016/j.jcmg.2008.01.013

9. Ahmad T, Pencina MJ, Schulte PJ, et al. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. J Am Coll Cardiol 2014;64(17):1765-74. DOI: 10.1016/j.jacc.2014.07.979

10. Verdonschot JAJ, Merlo M, Dominguez F, et al. Phenotypic clustering of dilated cardiomyopathy patients highlights important pathophysiological differences. Eur Heart J 2021;42(2):162-174. doi: 10.1093/eurheartj/ehaa841

11. Fereshtehnejad SM, Romenets SR, Anang JBM, Latreille V, Gagnon JF, Postuma RB. New clinical subtypes of Parkinson disease and their longitudinal progression a prospective cohort comparison with other phenotypes. JAMA Neurol 2015;72:863–873. oi: 10.1001/jamaneurol.2015.0703

12. El-Tallawi KC, Kitkungvan D, Xu J, et al. Resolving the disproportionate left ventricular enlargement in mitral valve prolapse due to Barlow disease: insights from cardiovascular magnetic resonance. J Am Coll Cardiol Img 2021;14(3):573–584. doi: 10.1016/j.jcmg.2020.08.029.

13. Yang LT, Ahn SW, Li Z, et al. Mitral valve prolapse patients with less than moderate mitral regurgitation exhibit early cardiac chamber remodeling. J Am Soc Echocardiogr 2020;33(7):815–825.e2. DOI: 10.1016/j.echo.2020.01.016

14. Schulz-Menger J, Bluemke DA, Bremerich J, et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update : Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing. J Cardiovasc Magn Reson 2020;22:19. doi: 10.1186/s12968-020-00610-6. doi: 10.1186/s12968-020-00610-6

15. Zoghbi WA, Adams D, Bonow RO, et al. Recommendations for non-invasive evaluation of native valvular regurgitation: a report from the American Society of Echocardiography developed in collaboration with the Society for Cardiovascular Magnetic Resonance. J Am Soc Echocardiogr 2017;30(4):303–371. doi: 10.1016/j.echo.2017.01.007

16. Flett AS, Hasleton J, Cook C, et al. Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance. J Am Coll Cardiol Img 2011;4(2):150–156. doi: 10.1016/j.jcmg.2010.11.015

17. Hourdain J, Clavel MA, Deharo JC, et al. Common phenotype in patients with mitral valve prolapse who experienced sudden cardiac death. Circulation 2018;138 (10):1067–1069. doi: 10.1161/CIRCULATIONAHA.118.033488

18. Charrad M, Ghazzali N, Boiteau V, NbClust NiknafsA. An R package for determining the relevant number of clusters in a data set. J Stat Softw 2014; 61 :1–36.

19. Peterson AD, Ghosh AP, Maitra R. Merging K-means with hierarchical clustering for identifying general shaped groups. Stat (International Statistical Institute) 2018; 7: e172. doi: 10.1002/sta4.172

20. Lundberg SM, Erion G, Chen H, et al. From Local Explanations to Global Understanding with Explainable AI for Trees, Nat. Mach. Intell. 2 (2020) 56. doi: 10.1038/s42256-019-0138-9

21. Kligfield P, Levy D, Devereux RB, Savage DD. Arrhythmias and sudden death in mitral valve prolapse. Am Heart J 1987;113(5):1298–1307. doi: 10.1016/0002-8703(87)90958-6

22. Martínez-Rubio A, Schwammenthal Y, Schwammenthal E, et al. Patients with valvular heart disease presenting with sustained ventricular tachyarrhythmias or syncope: results of programmed ventricular stimulation and long-term follow-up. Circulation 1997;96(2):500–508. doi: 10.1161/01.cir.96.2.500

23. Grigioni F, Enriquez-Sarano M, Ling LH, et al. Sudden death in mitral regurgitation due to flail leaflet. J Am Coll Cardiol 1999;34(7):2078–2085. doi: 10.1016/s0735-1097(99)00474-x.

24. Franz, MR. Mechano-electrical feedback. Cardiovasc Res 2000;45(2):263-6. doi: 10.1016/s0008-6363(99)00390-9

25. Lab MJ. Mechanoelectric feedback (transduction) in heart: concepts and implications. Cardiovasc Res 1996, https://doi.org/10.1016/S0008-6363(96)00088-0

26. Morningstar JE, Gensemer C, Moore R, et al. Mitral Valve Prolapse Induces Regionalized Myocardial Fibrosis. J Am Heart Assoc 2021;10(24):e022332. doi: 10.1161/JAHA.121.022332

27. Maruyama T, Fukata M. Increased coupling interval variability -- mechanistic, diagnostic and prognostic implication of premature ventricular contractions and underlying heart diseases. Circ J 2015;79(11):2317-9. doi: 10.1253/circj.CJ-15-0963

28. Zugwitz D, Fung K, Aung N, et al. Mitral Annular Disjunction Assessed Using CMR Imaging: Insights From the UK Biobank Population Study. J Am Coll Cardiol Img 2022;15:1856-1866. doi: 10.1016/j.jcmg.2022.07.015

29. Essayagh B, Sabbag A, Antoine C, et al. The Mitral Annular Disjunction of Mitral Valve Prolapse: Presentation and Outcome. J Am Coll Cardiol Img 2021;14:2073-2087. doi: 10.1016/j.jcmg.2021.04.029

30. Mantegazza V, Volpato V, Gripari P, et al. Multimodality imaging assessment of mitral annular disjunction in mitral valve prolapse. Heart 2021;107(1):25-32. DOI: 10.1136/heartjnl-2020-317330

31. Carmo P, Andrade MJ, Aguiar C, Rodrigues R, Gouveia R, Silva JA. Mitral annular disjunction in myxomatous mitral valve disease: a relevant abnormality recognizable by transthoracic echocardiography. Cardiovasc Ultrasound 2010:8:53. doi: 10.1186/1476-7120-8-53

32. Demolder A, Timmermans F, Duytschaever M, Muiño-Mosquera L, De Backer J. Association of Mitral Annular Disjunction With Cardiovascular Outcomes Among Patients With Marfan Syndrome. JAMA Cardiol 2021;6(10):1177-1186. doi: 10.1001/jamacardio.2021.2312

33. Dina C, Bouatia-Naji N, Tucker N, et al. Genetic association analyses highlight biological pathways underlying mitral valve prolapse. Nat Genet 2015;47:1206-11. doi: 10.1038/ng.3383

34. Pavon AG, Monney P, Schwitter J. Mitral Valve Prolapse, Arrhythmias, and Sudden Cardiac Death: The Role of Multimodality Imaging to Detect High-Risk Features. Diagnostics 2021;11:683. doi: 10.3390/diagnostics11040683

35. Pavon AG, Arangalage D, Pascale P, et al. Myocardial extracellular volume by T1 mapping: a new marker of arrhythmia in mitral valve prolapse. J Cardiovasc Magn Reson 2021;23(1):102. doi: 10.1186/s12968-021-00797-2

36. Guglielmo M, Fusini L, Muscogiuri G, et al. T1 mapping and cardiac magnetic resonance feature tracking in mitral valve prolapse. Eur Radiol 2021;31(2):1100-1109. doi: 10.1007/s00330-020-07140-w

37. Chivulescu M, Aabel EW, Gjertsen E, et al. Electrical markers and arrhythmic risk associated with myocardial fibrosis in mitral valve prolapse. Europace 2022;24(7):1156-1163. doi: 10.1093/europace/euac017

