scientific reports

Check for updates

OPEN

Four-dimensional left ventricular motion clustering reveals cardiovascular phenotypes at population scale

Pierre-Raphael Schiratti[1], Soodeh Kalaie[1,2], Jin Zheng[1], Paolo Inglese[1], Kathryn A. McGurk[1,3], Sean L. Zheng[1,3], James S. Ware[1,3,4,5,6], Wenjia Bai[2,7] & Declan P. O'Regan[1]✉

Characterisation of the motion dynamics of the left ventricle is key to understanding pathophysiological mechanisms and transitions from health to disease. Conventional volumetric assessments of the heart using imaging represent mainly aggregate global features of function that are poorly discriminating. Here we present a novel approach to quantify and visualise how the left ventricle is affected by cardiovascular risk factors through efficient representations of motion trajectories. We use computer vision to survey four-dimensional cardiac motion traits using densely sampled point clouds of the left ventricle in over 20,000 participants of UK Biobank. We developed a computational framework for dimensionality reduction of spatiotemporal information to derive a human-interpretable signature summarising variation in complex patterns of motion. We found six phenogroups representing a novel classification of heterogeneous motion phenotypes with differential enrichment of cardiovascular outcomes and genetic risk. Low dimensional representations of motion are visualised as a simple spatial signature capturing deviation from an average state. Discovering compact cardiac motion signatures of health and disease from dynamic point clouds enables efficient classification of patient risk and predisposing polygenic factors.

The motion dynamics of the beating heart comprise a complex rhythmic pattern of non-linear trajectories regulated by molecular, electrical and biophysical processes that are susceptible to disease[1]. While advanced functional imaging modalities provide localised assessment, conventional approaches for assessing cardiovascular function have relied on simple volumetric measures obtained using imaging however such approaches have significant limitations as they capture only composite global markers of function that are insensitive to complex traits of contractile coordination, fibre geometry and active relaxation[2]. The measurement and interpretation of conventional parameters also relies on expertise and is constrained by one-dimensional dichotomous thresholds for diagnosis[3]. Cardiomyopathies for instance are characterised by complex regional patterns of remodelling and early signs of disease may not be apparent using only volumetric parameters[4].

Computer vision approaches now enable three-dimensional cardiac motion estimation from sparse medical imaging of the heart. These approaches transform planar cross-sectional imaging into a digital representation of whole heart motion that can be aggregated across large and diverse populations[5,6]. While rich representations of motion contain latent prognostic information[1], their value in classifying disease sub-types is not yet known. Unsupervised deep learning approaches have efficiently classified patients from controls using complex traits[7], but have so far been limited to static geometries and have not provided interpretable temporal features. Here, we present a novel framework for characterising diversity in cardiac motion using spatio-temporal point clouds derived from cardiac magnetic resonance (CMR) imaging. We developed a convolutional variational autoencoder to embed these complex cardiac models in a low-dimensional latent space. Unlike supervised approaches that are limited to predicting pre-defined clinical labels, we employed an unsupervised clustering framework. This

[1]MRC Laboratory of Medical Sciences, Imperial College London, London, UK. [2]Biomedical Image Analysis Group, Department of Computing, London, UK. [3]National Heart and Lung Institute, Imperial College London, London, UK. [4]Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, USA. [5]Department of Cardiology, Imperial College Healthcare NHS Trust, London, UK. [6]Royal Brompton and Harefield Hospitals, Guy's and St. Thomas' NHS Foundation Trust, London, UK. [7]Department of Brain Sciences, Imperial College London, London, UK. ✉email: declan.oregan@imperial.ac.uk

allows for the discovery of intrinsic, data-driven motion phenotypes that may not map directly with current diagnostic categories, thereby revealing novel stratifications of left ventricular motion phenotypes. Temporal signatures were clustered on latent features of motion to define six phenogroups with differential enrichment of cardiovascular risk factors, polygenic scores and clinical outcomes. A human interpretable signature allows visualisation of how each three dimensional motion state deviates from a normative reference.

Results

Study overview

The UK Biobank study recruited 500,000 participants aged 40 to 69 years between 2006 and 2010[8]. Genotyping array and exome sequencing were available for over 450,000 participants. A sub-study recalled participants for CMR[9], which included retrospectively-gated short axis cine scans, and here we analysed data from 20,318 participants. Distributions of demographic and image-derived features in the population are reported in Table 1. A flowchart (Fig. 1) describes each analysis step. Computer vision techniques transformed the two dimensional cine images of the left ventricle into a time-resolved three-dimensional point cloud of the endo and epicardial surfaces of the left ventricle (Fig. 2). Each model of the left ventricle comprised 40,616 points (downsampled to 1406) over 50 time frames throughout the cardiac cycle. Atlas registration ensured that the number and position of points was consistent between participants as well as at each cardiac phase[10,11]. The term "point cloud" refers to the ordered set of mesh vertices and because atlas registration enforces a fixed vertex indexing, each vertex consistently represents the same anatomical location across participants and cardiac phases, preserving the full topological information of the underlying mesh.

To capture spatio-temporal patterns throughout the cardiac cycle, sequences of consecutive point cloud frames were constructed using an overlapping sliding window approach. Dimensionality reduction was performed using a convolutional variational autoencoder followed by a principal component analysis (PCA). Clustering of spatio-temporal patterns was performed with K-means for time series data. Six clusters (phenogroups) of motion traits were summarised by enrichment for prevalent disease, future events and polygenic risk scores. Each phenogroup was compared to the population average to define a visually interpretable signature of motion diversity.

Dimensionality reduction of cardiac motion

Three dimensional time-resolved point clouds capturing cardiac motion are spatially and temporally correlated. Dimensionality reduction was performed to extract a low-dimensional abstract representation capturing the

Trait	Overall	Female	Male
n	20318	10695	9623
Age, years	63.0 (7.6)	62.4 (7.5)	63.7 (7.7)
Body surface area, m^2	1.9 (0.2)	1.7 (0.2)	2.0 (0.2)
Ethnic background			
British	18565 (91.4)	9769 (91.3)	8796 (91.4)
Any other white background	557 (2.7)	323 (3.0)	234 (2.4)
Irish	551 (2.7)	290 (2.7)	261 (2.7)
Indian	147 (0.7)	56 (0.5)	91 (0.9)
Other ethnic group	90 (0.4)	47 (0.4)	43 (0.4)
Diastolic blood pressure, mmHg	78.7 (10.5)	77.0 (10.5)	80.5 (10.2)
Systolic blood pressure, mmHg	139.4 (19.4)	136.3 (20.1)	142.7 (18.0)
MRI measures			
Circumferential strain, %	-22.3 (3.4)	-23.4 (3.1)	-21.1 (3.2)
Radial strain, %	45.0 (8.3)	47.5 (7.9)	42.2 (7.8)
Longitudinal strain, %	-18.5 (2.8)	-19.1 (2.7)	-17.8 (2.7)
LV end-diastolic volume, mL	147.9 (33.5)	129.4 (22.5)	168.4 (31.6)
LV end-systolic volume, mL	60.4 (19.0)	50.5 (12.7)	71.4 (18.9)
LV ejection fraction, %	59.6 (6.1)	61.1 (5.6)	57.8 (6.1)
LV mass, g	86.0 (22.1)	70.9 (12.3)	102.7 (18.3)
LV maximum wall thickness, mm	9.4 (1.8)	8.6 (1.4)	10.3 (1.7)
RV end-diastolic volume, mL	156.5 (37.3)	134.2 (24.2)	181.3 (33.4)
RV systolic volume, mL	67.6 (21.3)	54.9 (13.3)	81.6 (19.5)
RV ejection fraction, %	57.3 (6.2)	59.2 (5.7)	55.1 (5.9)

Table 1. Description of UK Biobank participants. Principal characteristics of UK Biobank participants with motion cardiac imaging data stratified by sex. Data on demographics, clinical measurements and MRI-derived measures were available. Statistics are described in total count and frequency in brackets for discrete features or mean and standard deviation in brackets for continuous features. mmHg, millimetre of mercury; CMR, cardiac magnetic resonance; LV, left ventricle; RV, right ventricle.

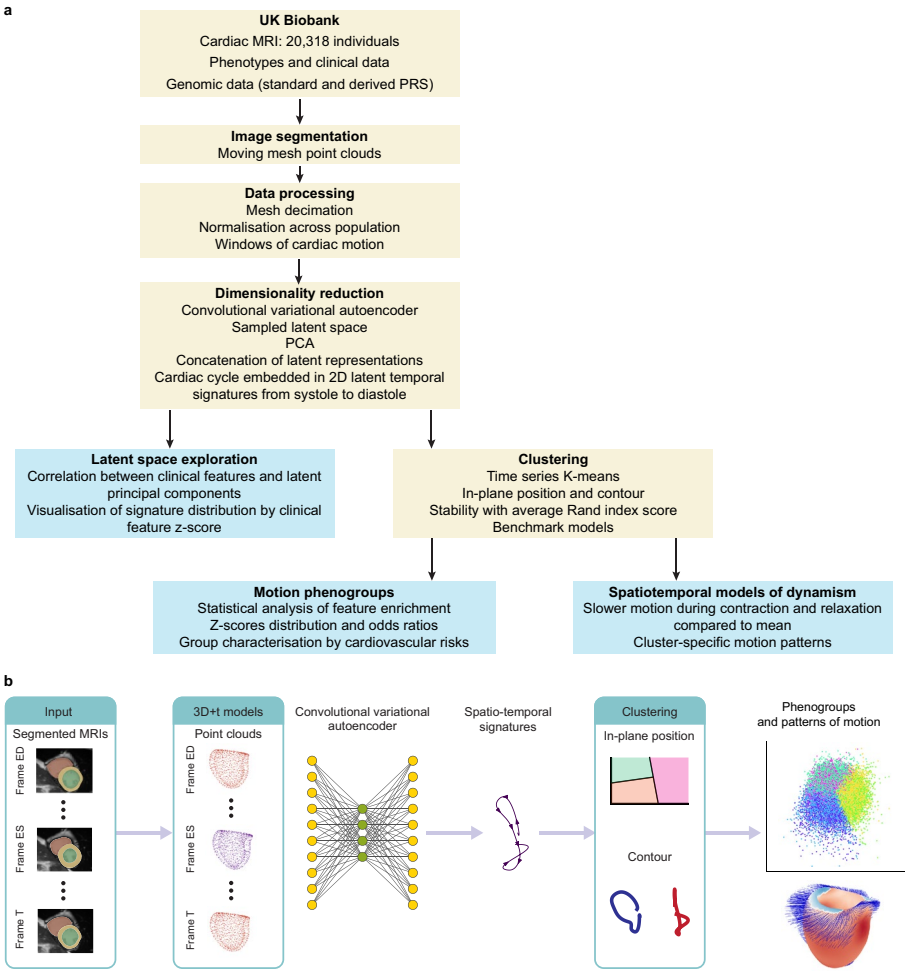


Fig. 1. Flowchart of the analysis. **(a)** Main steps of the generation of cardiac motion signatures and analysis of clinically relevant novel phenogroups. **(b)** Outline of the transformation of cardiac motion data into human-interpretable temporal signatures. Cardiovascular risk groupings and their associated cardiac motion patterns were derived from these signatures. MRI, magnetic resonance imaging; PRS, polygenic risk score; PCA, principal component analysis.

variance in motion traits over time. The entire pipeline from high-dimensional cardiac motion representation to low-dimensional interpretable signatures is described in Figure 3.

We used a variational autoencoder (VAE) with four convolutional layers in both the encoder and decoder. This network encodes a sequence of downsampled 3D left ventricular surface point cloud positions within a time window into a latent vector and then decodes the vector into a reconstruction of that sequence. Training the network to minimise the error between the reconstructed and input data ensures that the latent vector captures the complexity of the cardiac motion. However, the CVAE does not explicitly penalise redundancy in the latent space, and correlations between latent dimensions may persist. To address this, PCA was applied post-training to produce uncorrelated principal component scores ordered by explained variance, providing a principled way to identify and discard redundant components.

