ESC European Society of Cardiology

European Journal of Heart Failure (2025) **27**, 3205–3218
doi:10.1002/ejhf.70040

RESEARCH ARTICLE

Impact of genotype–phenotype associations on prognosis in dilated cardiomyopathy

Sophie L.V.M. Stroeks[1,2,3,4], Ping Wang[3,4], Marco Merlo[4,5], Steven Muller[4,6,7], Alessia Paldino[4,5], Nerea Mora-Ayestaran[4,8,9], Max Jason[10], Matteo dal Ferro[4,5], Carola Pio Loco detto Gava[4,5], Fernando Dominguez[4,8,9], Esther Gonzalez-Lopez[4,8,9], Arthur van den Wijngaard[3], Max F.G.H.M. Venner[1], Maurits Sikking[1], Michiel Minten[1], Bastien Nihant[1], Nina Beelen[1], Sharon Graw[10], Kristen Medo[10], Bart de Koning[3,4], Matthew Taylor[10], J. Peter van Tintelen[4,11], Luisa Mestroni[10], Gianfranco Sinagra[4,5], Anneline S.J.M. te Riele[4,6,7], Pablo Garcia-Pavia[4,8,9], Stephane Heymans[1,2,4], and Job A.J. Verdonschot[1,3,4]*

[1]Department of Cardiology, Maastricht University, Cardiovascular Research Institute Maastricht (CARIM), Maastricht, The Netherlands; [2]KU Leuven, Cardiovascular Sciences, Leuven, Belgium; [3]Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands; [4]Members of the European Reference Network for rare, low prevalence and complex diseases of the heart (ERN GUARD-Heart); [5]Department of Cardiology, Azienda Sanitaria Universitaria Giuliano-Isontina (ASUGI), University of Trieste, Trieste, Italy; [6]Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands; [7]Netherlands Heart Institute, Utrecht, The Netherlands; [8]Department of Cardiology, Hospital Universitario Puerta de Hierro, IDIPHISA, CIBERCV, Madrid, Spain; [9]Centro Nacional de Investigaciones Cardiovasculares (CNIV), Madrid, Spain; [10]CU Cardiovascular Institute, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; and [11]Department of Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

Received 23 May 2025; revised 5 August 2025; accepted 22 August 2025; online publish-ahead-of-print 12 September 2025

Aims Dilated cardiomyopathy (DCM) has a monogenic aetiology in up to 40% of patients. Understanding the spectrum of genotype–phenotype associations in DCM is crucial for risk stratification and personalized treatment. We aimed to (i) characterize genotype-specific features, (ii) evaluate whether phenotype-based clustering reflects underlying genotype, and (iii) compare the prognostic value of genotype- versus phenotype-based approaches.

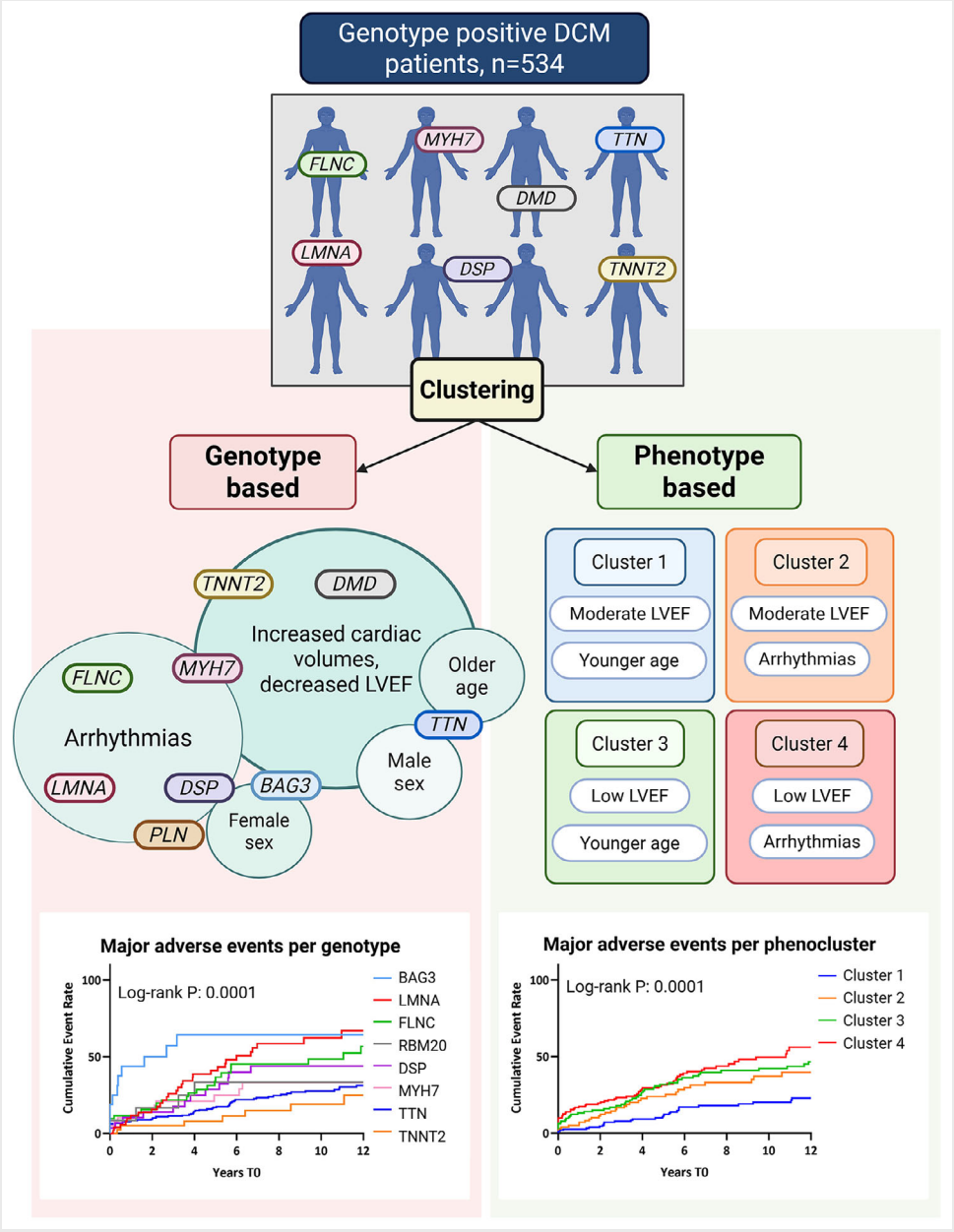
Methods and results A multicentre cohort of 534 DCM patients with a (likely) pathogenic variant were grouped by genotype (genotype-first approach) and clustered by clinical phenotype (phenotype-first approach). We compared clinical characteristics, identified genotype–phenotype associations, and evaluated outcomes, including all-cause mortality, heart failure hospitalization, heart transplantation, and malignant ventricular arrhythmias. Using the genotype-first approach, significant genotype–phenotype associations were found for 10 genes. *FLNC, LMNA, DSP*, and *PLN* variants were linked to arrhythmias. *BAG3, TNNT2, DMD*, and *TTN* were associated with increased cardiac volumes and decreased left ventricular ejection fraction (LVEF). Clustering identified four phenotypic clusters: (1) young, moderately reduced LVEF; (2) arrhythmias, moderate reduced LVEF; (3) low LVEF; (4) arrhythmias, low LVEF. There were no clear correlations between phenotypic clusters and genotype. The genotype-first approach showed that *LMNA, FLNC,* and *BAG3* variants had the highest risk for heart failure and arrhythmogenic adverse outcomes. The phenotype-first approach indicated that clusters 3 and 4 were associated with the worst prognosis. Overall, genotype was the strongest predictor of outcome.

Conclusions Patients with a genetic form of DCM exhibit clinical and genetic heterogeneity. Genotype-based risk stratification is more accurate compared to a phenotype-first approach, highlighting the importance of broad genetic screening among patients with DCM. Additionally, gene-specific risk prediction should become more prominent in current guidelines on management of genetic DCM patients.

*Corresponding author. Department of Cardiology, Maastricht University, Cardiovascular Research Institute Maastricht (CARIM), Maastricht, The Netherlands. Tel: +31 43 3884083, Email: job.verdonschot@mumc.nl

[Correction added on 23 September 2025, after first online publication: Carola Pio Loco detto Gava's last name and linked affiliations have been corrected in this version.]

Graphical Abstract



Prediction of clinical outcomes in genetic dilated cardiomyopathy (DCM). LVEF, left ventricular ejection fraction.

Keywords Dilated cardiomyopathy • Genetics • Prognosis • Risk prediction

Introduction

Dilated cardiomyopathy (DCM) is a severe cardiac disease characterized by systolic dysfunction and (bi-)ventricular dilatation unexplained solely by abnormal loading conditions or coronary artery disease.[1] Up to 40% of patients with DCM have a monogenetic aetiology, in which truncating variants in *TTN* are the most prevalent.[2] The finding of a genetic cause is important, as it provides the possibility for gene-specific clinical care in patients with DCM, including risk prediction and intervention, and cascade screening.[3–5] The

latest guidelines from the European Society of Cardiology (ESC) and American Heart Association (AHA) both include recommendations for preventive implantable cardioverter-defibrillator (ICD) therapy in patients with DCM and a disease-causing variant in a high-risk gene such as *LMNA, FLNC, RBM20, PLN* and *DSP*.[6,7]

Patient classification solely based on genotype showed stronger predictive value for adverse outcome, compared to classified patients based on their clinical diagnosis.[8] However, significant heterogeneity in clinical characteristics exists even within the same (high-risk) genotype, affecting disease trajectory and risk stratification. This variability suggests that phenotype-based approaches, such as clustering, might provide additional insights into clinical subgroups when performed within the group of patients with a genetic form of DCM. Additional knowledge on the role of the phenotype in genetic high-risk subgroups, can provide the opportunity to optimize clinical decision-making such as device implantation.