Table 1. Patient Baseline Characteristics Stratified by Clusters

	Total (n = 474)	Cluster 1 (n=275/474; 58.0%)	Cluster 2 (n=199/474; 42.0%)	*p*-values
Age (years)	47.0 ± 15.9	46.6 ± 16.0	47.5 ± 15.8	0.660
Sex				0.633
Female	244/474 (51.5)	139/275 (50.6)	105/199 (52.8)	
Male	230/474 (48.5)	136/275 (49.5)	94/199 (47.2)	
Body mass index (kg/m^2)	23.1 ± 3.6	23.5 ± 3.7	22.4 ± 3.5	0.001
Active smoking	19/474 (4.0)	8/275 (2.9)	11/199 (5.5)	0.151
Diabetes	5/474 (1.1)	3/275 (1.1)	2/199 (1.0)	0.928
Family history of SCD	20/474 (4.2)	17/275 (6.2)	3/199 (1.5)	0.012
Hyperlipidemia	60/474 (12.7)	45/275 (16.4)	15/199 (7.5)	0.004
Hypertension	72/474 (15.2)	51/275 (18.6)	21/199 (10.6)	0.017
Mitral valve prolapse				<0.001
Single leaflet	249/474 (52.5)	223/275 (81.1)	26/199 (13.1)	
Bi leaflet	225/474 (47.5)	52/275 (18.9)	173/199 (86.9)	
Anterior leaflet displacement (mm)	2 (0 – 4.5)	0 (0 – 2)	4.9 (3 – 7)	<0.001
Posterior leaflet displacement (mm)	4 (3 – 7)	3.1 (2.5 – 4)	7 (4.4 – 9.6)	<0.001
Left atrium area *i* (cm^2/m^2)	12.6 ± 3.3	11.5 ± 2.6	14.1 ± 3.6	0.0001
LGE papillary only	6/474 (1.3)	0	6 (3.0)	0.004
LGE pattern				<0.001
None	370/474 (78.1)	254/275 (92.4)	116/199 (58.3)	
Mid-wall	68/474 (14.4)	14/275 (5.1)	54/199 (27.1)	
Subendocardial/transmural	7/474 (1.5)	0	7/199 (3.5)	
Patchy (excluding RV/LV insertions)	29/474 (6.1)	22/275 (11.1)	7/199 (2.6)	
LGE extent (% LV mass)	0	0	0 (0 – 1.9)	<0.001
LVEDV*i* (ml/m^2)	88.4 ± 22.1	79.7 ± 18.1	100.5 ± 21.4	<0.001
LVEF (%)	59.9 ± 7.1	61.1 + 7.1	58.3 ± 6.8	<0.001
LVM*i* (g/m^2)	56.7 ± 15.9	55.0 ± 14.6	59.0 ± 17.3	0.027
MAD presence	322/474 (67.9)	174/275 (63.3)	148/199 (74.4)	0.011
MAD extent (mm)	3.5 (0 – 6)	2.7 (0 – 4.1)	6 (0 – 8)	<0.001
Right atrium area *i* (cm^2/m^2)	11.0 ± 2.6	10.4 ± 2.5	11.7 ± 2.5	<0.001
RVEDV*i* (ml/m^2)	83.6 ± 19.2	77.4 ± 17.2	92.2 ± 18.6	<0.001
RVEF (%)	62.1 ± 6.9	62.4 ± 7.2	61.6 ± 6.4	0.191
Symptoms				<0.001
Asymptomatic	186/474 (39.2)	91/275 (33.1)	95/199 (47.4)	
Chest Pain	49/474 (10.3)	39/275 (14.2)	10/199 (5.0)	
Palpitations	177/474 (37.3)	113/275 (41.1)	64/199 (32.2)	
Dyspnea (NYHA-class: II)	27/474 (5.7)	12/275 (4.4)	15/199 (7.5)	
Unexplained syncope	35/474 (7.4)	20/275 (7.3)	15/199 (7.5)	

Note.—Continuous data reported as mean ± SD or median (IQR), categorical data reported as proportion (percentage). *p*-values obtained from Kruskal-Wallis and χ-squared tests. *i*: indexed for body surface area; LGE: late-gadolinium-

enhancement; LVEDV: left ventricular end-diastolic-volume; LVEF: left ventricular ejection-fraction; LVM: left ventricular mass; MAD: mitral annulus disjunction; RVEDV: right ventricular end-diastolic-volume; RVEF: right ventricular ejection-fraction; SCD: sudden cardiac death.

Table 2 Association between the phenotypic clusters and risk of the study endpoint

Study endpoint	Number of events	Incidence rate (95% CI)	Unadjusted HR (95% CI)	p-value	Adjusted HR (95% CI)	p-value
Total (Cluster 1 & 2)	18	12.0 (7.6 – 19.1)				
Cluster 1	6	6.4 (2.9 – 14.2)	Reference		Reference	
Cluster 2	12	21.5 (12.2 – 37.9)	5.30 (1.79 – 15.74)	0.003	3.79 (1.19 – 12.12)	0.024

Note.— Incidence rate per 1,000 person-years. Cox proportional hazard regression model with hazard ratio (HR) reported. Adjusted for late gadolinium enhancement extent (percentage of left ventricular mass). The study endpoint was defined as a composite of either aborted sudden cardiac death (SCD) or SCD, sustained ventricular tachycardia and unexplained syncope at follow-up.

Figure Legends

Figure 1. Plots showing the grouping of patients based on cluster analysis.

a. Dendrogram showing the hierarchical grouping of patients (blue represents cluster 1 and red represents cluster 2).

b. Plot showing the grouping of patients by principal components (blue represents cluster 1 and red represents cluster 2).

Figure 2. Variable importance plot for identifying the phenotypic clusters. a) Cluster 1. b) Cluster 2. SHapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Variables are stacked vertically in descending order of importance. Each row plot is a summary of the SHAP dependence plot of each variable. Each dot represents a patient's SHAP value plotted horizontally. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low (blue) to high (red). If red points are plotted on the lower side and blue dots are plotted on the higher side, then the risk becomes higher as the value increases.

i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death.

Figure 3. Kaplan-Meier plot for the study endpoint according to clusters.

Survival free from composite endpoint of sustained ventricular tachycardia, (aborted) sudden cardiac death, or unexplained syncope. P =0.001 by log-rank test.

Figure 4. Underpinnings of Arrhythmic Mitral Valve Prolapse by Unsupervised Machine Learning.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Among 474 patients with isolated mitral valve prolapse (MVP) undergoing late gadolinium enhancement (LGE) cardiac MRI (CMR), unsupervised machine learning identified two phenotypic clusters (*left*). Cluster-2 patients showed a higher prevalence of bi-leaflet MVP, greater mitral leaflets displacements, ventricular and atrial sizes, and LGE and mitral annulus disjunction (MAD) extents (*center*). Cluster-2 patients had a 4-fold increased risk of sustained ventricular tachycardia, (aborted) sudden cardiac death (SCD), or unexplained syncope at follow-up (*right*).

SupplementaryNote: Figure 1. Correlations among baseline variables.

Twenty-five most relevant ranked cross-correlations with p-value < 0.05. Negative correlations are represented in red and positive correlations in blue.

Supplementary Figure 2. Visualizing the optimal number of clusters.

A) Optimal number of clusters based on *NbClust* package. B) Optimal number if clusters based on Gap statistic method. C) Optimal number of clusters based on Elbow method.

Supplementary Figure 3. Screen plot.

Principal components 1 and 2 explains 24.76% of the variability in the data.

Supplementary Figure 4. Variable importance plot for identifying the phenotypic clusters in patients without malignant ventricular arrhythmias at baseline.

a) Cluster 1. b) Cluster 2. Shapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Each point on the summary plot is a Shapley value for an individual. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low to high. The variables (features) are ordered according to importance.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death; Ventricular arrhythmia: Ventricular ectopic beats>10,000/24 h and/or non-sustained ventricular tachycardia at baseline.