To validate that the intermediate processing steps preserve geometric fidelity, we report quality metrics at each stage of the pipeline. The atlas registration achieved Dice similarity coefficients above 0.90 for LV and RV cavities with mean surface distances below 2 mm[10], and the segmentation pipeline produced landmark errors within inter-observer variability[11] (see Methods for details). Post-training reconstruction loss, or relative mean absolute error, expressed in percentage of absolute range in each dimension was: 4.09±1.03% for X, 4.09±0.97% for Y, and 2.29±0.96% for Z coordinates. These values are comparable to observed errors in automated cardiac segmentation algorithms[12]. Reconstruction was also evaluated post-training using a Chamfer distance (0.09±0.01) in normalised coordinates, which is comparable to the prediction error for advanced point cloud reconstruction algorithms[13,14].

Phenotypes associated with latent motion signatures

The latent space of the VAE is transformed using PCA. Reduced latent representations of consecutive time windows were concatenated into visually-interpretable temporal signatures or trajectories, characterising

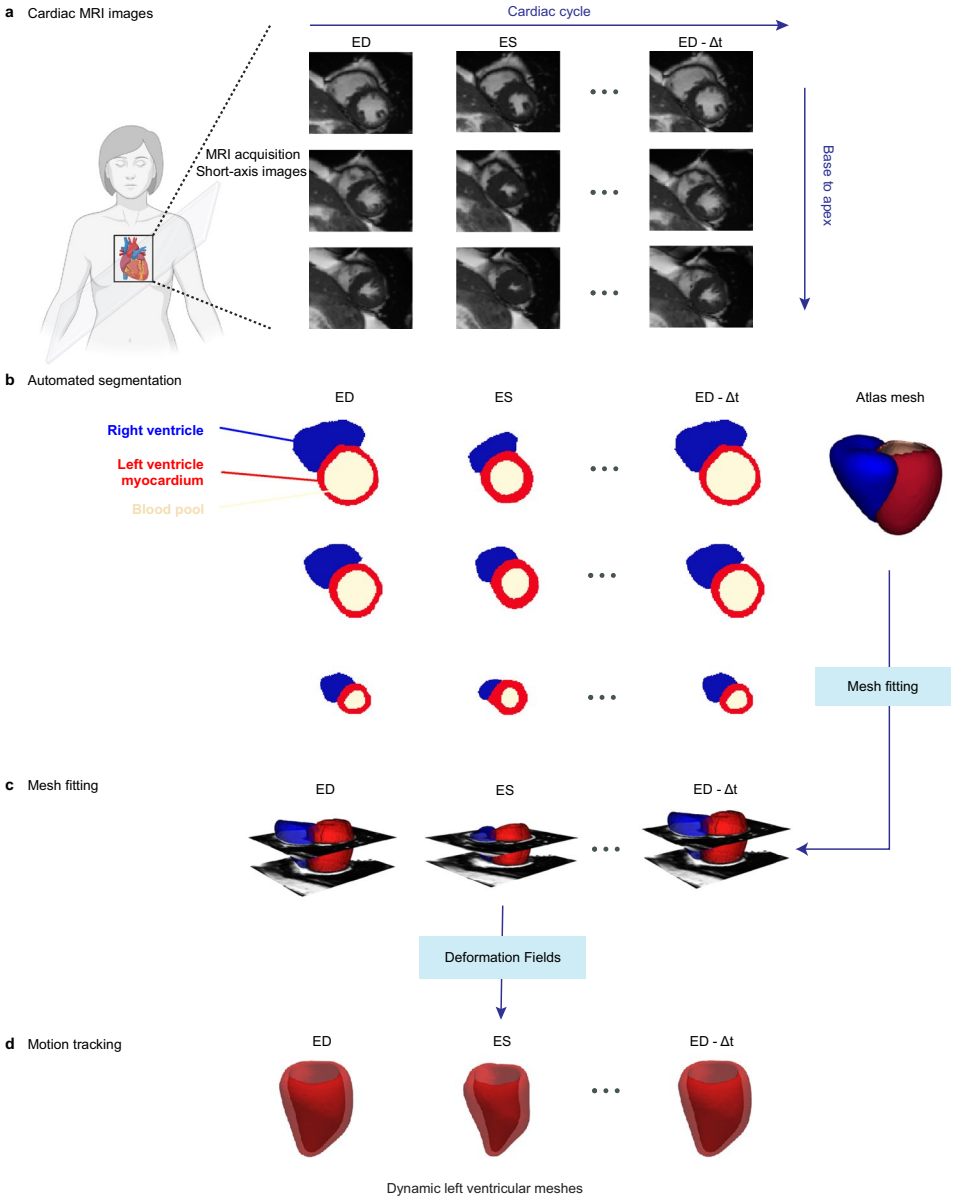


Fig. 2. Cardiac imaging segmentation and spatio-temporal models. **(a)** Cardiac MRI images were acquired for each participant. They capture the anatomy of the heart from base to apex of ventricles and the temporal data throughout the cardiac cycle from diastole to systole. **(b)** These images were then automatically segmented with our algorithm to locate the right ventricle and the left ventricle myocardium and blood pool. **(c)** These segmentation masks were then fitted and registered to a common atlas space between participants to develop a consistent three-dimensional model of the left ventricle. **(d)** Deformation fields were used to generate the entire cardiac cycle of surface meshes.

variations in complex cardiac motion traits. Such individual-level temporal signatures are low-dimensional uncorrelated abstractions of complex and highly-correlated moving point clouds.

The latent space was explored for associations with structural and functional features as well as circulating biomarkers (Supplementary Fig. 1). The first principal component (PC1) was negatively correlated with body surface area, ventricular and atrial volumes, left ventricular mass, and wall thickness, and positively correlated with ejection fraction and high-density lipoprotein (HDL) cholesterol. PC2 was positively correlated with body surface area, ventricular volume, and mass. PC3 was positively correlated with age, and negatively correlated with ventricular and atrial volumes, left ventricular mass and wall thickness. PC4 was positively correlated with circumferential strain, ventricular volumes and mass and ejection fraction. These findings show how biophysical, structural and functional traits are embedded in a latent representation of left ventricle motion. These associations serve as a validation that the latent space encodes physiologically meaningful information: PC1 corresponds broadly to global contractile function, PC2 to bodily variations and chamber dimensions,

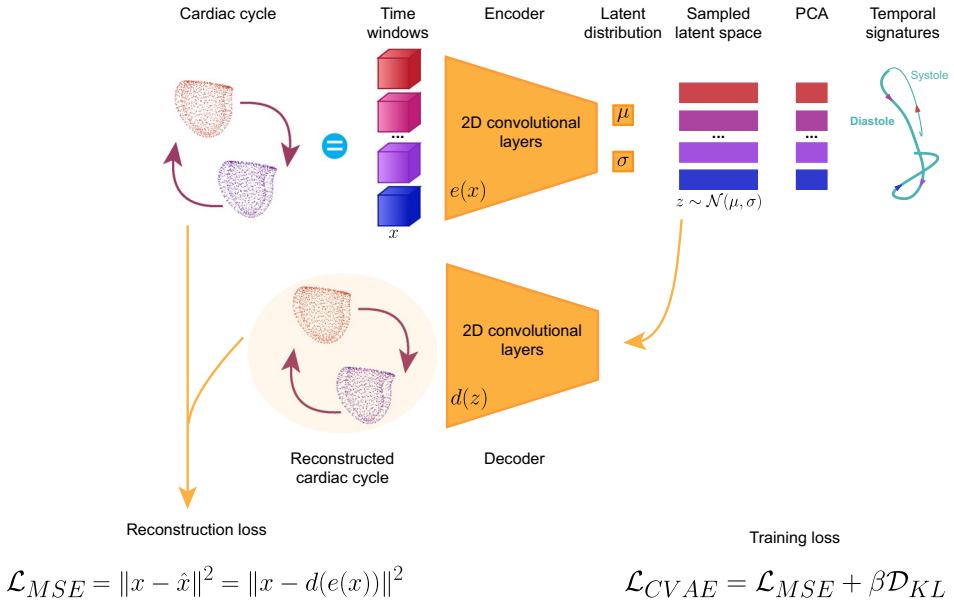


Fig. 3. Convolutional variational autoencoder (CVAE) and temporal signatures. The cardiac cycle was fed in the model as a collection of time windows. The encoder and decoder were made of two-dimensional convolution layers. The encoder took as input the collection of windows and encoded a latent distribution of mean μ and standard deviation σ. The latent space z was sampled from that distribution and was fed back into the model through the decoder. The decoder aimed to reconstruct the input cardiac cycle. The reconstruction loss between input and model reconstruction was assessed using a mean squared error (MSE). The model was trained with the weighted sum of MSE and Kullback–Leibler (KL) divergence, which ensured the coherent continuity of the latent space. After training, the dimensionality of the latent space was further reduced with principal component analysis (PCA). Subsequent windows were concatenated and visualised in the reduced space in the form of temporal signatures describing the input cardiac cycle. The detailed layer-by-layer architecture of the CVAE is provided in Supplementary Fig. 5.

and PC3 to age-related remodelling. This interpretive anchoring is a prerequisite for any downstream clinical application of the framework.

A low dimensional motion signature was visualised using the first two components at each time sliding window of the cardiac cycle. The position and morphology of the signature were represented across different z-scores of a given feature of interest in the population distribution (Fig. 4). For instance, aging caused an in-plane translation of the signature towards lower values of PC1 and PC2, and a larger contour area. We observed opposing effects for hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM) polygenic risk scores in the latent space projection consistent with findings from studies assessing static morphological traits[15,16]. Maximum wall thickness increased for low values of PC1, high values of PC2, and larger contour areas. The same trend was observed for systolic blood pressure along with greater differences in contour during systole (beginning of the signature), demonstrating that our model preserves temporal characteristics during dimensionality reduction of the cardiac cycle. HDL cholesterol increased for high values of PC1. Glycated hemoglobin (HbA1c), a marker of average blood sugar levels, rose with decreasing values of PC1 and PC2.

Cardiovascular risk factors such as hypertension, obesity and diabetes were associated with low values of PC1. Adverse prevalent outcomes, including myocardial infarction, heart failure and stroke, resulted in a similar motion signature. Critical future outcomes, such as cardiac arrest, showed the lowest PC1 values and more extreme motion signatures. Incident HCM and incident myocardial infarction signatures were found to share the same in-plane position at low values of PC1. The average signature for future DCM was the furthest from the population average signature, with low values of both PC1 and PC2. These feature distributions showed how both the in-plane position and contour characteristics of the signature reflect the phenotype and genotype of participants, as well as prevalent and incident events.

Clustering of motion phenogroups

Signatures were clustered on their PCA projection: first by contour characteristics and second by in-plane position as seen in Fig. 5. K-means for time series data was used, and contour clustering identified two phenogroups, while in-plane position clustering revealed three, resulting in six clusters in total. Their relative in-plane positions are visualised in Fig. 6a and their centred contours in Fig. 6b. The most and least represented phenogroups represented 22.7% and 12.9% of the initial study population.