While genotype-specific models, such as those for *LMNA* or *PLN*, have improved risk prediction, these tools are available for only a few genes and are typically not generalizable across genotypes, due to differences in disease expression, arrhythmia burden, and structural remodelling. For the majority of genes, especially rare genotypes, such models are unavailable and unlikely to be developed due to limited cohort sizes. In this context, phenotype-based clustering may provide complementary value by identifying shared clinical profiles that transcend genotype, offering an alternative framework for risk stratification in under-characterized or ultra-rare genetic subtypes. This approach may ultimately support more accurate prognosis and treatment planning where genotype-based calculators are not available.

Clustering, an unsupervised machine learning method, can be used to identify patterns in clinical data and groups of patients based on phenotypic similarities.[9–11] We have previously explored phenotypic clustering in mixed DCM cohorts, including both genetic and non-genetic cases, demonstrating that clinical heterogeneity exists across the disease spectrum but patients subgroups can be identified.[12] However, no study has systematically assessed the utility of a phenotype-first approach in an exclusively genotype-positive DCM cohort, despite the clinical differences within specific genotypes.

By utilizing a large, multicentre, genotype-positive cohort we compared a genotype-first approach (assessing genotype–phenotype associations) and a phenotype-first approach (clustering patients based on clinical characteristics) with the following three aims: (i) to characterize genotype-specific clinical features, (ii) to evaluate whether an unsupervised phenotype-based clustering can effectively predict the underlying genotype, and (iii) to compare the prognostic utility of genotype-based versus phenotype-based clustering.

Methods

Study population

This study was a multicentre, observational, longitudinal investigation focusing on individuals diagnosed with genotype positive DCM. Participants were enrolled from inherited cardiac diseases and heart failure units at five international hospitals (Maastricht University Medical Centre [MUMC+], Maastricht, The Netherlands; Hospital Universitario Puerta de Hierro Majadahonda, Majadahonda, Madrid, Spain; United Hospitals of Trieste University Hospital, Trieste, Italy; Utrecht University Medical Center, Utrecht, The Netherlands; and University of Colorado Anschutz Medical Campus, Aurora, CO, USA).

The cohort from Maastricht, The Netherlands, was derived from the Maastricht Cardiomyopathy Registry which actively enrols individuals displaying symptoms of heart failure or those who have undergone cardiac and/or genetic evaluation. The cohort from Utrecht, The Netherlands, was derived from the UNRAVEL registry of cardiomyopathy patients who underwent genetic evaluation.[13] The Italian participants were included from the Heart Muscle Disease Registry of the Trieste Registry, the Spanish cohort was selected from the Hospital Puerta de Hierro's Inherited Cardiac Diseases Unit in Madrid, and the cohort from the United States was identified from the Familial Cardiomyopathy Registry.

Diagnosis of DCM followed the criteria set by the World Health Organization and the most recent suggestions from the ESC.[6] Eligible patients showed a left ventricular ejection fraction (LVEF) of <50%, without significant coronary artery occlusion on coronary angiography, absence of pericardial, congenital heart, or pulmonary heart disease, and no active myocarditis. Where applicable, patients were treated according to the current ESC and AHA/American College of Cardiology (ACC) guidelines,[6,14] which included guideline-directed medication at the maximum tolerated doses and device therapy. Only patients aged >16 years at the time of diagnosis were included. All included patients were probands. Data at participating centres were extracted from clinical records using a uniform methodology. The study was performed according to the Declaration of Helsinki and was approved by the Medical Ethics Committee of all participating institutions. All individuals provided written informed consent.

Clinical evaluation of patients with dilated cardiomyopathy

Baseline evaluation was defined as the first outpatient clinic visit of a genetic DCM patient. In alignment with the diagnostic procedures at each participating medical centre, these individuals underwent comprehensive genetic testing following genetic counselling. The evaluation also included a thorough physical examination, blood tests, a 12-lead electrocardiogram (ECG), 24-h Holter monitoring, an echocardiogram, cardiac magnetic resonance (CMR) and planned regular follow-ups. Key information collected at first evaluation included demographic details, the patient's age when diagnosed, their New York Heart Association (NYHA) functional classification, systolic and diastolic blood pressure, and any family history of DCM. Additionally, data regarding peripartum cardiomyopathy and skeletal myopathy were collected.

Genetic testing and variant classification

Participating individuals received genetic testing using targeted next-generation sequencing panels. All patients underwent genetic testing from 2014 onwards. All patients were at least tested for 19 genes most robustly associated with DCM (including *ACTC1*, *BAG3*, *DMD*, *DES*, *DSP*, *EMD*, *FLNC*, *LMNA*, *MYH7*, *NEXN*, *PLN*, *RBM20*, *SCN5A*, *TTN*, *TNNC1*, *TNNI3*, *TNNT2*, *TPM1*, and *VCL*). The specific gene panels used by each centre per time period are depicted in online supplementary *Methods*.

The classification of all genetic variants followed the standards set by the American College of Medical Genetics and Genomics (ACMG) guidelines.[15] To guarantee uniformity and accuracy in variant

Table 1 Phenotypic domains and corresponding clinical variables

Phenotypic domain	Clinical variables
Demographics	Age of diagnosis[a], sex[a]
Disease modifiers	Genetic LP/P variant[a], familial disease[a], BMI[b,c], alcohol comsumption[b,c], hypertension[b,c], diabetes[b,c], thyroid disease[b,c], COPD[b,c], anthracycline/ChemoRx[b,c], exercise[b,c], skeletal myopathy[b], peripartum cardiomyopathy[b]
Physical characteristics	NYHA class[a], heart rate[a,c], systolic blood pressure[a], diastolic blood pressure[b]
Echocardiography	LV ejection fraction[a], LV end-diastolic diameter[a,c], LV end-diastolic diameter index[b], LV end-systolic diameter[a], left atrial volume index[e], LV mass index[e], E/A ratio[d], E/e' ratio[d], posterior wall thickness[a]
ECG/Holter	Atrial fibrillation[a], non-sustained ventricular tachycardia[a], left bundle branch block[a], out-of-hospital cardiac arrest[a], atrioventricular block[a], PVC[a], low ECG voltage in the peripheral leads[a]
Cardiac magnetic resonance	LV mass index[e], LV end-diastolic volume index[e], LV end-systolic volume index[e], LV stroke volume index[e], LV ejection fraction[d], late gadolinium enhancement[d]

BMI, body mass index; ChemoRx, chemoradiation therapy; COPD, chronic obstructive pulmonary disease; ECG, electrocardiogram; LP/P, likely pathogenic/pathogenic; LV, left ventricular; NYHA, New York Heart Association; PVC, premature ventricular complex.
[a]Clinical variables used in the missing data imputation and the clustering.
[b]Clinical variables used in the missing data imputation, but not in the clustering.
[c]Clinical variables not present in details (*Tables 3* and *4* and *v*-test results).
[d]Clinical variables had large missing data, neither used in missing data imputation, nor for clustering, but reported descriptive summary on raw data with missing anyway.
[e]Clinical variables had large missing data, neither used in missing data imputation, nor for clustering, and not reported in detail.

interpretation, specialists in molecular genetics based in Maastricht conducted additional reviews of all variants. Only those individuals whose genetic testing revealed pathogenic or likely pathogenic variants (classified as P/LP; class 4 or 5) were considered to have received a genetic diagnosis and were thus identified as genotype-positive.

Outcomes

To compile data on overall mortality, heart transplantation (HTx), episodes of malignant ventricular arrhythmia (MVA), and heart failure hospitalizations (HFH), the research utilized information from the patients' health records, municipal demographic databases, or calls with primary care providers. Patients who did not experience any of these events were marked as censored on the date of their most recent follow-up. End date of follow-up was 1 February 2024. MVAs were specifically identified as instances of non-lethal ventricular fibrillation (with or without the intervention of an ICD shock), haemodynamically unstable prolonged ventricular tachycardia (VT), prolonged VT that triggered an appropriate ICD shock, or sudden cardiac death (SCD). Prolonged VT was defined as sustained VT lasting ≥30 s or requiring termination due to haemodynamic instability. The study primary objective was to assess the combined incidence of all-cause mortality, HFH, HTx, and MVA. Secondary outcomes were focused on heart failure incidents (including HFH and HTx) and arrhythmic events (MVA exclusively).

Data processing

Table 1 includes all clinical parameters which were measured in the genotype-positive DCM patients. We selected clinical parameters with less than 25% missing data to create a complete dataset for clustering and *v*-test, with missing values imputed by the principal component method 'factorial analysis for mixed data' (FAMD).[16]

Phenotypic clustering

To construct a reliable clustering model, we selected 17 clinical variables with low multicollinearity and less than 25% missingness (*Table 1*). Excluded variables included for example some CMR and ECG parameters. Missing data were imputed using FAMD, which accommodates both continuous and categorical variables by extracting principal components while preserving the underlying data structure. Phenotypic clusters were then defined using hierarchical clustering on principal components (HCPC), as in our previous work.[17] In this method, principal components of the clinical data were derived using FAMD, ensuring a balanced influence between continuous and categorical variables. The optimal dimensions of FAMD to be used was determined by cross-validation, i.e. 2. The optimal number of clusters was determined by evaluating the gain in within-inertia (within-group variance) by HCPC and using the NbClust package.[18] The hierarchical clustering used the Manhattan distance and complete linkage. The clustering process was conducted with the FactoMineR package v2.9[19] in R v4.3.1 environment. The relative importance of variables in Cox regression model was accessed by the proportion of explainable log-likelihood statistics (likelihood ratio χ^2) that is explained by each variable, using R rms package (https://cran.r-project.org/web/packages/rms/index.html). To assess the concordance between genotype-based and phenotype-based clustering, we calculated the adjusted Rand index as a measure of overall similarity between the two classifications.