Supplementary Figure 5. Kaplan-Meier plot for the study endpoint according to clusters in patients without malignant ventricular arrhythmias at baseline

Survival free from composite endpoint of sustained ventricular tachycardia, (aborted) sudden cardiac death, or unexplained syncope in patients without malignant ventricular arrhythmias at baseline. P =0.005 by log-rank test.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Supplementary Table 1. List of baseline variables selected in the study population

Variable name	Variable definition
Active Smoking	Active smoker
Age	Age at baseline (years)
Anterior leaflet displacement	Systolic atrial displacement of the anterior mitral leaflet (mm)
Anterior leaflet length	Length of the anterior mitral leaflet (mm)
BMI	Body mass index (kg/m^2)
Diabetes	History of diabetes mellitus
Family history of SCD	Family history of sudden cardiac death
Hyperlipidemia	History of hyperlipidaemia
Hypertension	History of hypertension
Leaflets involved*	Posterior, Anterior or bi-leaflet prolapse
Left atrium area*	Left atrial area (cm^2)
Left atrium area *i*	Left atrial area indexed for BSA (cm^2/m^2)
LGE extent	Late gadolinium enhancement extent (% of left ventricular mass)
LGE papillary only	Late gadolinium enhancement only in the papillary muscles
LGE pattern	Late gadolinium enhancement - pattern of distribution
LGE segments*	Number of AHA segments showing late gadolinium enhancement
LVEDV*i*	Left ventricular end-diastolic volume indexed for BSA (ml/m^2)
LVEF	Left ventricular ejection fraction (%)
LVESV*i**	Left ventricular end-systolic volume indexed for BSA (ml/m^2)
LVM*i*	Left ventricular mass indexed for BSA (g/m^2)
MAD presence	Presence of mitral annulus disjunction
MAD extent	Longitudinal extent of mitral annulus disjunction (mm)
MVP (bi-leaflet vs. single leaflet)	Bi-leaflet mitral valve prolapse
Posterior leaflet displacement	Systolic atrial displacement of the posterior mitral leaflet (mm)
Posterior leaflet lenght	Posterior mitral leaflet length (mm)
Right atrium_area *i*	Right atrial area indexed for BSA (cm^2/m^2)
Right atrium_area*	Right atrial area (cm^2)
RVEDV*i*	Right ventricular end-diastolic volume indexed for BSA (ml/m^2)
RVEF	Right ventricular ejection fraction (%)
RVESV*i*	Right ventricular end-systolic volume indexed for BSA (ml/m^2)
Sex	Sex
Symptoms	Asymptomatic, chest pain, palpitations, New York Heart Association ≥2

* Excluded from cluster analysis; BSA: body surface area (m^2)

Supplementary Table 2. Baseline characteristics stratified by clusters in the sensitivity analysis

	Total (n = 456)	**Cluster 1** (n=273; 59.9%)	**Cluster 2** (n=183; 40.1%)	***p*-values**
Age (years)	47.0 ± 15.9	46.7 ± 16.0	47.3 ± 15.6	0.821
Sex, female [n, (%)]	220 (48.3)	136 (49.8)	84 (45.9)	0.412
Body mass index (kg/m^2)	23.0 ± 3.6	23.5 ± 3.6	22.2 ± 3.4	<0.001
Active smoking	18 (4.0)	9 (3.3)	9 (4.9)	0.383
Diabetes	5 (1.1)	3 (1.1)	2 (1.1)	0.995
Family history of SCD	19 (4.2)	16 (5.9)	3 (1.6)	0.027
Hyperlipidemia	59 (12.9)	45 (16.5)	14 (7.7)	0.006
Hypertension	67 (14.7)	48 (17.6)	19 (10.4)	0.033
Mitral valve prolapse				<0.001
Single leaflet	243 (53.3)	214 (78.4)	29 (15.9)	
Bi leaflet	213 (46.7)	59 (21.6)	154 (84.2)	
Anterior leaflet displacement (mm)	2 (0 – 4.5)	0 (0 – 2)	4.9 (3 – 7)	<0.001
Posterior leaflet displacement (mm)	4 (3 – 7)	3.1 (2.5 – 4)	7 (4.2 – 9.6)	<0.001
Left atrium area ***i*** (cm^2/m^2)	12.6 ± 3.3	11.5 ± 2.6	14.1 ± 3.6	<0.001
LGE papillary only	6 (1.3)	0	6 (3.3)	0.003
LGE pattern				<0.001
None	362 (79.4)	251 (91.9)	111 (60.7)	
Mid-wall	59 (12.9)	14 (5.1)	45 (24.6)	
Subendocardial/transmural	7 (1.5)	0	7 (3.8)	
Patchy (excluding RV/LV insertions)	28 (6.1)	8 (2.9)	20 (10.9)	
LGE extent (% LV mass)	0	0	0 (0 – 1.5)	<0.001
LV-EDV*i* (ml/m^2)	88.3 ± 22.0	79.2 ± 17.7	101.7 ± 20.8	<0.001
LV-EF (%)	60.0 ± 7.1	61.3 ± 7.1	58.1 ± 6.7	<0.001
LVM***i***	56.6 ± 15.7	54.4 ± 14.2	60.0 ± 17.2	0.002
MAD presence	309 (67.8)	173 (63.4)	136 (74.3)	0.014
MAD extent (mm)	3.5 (0 – 6)	2.7 (0 – 4.1)	6 (0 – 8)	<0.001
Right atrium area ***i*** (cm²/m²)	11.0 ± 2.5	10.4 ± 2.5	11.8 ± 2.4	<0.001
RV-EDV*i* (ml/m^2)	83.6 ± 19.2	76.9 ± 16.7	93.6 ± 18.3	<0.001
RV-EF (%)	62.1 ± 7.0	62.6 ± 7.1	61.4 ± 6.6	0.047
Symptoms				0.001
Asymptomatic	179 (39.3)	90 (33.0)	89 (48.6)	
Chest Pain	48 (10.5)	39 (14.3)	9 (4.9)	
Palpitations	172 (37.7)	112 (41.0)	60 (32.8)	
Dyspnea (NYHA-class: II)	26 (5.7)	15 (5.5)	11 (6.0)	
Unexplained syncope	31 (6.8)	17 (6.2)	14 (7.7)	
Ventricular arrhythmias at ECG monitoring (VEBs >10,000/24 h and/or NSVT)	85 (18.6)	46 (16.9)	39 (21.3)	0.230

EDV: end-diastolic volume; EF: ejection fraction; *i:* indexed for body surface area; LGE: late-gadolinium-enhancement; MAD: mitral annulus disjunction; LV: left ventricle; LVM – left ventricular mass; NSVT: non-sustained ventricular tachycardia; RV: right ventricle; SCD: sudden cardiac death; VEBs: ventricular ectopic beats.

p-values obtained from Kruskal-Wallis and χ-squared tests.

820 Jorie Blvd., Suite 200, Oak Brook, IL, 60523, 630-481-1071, rcti@rsna.org

Supplementary Table 3. The association between the phenotypic clusters and risk of the study endpoint in patients without malignant ventricular arrhythmias at baseline

Study endpoint	Number of events	Incidence rate	Unadjusted HR (95% CI)	p-value	Adjusted HR (95% CI)	p-value
Total (Cluster 1 & 2)	13	8.97 (5.21 – 15.45)				
Cluster 1	5	5.33 (2.22 – 12.81)	Reference		Reference	
Cluster 2	8	15.65 (7.82 – 31.28)	5.18 (1.47 – 18.19)	0.010	4.03 (1.06 – 15.32)	0.041

CI – confidence interval

Incidence rate per 1,000 person-years.