Individual-level features were tested for significant differences in the six phenogroups. Principal findings are summarised in Fig. 7 for significantly different pairwise comparisons of continuous traits, and in Fig. 8 for discrete variables and associated odds ratios (OR). Each odds ratio is derived from the contingency table of

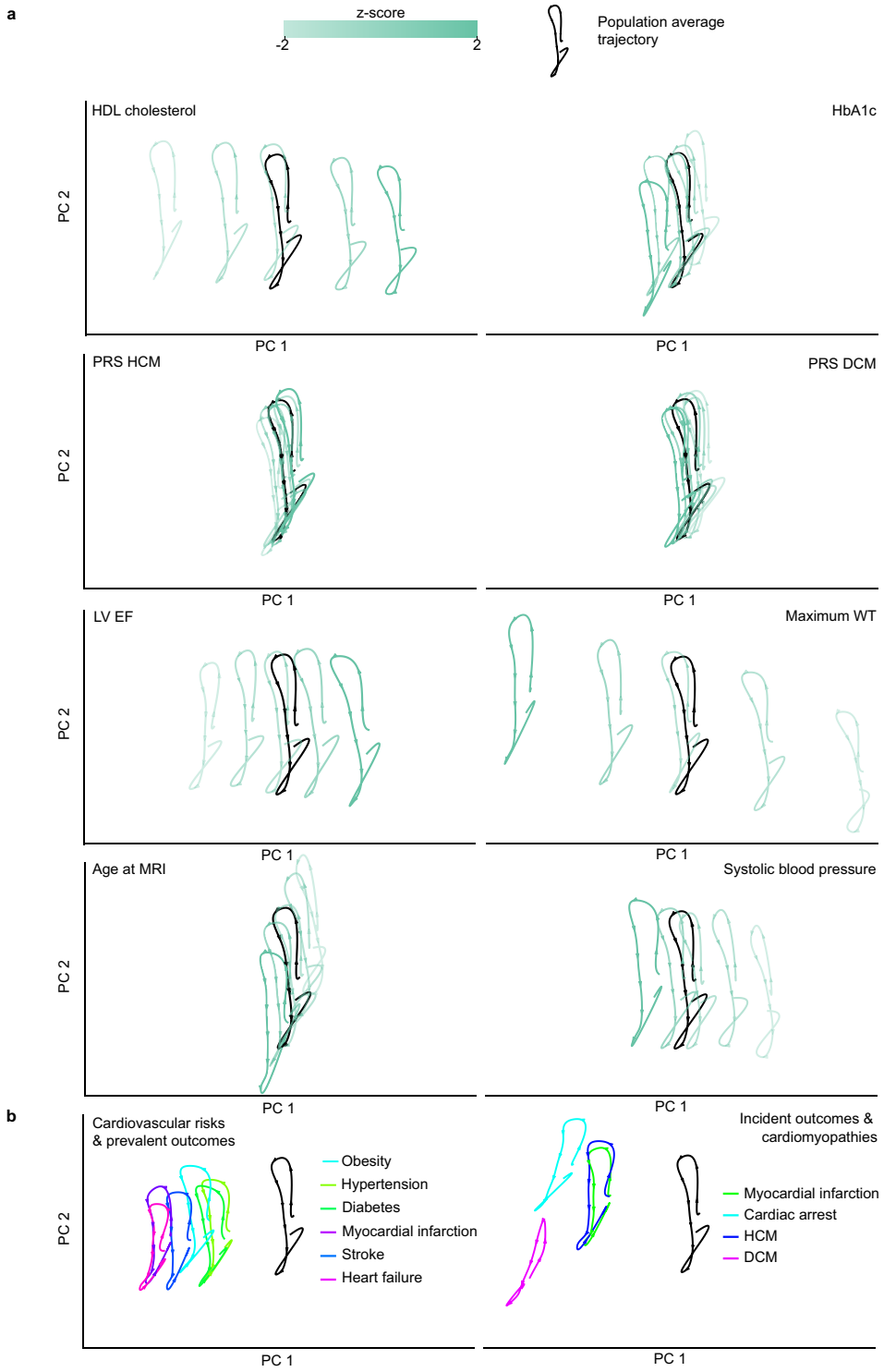


Fig. 4. Signature distribution by feature z-score and signature average by binary outcome, on latent space position. **(a)** Temporal signatures are plotted for relevant variables grouped by increasing z-scores. For each variable, individuals were stratified into five evenly spaced z-score bins within the range $[-2, 2]$, and the mean latent signature was computed for each group. Feature variations shift signatures from the population average signature, in black on each plot. **(b)** The average signature is plotted for each cardiovascular risk, prevalent and incident outcome. HDL, high-density lipoprotein; HbA1c, glycated hemoglobin; PRS, polygenic risk score; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; LV, left ventricle; EF, ejection fraction; WT, wall thickness; MRI, magnetic resonance imaging.

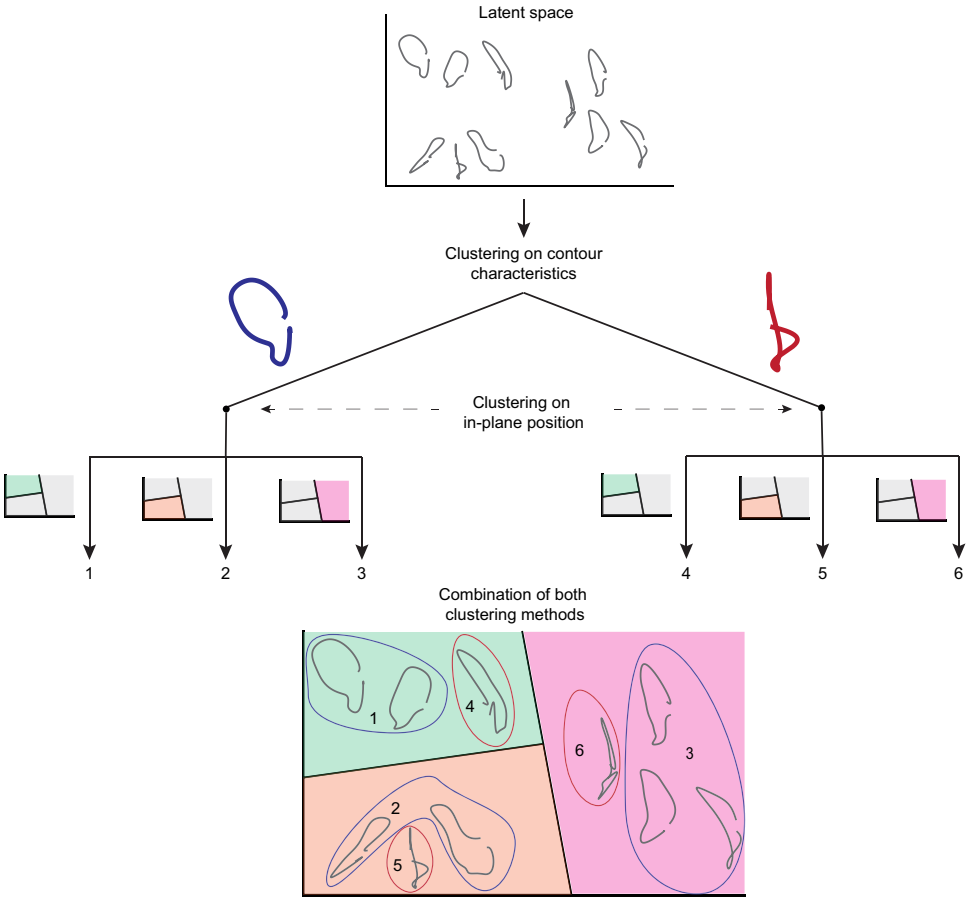


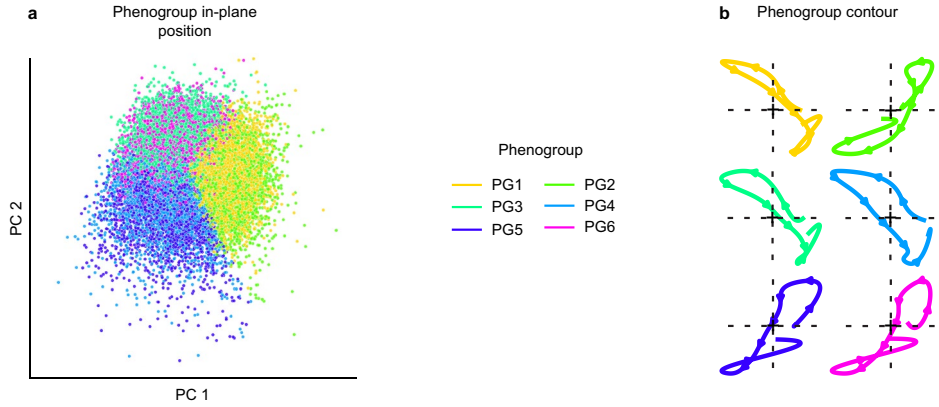
Fig. 5. Combination of clustering on in-plane position and contour. Temporal signatures were clustered in two stages. A few individual examples are shown here to simplify the figure. First, the signatures were centred around the origin and clustered on their relative coordinates, capturing the contour of the trajectories, resulting in two clusters. Second, clustering was done on the original, uncentred coordinates of the temporal signatures, capturing their in-plane position, and revealing three clusters. The results of both methods were combined to give a dual clustering of the latent space, giving six clusters numbered from 1 to 6.

the χ^2 test performed to assess the enrichment of the observed discrete feature in one phenogroup against the rest of the population. A summary of findings for numerical features is found in Supplementary Fig. 2. All the statistical test results are detailed in the Supplementary Data.

Associations with risk factors and outcomes allowed us to construct an interpretative table of phenogroups from PG1 to PG6 ranked by increasing burden of cardiovascular risks (Fig. 6c). Motion traits in PG1 were associated with reduced cardiovascular risk factors including hypertension (OR: 0.53; 95% confidence interval [CI]: 0.47, 0.59), diabetes (OR: 0.63; 95% CI: 0.54, 0.73), and obesity (OR: 0.34; 95% CI: 0.24, 0.49). It also exhibited the highest contractile function (radial strain (47.5%; 95% CI: 47.2, 47.8) and ejection fraction (61.8%; 95% CI: 61.6, 62)). In PG2 cardiometabolic risks were slightly higher (hypertension (OR: 0.71; 95% CI: 0.63, 0.80), diabetes (OR: 0.61; 95% CI: 0.51, 0.73), and obesity (OR: 0.57; 95% CI: 0.40, 0.81)). The least differentiated group for risk factors was PG3 with no significant associations. PG4 had the highest risk of diabetes (OR 1.75; 95% CI: 1.50, 2.04), and was also enriched for smoking (OR: 1.19; 95% CI: 1.10, 1.29) and hypertension (OR: 1.37; 95% CI: 1.23, 1.52). This group also had a significant risk for prevalent coronary artery disease (CAD) (OR: 1.74; 95% CI: 1.42, 2.12) and myocardial infarction (OR: 1.43; 95% CI: 1.07, 1.92).

PG5 had the highest risk for hypertension (OR: 1.65; 95% CI: 1.48, 1.83), and was also associated with diabetes (OR: 1.51; 95% CI: 1.29, 1.78) and obesity (OR: 1.58; 95% CI: 1.22, 2.07). This group also had lower ejection fraction (58.2%; 95% CI: [57.9, 58.5]) and radial strain (42.4; 95% CI: [42, 42.8]). PG5 had the most enrichment for stroke (OR: 1.93; 95% CI: 1.16, 3.21), prevalent (OR: 2.52; 95% CI: 1.62, 3.93) and incident (OR: 2.13; 95% CI: 1.48, 3.06) heart failure, and a future diagnosis of DCM (OR: 5.09; 95% CI: 1.77, 14.68).