Statistics

The link between clinical variables and the clusters was analysed by ANOVA for numeric after logarithm transformation and χ^2 test for categorical variables. The over- or underrepresentation of variables within each cluster or genotype was analysed using the *v*-test, which transforms the *p*-value into a quantile, with the sign indicating over- or underrepresentation within the cluster, using the catdes function of FactoMineR package. The cumulative incidence Kaplan–Meier curve for outcome was used and differences were assessed by the log-rank test. The 95% confidence intervals (CIs) for risk probabilities at a 5-year time point were estimated around the cumulative incidence proportions. Cox proportional hazards regression analysis was performed to assess the association of genotype or phenotype status with event-free survival. To confirm the Cox model's assumption of proportional hazards, tests based on Schoenfeld residuals revealed

Table 2 Distribution of genotypes across the five international centres

Genotype	Centre					Total
	Denver (USA)	Madrid (ESP)	Maastricht (NL)	Trieste (IT)	Utrecht (NL)	
ACTC1	0	1	0	2	0	3
TPM1	0	0	3	1	0	4
PLN	0	0	2	1	2	5
NEXN	1	0	1	4	0	5
TNNC1	0	0	3	3	0	6
DMD	1	0	0	6	0	6
SCN5A	0	0	5	3	1	8
RBM20	0	2	6	4	0	12
BAG3	1	8	2	2	3	16
DSP	1	14	5	9	2	31
MYH7	1	10	12	13	2	38
TNNT2	1	5	6	29	1	42
FLNC	0	11	7	31	3	52
LMNA	8	4	21	21	2	56
TTN	8	43	77	112	7	247
Total	**22**	**98**	**150**	**241**	**23**	**534**

no significant deviations. Multivariate mixed-effect Cox modelling was used to adjust the fixed effect of age, sex, and baseline LVEF. Statistical significance was defined as the null hypothesis (hazard ratio = 1) did not fall within 95% CI, equivalent to $p < 0.05$. Statistical analysis was performed using SPSS 23.0 (IBM Corp., Armonk, NY, USA) and R version 4.3.1, and visualized using GraphPad Prism version 10.2.0.

Results

Cohort

In total, we included 534 patients with DCM and a P/LP variant from five different centres. Although *TTN* variants were prevalent among all centres, we noticed differences in the genetic spectrum per centre (*Table 2*). The cohort from Maastricht and Denver were mostly enriched with *TTN* and *LMNA* patients, while the cohort from Madrid and Trieste also had a considerable number of *FLNC* and *MYH7* patients. Geographic differences were expected for *PLN* prevalence, which is primarily due to a founder mutation (p.Arg14del) highly prevalent in the northern Netherlands.[20] Additionally, a large cohort of *BAG3* patients was present in the Madrid cohort while patients with a specific *TNNT2* variant were enriched in the Trieste cohort.

Genotype–phenotype associations: genotype-first approach

Using a genotype-first approach, we grouped all patients based on the genotype (i.e. *TTN, LMNA, DSP*, etc.) irrespective of the clinical information (*Table 3*). We identified clinical variables that were characteristic for a genotype using *v*-test analysis. This provided significant results for 10 genes (*ACTC1*, *BAG3*, *DMD*, *DSP*, *FLNC*, *LMNA*, *MYH7*, *PLN*, *TNNT2* and *TTN*; *Figure 1* depicts the nine genotypes with more than five patients). Overall, *LMNA* and *PLN* were associated with arrhythmias and conduction disorders, including atrial fibrillation (AF), atrioventricular block, left bundle branch block, non-sustained VT (NSVT), premature ventricular complexes, out-of-hospital cardiac arrest, and low peripheral voltages, while *BAG3*, *TNNT2*, *DMD* and *TTN* were mostly associated with increased cardiac volumes and decreased LVEF in the absence of arrhythmias. *DSP*, *BAG3*, and *PLN* were associated with female sex, while *TTN* was associated with male sex and an older age at diagnosis compared to the other genotypes. As expected, muscular involvement was characteristic for *DMD*.

Clustering of clinical parameters: phenotype-first approach

We performed clustering of patients using only the clinical parameters, independent of the genotype. Age of onset, LVEF, left ventricular end-diastolic and end-systolic diameter, AF, and family history were the major contributing variables for FAMD and clustering (online supplementary *Figure S1*). We identified four clusters: cluster 1 (C1 = young age of onset and primarily a NYHA class I) and cluster 2 (C2 = old age of onset with a high prevalence of arrhythmias) both had a moderate reduction in LVEF. In contrast, cluster 3 (C3 = primarily NYHA class III and NSVT) and cluster 4 (C4 = older age of onset with prevalent arrhythmias) both have a low LVEF (*Figure 2*, *Table 4*).

Next, we added the genetic information to the generated clusters. Although specific genotypes were significantly overrepresented in clusters (*FLNC* and *ACTC1* in C1, *LMNA* and *PLN* in C2, *BAG3* in C3, and *TTN* in C4), there was no strong correlation between the clusters and genotypes. No other genotypes were exclusively assigned to any of the four clusters with the exception of *ACTC1*, which had a small sample number ($n = 3$) (*Figure 2*

Table 3 Baseline characteristics of genotype-positive dilated cardiomyopathy patients stratified per genotype

	ACTC1 (*n* = 3)	*BAG3* (*n* = 16)	*DMD* (*n* = 7)	*DSP* (*n* = 31)	*FLNC* (*n* = 53)	*LMNA* (*n* = 56)	*MYH7* (*n* = 40)	*NEXN* (*n* = 6)	*PLN* (*n* = 5)	*RBM20* (*n* = 12)	*SCN5A* (*n* = 9)	*TNNC1* (*n* = 6)	*TNNT2* (*n* = 40)	*TPM1* (*n* = 4)	*TTN* (*n* = 247)
Demographics															
Age of diagnosis (years), median, IQR	17 (16–51)	45 (35–50)	40 (25–50)	46 (29–58)	43 (36–56)	50 (38–58)	47 (30–55)	50 (42–54)	63 (36–70)	49 (40–60)	53 (24–71)	39 (35–52)	46 (32–55)	51 (38–63)	52 (42–60)
Female sex, *n* (%)	0 (0)	9 (56)	1 (14)	19 (61)	20 (38)	15 (27)	11 (28)	3 (50)	5 (100)	5 (42)	5 (56)	3 (50)	14 (33)	1 (25)	59 (24)
Disease modifiers, *n* (%)															
Family history of DCM	2 (67)	10 (63)	4 (57)	13 (42)	35 (66)	31 (55)	20 (50)	3 (50)	1 (20)	8 (67)	5 (56)	5 (83)	28 (67)	2 (50)	108 (44)
Skeletal myopathy	0 (0)	0 (0)	4 (57)	0 (0)	0 (0)	3 (5)	1 (3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (1)
Peripartum cardiomyopathy	0 (0)	1 (6)	0 (0)	1 (3)	5 (9)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (17)	2 (5)	0 (0)	2 (1)
Physical characteristics															
NYHA class ≥III, *n* (%)	0 (0)	6 (38)	2 (29)	5 (16)	9 (17)	12 (21)	12 (30)	1 (17)	3 (60)	2 (17)	1 (11)	2 (5)	8 (19)	0 (0)	59 (24)
Systolic blood pressure (mmHg), mean (SD)	113 (3)	116 (14)	121 (15)	118 (15)	127 (16)	124 (24)	119 (14)	119 (11)	122 (25)	133 (13)	130 (22)	123 (22)	126 (27)	121 (6)	123 (18)
Diastolic blood pressure (mmHg), mean (SD)	63 (12)	80 (10)	77 (8)	74 (11)	78 (11)	77 (9)	76 (11)	76 (11)	77 (10)	84 (7)	81 (8)	75 (18)	79 (12)	76 (7)	78 (12)
Echocardiography															
LVEF (%), mean (SD)	47 (11)	25 (7)	32 (8)	36 (12)	37 (12)	40 (15)	36 (12)	35 (12)	35 (9)	32 (12)	43 (16)	37 (14)	35 (15)	35 (14)	33 (12)
LVEDDi (mm/m^2), mean (SD)	31 (4)	36 (5)	34 (5)	34 (5)	33 (5)	30 (4)	32 (7)	32 (4)	29 (1)	33 (5)	33 (9)	31 (5)	35 (7)	30 (8)	32 (5)
LVESD (mm), mean (SD)	42 (5)	61 (6)	57 (10)	48 (11)	49 (10)	45 (10)	49 (12)	50 (13)	43 (2)	52 (9)	49 (11)	53 (5)	50 (12)	51 (9)	50 (10)
E/A ratio, median (IQR)	1 (0–2)	2 (1–1)	2 (0–5)	1 (0–6)	1 (0–9)	2 (1–2)	2 (0–9)	1 (0)	1 (0–4)	1 (0–6)	2 (0–9)	9 (15)	2 (1–1)	1 (0–6)	2 (1–0)
E/e′ ratio, median (IQR)	7 (2)	21 (13)	12 (3)	10 (3)	11 (7)	12 (6)	13 (9)	17 (0)	10 (2)	9 (3)	7 (2)	12 (8)	17 (11)	18 (7)	13
PWT (mm), mean (SD)	8 (2)	9 (1)	9 (2)	9 (2)	8 (1)	8 (1)	9 (2)	10 (2)	9 (3)	9 (1)	9 (1)	8 (2)	9 (2)	9 (3)	9 (2)
Arrhythmias, *n* (%)															
AF	0	0	0	2 (6)	8 (15)	28 (50)	12 (30)	1 (17)	1 (20)	2 (17)	1 (11)	0 (0)	1 (2)	2 (50)	49 (20)
NSVT	0 (0)	7 (44)	3 (43)	17 (55)	32 (50)	36 (64)	11 (28)	5 (83)	0 (0)	7 (58)	3 (33)	1 (17)	11 (26)	0 (0)	101 (41)
LBBB	0 (0)	2 (13)	1 (14)	2 (6)	5 (10)	19 (34)	6 (15)	1 (17)	0 (0)	3 (25)	1 (11)	0 (0)	6 (14)	2 (50)	24 (10)
OHCA	0 (0)	0 (0)	0 (0)	1 (3)	1 (2)	3 (5)	1 (3)	1 (17)	2 (40)	1 (8)	1 (11)	0 (0)	0 (0)	0 (0)	2 (1)
AV block	0 (0)	1 (6)	0 (0)	1 (3)	2 (4)	23 (41)	6 (15)	2 (33)	0 (0)	2 (17)	1 (11)	0 (0)	3 (7)	0 (0)	23 (9)
PVC	0 (0)	0 (0)	1 (14)	3 (10)	4 (8)	9 (16)	3 (8)	1 (17)	0 (0)	0 (0)	0 (0)	0 (0)	2 (5)	0 (0)	12 (5)
Low ECG voltage in the peripheral leads	1 (33)	4 (25)	0 (0)	9 (29)	14 (26)	10 (18)	6 (15)	1 (17)	2 (40)	1 (8)	0 (0)	1 (17)	2 (5)	1 (25)	41 (17)