Cox proportional hazard regression model with hazard ratio (HR) reported.

Adjusted for late gadolinium enhancement extent (percentage of left ventricular mass).

The study endpoint was defined as a composite of either aborted sudden cardiac death (SCD) or SCD, sustained ventricular tachycardia and unexplained syncope at follow-up.

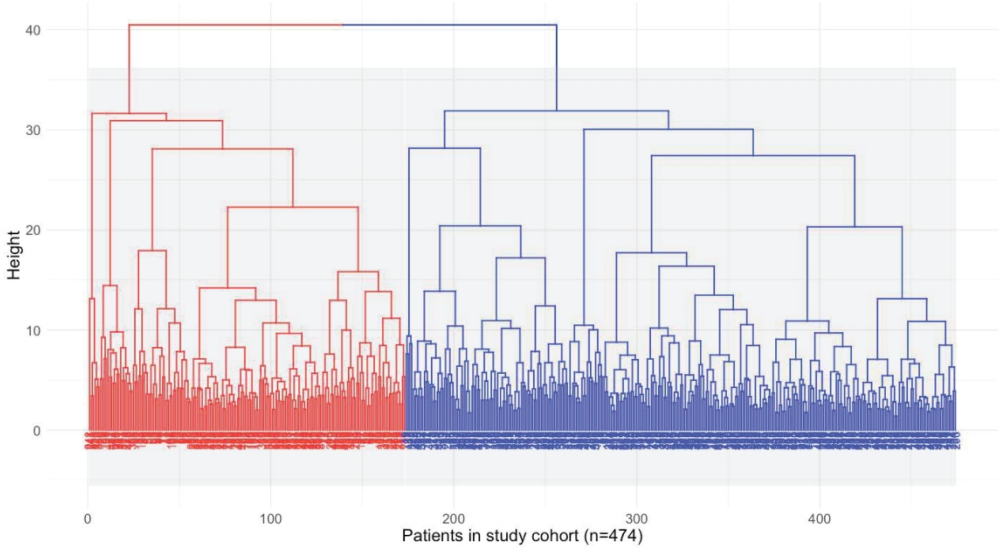


Figure 1a. Plots showing the grouping of patients based on cluster analysis.
a. Dendrogram showing the hierarchical grouping of patients (blue represents cluster 1 and red represents cluster 2).

497x286mm (300 x 300 DPI)

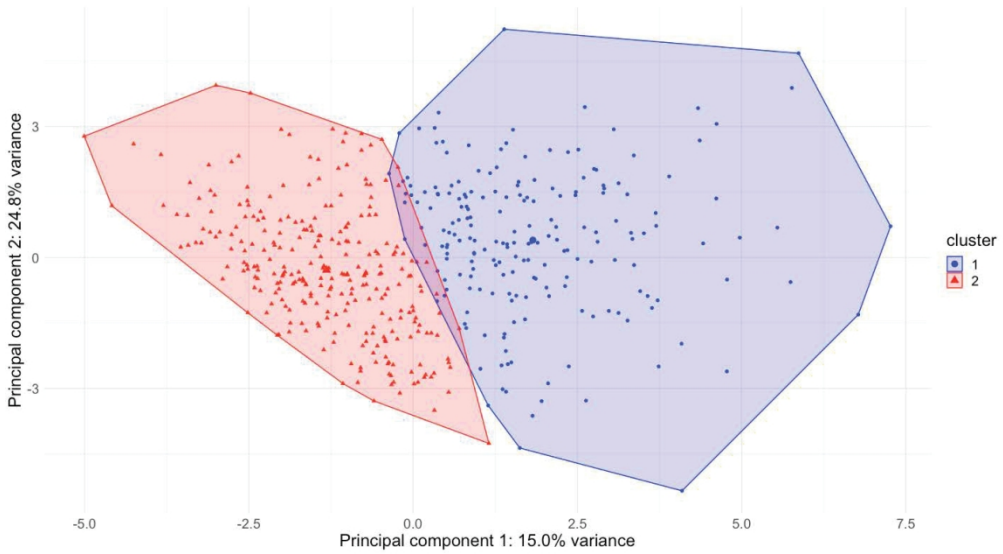


Figure 1b. Plots showing the grouping of patients based on cluster analysis.
b. Plot showing the grouping of patients by principal components (blue represents cluster 1 and red represents cluster 2).

496x285mm (300 x 300 DPI)

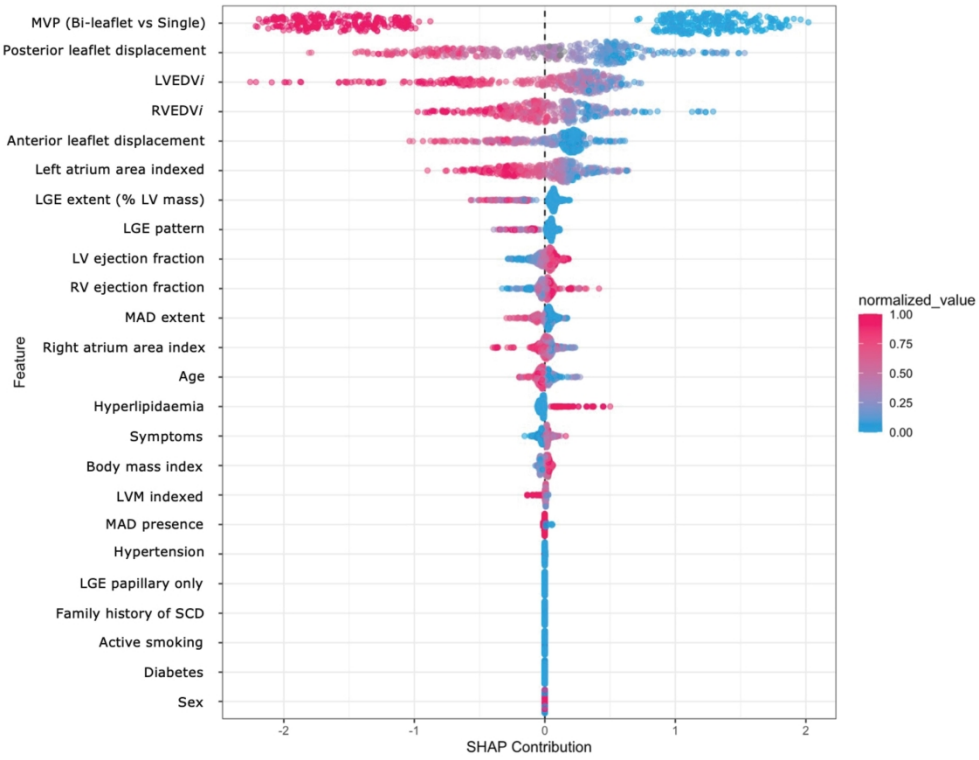


Figure 2a. Variable importance plot for identifying the phenotypic clusters. a) Cluster 1. SHapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Variables are stacked vertically in descending order of importance. Each row plot is a summary of the SHAP dependence plot of each variable. Each dot represents a patient's SHAP value plotted horizontally. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low (blue) to high (red). If red points are plotted on the lower side and blue dots are plotted on the higher side, then the risk becomes higher as the value increases.

i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death.