Lastly, PG6 was linked to genetic risk factors including polygenic scores for hypertension and DCM. This group had the lowest ejection fraction (56.9%; 95% CI: [56.6, 57.1]) and radial strain (41; 95% CI: [40.7, 41.3]), as well as the highest maximum wall thickness (10.1 mm; 95% CI: [9.99,10.14]). PG6 was enriched for conventional risk factors, as well as prevalent adverse cardiovascular outcomes: CAD (OR: 1.34; 95% CI: 1.10, 1.64), heart failure (OR: 1.76; 95% CI: 1.13, 2.76), and myocardial infarction (OR: 1.85; 95% CI: 1.43, 2.39). PG6 had the most associations with incident outcomes of all the phenogroups: CAD (OR: 1.68; 95% CI: 1.30, 2.18), heart



c

		Phenogroup 1 Lowest risk	Phenogroup 2 Low risk	Phenogroup 3 Undifferentiated	Phenogroup 4 Multiple risks	Phenogroup 5 High risk with events	Phenogroup 6 Highest risk with events
Cardiovascular & genetic risks	Ever smoked (OR)	0.82 (0.77,0.88)	0.83 (0.77,0.89)	1.09 (1.01,1.17)	1.19 (1.10,1.29)	ns	1.19 (1.10,1.28)
	PRS DCM	-4.37 (-4.38,-4.35)$_{5,6}$	-4.35 (-4.37,-4.33)^{ns}	-4.34 (-4.36,-4.32)^{ns}	-4.34 (-4.36,-4.32)^{ns}	-4.33 (-4.35,-4.32)^{1}	-4.33 (-4.35,-4.31)^{1}
	PRS HT	-0.17 (-0.21,-0.14)$_{2,4,6}$	-0.10 (-0.15,-0.06)$_{6}^{1}$	-0.07 (-0.11,-0.03)^{ns}	-0.10 (-0.14,-0.05)$_{6}^{1}$	-0.08 (-0.12,-0.03)^{ns}	-0.03 (-0.07,0.006)^{1,2,4}
	Hypertension (OR)	0.53 (0.47,0.59)	0.71 (0.63,0.80)	ns	1.37 (1.23,1.52)	1.65 (1.48,1.83)	1.19 (1.07,1.32)
	Diabetes (OR)	0.63 (0.54,0.73)	0.61 (0.51,0.73)	ns	1.75 (1.50,2.04)	1.51 (1.29,1.78)	1.19 (1.02,1.39)
	Obesity (OR)	0.34 (0.24,0.49)	0.57 (0.40,0.81)	ns	ns	1.58 (1.22,2.07)	1.59 (1.25,2.03)
Cardiac function	EF (%)	61.8 (61.6,62.0)*	59.9 (59.7,60.2)*	59.5 (59.3,59.8)*	60.7 (60.5,61.0)*	58.2 (57.9,58.5)*	56.9 (56.6,57.1)*
	Max LV WT (mm)	8.29 (8.24,8.34)*	8.54 (8.47,8.61)*	9.65 (9.58,9.72)*	9.45 (9.38,9.53)*	9.71 (9.64,9.79)*	10.06 (9.99,10.14)*
	Radial strain (%)	47.5 (47.2,47.8)*	45.3 (45.0,45.7)*	44.4 (44.1,44.7)*	46.1 (45.7,46.5)*	42.4 (42.0,42.8)*	41.0 (40.7,41.3)*
Prevalent outcomes	Stroke (OR)	0.30 (0.14,0.65)	ns	ns	ns	1.93 (1.16,3.21)	ns
	CAD (OR)	0.39 (0.30,0.51)	0.48 (0.36,0.64)	ns	1.74 (1.42,2.12)	1.65 (1.34,2.03)	1.34 (1.10,1.64)
	Heart failure (OR)	0.22 (0.09,0.49)	0.40 (0.18,0.86)	ns	ns	2.52 (1.62,3.93)	1.76 (1.13,2.76)
	MI (OR)	0.27 (0.18,0.42)	0.55 (0.38,0.81)	ns	1.43 (1.07,1.92)	1.74 (1.31,2.31)	1.85 (1.43,2.39)
Incident outcomes	CAD (OR)	0.36 (0.25,0.52)	0.69 (0.49,0.98)	ns	ns	1.39 (1.03,1.88)	1.68 (1.30,2.18)
	HF (OR)	0.32 (0.18,0.55)	ns	ns	ns	2.13 (1.48,3.06)	1.62 (1.13,2.32)
	MI (OR)	0.37 (0.24,0.56)	0.66 (0.44,0.98)	ns	ns	ns	2.08 (1.57,2.74)
	Cardiac arrest (OR)	0.15 (0.02,1.10)	ns	ns	ns	ns	7.03 (3.12,15.84)
	DCM (OR)	ns	ns	ns	ns	5.09 (1.77,14.68)	3.76 (1.30,10.83)

Legend

Report for continuous features

median (CI lower bound, CI upper bound)$^{A}_{B}$

A = * significant differences with all pairs of groups
= ns no significant difference with any other pair
= i,j significantly higher than superscripted groups i and j
B = k,l significantly lower than subscripted groups k and l

Report for discrete features

odds ratio (CI lower bound, CI upper bound)
ns = no stastistical enrichment

Fig. 6. Clinical interpretation of motion-based phenogroups. **(a)** Barycentre position of all individual trajectories, coloured by their assigned phenogroups, in the latent space of the two principal components. **(b)** Average signatures for each phenogroup centred around the origin of the latent space. **(c)** Summary of findings characterising the cardiovascular risks and disease status of the phenogroups. Participant characteristics and diagnostics were used to highlight the differences between groups. Findings are reported as the median confidence interval for continuous variables (Fig. 7) and odds ratio and direction of association for discrete features (Fig. 8) following statistical testing. A Kruskal-Wallis test followed by a post-hoc Dunn test was used for continuous variables, while a χ^2 test followed by a post-hoc Fisher test was performed for discrete variables. Multiple testing was accounted for by the Benjamini-Hochberg procedure, and the level of significance was $P < 0.05$. Sample size was n = 20,318. PRS DCM has been multiplied by 10^7 for readability. PRS, polygenic risk score; DCM, dilated cardiomyopathy; HT, hypertension; EF, ejection fraction; LV, left ventricle; WT, wall thickness; mm, millimeters; CAD, coronary artery disease; MI, myocardial infarction; CI, confidence interval.

failure (OR: 1.62; 95% CI: 1.13, 2.32), myocardial infarction (OR: 2.08; 95% CI: 1.57, 2.74), cardiac arrest (OR: 7.03; 95% CI: 3.12, 15.84) and a diagnosis of DCM (OR: 3.76; 95% CI: 1.30, 10.83).

Method stability and benchmark models

A stability analysis was performed to assess the robustness of the clustering method. The analysis also revealed the optimal number of clusters as per the maximum average Rand index (ARI) score[17], which was three for the in-plane position clustering (Supplementary Table 1) and two for the trajectory contour clustering (Supplementary Table 2). The ARI scores range from -1 to 1, with 1 being a perfect match between two clustered sets.

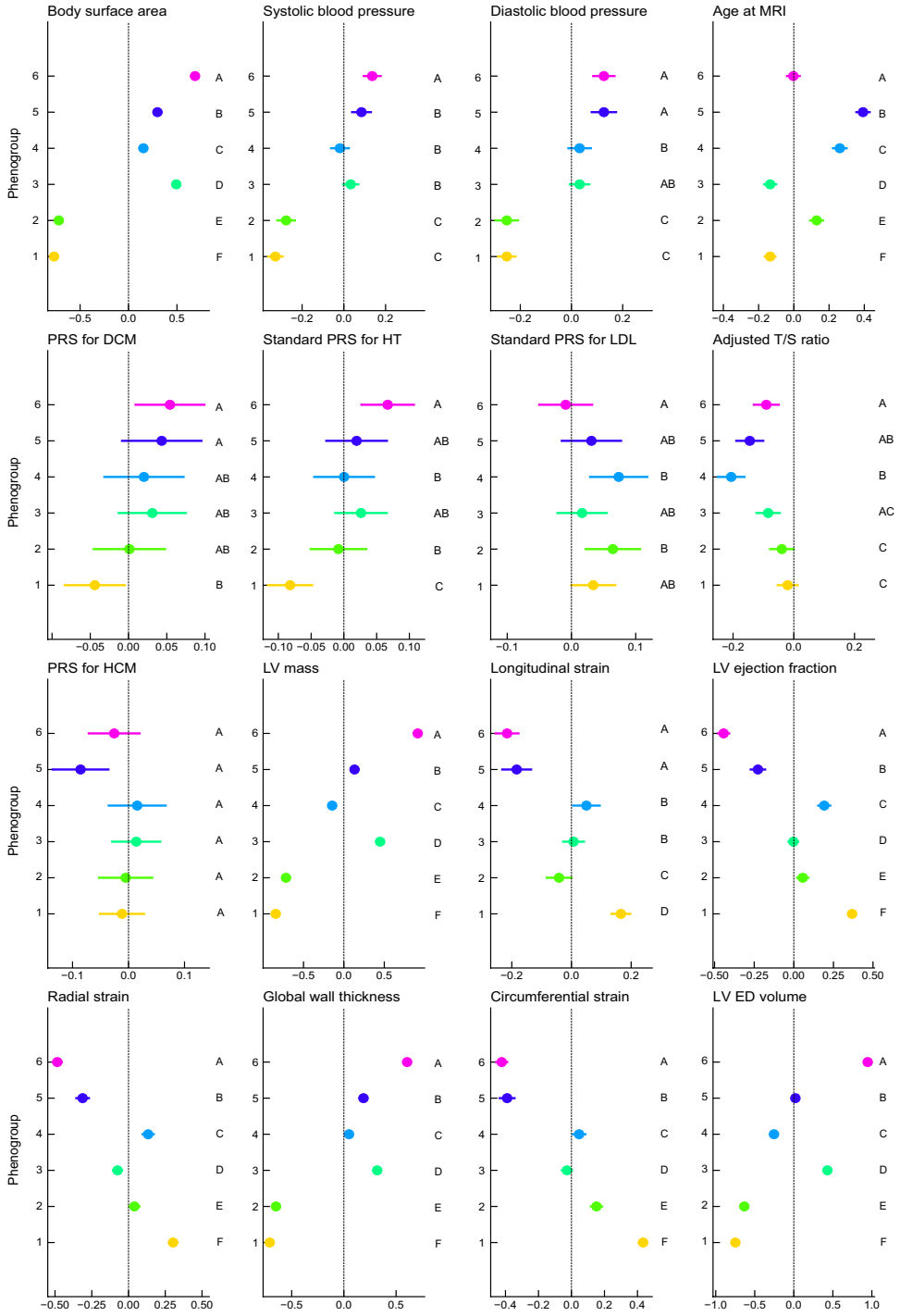


Fig. 7. Numerical features significantly associated with motion-based phenogroups. Numerical features were tested for significant associations with the phenogroups. A Kruskal-Wallis test was performed, followed by a post-hoc Dunn test revealing pairs of phenogroups sharing significant differences in variable distribution. Benjamini-Hochberg correction for multiple testing was done, taking into consideration discrete features as well (see Supplementary Data for all tested features). The significance level was $P < 0.05$ and the sample size was n = 20,318. The median of the z-score and its confidence interval are displayed for each phenogroup. For a given feature, groups share a letter if there is no significant difference between them for that feature. MRI, magnetic resonance imaging; DCM, dilated cardiomyopathy; PRS, polygenic risk score; HT, hypertension; LDL, low-density lipoprotein; T/S, telomere to single copy gene; HCM, hypertrophic cardiomyopathy; LV, left ventricle; ED, end-diastolic.