AF, atrial fibrillation; AV, atrioventricular; DCM, dilated cardiomyopathy; ECG, electrocardiogram; IQR, interquartile range; LBBB, left bundle branch block; LVEDDi, left ventricular end-diastolic diameter index; LVEF, left ventricular ejection fraction; LVESD, left ventricular end-systolic diameter; NSVT, non-sustained ventricular tachycardia; NYHA, New York Heart Association; OHCA, out-of-hospital cardiac arrest; PWT, posterior wall thickness; SD, standard deviation.

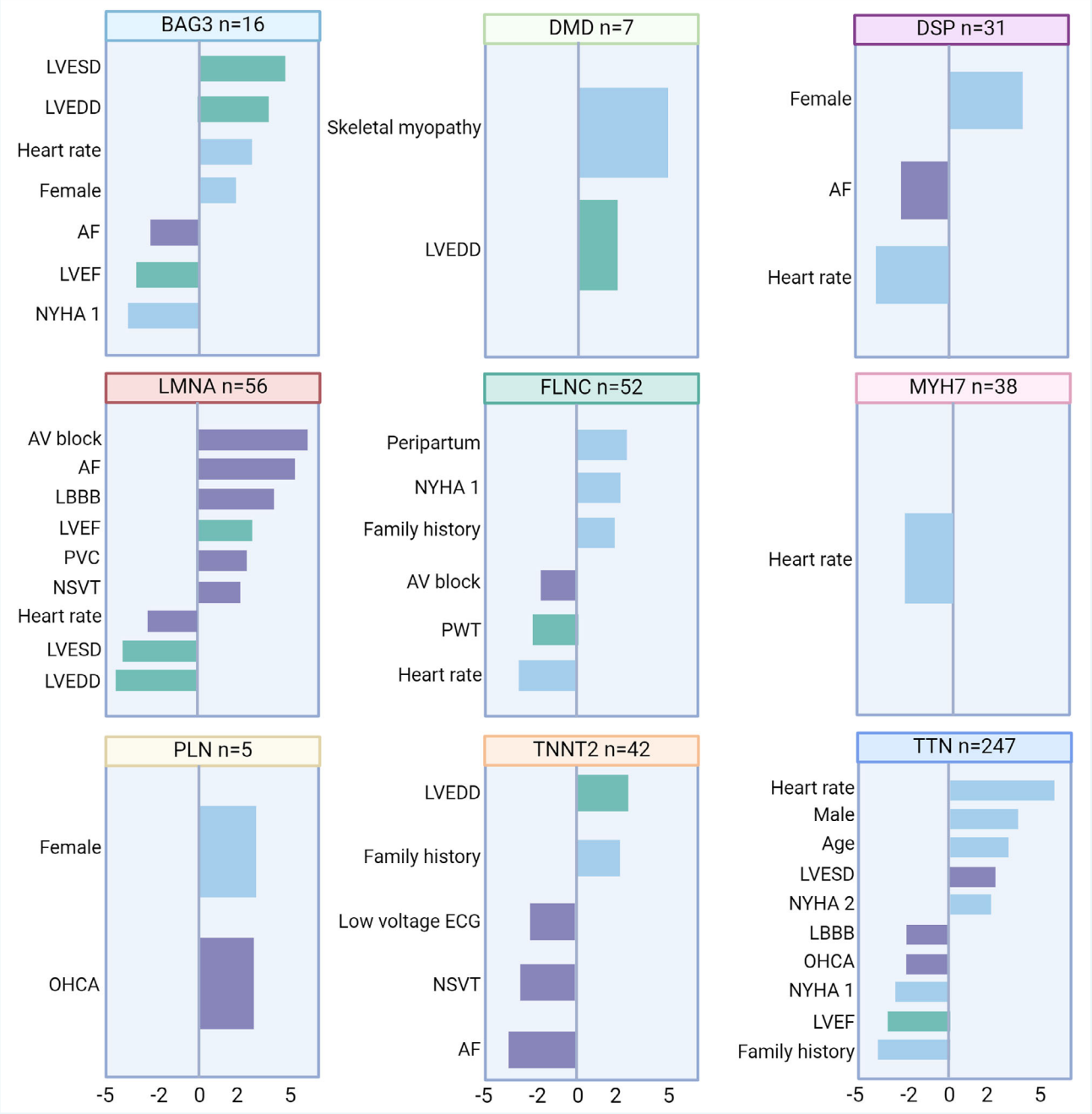


Figure 1 Phenotypic characteristic plots of the nine genotypes ($n \geq 5$ patients) including their most representative clinical variables. The over- or underrepresentation of a variable within a cluster was analysed by *v*-test. A positive value indicates overrepresentation of this variable in the applicable clusters, a negative value indicates underrepresentation of the corresponding variable. AF, atrial fibrillation; AV, atrioventricular; ECG, electrocardiogram; LBBB, left bundle branch block; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; LVESD, left ventricular end-systolic diameter; NSVT, non-sustained ventricular tachycardia; NYHA, New York Heart Association; OHCA, out-of-hospital cardiac arrest; PVC, premature ventricular complex; PWT, posterior wall thickness.

and online supplementary *Figure S1*). The adjusted Rand index between genotype and phenotype clustering was 0.024, indicating that phenotypic clustering does not reliably reflect underlying genetic architecture. Because patients with the same genotype can have a very different clinical presentation, their phenotype provides limited information about the underlying genetic cause.

Prognostic relevance of genotype- versus phenotype-first approach

In the overall cohort, 175 patients (33%) reached the combined endpoint (all-cause mortality, HFH, HTx, or MVA) after a median follow-up of 7.6 years (interquartile range 4.2–13.3 years) (online

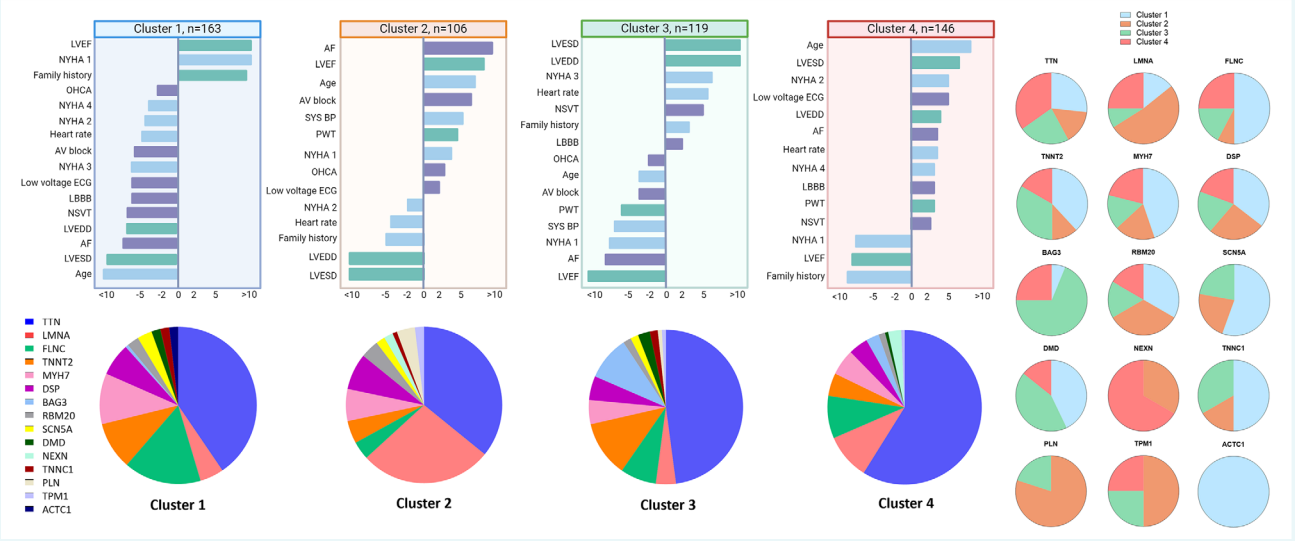


Figure 2 Four mutually exclusive phenoclusters as determined by hierarchical clustering of principal component using phenotypical information as input. Characteristic plots of the four proposed phenoclusters including their most representative clinical variables. A positive value indicates overrepresentation of this variable in the applicable phenoclusters, a negative value indicates underrepresentation of the corresponding variable. The phenoclusters were highly heterogeneous with regard to the genotype, with various genotypes distributed across all groups. AF, atrial fibrillation; AV, atrioventricular; ECG, electrocardiogram; LBBB, left bundle branch block; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; LVESD, left ventricular end-systolic diameter; NSVT, non-sustained ventricular tachycardia; NYHA, New York Heart Association; OHCA, out-of-hospital cardiac arrest; PWT, posterior wall thickness; SYS BP, systolic blood pressure.

supplementary *Table S1*). The secondary endpoints (occurrence of heart failure events [HFH and HTx] or arrhythmogenic events [MVA] during follow-up) occurred in 112 (21%) and 67 patients (13%), respectively.