184x150mm (600 x 600 DPI)

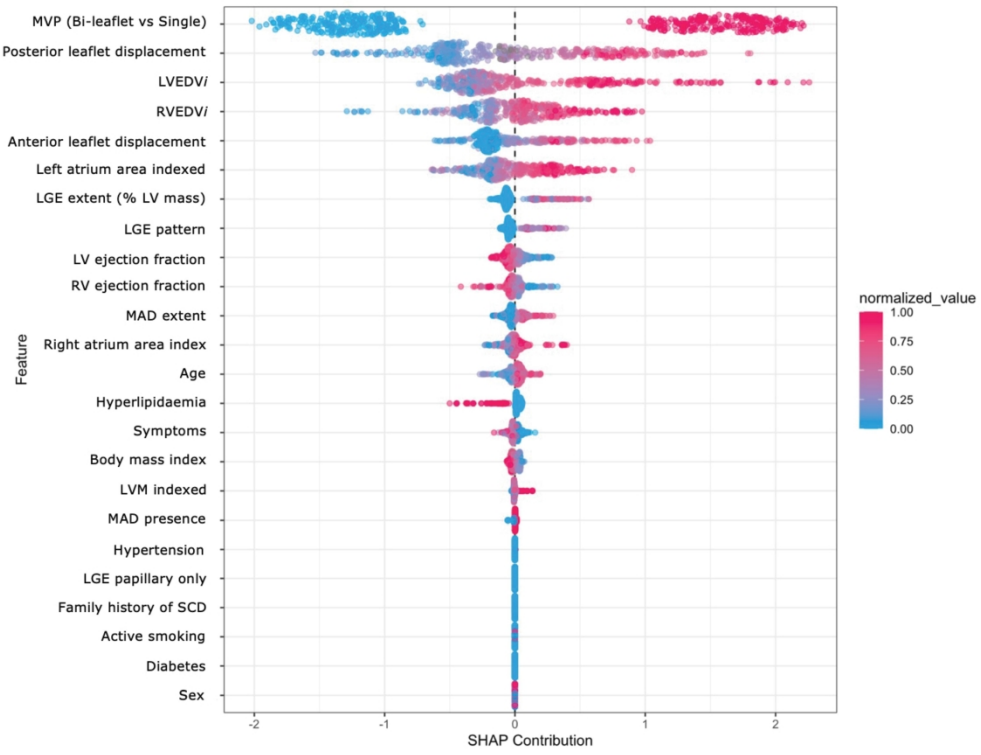


Figure 2 (b). Variable importance plot for identifying the phenotypic clusters. b) Cluster 2. SHapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Variables are stacked vertically in descending order of importance. Each row plot is a summary of the SHAP dependence plot of each variable. Each dot represents a patient's SHAP value plotted horizontally. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low (blue) to high (red). If red points are plotted on the lower side and blue dots are plotted on the higher side, then the risk becomes higher as the value increases.

i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death.

184x149mm (600 x 600 DPI)

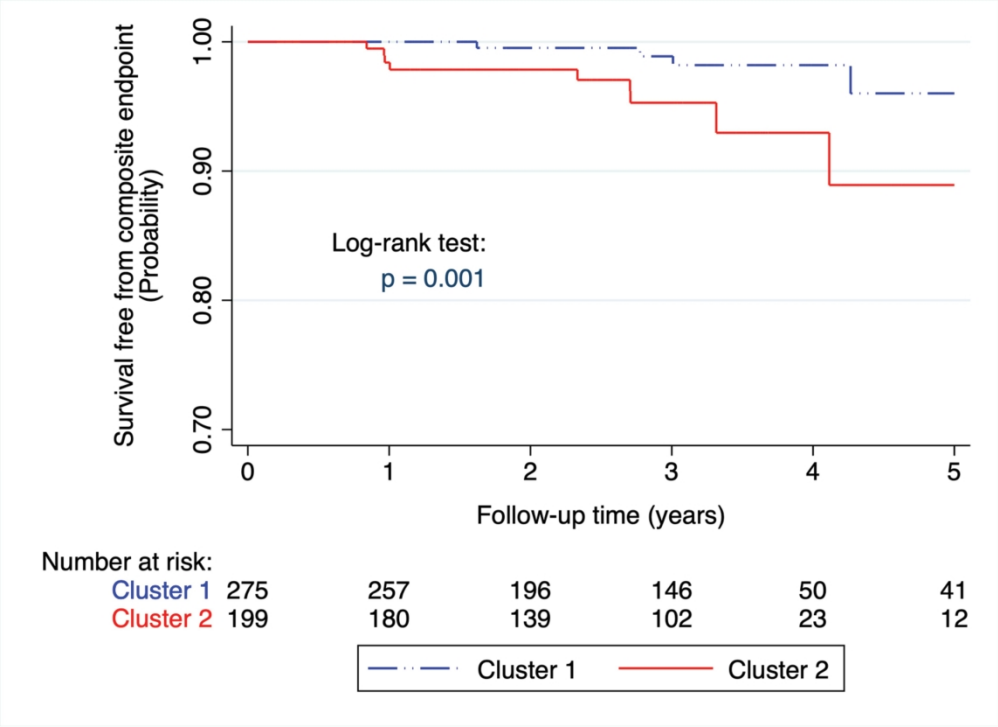


Figure 3. Kaplan-Meier plot for the study endpoint according to clusters.
Survival free from composite endpoint of sustained ventricular tachycardia, (aborted) sudden cardiac death, or unexplained syncope. P =0.001 by log-rank test.

184x134mm (600 x 600 DPI)

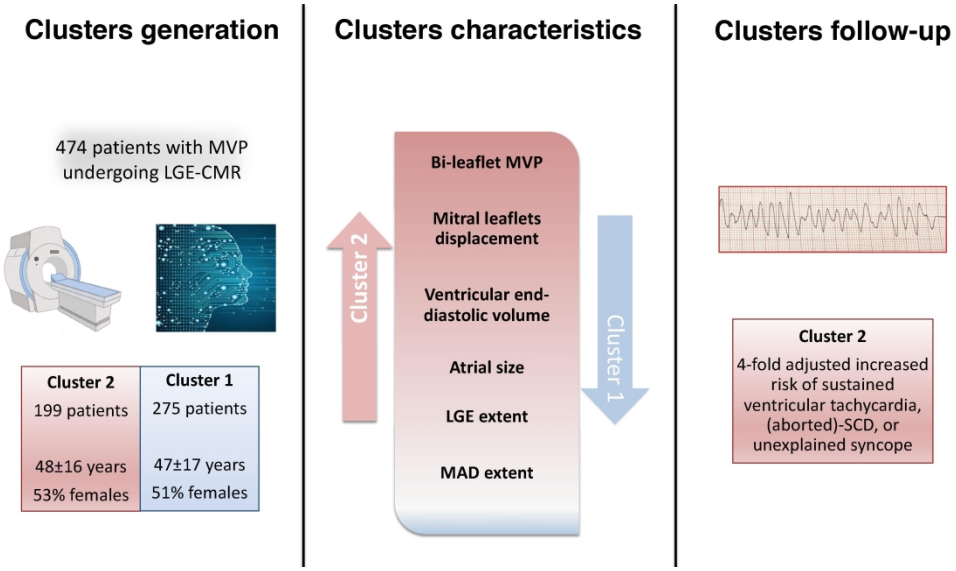
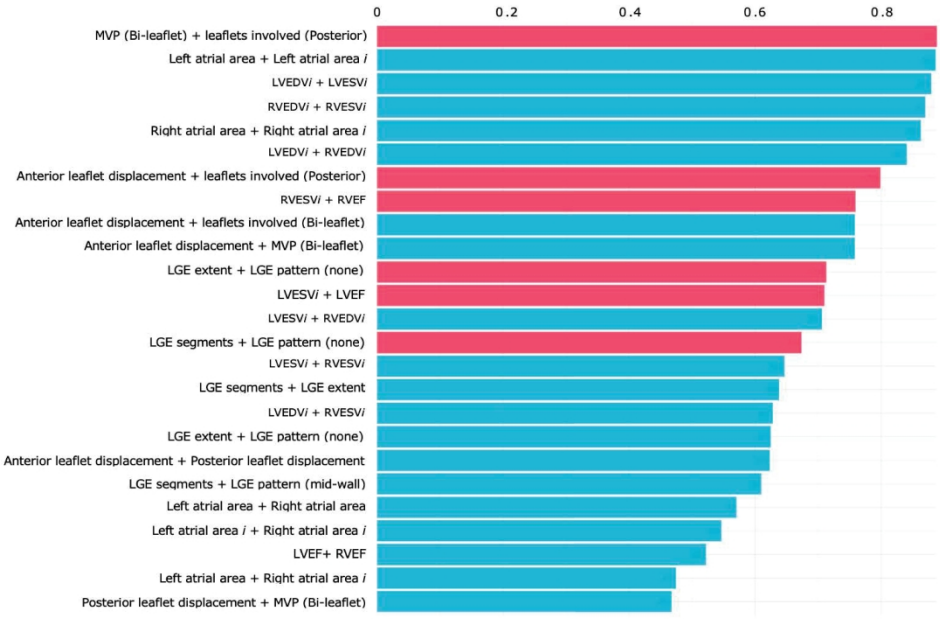