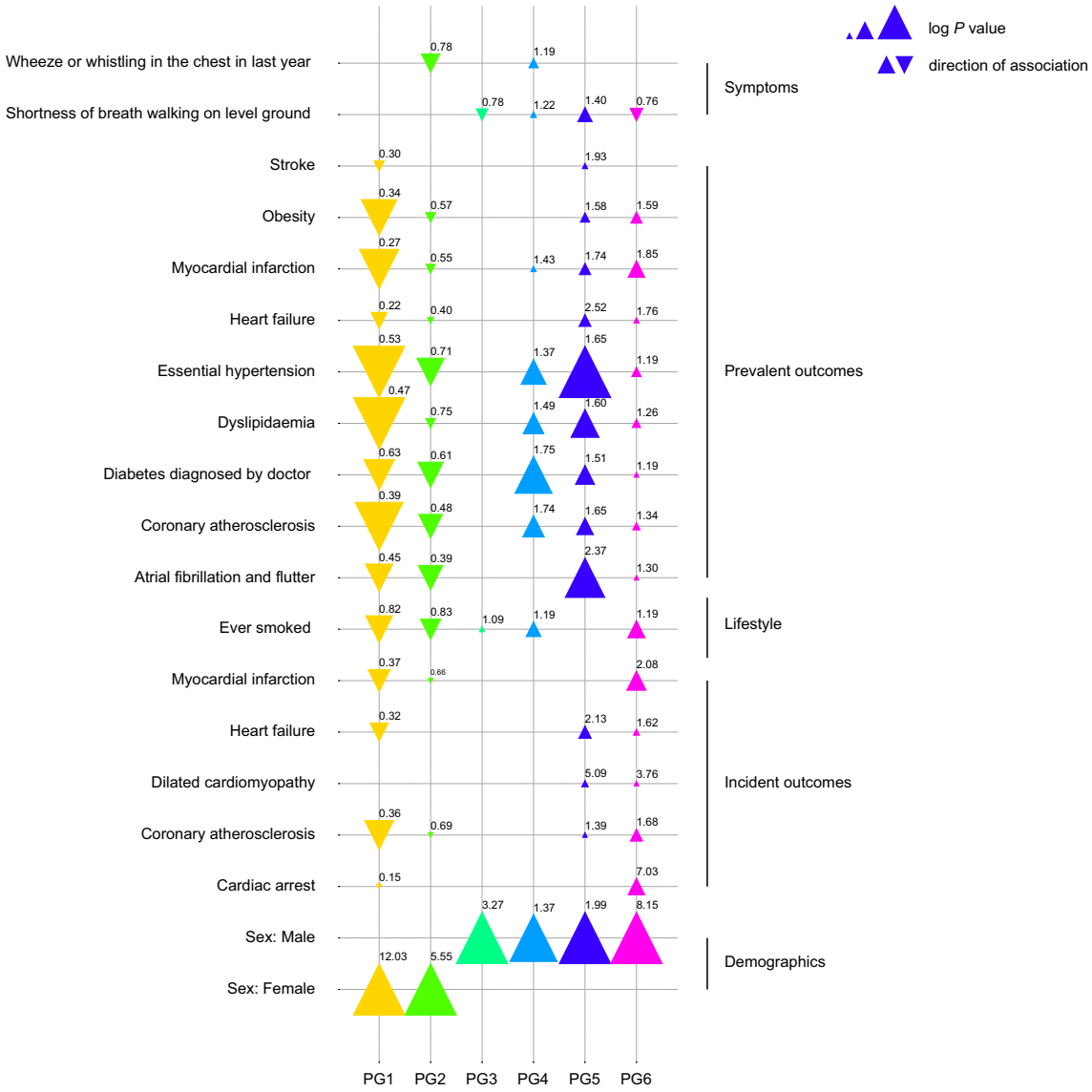


Fig. 8. Discrete and binary features significantly associated with motion-based phenogroups. Discrete features were tested for significant associations with the phenogroups. A χ^2 test was used, followed by a post-hoc Fisher's unit test which highlighted significant enrichment for discrete features. Benjamini-Hochberg correction for multiple testing was done, taking into consideration numerical features as well (refer to Data for all tested features and detailed statistical results). The significance level was $P < 0.05$ and the sample size was n = 20,318. The size of the triangle reflects the logarithm of the P value obtained with each statistical test while its direction is of the association. The odds ratio is also displayed for each significant association.

Our method was compared to two benchmark models. First, we swapped the latent encodings of high-dimensional point clouds for 28 different CMR-derived tabular measures spanning volumes, mass and strain (Supplementary Table 4) to show that latent representations of motion yield more statistically robust and stable phenotypic clusters. Second, we changed our dimensionality reduction method to a simple linear PCA applied directly on the cardiac point clouds.

The comparison of our framework to the two benchmark models is summarised in Supplementary Table 5 with our framework having an ARI of 0.91. Using CMR-derived clinical measures as input, UMAP (uniform manifold approximation and projection)[18] was applied to reduce the dataset dimension to 16, the same as our model bottleneck, followed by PCA. A K-means algorithm was used to cluster the resulting principal components. This resulted in a lower ARI of 0.82. Using simple PCA to reduce dimension, keeping everything else equal, gave an ARI of only 0.50 and 0.85 for clustering by in-plane position and contour, respectively, giving a less stable method when comparing with our scores of 0.914 and 0.904.

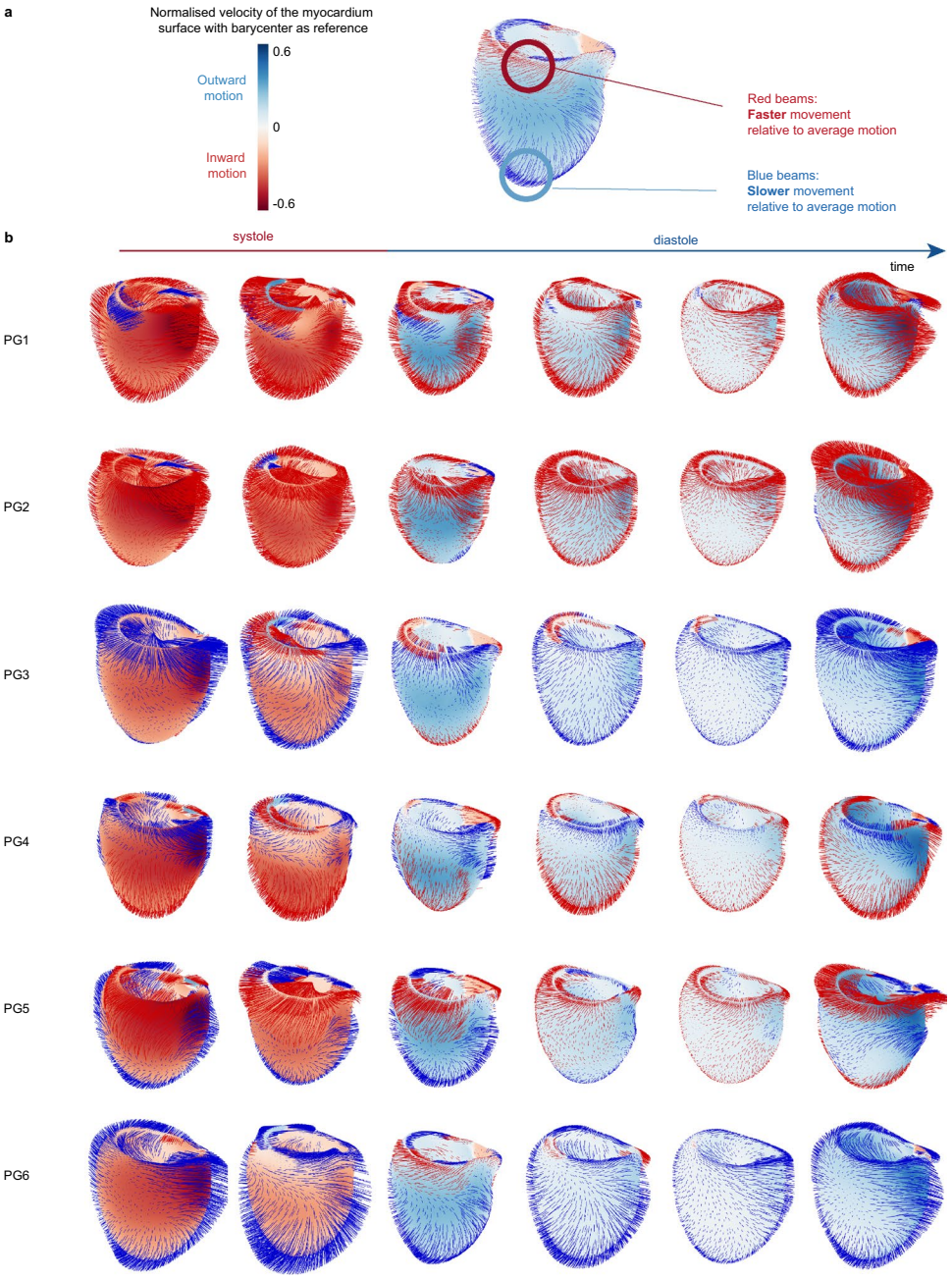


Fig. 9. Left ventricle motion velocity by phenogroup - inferolateral view. **(a)** Zoomed in examples of the left ventricle during systole and diastole. The velocity of the ventricle wall is displayed on the surface, relative to its barycentre. The velocity is positive during relaxation, or diastole when the ventricle is filling with blood, while it is negative during contraction, or systole when ventricular ejection occurs. Red beams indicate areas of faster movement relative to the population average motion, while blue beams show areas of relative slower movement. **(b)** The left ventricle motion was compared to the population's average motion for each phenogroup. The six frames represent one entire cardiac cycle from systole (contraction of the ventricle) to diastole (relaxation of the ventricle). The velocity of the ventricle was calculated as the local gradient of change between two consecutive frames, averaged out over multiple frames to reduce the number of displayed frames, and normalized for the size of the ventricle.

Visualising phenogroup motion patterns

Phenogroup motion patterns were computed by comparing average point velocity of the ventricular surface to the population average (inferolateral: Fig. 9b and anteroseptal: Supplementary Fig. 3). In the higher risk groups we observed characteristic changes in patterns of cardiac motion. In PG4, where cardiometabolic risks were highest, we observed reduced velocities in the basal regions of the left ventricle in both systole and diastole suggesting there are regionally localised changes. For PG5 reduced velocities were more globally distributed in

the left ventricle but confined predominately to systole and early diastole. In contrast, for PG6 with the highest burden of poor outcomes, reduced function extended throughout diastole as well. The lowest risk phenogroups show comparatively faster systolic velocities especially in the basal ventricle as well as more rapid diastolic relaxation. These findings show how spatio-temporal patterns of motion vary between the discovered clusters.

Discussion

The heart is a highly complex electromechanical organ system but conventional assessment of dynamic contraction relies on pre-defined volumetric, deformational and flow parameters that may not provide efficient evaluation of early dysfunction or advancing disease. Our study shows there is latent information in point cloud data representing patterns of cardiac motion that predict future events, capture the effect of risk factors on function, and provide a new classification of motion-based traits in an adult population. This approach also produces human-interpretable signatures to assist in understanding how function deviates across multiple dimensions of health.

Ejection fraction is an established marker of global ventricular function which is representative of volumetric approaches for assessing cardiac status and defining a diagnosis of heart failure[3,19]. It is also poorly predictive of symptoms or treatment response with various thresholds used to define putative underlying mechanisms[20,21]. Myocardial strain is a more sensitive marker of contractile function, can be measured globally or regionally, and is independently predictive of outcomes[22]. Other approaches include tissue Doppler imaging for regional velocity assessment and speckle tracking for two-dimensional deformation analysis, both widely available in echocardiography. Tagged MRI provides transmural deformation data but requires dedicated acquisition sequences rarely available at population scale. Displacement Encoding with Stimulated Echoes (DENSE) MRI offers direct measurement of myocardial displacement but remains largely confined to research settings[23,24]. However, while strain is formally a second-order tensor encompassing normal and shear components, in clinical practice it is typically reported as scalar measures of myocardial shortening or thickening[24], offering only a narrow assessment of cardiac mechanics. Statistical shape modelling of the heart hints at the potential for even static whole-heart features to encode latent information on genetic regulation[7,25,26]. Here we use a strategy of creating a dynamic three dimensional point cloud, densely sampling the left ventricle, to capture a rich feature space of motion throughout the whole cardiac cycle. This concept applied to patients with heart failure is strongly predictive of outcomes[1], but has not previously been used in large populations or for discovery of disease mechanisms. Clustering of conventional image-derived phenotypes has found mechanistically-distinct groups in patients with DCM[27], and congenital heart disease for instance[28]. Clustering of static three-dimensional shape data can also reveal mechanistic and prognostic patterns in HCM[29], but there is limited knowledge of undertaking this on moving point clouds. To achieve this we created a latent space of cardiac motion using a task-specific VAE framework and visualised this as a low dimensional signature that could be used for stable clustering.