Using a genotype-first approach, the highest long-term risk for the combined outcome was observed in *LMNA*, *FLNC* and *BAG3* patients (*Figure 3*; only genotypes with $n = 10$ or more are included in the analysis). The highest risk for heart failure-related events was observed in *BAG3*, *LMNA*, and *RBM20*, while the highest risk for MVA was in patients with a P/LP variant in *LMNA*, *DSP*, *FLNC*, and *BAG3*. These differences were also observed after adjustment for age, sex and baseline LVEF, when the genotypes were compared to *TTN* (*Figure 4*).

Using a phenotype-first approach, the lowest risk was observed in patients in phenocluster 1 (moderate LVEF and young age; *Figure 3*). While patients in phenocluster 4 (low LVEF and baseline arrhythmias) have the highest risk for heart failure events, those in phenocluster 3 (low LVEF and younger age) had the highest risk for MVA. Interestingly, the genotypes with the highest risk in the genotype-first approach largely clustered within the phenoclusters with lower risk, with overrepresentation of *FLNC* in phenocluster 1, and *LMNA* in phenocluster 2 (moderate LVEF and baseline arrhythmias). In contrast, the genotype with the lowest risk was significantly overrepresented in the phenocluster with the highest risk (*TTN* in C4). To understand these differences, we looked into the outcome per genotype in each of the four phenoclusters (online supplementary *Figure S2*). We identified that *TTN* patients overall had the lowest risk contribution per phenocluster while *LMNA*, *FLNC*, and *BAG3* had the highest risk contribution in each of the phenoclusters. Similar outcome of the prognostic impact of the phenoclusters was observed after removing *TTN* patients (online supplementary *Figure S3*), suggesting that the clustering association with outcome were not solely driven by *TTN*-related clinical variation. Furthermore, when we compared the phenoclusters for each individual genotype, we also noted prognostic differences (online supplementary *Figure S4*). Significant differences in outcome among phenoclusters were observed for *TTN*, *MYH7*, *LMNA*, and *RBM20*.

To determine whether genotype or phenotype contributes more significantly to the outcome in genetic DCM patients, we used the proportion of explainable log-likelihood statistics (likelihood ratio χ^2). Genotype consistently emerges as the most significant predictor in all outcome measures (major adverse events, HFH/HTx events, and MVA events), indicating its robust predictive importance in genetic DCM. Phenocluster contributes to a lesser extent, and LVEF, age, and sex have relatively low contributions (*Figure 5* and online supplementary *Figure S5*). These findings suggest that the specific genotype has the greatest role in predicting outcomes in genetic DCM patients.

One possible explanation for the observed difference in prognosis between a genotype- versus phenotype-first approach is that the clinical parameters that influence prognosis significantly differ per genotype. As a proof of principle, we used the gene-specific risk calculator for *LMNA*.[21] We compared this calculated MVA risk with both our geno- and phenotype-first approaches for *LMNA* patients (online supplementary *Figure S6*). There was no significant difference in the C-statistics of the geno- versus phenotype-first approach (area under the curve [AUC]: 0.691 (0.632–0.749) vs.

Table 4 Baseline characteristics of genotype-positive dilated cardiomyopathy patients stratified per phenocluster

	Cluster 1 (*n* = 163)	Cluster 2 (*n* = 106)	Cluster 3 (*n* = 119)	Cluster 4 (*n* = 146)
Demographics				
Age of diagnosis (years), median (IQR)	38 (27–48)	57 (49–65)	43 (34–50)	55 (50–62)
Female sex, *n* (%)	54 (33)	37 (35)	33 (28)	45 (31)
Disease modifiers				
Family history of DCM, *n* (%)	130 (81)	33 (31)	76 (35)	34 (23)
Skeletal myopathy, *n* (%)	5 (3)	2 (2)	3 (3)	0 (0)
Peripartum cardiomyopathy, *n* (%)	7 (13)	0 (0)	2 (6)	3 (7)
Physical characteristics				
NYHA class ≥III, *n* (%)	7 (4)	15 (14)	54 (45)	45 (31)
Systolic blood pressure (mmHg), median (SD)	124 (17)	131 (18)	114 (16)	123 (21)
Diastolic blood pressure (mmHg), median (SD)	76 (10)	81 (11)	76 (12)	77 (12)
Echocardiography				
LVEF (%), mean (SD)	43 (9)	44 (12)	24 (7)	27 (8)
LVEDDi (mm/m^2), mean (SD)	31 (4)	28 (3)	38 (5)	33 (4)
LVESD (mm), mean (SD)	44 (7)	41 (6)	62 (6)	54 (5)
E/A ratio, median (IQR)	1 (0.9–1.5)	1 (0.8–1.6)	1.7 (1.2–3)	1.2 (0.8–2.3)
E/e′ ratio, median (IQR)	8 (6–10)	10 (8–15)	15 (12–24)	14 (10–18)
PWT (mm), mean (SD)	8.5 (2)	9.4 (2)	8 (2)	9.1 (1)
Arrhythmias, *n* (%)				
AF	6 (4)	57 (55)	0 (0)	43 (30)
NSVT	40 (28)	49 (49)	72 (67)	73 (55)
LBBB	2 (1)	15 (15)	23 (20)	30 (21)
OHCA	0 (0)	9 (9)	0 (0)	4 (3)
AV block	2 (1)	34 (33)	4 (4)	23 (16)
PVC	10 (7)	11 (11)	5 (5)	9 (8)
Low ECG voltage in the peripheral leads	6 (4)	26 (26)	16 (14)	45 (31)

AF, atrial fibrillation; AV, atrioventricular; DCM, dilated cardiomyopathy; ECG, electrocardiogram; IQR, interquartile range; LBBB, left bundle branch block; LVEDDi, left ventricular end-diastolic diameter index; LVEF, left ventricular ejection fraction; LVESD, left ventricular end-systolic diameter; NSVT, non-sustained ventricular tachycardia; NYHA, New York Heart Association; OHCA, out-of-hospital cardiac arrest; PWT, posterior wall thickness; SD, standard deviation.

0.634 (0.570–0.699), $p = 0.22$) although the risk calculator (which integrates both genotype and phenotype) was significantly stronger compared to the phenotype approach (AUC: 0.737 (0.607–0.866), $p = 0.02$), but not to the genotype approach.

Discussion

This large, multicentre cohort study of genotype-positive DCM patients combined and compared genotype and phenotype approaches in recognizing clinical characteristics unique for specific genetic subgroups. First, genotype-specific clinical characteristics play a key role in disease expression, with distinct clinical features observed among genetic subgroups. *LMNA* and *PLN* variants were associated with conduction abnormalities and arrhythmias, whereas *BAG3*, *TTN*, and *TNNT2* variants were linked to increased cardiac volumes and reduced LVEF. Second, phenotypic clustering in a cohort of genetic DCM patients did not predict the underlying genotype. While certain genotypes were overrepresented in specific clusters, no genotype was exclusively confined to a single group. Third, genotype emerged as the strongest predictor of adverse outcomes, with *LMNA*, *FLNC*, and *BAG3* conferring the highest risk of adverse events. Phenotypic clusters, in contrast, showed significant overlap in risk profiles, indicating that clinical variables do not reliably stratify a diverse cohort of patients with a genetic form of DCM (*Graphical Abstract*). The clinical predictors of an adverse outcome are likely to be different for each individual genotype, emphasizing the importance to develop gene-specific risk calculators *(Graphical Abstract)*.

Phenotypic heterogeneity within specific genotypes

By comparing the genetic yield from five expert centres that routinely perform genetic testing in DCM patients, we observed a different spectrum of gene variants possibly related to overrepresentation of specific genotypes due to a founder effect in geographical regions. Insight in the presence of regional founder variants in specific genes is important to estimate a prior risk of a genetic variant and constitution of gene panels (e.g., *BAG3* variants in Madrid).