Figure 4. Underpinnings of Arrhythmic Mitral Valve Prolapse by Unsupervised Machine Learning.

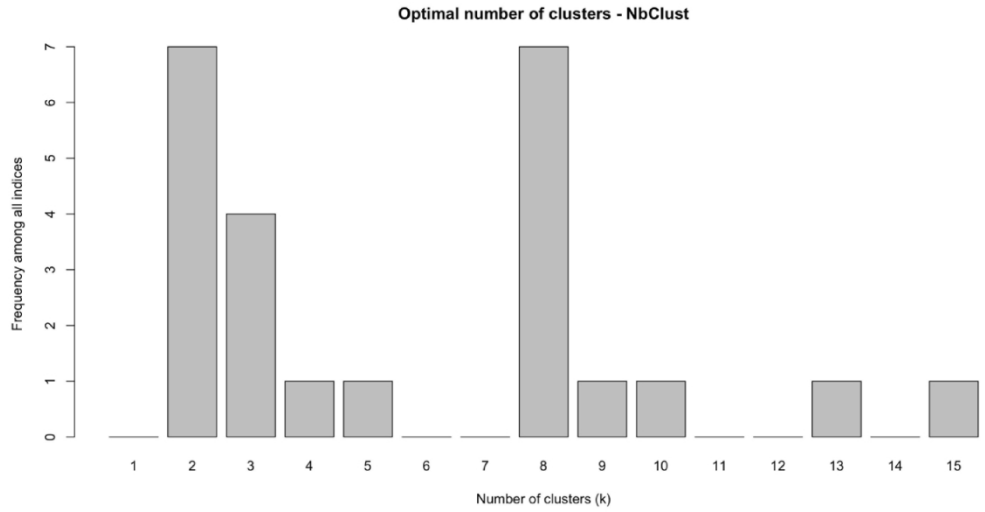
Among 474 patients with isolated mitral valve prolapse (MVP) undergoing late gadolinium enhancement (LGE) cardiac MRI (CMR), unsupervised machine learning identified two phenotypic clusters (left). Cluster-2 patients showed a higher prevalence of bi-leaflet MVP, greater mitral leaflets displacements, ventricular and atrial sizes, and LGE and mitral annulus disjunction (MAD) extents (center). Cluster-2 patients had a 4-fold increased risk of sustained ventricular tachycardia, (aborted) sudden cardiac death (SCD), or unexplained syncope at follow-up (right).

159x89mm (600 x 600 DPI)



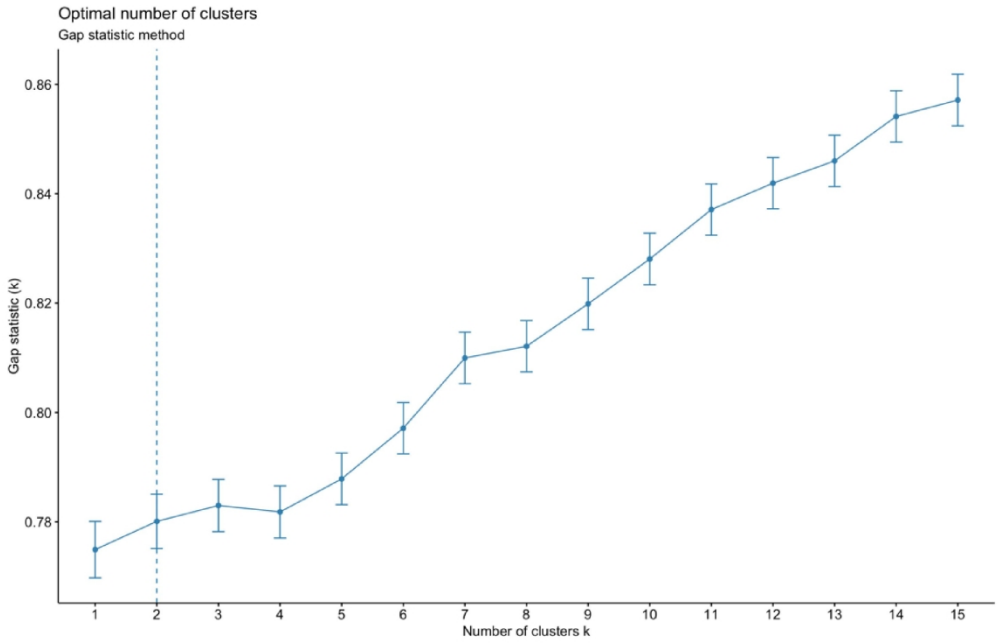
Supplementary Figure 1. Correlations among baseline variables.
Twenty-five most relevant ranked cross-correlations with p-value < 0.05. Negative correlations are represented in red and positive correlations in blue.

288x189mm (600 x 600 DPI)



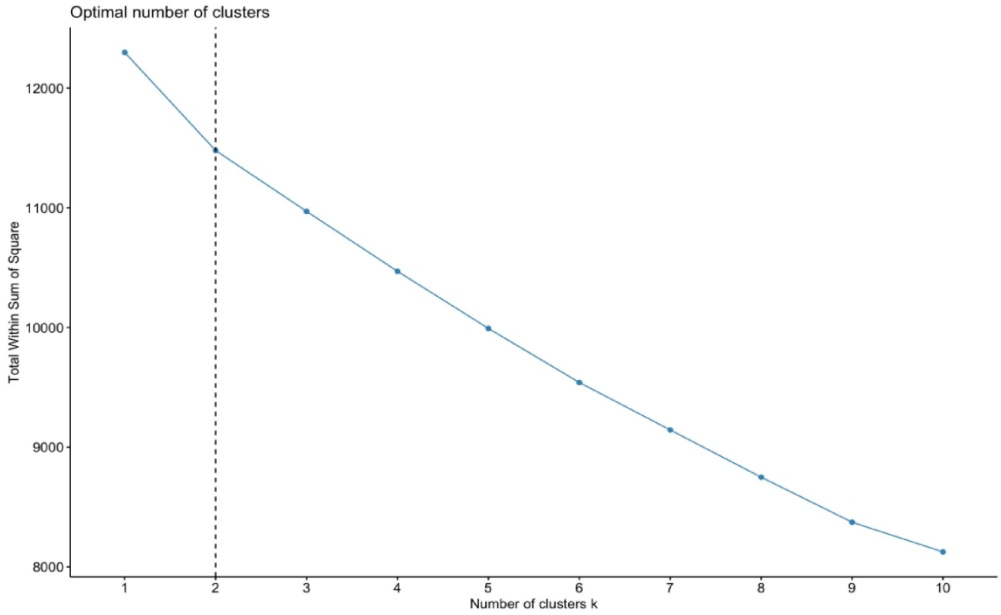
Supplementary Figure 2 (a). Visualizing the optimal number of clusters.
A) Optimal number of clusters based on NbClust package.