We found that a low dimensional representation of whole-ventricle motion captured associations with circulating biomarkers for diabetes and coronary artery disease. Established diabetes has been causally linked to diastolic dysfunction[30], suggesting that our models are sensitive to changes in myocardial stiffness without relying on predefined measures of relaxation. We also assessed polygenic scores that estimate predisposition to cardiomyopathies that are defined by static phenotypes, finding that they also have an effect on cardiac motion. The association between common genetic variants and heart motion in the adult population is not well understood but abnormal contractility, diastolic dysfunction and dyssynchrony are recognised in established disease[31]. Conventional cardiovascular risk factors such as obesity and hypertension were associated with motion traits and we found enrichment for future cardiac events in the latent space. We also observed sex-related differences in motion traits highlighting the potential for three-dimensional analysis in revealing sex-specific functional characteristics that may underlie variance in disease penetrance[32]. Taken together, this shows the potential for unsupervised motion analysis to discover features that relate to a range of environmental and genetic risk factors which are also prognostically important.

From a translational perspective, we envision two pathways for clinical deployment. First, as a risk stratification tool: once phenogroups are defined on a reference population, a new patient's CMR data can be projected into the trained latent space and assigned to the nearest phenogroup through a computationally inexpensive forward pass, analogous to mapping genomic data onto polygenic risk categories. Second, as a discovery platform: the framework reveals that 4D motion data contains prognostically important information beyond conventional scalar parameters, motivating the development of simpler motion-derived biomarkers (e.g. a continuous risk score from latent space position) that could complement existing clinical metrics such as ejection fraction and strain. We emphasise that the current work represents a research framework and deployment in clinical decision-making would require prospective validation, regulatory evaluation, and integration into existing imaging workflows.

To facilitate clinical interpretation, the principal components of the motion signatures can be understood as distinct modes of deformation. As illustrated in Fig. 4, traversing PC1 correlates strongly with global contractility (ejection fraction), while PC2 captures variations in temporal dyssynchrony and regional wall motion abnormalities. Clinicians can interpret a patient's location in this space as a data point and as a composite metric of cardiac function. For instance, outliers in the lower-right quadrant consistently represent impaired cardiovascular function characterised by incident outcomes and cardiomyopathies.

Clustering identified six phenogroups with varying enrichment for cardiovascular risk factors and outcomes. The lowest risk groups (PG1 and PG2) had motion traits associated with the least risk for obesity, diabetes and hypertension, as well as both prevalent and incident outcomes. This suggests that healthy individuals with a low risk of adverse outcomes can be identified through motion analysis, and that the approach is not limited to those with established or sub-clinical disease. The cluster most strongly enriched for cardiometabolic disease was

PG4. Diabetic cardiomyopathy is related to hypertrophy and fibrosis, early-onset diastolic dysfunction and late-onset systolic dysfunction[33]. Here we show even in a broad population of community-living adults that motion traits can be associated with metabolic risk. The last two clusters (PG5 and PG6) had the strongest enrichment for both prevalent and incident endpoints, as well as highest polygenic risks. PG6 was the only cluster marked by higher incident myocardial infarction and cardiac arrest, suggesting that motion traits are predictive of fatal outcomes. Conventional parameters including ejection fraction and radial strain were lowest in PG5 and PG6 but we found that even a broad range of CMR parameters did not result in stable clustering. While these phenogroups recapitulate established cardiovascular risk patterns, supporting the physiological validity of an entirely unsupervised analysis, several findings extend beyond what is accessible through conventional volumetric or strain-based assessment: notably the enrichment of PG6 for incident cardiac arrest (OR: 7.03; 95% CI: 3.12, 15.84) and the differential polygenic risk profiles distinguishing PG5 (hypertension, DCM) from PG6 (HCM, DCM), neither of which is captured by ejection fraction or scalar strain alone.

Our convolutional VAE generates temporal embeddings from time windows of cardiac motion data, and avoids flattening temporal data before passing it to the model. Our model also provides interpretability by preserving temporal characteristics of the data in its output. This methodology could be extended to longitudinal imaging datasets to reveal temporal trajectories of disease progression, further improving risk modelling and targeted clinical interventions.

There are some limitations to consider. First, regarding the scope of the input data: this framework currently models only left ventricular motion. Pathologies affecting the right ventricle or atria are not captured, which is a limitation for conditions with biventricular involvement. Extensions to multi-chamber point clouds are planned. Additionally, the point clouds represent the endocardial and epicardial surfaces rather than the full transmural myocardium, capturing intra-myocardial motion gradients and shear would require tagged MRI, which is not available at this scale. While the CMR contained 50 temporal frames it does not have the temporal resolution of echocardiography. It is also not as widely available for inference as echocardiography.

Second, there are methodological considerations in the processing pipeline. The reference atlas was constructed from over 1000 healthy adults, which may introduce a regularisation bias: while non-rigid registration allows substantial deformation from the atlas template, extreme pathological morphologies (e.g. aneurysmal dilation) may be partially attenuated during the fitting process. Independent frame-by-frame validation of the motion-tracked meshes against per-phase segmentations was not feasible at this scale due to the absence of ground-truth annotations at all 50 temporal frames; however, the motion estimation method employs forward-backward tracking with sparsity regularisation to minimise accumulated alignment errors[10]. Additionally, although the cardiac cycle is physiologically periodic, the latent space trajectories may exhibit slight discontinuities between the first and last timeframes due to the absence of an explicit cyclic constraint in the CVAE loss function and minor interpolation artefacts at the temporal boundaries. Future work could enforce periodicity via a cyclic loss term.

Third, regarding the model architecture: the convolutional approach leverages the fixed vertex ordering enforced by atlas registration. Graph neural network (GNN) architectures that explicitly model mesh topology represent a promising alternative, particularly for multi-chamber models where consistent vertex correspondence may be harder to enforce.

The population being studied also has limitations including under-representation of older age groups and persons living in less socioeconomically deprived areas. The UK Biobank population is also predominantly European and further work is required in people of diverse ancestries and social groups. We performed a robust analysis of model stability and comparison to a benchmark of conventional parameters but there is no equivalent data available at scale for external validation. A complementary stability analysis in which PCA was recomputed on cohort subsets showed that the in-plane position partition was highly reproducible (ARI: 0.91 ± 0.05) while the contour partition was more sensitive to PCA recomputation (ARI: 0.59 ± 0.23), and conclusions drawn specifically from contour-based features should be interpreted with that variability in mind.

In conclusion, we show how time-resolved three dimensional representations of cardiac motion contain rich information on cardiovascular risks, disease mechanisms and outcomes while providing human interpretable signatures.

Methods

UK Biobank participants

UK Biobank is a large cohort study on half a million individuals in the UK aged 40–69 at recruitment between 2006 and 2010[34]. All participants provided written informed consent for the study, which was also approved by the National Research Ethics Service, North West Multi-centre Research Ethics Committee (11/NW/0382). All research was performed in accordance with relevant guidelines and regulations in accordance with the Declaration of Helsinki. Our study was conducted under terms of access approval number 28807 and 40616. A collection of a wide range of phenotypic and genotypic attributes, risk factors and physical measures is available. They were obtained by touch screen questionnaires, interview, biophysical measurement, hospital episode statistics, and primary care data. CMR images were also acquired for a subset of participants. Details of how each phenotype was acquired are available on the UK Biobank Showcase (http://biobank.ctsu.ox.ac.uk/crystal/).

Cardiovascular diseases and outcomes were derived from groupings of ICD-10 codes as described in Supplementary Table 6. Outcomes were classified as prevalent or incident if they happened before or after the imaging visit. Polygenic risk scores (PRS) for HCM and DCM were derived from genetic sequencing data identifying rare pathogenic and likely pathogenic variants in the UK Biobank population and other cohorts[35–37]. All the other mentioned PRS were accessed through the UK Biobank PRS Release[38] .

Cardiac imaging segmentation and strain analysis

CMR imaging was performed on participants to capture two-dimensional retrospectively-gated cine imaging on a 1.5T magnet (Siemens Healthineers, Erlangen, Germany)[39]. Left ventricular short-axis plane cine images from base to apex were acquired as well as long-axis cine images in the two and four-chamber views. All cine sequences comprised a constant number of 50 frames to normalise for heart rate variability, establishing a unit cardiac cycle. The original temporal resolution was 31.56 ms. It is important to note that standard retrospective ECG-gated cine MRI reconstructs a single representative cardiac cycle from data acquired over multiple heartbeats, therefore, beat-to-beat variability is averaged out during image reconstruction. While physiological states influence cardiac dynamics, this study utilises resting-state scans, minimising intra-subject physiological variance. Ascending and descending thoracic aorta images were also acquired with transverse cine imaging.

The two-dimensional cine images in the short and long-axis were segmented using fully convolutional networks, with segmentation quality equivalent to expert human readers[9]. Derived measurements of the left and right ventricles were determined from the two-dimensional images: end-diastolic and end-systolic volumes, stroke volume, and ejection fraction. A density of 1.05 g.ml^{-1} for the left ventricular myocardium was used to derive its mass from its volume. With A_{2Ch} and A_{4Ch} as the atrial areas on the two and four-chamber cine views respectively, and L as the averaged longitudinal diameter across two views, the atrial volumes were derived as per the following formula on biplane area-length: $V = \frac{8}{3\pi} \cdot \frac{A_{2Ch} \cdot A_{4Ch}}{L}$ [40,41]. CMR-derived measurements were indexed to body surface area (BSA) calculated with the Du Bois formula: BSA = $0.007184 \times \text{Height}^{0.725} \times \text{Weight}^{0.425}$ with height in cm and weight in kg. The wall thickness of the left ventricular myocardium was measured as the distance between the entire segmented epicardium and endocardium at end-diastole.

Circumferential and radial strains were calculated with cine short-axis images as per the formula $E_{dir} = \frac{\Delta L_{dir}}{L_{dir}}$ with *dir* the circumferential or radial direction, L_{dir} the length of a line segment in this direction and ΔL_{dir} the change of length over time. Longitudinal strain was derived from long-axis four-chamber motion tracking.

A spatio-temporal neural network was used to segment the aorta cine images[42]. Maximum and minimum cross-sectional areas were derived and aorta distensibility was calculated with central blood pressure obtained using peripheral pulse-wave analysis (Vicorder, Wuerzburg, Germany)[43].

Three dimensional motion capture

Motion was represented in a three-dimensional biventricular model by fitting an atlas-derived mesh to a stack of 2D segmentations and aligning these to a reference atlas. This atlas was defined from over 1000 healthy adults which enforced spatial and temporal consistency[10]. The processing pipeline consists of three main steps, as illustrated in Fig. 2.

Segmentation: a multi-task fully convolutional network[11] was used to simultaneously predict bi-ventricular segmentation labels and six anatomical landmarks from each CMR volume. The network was trained with a Dice loss for segmentation and a class-balanced cross-entropy loss for landmark detection. To ensure topological consistency and point correspondence across the cohort, a shape refinement step was applied using atlas propagation from a cohort of high-resolution atlases constructed from over 1000 healthy subjects[10]. The predicted landmarks were used to compute affine transformations between each subject and the atlas space, and the five most similar atlases were selected based on normalised mutual information. Each selected atlas was non-rigidly registered to the target using B-spline free-form deformations with a label consistency objective, implemented with the IRTK software package[10,11]. Final segmentations were obtained by non-local patch-based label fusion, producing anatomically smooth 3D models. In the original validation, atlas-to-subject registration achieved Dice similarity coefficients of 0.950 ± 0.024 (LV cavity) and 0.824 ± 0.062 (myocardium) with mean surface distances below 2 mm[10], and the segmentation pipeline achieved Dice similarity coefficients above 0.90 for LV and RV cavities with landmark errors within inter-observer variability[11].