A key finding of our study is that baseline clinical characteristics significantly differ between genotypes which is important for risk prediction. Compared to the other genes, *LMNA* and *PLN* had the highest prevalence of arrhythmias and conduction disorders, while LVEF and volumes of the heart were relatively normal (most

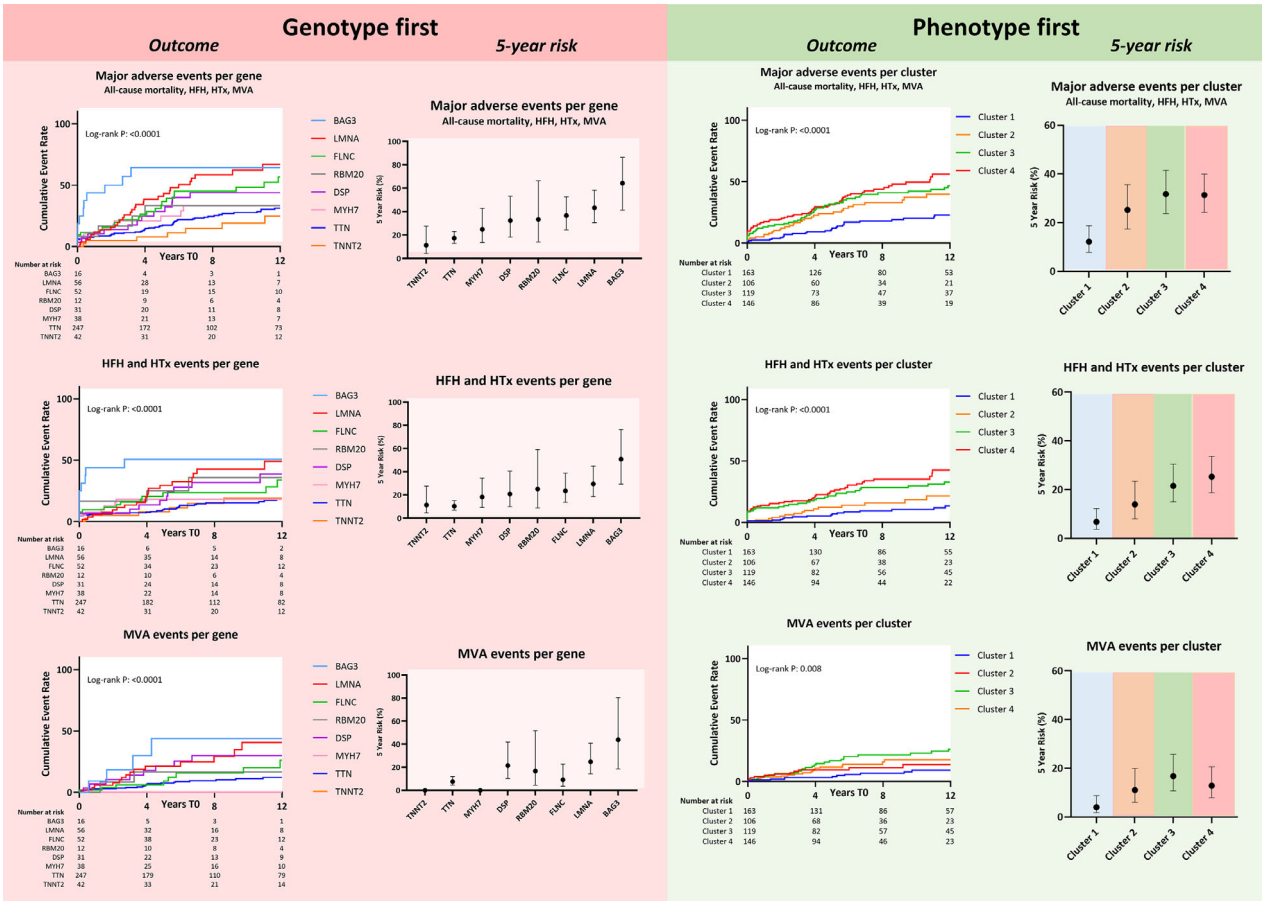


Figure 3 Outcome in genetic dilated cardiomyopathy using the genotype-first and phenotype-first approaches. Plots showing cumulative event rates and 5-year risk for major adverse events (all-cause mortality, heart failure hospitalization [HFH], heart transplantation [HTx], and malignant ventricular arrhythmias [MVA]), heart failure-related events (combining HFH and HTx) or MVA.

pronounced in *LMNA*). The association of *BAG3*, *TNNT2*, *DMD*, and *TTN* variants with increased cardiac volumes and decreased LVEF, in the absence of significant arrhythmias at baseline, suggests that these patients may primarily suffer from structural and functional cardiac deterioration prior to any electrical disturbances. The observed association between *BAG3* variants and atrioventricular block warrants further investigation. Prospective natural history studies are currently running for patients with *BAG3* variants. Given the two subgroups of baseline dilatation and baseline arrhythmias, it was hypothesized that patients in the arrhythmogenic subgroup suffer from more MVA during follow-up, while the dilated group might have an increased risk for heart failure-related events.[22,23] Interestingly, arrhythmogenic genotypes are also associated with a worse prognosis for heart failure-related events, while *BAG3* had the worst overall outcome. These results could indicate that arrhythmogenic genotypes have a more severe disease progression and that genotype-specific baseline clinical characteristics could play a crucial role in refining gene-specific risk calculators. Importantly, our large cohort included only genotype-positive DCM patients with detailed clinical information. In contrast to previous studies, this provided us the opportunity to perform advanced unsupervised clustering methods and report results at the level of individual genes.

Genetic heterogeneity within identified phenoclusters

Clustering based on the phenotype of patients demonstrated four heterogeneous clusters: C1 and C2 with a moderate reduced LVEF where C1 had a young age of onset and a low prevalence of arrhythmias, and C3 and C4 with a severe reduced LVEF where C3 had a young age of onset. In line with previous phenoclustering studies, the outer cluster (C1 and C4) are mainly defined by low and high LVEF.[11,12] However, the explainable log-likelihood analysis shows that genotypes is a stronger predictor for outcome compared to LVEF in a cohort of patients with a genetic form of DCM. Basically all genotypes were divided among the four phenoclusters, indicating a diverse clinical presentation across the whole phenotypic spectrum for each genotype. The clinical expression of genetic DCM varies widely, limiting the usability of a genotype-first approach at the individual level. However, information from the genotype can provide important information

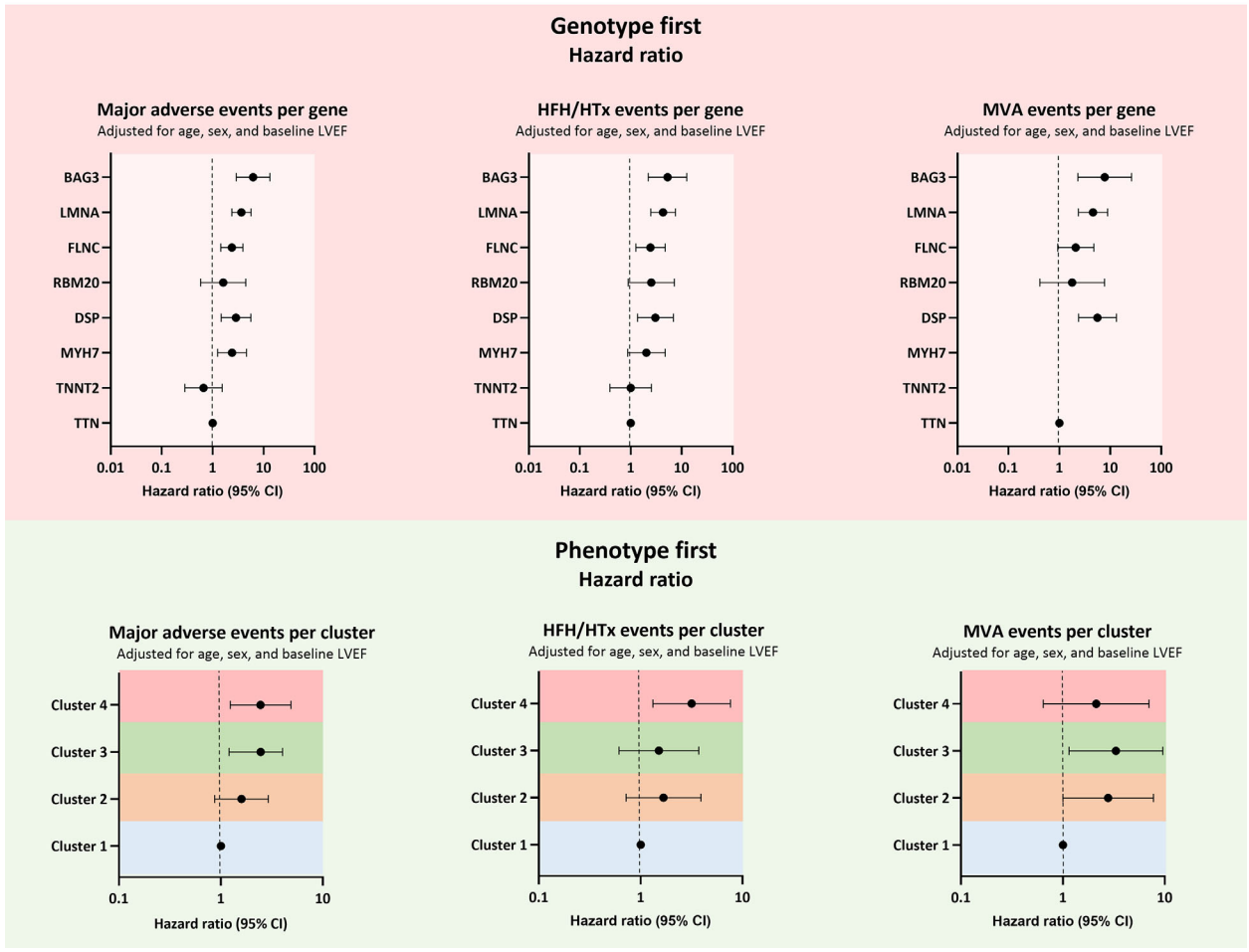


Figure 4 Adjusted hazard ratios for outcome in genetic dilated cardiomyopathy using the genotype-first and phenotype-first approaches. Plots showing hazard ratios adjusted for major adverse events (all-cause mortality, heart failure hospitalization [HFH], heart transplantation [HTx], and malignant ventricular arrhythmias [MVA]), heart failure-related events (combining HFH and HTx) or MVA. CI, confidence interval; LVEF, left ventricular ejection fraction.

for risk prediction. For example, patients with a *FLNC* variant that presented with a relative mild clinical phenotype (i.e., high LVEF and low prevalence of arrhythmias, enriched in C1) still had a high risk for major adverse events. This observation is probably due to the fact that clinical parameters that predict risk are genotype-specific and may differ between genes. Indeed, LVEF is not a predictor for outcome for patients with *FLNC*. Therefore, patient stratification based on phenotype only will miss important information.[24] Integrating deep phenotyping in genotype-specific cohorts to develop risk calculators seems therefore the way forward towards a more precision medicine approach.