139x75mm (600 x 600 DPI)



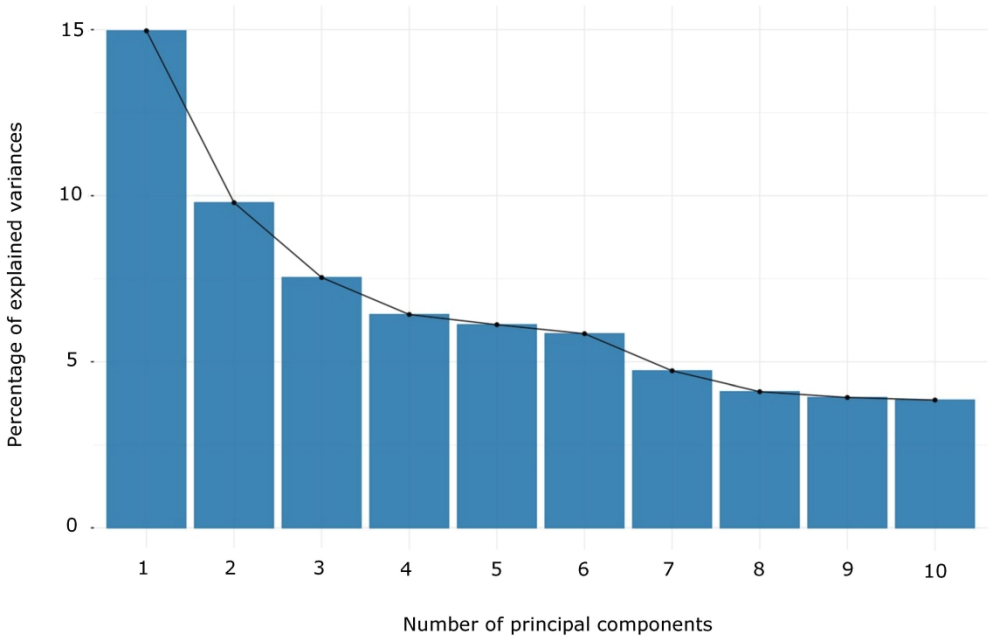
Supplementary Figure 2 (b). Visualizing the optimal number of clusters.
B) Optimal number if clusters based on Gap statistic method.

122x79mm (600 x 600 DPI)



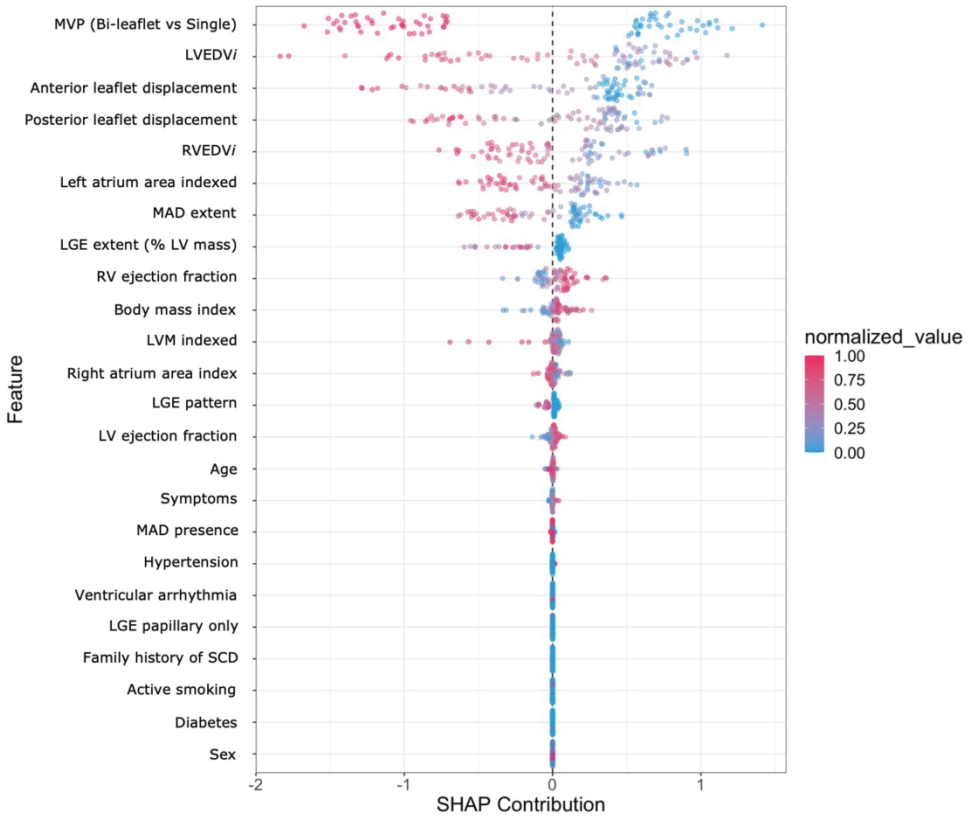
Supplementary Figure 2 (c). Visualizing the optimal number of clusters.
C) Optimal number of clusters based on Elbow method.

129x80mm (600 x 600 DPI)



Supplementary Figure 3. Screen plot.
Principal components 1 and 2 explains 24.76% of the variability in the data.

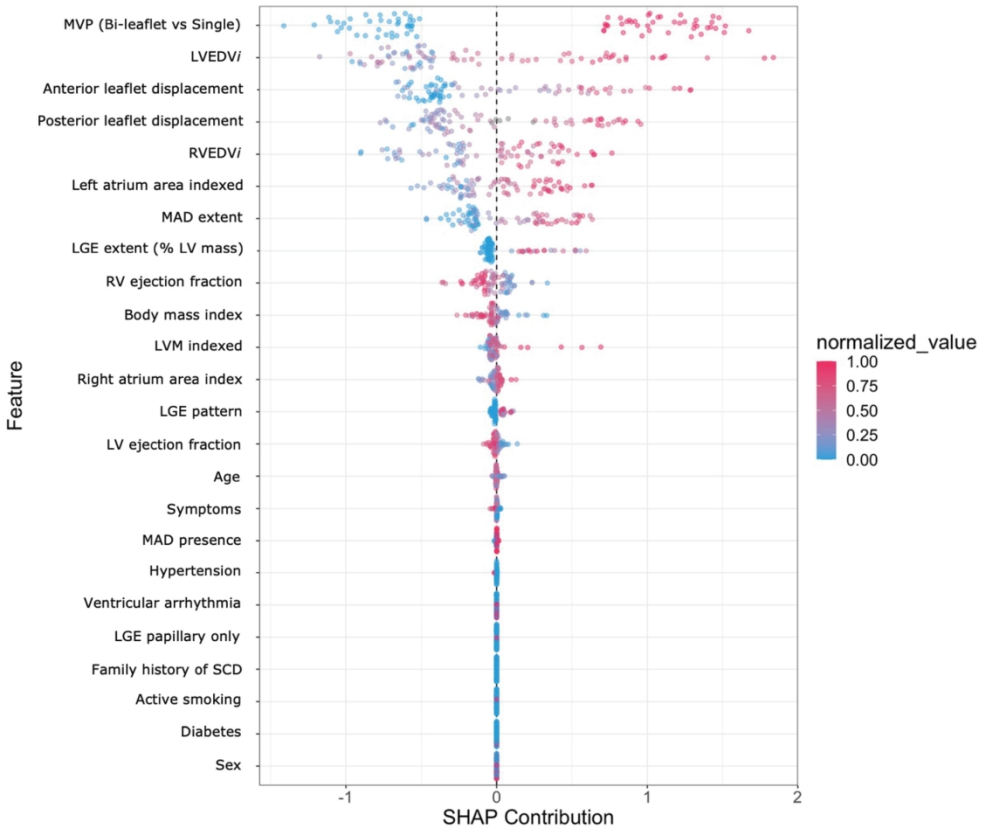
158x105mm (600 x 600 DPI)