Mesh fitting: after segmentation, each participant's refined segmentation was used to generate a subject-specific end-diastolic mesh by non-rigidly fitting the atlas template mesh to the segmentation surface, reconstructed using a marching-cubes algorithm. B-spline non-rigid registration was regularised implicitly through the smoothness of the basis functions. This process ensured spatial consistency across the cardiac cycle and preserved anatomical correspondence of mesh vertices over time, such that each vertex index consistently represents the same anatomical location across all participants.

Three dimensional time-resolved motion estimation: cardiac motion was estimated using a "four dimensional" spatio-temporal free-form deformation applied to the cine images with sparsity regularisation[10]. Displacement fields were computed between successive time frames in both forward and backward directions to reduce alignment errors. These motion fields were applied to the subject-specific mesh to generate a sequence of 50 deformed surface meshes each corresponding to a distinct time point in the cardiac cycle. Temporal normalisation was performed to ensure consistency across frames resulting in meshes with identical vertex counts that maintain anatomical correspondence both temporally and spatially across participants.

Data pre-processing

Point clouds were extracted from the cardiac meshes as ordered collections of vertices with fixed anatomical correspondence enforced by atlas registration. The point clouds described the movement of the endocardial and epicardial surfaces of the left ventricle during one periodic cardiac cycle. Point clouds were down-sampled by a 97% factor using quadratic clustering in Python using the `vtk` package, from 40,616 to 1406 points while preserving geometric topology and curvature features. This does not affect any subsequent analysis as demonstrated in a previous study using cardiac motion data for prediction and risk-stratification[1]. Point clouds were then transformed to fit in a unit sphere: they were centred around the origin by subtracting their centroid

and normalised by isotropic rescaling of the physical extent, dividing the coordinates by the distance of the most distant point to the centroid across the whole dataset. This pre-processing preserves intra-population size variability because a single global scaling factor, derived from the maximum extent across the entire dataset, is applied uniformly to all subjects rather than normalising each heart individually. All coordinates are mapped into the $[-1, 1]$ range, which is needed before feeding the data into the neural network. Prior to the construction of time windows, a semi-automated quality control procedure was applied. The Hausdorff distance between each frame's mesh and a reference shape profile was computed, and subjects exceeding empirically calibrated thresholds were removed. A PCA-based outlier detection step on the mesh coordinates identified additional cases with non-physiological shapes. Thresholds for both steps were determined by visual inspection of flagged cases to distinguish genuine anatomical variation from processing artefacts. Subjects for whom the mesh decimation pipeline failed due to upstream segmentation errors were also excluded.

Windows of cardiac motion were constructed along the cardiac cycle using a sliding window of 16 frames, a stride of 4 frames and padding at the end to ensure periodicity (Supplementary Fig. 4a). These parameters were chosen to capture an accurate representation of the cardiac cycle in each window while limiting the use of computational power. Consecutive windows overlap each other to ensure the continuity of the cardiac cycle. This resulted in 13 windows for each cardiac cycle as per the ceiling value of the number of total frames divided by the stride. The 1406 vertices were downsampled to 1024 using bilinear interpolation along the vertex dimension of the data tensor. Because the fixed vertex ordering from atlas registration ensures that adjacent indices correspond to spatially proximate anatomical locations, this interpolation preserves spatial continuity and is consistent with the downstream treatment of the vertex dimension as a spatial axis in the CVAE. The target size of 1024 facilitates feature extraction consistent with powers of two (2^{10}).

Convolutional variational autoencoder

The input x of the model is a 3-channel tensor of size $3 \times 16 \times 1024$, where the three channels represent the X, Y and Z coordinates, 16 is the number of time points in a window, and 1024 is the number of points in each point cloud (Supplementary Fig. 4b).

A convolutional variational autoencoder (CVAE) was implemented using PyTorch to encode these spatio-temporal matrices into a compact latent representation. The encoder and decoder were made of successive 2D convolutional and transposed convolutional layers, respectively. The input to the CVAE was formatted as a 2D tensor of shape $(3 \times T \times N)$, where $T = 16$ represents timeframes and $N = 1024$ represents vertices. The (x, y, z) coordinates were treated as channels. Consequently, 2D convolutional layers were employed to extract features across both the temporal and vertex domains simultaneously, enabling the model to learn local spatial and temporal features hierarchically. The encoder used four 2D convolutional layers with kernel size 4, stride 2 and zero padding. The number of feature channels was progressively doubled after each convolution, starting from 8 channels in the first layer. This hierarchical structure allowed the model to capture increasingly abstract features. The architecture of the model is described in detail in Supplementary Fig. 5. Following the convolutional layers, an adaptive average pooling layer reduced each feature map to a single value, irrespective of its input size. This operation produced a fixed-size feature vector that is invariant to small changes in input dimensions, facilitating processing by subsequent fully connected layers. Adaptive pooling is suitable in our case as it standardises feature map sizes before flattening, simplifying the subsequent latent space parametrisation. The pooled feature vector was then passed through two parallel fully connected layers to produce the latent parameters: the mean vector μ and the standard deviation vector σ of a multivariate Gaussian distribution. The latent vector z was then sampled using the reparametrisation trick:

$$z = \mu + \sigma \odot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \tag{1}$$

where $\odot$ denotes element-wise multiplication. This technique allows the gradient to be propagated through the stochastic sampling operation by expressing the random variable z as a deterministic function of μ, σ and a noise term ϵ[44]. The decoder mirrored the architecture of the encoder, beginning with a fully connected layer that projected the latent vector back to a higher-dimensional feature space. The output was reshaped and successively upsampled using four 2D transposed convolutional layers with the same kernel size, stride, and padding settings as the encoder, progressively halving the number of channels while restoring the original matrix dimensions. This design choice allowed the model to reconstruct the spatial configuration of the cardiac geometry while leveraging the hierarchical features captured by the encoder.

The CVAE was optimised using a loss function combining a reconstruction loss and a Kullback-Leibler (KL) divergence term as per the β-VAE framework[45,46]:

$$\mathcal{L}_{\text{CVAE}} = \mathcal{L}_{\text{MSE}} + \beta \mathcal{D}_{\text{KL}} \tag{2}$$

The reconstruction loss, $\mathcal{L}_{MSE}$, was computed as the mean squared error (MSE) between the input point cloud x and $\hat{x}$: $\mathcal{L}_{MSE} = \|x - \hat{x}\|^2$. We selected MSE as it directly penalised large reconstruction errors and is suited for continuous numerical data such as point coordinates, and corresponds to assuming a Gaussian likelihood on the reconstruction error, which is consistent with the probabilistic framework of variational autoencoders. The KL divergence term, $\mathcal{D}_{KL}$, measures the divergence between the approximate posterior parametrised by μ and σ, and a multivariate standard normal prior:

$$\mathcal{D}_{\text{KL}} = \mathcal{D}_{\text{KL}}(\mathcal{N}(\mu(x), \sigma(x)), \mathcal{N}(0, 1)) = \frac{1}{2}\mu(x)^2 + \sigma^2(x) - \log \sigma^2(x) - 1 \tag{3}$$

This regularisation encourages the latent space to conform to a known and structured distribution. This promoted generalisability to avoid overfitting and enabled interpolation between latent representations.[47] The weighting parameter β controlled the trade-off between reconstruction accuracy and latent space regularisation. A higher β enforces stronger latent space regularisation, potentially at the cost of reconstruction quality, while a lower β prioritises reconstruction. We selected $\beta = 0.001$ through a grid search over $\{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$. For each value, we evaluated the trade-off between reconstruction quality (relative MAE and Chamfer distance on the validation set) and latent space regularity (smooth interpolation and phenotypic separation in the 2D PCA projection). Values below 10^{-3} yielded negligibly better reconstruction at the cost of a less structured latent space, while values above 10^{-3} over-regularised the space and degraded reconstruction fidelity. Finally, the model was trained on an NVIDIA RTX A6000 with an Adam optimiser, with an 80–20 split between training and validation, using a batch size of 256 and a learning rate of 0.001.

The quality of reconstruction of the cardiac point cloud time windows was quantitatively assessed post-training using two complementary metrics: the relative mean absolute error (MAE) and the Chamfer distance. For each time window, the absolute error between the input point cloud x and its reconstruction $\hat{x}$ was computed as:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \tag{4}$$

where *N* is the total number of points in the point cloud, $|\cdot|$ is the Euclidean distance, and x_i, $\hat{x}_i$ represent the 3D coordinates of the *i*th point in the input and reconstruction, respectively. To enable interpretable error comparisons across participants with potentially different cardiac sizes and anatomical ranges, the MAE was normalised by the absolute range of the data in each dimension specific to each individual participant. This was computed as:

$$\text{Relative MAE}_p = \frac{\mathrm{MAE}_p}{\mathrm{range}(x_p)} \tag{5}$$

where $\mathrm{range}(x_p)$ represents the range of the input point cloud for participant *p*, computed as the difference between the maximum and minimum values along each spatial dimension, aggregated across all time windows for that participant. This normalisation gave a relative error expressed as a percentage of the anatomical data range of each participant, preserving inter-individual variability while ensuring interpretability and direct comparison across the cohort of the reconstruction error.

Additionally, we computed the Chamfer distance to further characterise the geometric similarity between the input and reconstruction as this metric is widely used for measuring distances between two point sets. Denoting the input point cloud *P* and the reconstructed point cloud *R*, the Chamfer distance is defined as[48]:

$$d_{\mathrm{Chamfer}}(P,R) = \frac{1}{2N}\left(\sum_{x\in P}\min_{\hat{x}\in R}|x-\hat{x}|^2 + \sum_{\hat{x}\in R}\min_{x\in P}|\hat{x}-x|^2\right) \tag{6}$$

where $|\cdot|$ is the Euclidean distance, and N is the constant number of points in the point clouds. The metric computed the average squared nearest-neighbour distance between points in one cloud to the other, in both directions. The Chamfer distance captures both global and local discrepancies in the point cloud structure, independent of point ordering. This makes it particularly suitable for evaluating 3D anatomical reconstructions where point correspondence is explicitly enforced via atlas registration. Unlike purely point-wise error metrics, Chamfer distance evaluated the overall spatial distribution of the reconstructed geometries. By combining participant-relative MAE and Chamfer distance, we obtain a complementary quantitative assessment of the performance of our model: the relative MAE is sensitive to normalised point-wise errors within the anatomical context while the Chamfer distance captures the global shape consistency of the reconstructions.