Combining phenotype and genotype towards more accurate risk prediction

A key finding of our study is that phenotypic clustering, when applied independently of genetic information, does not strongly correlate with genotype and has limited prognostic utility. This suggests that phenotypic variability within each genotype is substantial, making a phenotype-first approach insufficient for risk stratification in genetic forms of DCM. However, our study also suggests that clinical factors contributing to prognosis may be specific for individual genotypes, underscoring the need for gene-specific risk calculators rather than broad phenotype-based classification.

Recent literature suggests that certain clinical features may be more typical, or even pathognomonic, for specific genotypes. For example, *DSP* variants are increasingly associated with recurrent myocardial inflammation, often described as 'hot phases' resembling myocarditis.[25–27] These episodes, frequently preceding overt systolic dysfunction, are thought to contribute to arrhythmogenic risk and may represent a key early disease phenotype. Similarly, *FLNC* truncating variants have been linked to a fibrotic or inflammatory cardiomyopathy phenotype, often with ventricular arrhythmias and diffuse myocardial fibrosis even in the presence of preserved LVEF.[29–30] These observations underscore the value

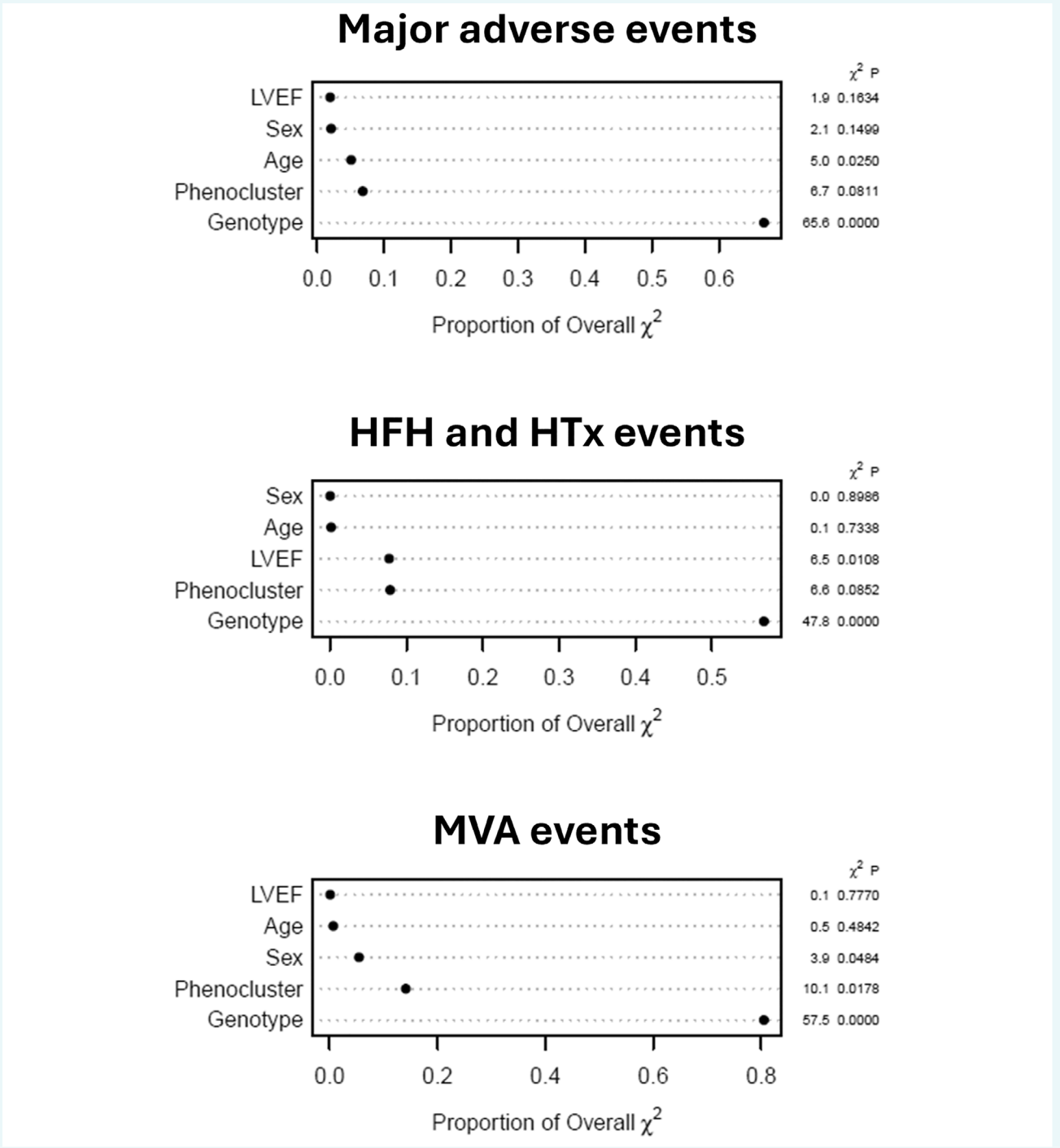


Figure 5 Relative contribution of genotype, phenocluster, and clinical variables to outcome prediction in genetic dilated cardiomyopathy patients using Cox proportional hazards models. Plots showing predictions of imputed data for major adverse events (all-cause mortality, heart failure hospitalization [HFH], heart transplantation [HTx], and malignant ventricular arrhythmias [MVA]), heart failure-related events (combining HFH and HTx) or MVA. LVEF, left ventricular ejection fraction.

of gene-informed phenotyping. Capturing such gene-specific features in future clustering frameworks or risk prediction models may improve both clinical interpretability and prognostic accuracy. Gene- or even variant (type)-specific risk calculators are created by leveraging the extensive data from large gene-specific patient cohorts. However, genetic forms of DCM are still considered rare, meaning that the availability of large gene-specific cohorts to develop risk calculators can only be realized by international

collaboration. Currently, there are two validated risk calculators included in the guidelines: one for *LMNA* and another for the *PLN* founder variant (p.Arg14del).[20,21] Additionally, a new risk calculator for *DSP* has just been published,[31] and a calculator for *FLNC* is on the way.[24] The *LMNA*-specific risk calculator outperformed the phenotype-first approach, while it was comparable to the genotype-first approach. This indicates that the genotype is a crucial prognostic predictor, and incorporating phenotype can further improve predictive accuracy. Although both genotype and phenotype appear relevant for risk stratification, the current study was not designed to formally assess their combined predictive performance. The clinical variables included in the available calculators (*LMNA*, *PLN*, and *DSP*) are different per gene, also confirming that an integrative approach of genotype and phenotype is the best method in predicting outcome in patients with a genetic form of DCM. Our findings have important implications for clinical practice. It emphasizes the importance of genetic testing in patients with DCM, as the genotype is almost intrinsically linked to risk prediction for the individual and their relatives.

Limitations

The current study was conducted in five tertiary referral centres from three different countries. These geographic differences might introduce biases related to regional (founder) genetic predispositions and healthcare practices. Consequently, the included patients in this study might not completely represent the entire DCM spectrum and these results should only be extrapolated to similar cohorts. Acknowledging the retrospective design of our study, we understand the potential for sampling bias. Our clustering analysis was limited by data availability, particularly missingness in certain clinically relevant variables. Arrhythmic phenotyping was based only on baseline NSVT and AF, potentially underestimating risk in genotypes with intermittent electrical abnormalities. CMR parameters were excluded due to incomplete data, which may have limited detection of scar- or inflammation-driven phenotypes, especially in genotypes like *DSP* and *FLNC*. Recurrent myocarditis, increasingly recognized in *DSP* carriers as an early arrhythmogenic feature,[27] could not be evaluated but may have prognostic relevance. Also, the low number of some rare genotypes limited the possibility to include variant-specific effects on outcome (e.g. *RBM20*). The sample sizes for certain genetic subgroups were relatively small, mainly for *ACTC1*, *TPM1*, *PLN*, *NEXN*, *TNNC1*, *DMD*, and *SCN5A*. This may limit the statistical power and accuracy to detect significant associations and differences. In addition, their prognosis was not analysed in this paper. The low number of *PLN* patients in our cohort included from the Dutch centres is primarily due to geographic factors. The *PLN*-Arg14del founder mutation is highly prevalent in the northern Netherlands, whereas our Maastricht and Utrecht cohorts predominantly include patients from the southern and central regions, where prevalence is lower. However, genetic P/LP variants in these genes are very rare and thus genotype-specific multicentre cohort collaborations are necessary to conduct genotype-based risk stratification studies.

Conclusion

Patients with a genetic form of DCM have a broad clinical and genetic heterogeneity. Risk stratification based on genotype is more accurate compared to a phenotype-first approach. The identification of genotype-specific clinical risk factors appears to be a promising step towards a precision medicine approach for risk prediction and therapy in DCM patients.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Funding

We acknowledge the support of IMI2-CARDIATEAM, from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement N° 821508. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA; DCM NEXT, HORIZON-EIC-2022-PATHFINDERCHALLENGES 01 'CARDIOGENOMICS' from HORIZON European Innovation Council Grants/European Commission, under grant agreement 101115416; grant of Dutch Research Council Open Call ZonMW-Metacor. Member of the European Reference Network for rare low prevalence and complex diseases of the heart (ERN GUARD-Heart).