Supplementary Figure 4.a
Variable importance plot for identifying the phenotypic clusters in patients without malignant ventricular arrhythmias at baseline.
a). Cluster-1

Shapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Each point on the summary plot is a Shapley value for an individual. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low to high. The variables (features) are ordered according to importance.i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death; Ventricular arrhythmia: Ventricular ectopic beats>10,000/24 h and/or non-sustained ventricular tachycardia at baseline.

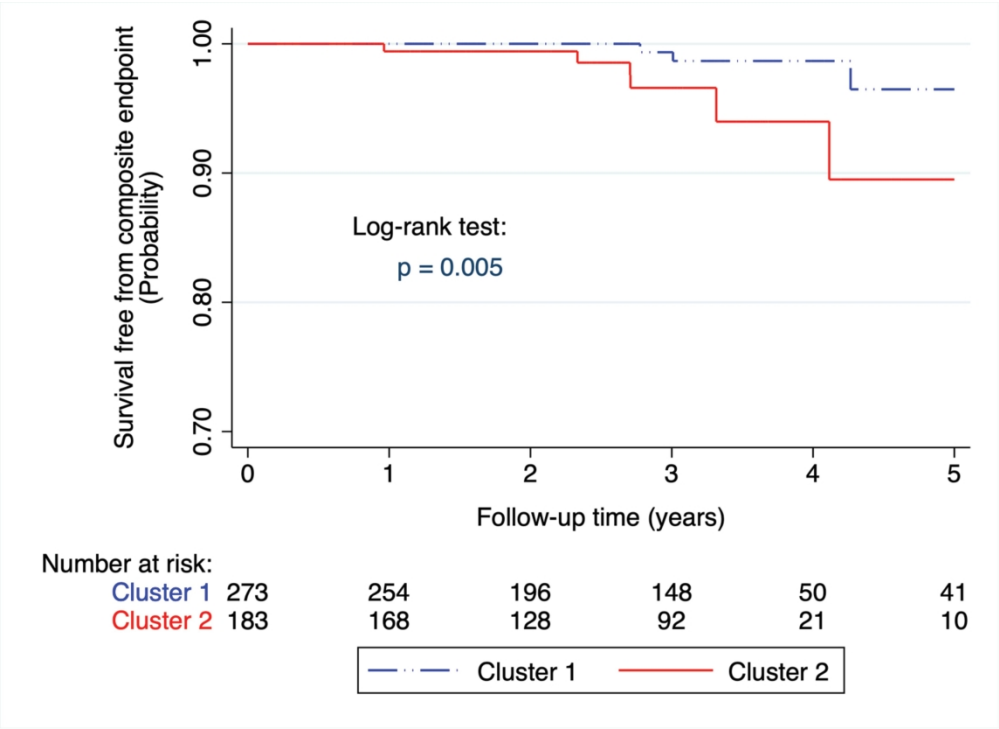
184x163mm (600 x 600 DPI)



Supplementary Figure 4b.
Variable importance plot for identifying the phenotypic clusters in patients without malignant ventricular arrhythmias at baseline.
b). Cluster-2

Shapley Additive exPlanations (SHAP) summary plot combines variable (feature) importance with variable effects. Each point on the summary plot is a Shapley value for an individual. The position on the y-axis is determined by the variable (feature) and on the x-axis by the Shapley value. The color represents the value from low to high. The variables (features) are ordered according to importance.i: indexed for body surface area; LGE: late-gadolinium-enhancement; LV: left ventricle; LVEDV: left ventricular end-diastolic-volume; LVM: left ventricular mass; MAD: mitral annulus disjunction; MVP: mitral valve prolapse: RV: right ventricle; RVEDV: right ventricular end-diastolic-volume; SCD: sudden cardiac death; Ventricular arrhythmia: Ventricular ectopic beats>10,000/24 h and/or non-sustained ventricular tachycardia at baseline.

184x168mm (600 x 600 DPI)



Supplementary Figure 5. Kaplan-Meier plot for the study endpoint according to clusters in patients without malignant ventricular arrhythmias at baseline
Survival free from composite endpoint of sustained ventricular tachycardia, (aborted) sudden cardiac death, or unexplained syncope in patients without malignant ventricular arrhythmias at baseline. P =0.005 by log-rank test.

184x134mm (600 x 600 DPI)

Arrhythmic Mitral Valve Prolapse Phenotype: An Unsupervised Machine Learning Analysis Using a Multicenter Cardiac MRI Registry

Key Result

In patients with mitral valve prolapse (MVP), two cardiac MRI-based phenotypic clusters with distinct arrhythmic outcomes were identified using unsupervised machine learning.

Patients:

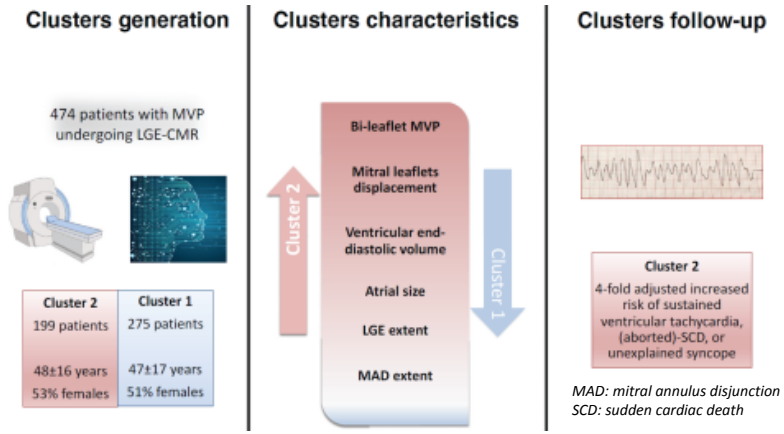
- 474 patients with MVP without significant mitral regurgitation or left ventricular dysfunction undergoing late gadolinium enhancement (LGE) cardiac MRI

Methods:

- An unsupervised data-driven hierarchical k-mean algorithm was used to identify phenotypic clusters.
- The association between clusters and the study endpoint (composite of sustained ventricular tachycardia, [aborted] sudden cardiac death, or unexplained syncope) was assessed.

Results

- Among the two phenotypic clusters identified, cluster-2 patients (n=199/474, 42%) had more severe mitral valve degeneration, left and right heart chamber remodeling, and myocardial fibrosis.
- Cluster-2 patients showed a higher risk of developing the arrhythmic endpoint (HR: 3.91) over a median follow-up of 3.3 years.



FIRST AUTHOR last name + first initials et al. Published Online: DATE, 2024
DOI
Radiology: Cardiothoracic Imaging