Conflict of interest: S.H. receives personal fees for independent scientific advice on early development in the field of heart failure for AstraZeneca, Ribocure, and CSL Behring, and receives research support from AstraZeneca and CSL Behring. All other authors have nothing to disclose.

References

1. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: The complexity of a diverse genetic architecture. *Nat Rev Cardiol* 2013;**10**:531–547. https://doi.org/10.1038/nrcardio.2013.105
2. Akhtar MM, Lorenzini M, Cicerchia M, Ochoa JP, Hey TM, Sabater Molina M, *et al.* Clinical phenotypes and prognosis of dilated cardiomyopathy caused by truncating variants in the TTN gene. *Circ Heart Fail* 2020;**13**:e006832. https://doi.org/10.1161/CIRCHEARTFAILURE.119.006832
3. Tayal U, Ware JS, Lakdawala NK, Heymans S, Prasad SK. Understanding the genetics of adult-onset dilated cardiomyopathy: What a clinician needs to know. *Eur Heart J* 2021;**42**:2384–2396. https://doi.org/10.1093/eurheartj/ehab286
4. Verdonschot JAJ, Heymans SRB, Van Linthout S. The next era of gene-specific clinical care in patients with dilated cardiomyopathy. *Eur J Heart Fail* 2024;**26**:2169–2172. https://doi.org/10.1002/ejhf.3446
5. Van Linthout S, Stellos K, Giacca M, Bertero E, Cannata A, Carrier L, *et al.* State of the art and perspectives of gene therapy in heart failure. A scientific statement of the Heart Failure Association of the ESC, the ESC Council on Cardiovascular Genomics and the ESC Working Group on Myocardial & Pericardial Diseases. *Eur J Heart Fail* 2025;**27**:5–25. https://doi.org/10.1002/ejhf.3516
6. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, *et al.*; ESC Scientific Document Group. 2023 ESC Guidelines for the management of cardiomyopathies. *Eur Heart J* 2023;**44**:3503–3626. https://doi.org/10.1093/eurheartj/ehad194
7. Bozkurt B, Colvin M, Cook J, Cooper LT, Deswal A, Fonarow GC, *et al.* Current diagnostic and treatment strategies for specific dilated cardiomyopathies: A scientific statement from the American Heart Association. *Circulation* 2016;**134**:e579–e646. https://doi.org/10.1161/CIR.0000000000000455
8. Paldino A, Dal Ferro M, Stolfo D, Gandin I, Medo K, Graw S, *et al.* Prognostic prediction of genotype vs phenotype in genetic cardiomyopathies. *J Am Coll Cardiol* 2022;**80**:1981–1994. https://doi.org/10.1016/j.jacc.2022.08.804
9. Sammani A, Baas AF, Asselbergs FW, Te Riele ASJM. Diagnosis and risk prediction of dilated cardiomyopathy in the era of big data and genomics. *J Clin Med* 2021;**10**:921. https://doi.org/10.3390/jcm10050921

10. Verdonschot JAJ, Wang P, Derks KWJ, Adriaens ME, Stroeks SLVM, Henkens MTHM, *et al.* Clustering of cardiac transcriptome profiles reveals unique: Subgroups of dilated cardiomyopathy patients. *JACC Basic Transl Sci* 2023;**8**:406–418. https://doi.org/10.1016/j.jacbts.2022.10.010
11. Tayal U, Verdonschot JAJ, Hazebroek MR, Howard J, Gregson J, Newsome S, *et al.* Precision phenotyping of dilated cardiomyopathy using multidimensional data. *J Am Coll Cardiol* 2022;**79**:2219–2232. https://doi.org/10.1016/j.jacc.2022.03.375
12. Verdonschot JAJ, Merlo M, Dominguez F, Wang P, Henkens M, Adriaens ME, *et al.* Phenotypic clustering of dilated cardiomyopathy patients highlights important pathophysiological differences. *Eur Heart J* 2021;**42**:162–174. https://doi.org/10.1093/eurheartj/ehaa841
13. Sammani A, Jansen M, Linschoten M, Bagheri A, de Jonge N, Kirkels H, *et al.* UNRAVEL: Big data analytics research data platform to improve care of patients with cardiomyopathies using routine electronic health records and standardised biobanking. *Neth Heart J* 2019;**27**:426–434. https://doi.org/10.1007/s12471-019-1288-4
14. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, *et al.* 2022 AHA/ACC/HFSA guideline for the management of heart failure: A report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2022;**145**:e895–e1032. https://doi.org/10.1161/CIR.0000000000001063
15. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;**17**:405–424. https://doi.org/10.1038/gim.2015.30
16. Audigier V, Husson F, Josse J. A principal component method to impute missing values for mixed data. *Adv Data Anal Classif* 2013;**10**:5–26. https://doi.org/10.1007/s11634-014-0195-1
17. Husson F, Josse J, Pagès J. Principal component methods – hierarchical clustering – partitional clustering: Why would we need to choose for visualizing data? 2010 http://factominer.free.fr/more/HCPC_husson_josse.pdf Accessed 22/05/25.
18. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R package for determining the relevant number of clusters in a data set. *J Stat Softw* 2014;**61**:1–36. https://doi.org/10.18637/jss.v061.i06
19. Lê S, Josse J, Husson F. FactoMineR: An R package for multivariate analysis. *J Stat Softw* 2008;**25**:1–18. https://doi.org/10.18637/jss.v025.i01
20. Verstraelen TE, van Lint FHM, Bosman LP, de Brouwer R, Proost VM, Abeln BGS, *et al.* Prediction of ventricular arrhythmia in phospholamban p.Arg14del mutation carriers – reaching the frontiers of individual risk prediction. *Eur Heart J* 2021;**42**:2842–2850. https://doi.org/10.1093/eurheartj/ehab294
21. Wahbi K, Ben Yaou R, Gandjbakhch E, Anselme F, Gossios T, Lakdawala NK, *et al.* Development and validation of a new risk prediction score for life-threatening ventricular tachyarrhythmias in laminopathies. *Circulation* 2019;**140**:293–302. https://doi.org/10.1161/CIRCULATIONAHA.118.039410
22. Lukas Laws J, Lancaster MC, Ben Shoemaker M, Stevenson WG, Hung RR, Wells Q, *et al.* Arrhythmias as presentation of genetic cardiomyopathy. *Circ Res* 2022;**130**:1698–1722. https://doi.org/10.1161/CIRCRESAHA.122.319835
23. Kim KH, Pereira NL. Genetics of cardiomyopathy: Clinical and mechanistic implications for heart failure. *Korean Circ J* 2021;**51**:797–836. https://doi.org/10.4070/kcj.2021.0154
24. Gigli M, Stolfo D, Barbati G, Graw S, Chen SN, Merlo M, *et al.*; Filamin C Registry Consortium. Arrhythmic risk stratification of carriers of filamin C truncating variants. *JAMA Cardiol* 2025;**10**:359–369. https://doi.org/10.1001/jamacardio.2024.5543
25. Bassetto G, Angriman F, Gava C, Paldino A, Perotto M, Bordignon L, *et al.* Hot phases cardiomyopathy: Pathophysiology, diagnostic challenges, and emerging therapies. *Curr Cardiol Rep* 2025;**27**:11. https://doi.org/10.1007/s11886-024-02168-6
26. Mayfield JJ, Bogomolovas J, Abraham MR, Sullivan K, Seo Y, Sheikh F, *et al.* Recurrent myocarditis in patients with desmosomal pathogenic variants: Is self antigen presentation the link? *JACC Clin Electrophysiol* 2023;**9**:2024–2033. https://doi.org/10.1016/j.jacep.2023.04.020
27. Bariani R, Rigato I, Cipriani A, Bueno Marinas M, Celeghin R, Basso C, *et al.* Myocarditis-like episodes in patients with arrhythmogenic cardiomyopathy: A systematic review on the so-called hot-phase of the disease. *Biomolecules* 2022;**12**:1324. https://doi.org/10.3390/biom12091324
28. Begay RL, Graw SL, Sinagra G, Asimaki A, Rowland TJ, Slavov DB, *et al.* Filamin C truncation mutations are associated with arrhythmogenic dilated cardiomyopathy and changes in the cell-cell adhesion structures. *JACC Clin Electrophysiol* 2018;**4**:504–514. https://doi.org/10.1016/j.jacep.2017.12.003
29. Celeghin R, Cipriani A, Bariani R, Bueno Marinas M, Cason M, Bevilacqua M, *et al.* Filamin-C variant-associated cardiomyopathy: A pooled analysis of individual patient data to evaluate the clinical profile and risk of sudden cardiac death. *Heart Rhythm* 2022;**19**:235–243. https://doi.org/10.1016/j.hrthm.2021.09.029
30. Hall CL, Akhtar MM, Sabater-Molina M, Futema M, Asimaki A, Protonotarios A, *et al.* Filamin C variants are associated with a distinctive clinical and immunohistochemical arrhythmogenic cardiomyopathy phenotype. *Int J Cardiol* 2020;**307**:101–108. https://doi.org/10.1016/j.ijcard.2019.09.048
31. Carrick RT, Gasperetti A, Protonotarios A, Murray B, Laredo M, van der Schaaf I, *et al.* A novel tool for arrhythmic risk stratification in desmoplakin gene variant carriers. *Eur Heart J* 2024;**45**:2968–2979. https://doi.org/10.1093/eurheartj/ehae409