Latent space generation

The resulting latent space was normalised and PCA was performed to remove any correlation between the latent vectors. PCA serves a complementary role to the CVAE: while the CVAE bottleneck dimension (16) controls the expressiveness of the generative model and cannot be reduced further without degrading reconstruction quality, PCA provides a post-hoc linear rotation that projects the latent vectors onto the orthogonal eigenvectors of the covariance matrix, producing uncorrelated principal component scores ordered by explained variance, enabling principled selection of the most informative components for downstream tasks. Formally, given the matrix of latent vectors $Z \in \mathbb{R}^{n\times d}$ (where *n* is the number of time windows and $d = 16$ is the latent dimension), PCA computes the eigendecomposition of the covariance matrix $\Sigma = \frac{1}{n}Z^\top Z$, giving eigenvectors $v_1, \ldots, v_d$ ordered by their eigenvalues $\lambda_1 \geq \ldots \geq \lambda_d$. The projected scores ZV are uncorrelated by construction, though the loadings within each eigenvector are not necessarily so. The use of PCA also facilitated further dimensionality reduction needed for visual representation and clustering of low-dimensional time series, as well as improved the interpretability of the latent features characterising the cardiac cycle. First, ten principal components were kept while preserving 99% of the initial variance. Consecutive latent representations were concatenated for each patient to form an abstract trajectory of embeddings, called temporal signatures (Fig. 3). The signatures can be visually inspected in the reduced space of the first and second principal components accounting for 82% of the variance. As the cardiac cycle frames are ordered from the beginning of systole to the end of diastole in our

dataset, signatures are similarly moving from systole to diastole. Only the first four principal components were kept for further analysis steps as they represented 92% of the variance, which was deemed an appropriate cut-off for representing the data while considering computational resources. The choice of principal components was task-dependent: 10 PCs were retained for quantitative correlation analysis to capture >99% of the cumulative explained variance, and 4 PCs were selected for clustering as a trade-off between information retention (92% variance) and computational cost of DTW-based time-series K-means. A Scree plot comparing the explained variance curves for the CVAE+PCA pipeline and the benchmark models is provided in Supplementary Figure 6.

Phenotype associations with temporal signatures

Temporal signatures are generated in an unsupervised way and are only abstract representations of the high-dimensional point clouds. Relevant participant features were tested for correlation with the latent principal components to explain the latent space variance. Data from demographics, body and clinical measurements, biomarkers and MRI-derived measures were accounted for. A subset of relevant features were also visually inspected with signature distribution stratified by z-score in the latent space of the first and second principal components: age, systolic blood pressure, maximum wall thickness, left ventricle ejection fraction, HDL cholesterol, HbA1c and PRS for HCM and DCM. For each feature, z-scores were computed as $z = \frac{x-\mu}{\sigma}$ where x is the feature value of the individual, μ is the mean, and σ is the standard deviation of that feature across the population. The trajectories were grouped into five evenly spaced z-score bins within the range $[-2, 2]$, and the average latent signature was computed within each bin. The average signature for selected risk factors and outcomes were also displayed in the latent space: hypertension, obesity, diabetes, stroke, heart failure, myocardial infarction, cardiac arrest, HCM and DCM.

Unsupervised clustering of signatures

The latent representations capture the dynamic evolution of cardiac point clouds over time. These temporal signatures contain information on both the in-plane position and contour geometry of the trajectories in latent space. Importantly, as the temporal signatures are sequences of low-dimensional latent coordinates, they can be assimilated to multivariate time series, enabling the use of special time series clustering techniques.

Clustering of the temporal signatures was performed using K-means applied to time series data, implemented with the `tslearn`[49] Python library. This algorithm is an extension of the classic K-means clustering method, adapted for temporal data by replacing the standard Euclidean distance with a time series-specific distance measure of dynamic time warping (DTW). DTW is particularly well-suited for time series data as it accounts for potential temporal misalignments between sequences by finding an optimal non-linear alignment[50,51].

We denote our set of n time series $s_1, s_2, \ldots, s_n$ for each participant, where each $s_i \in \mathbb{R}^{T\times d}$ represents a temporal signature of length $T = 13$ in $d = 4$-dimensional latent space (the first four principal components). The objective is to partition the time series into K clusters $C_1, C_2, \ldots, C_K$ to minimise the within-cluster sum of squared DTW distances:

$$\underset{\{C_k\}}{\operatorname{argmin}} \sum_{k=1}^{K} \sum_{s_i \in C_k} \mathrm{DTW}(s_i, \mu_k)^2 \tag{7}$$

where $\mu_k \in \mathbb{R}^{13\times 4}$ is the centroid of cluster C_k. DTW calculates the similarity between two time series by warping their time axis to achieve the best possible alignment. The metric is computed by finding the warping path $W = \{(t_1, t'_1), (t_2, t'_2), \ldots, (t_T, t'_T)\}$ which minimises the distance:

$$\mathrm{DTW}(s_i, \mu_j) = \min_{W} \sqrt{\sum_{l=1}^{L} |s_i(t_l) - \mu_j(t'_l)|^2} \tag{8}$$

where L represents the length of the time series sequence.

Signatures were clustered based on two properties: in-plane position and contour. First, temporal signatures were centred around the origin by subtracting their mean. This operation removed information about their absolute position in the latent space, isolating the relative contour of the trajectory. Clustering was then performed on these centred time series to group signatures with similar temporal evolution patterns, irrespective of their absolute position. After that, clustering was performed on the original, uncentred temporal signatures, which preserved information about the absolute in-plane position of the trajectories in latent space. This allowed grouping based on spatial positioning. The rationale behind this dual clustering approach lies in the hypothesis that both the contour geometry of the signature and its position in latent space convey clinically or physiologically distinct information. By combining the results of both methods, the final phenogroups integrate both geometrical and structural characteristics of the trajectories (Fig. 5).

Stability analysis was performed to assess the robustness and reproducibility of the clustering method. The initial dataset was partitioned into five equally-sized, non-overlapping sub-partitions. The clustering algorithm was independently performed on each of these five sub-partitions $\mathcal{T}^{(1)}, \mathcal{T}^{(2)}, \ldots, \mathcal{T}^{(5)}$. The cluster labels obtained from each sub-partition, noted $\mathcal{S}^{(m)}$, m ranging from 1 to 5, were then compared to the cluster labels obtained from clustering the full dataset, noted $\mathcal{S}^{\mathrm{ref}}$, treated as the reference clustering. The agreement between clustering solutions was evaluated with the adjusted Rand index (ARI) score[17,52]. The ARI score measures the similarity between two clustering solutions by considering all possible pairs of data points and evaluating whether each pair is consistently assigned to the same cluster or to different clusters in both the reference clustering and

the sub-partition clustering. For each sub-partition, we can denote a the number of pairs of temporal signatures that are assigned to the same cluster in both $\mathcal{S}^{(m)}$ and $\mathcal{S}^{ref}$ and b the number of pairs assigned to different clusters in both methods. The total number of possible pairs of temporal signatures in a sub-partition of size n is:

$$\binom{n}{2} = \frac{n(n-1)}{2} \tag{9}$$

The Rand index (RI) is then calculated as:

$$\mathrm{RI} = \frac{a+b}{\binom{n}{2}} \tag{10}$$

which gives the proportion of pairs on which the two clustering solutions agree. To adjust for random similarity agreement, the adjusted Rand index corrects the initial metric by subtracting the expected agreement between random clusterings and normalising by the maximum possible agreement:

$$\mathrm{ARI} = \frac{\mathrm{RI} - \text{Expected RI}}{\text{Max RI} - \text{Expected RI}} \tag{11}$$

where the expected RI is the expected value for random clusterings, and the max RI is the maximum possible agreement. ARI ranges from -1 for a complete disagreement to 1 for a perfect agreement, with 0 indicating agreement expected by random chance. For each sub-partition, the ARI was computed by comparing $\mathcal{S}^{(m)}$ and $\mathcal{S}^{ref}$. The final stability score for both clustering on in-plane position and contour characteristics was obtained by averaging the ARI scores across the five sub-partitions:

$$\text{Mean ARI} = \frac{1}{5}\sum_{k=1}^{5} \mathrm{ARI}^{(k)} \tag{12}$$

This analysis was repeated for different numbers of clusters K in the range of $2 \leq K \leq 7$. This also allowed us to find the number of optimal clusters K for the K-means clustering algorithm applied to time series to ensure the maximum stability of the dual clustering method as per the ARI scores.

To assess the sensitivity of the phenogroup assignments to the PCA decomposition itself, an additional stability analysis was performed. The dataset was randomly split into five stratified 80% subsets at the patient level (16,281 patients each). For each subset, the CVAE latent space was re-standardised, PCA was recomputed, and time-series K-means clustering with DTW was performed independently for both the in-plane position and centred-trajectory (contour) variants, using $K = 6$ to test stability of the final combined dual-clustering output (the original main analysis used $K = 3$ for position and $K = 2$ for contour separately, which combined to give six phenogroups). To prevent the arbitrary sign of eigenvectors from inflating distances on flipped axes, subset principal components were sign-aligned to the full-dataset components prior to clustering, and K-means was run with $n_{\text{init}} = 5$ to limit initialisation noise. Cluster labels from each subset were compared to those obtained from the full-dataset PCA on the overlapping patients using the adjusted Rand index. In-plane position clustering was highly stable under PCA recomputation, with a mean ARI of 0.91 ± 0.05 across the five subsets (range: 0.82–0.95), in close agreement with the original clustering-only stability analysis (ARI: 0.91). The contour clustering showed a mean ARI of 0.59 ± 0.23 with a bimodal distribution across subsets (0.93, 0.79, 0.43, 0.40, 0.37), indicating that the contour partition is more sensitive than the placement partition to recomputation of the higher-order principal components (PC3, PC4) on subsamples of the cohort. Per-subset values are reported in Supplementary Table 3.

Statistical testing for enrichment

Associations between phenogroups and a wide range of individual participant features were statistically tested. Features included demographics, lifestyle, risk factors, body and clinical measurements, MRI-derived measures, biomarkers, PRS, symptoms, outcomes and diagnosis. Longitudinal, radial and circumferential strains were considered as absolute values to simplify their interpretations. The full list of features can be found in the Supplementary Data. A Kruskal-Wallis test was used on continuous features to test for significant differences between the phenogroups. This was followed by a post-hoc Dunn test to find pairs of phenogroups with significant differences in feature distribution. A χ^2 test was used on binary and categorical features, followed by a post-hoc Fisher's test to find significant associations of phenogroups with specific values of a feature. The effect size of each significant association was calculated as the odds ratio derived from the contingency table of the χ^2 test. The odds ratio measures the association of the clinical traits (outcomes and risk factors) to the exposure of one phenogroup against the rest of the population. All resulting P values were adjusted for multiple comparisons under the Benjamini-Hochberg procedure, first for Kruskal-Wallis and χ^2 tests together, and then for post-hoc Dunn and Fisher tests together. Significant associations were reported for P value < 0.05. The Python packages used for the Kruskal-Wallis test, the post-hoc Dunn test and the Benjamini-Hochberg correction were respectively `scipy`, `scikit-posthocs`, and `statsmodels`. Additionally, a mass univariate group analysis was performed to compare geometry differences of temporal signatures between phenogroups. More details can be found in Supplementary Figure 7.

Modelling motion in phenogroups

The velocity gradient of the left ventricle myocardium surface between consecutive point cloud frames was computed for the average cardiac motion in phenogroup. The average velocity of each phenogroup was normalised by dividing it by the bounding box size of the average left ventricle for that group. Velocity differences were computed between each phenogroup and the population average cardiac motion. These variations were visually inspected across the cardiac cycle to identify unique motion patterns. The velocity fields were projected on the surface of the average motion by phenogroup to visualise outward and inward movements of the myocardium surface. Fields of beams were added onto the surface where the velocity of the myocardium was slower, for blue beams, or faster for red beams, than the population average velocity (Figure 9a). These four-dimensional models of motion differences helped identify distinctive motion patterns in each phenogroup.

Data availability

The analysis code repository is available on GitHub at: https://github.com/ImperialCollegeLondon/cardiac_motion_signatures/. All data is available from UK Biobank.

Received: 22 December 2025; Accepted: 29 May 2026
Published online: 11 June 2026

References

1. Bello, G. A. et al. Deep learning cardiac motion analysis for human survival prediction. *Nat. Mach. Intell.* **1**, 95–104. https://doi.org/10.1038/s42256-019-0019-2 (2019).
2. Cikes, M. & Solomon, S. D. Beyond ejection fraction: An integrative approach for assessment of cardiac structure and function in heart failure. *Eur. Heart J.* **37**, 1642–50. https://doi.org/10.1093/eurheartj/ehv510 (2016).
3. McDonagh, T. A. et al. ESC Scientific Document Group. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur. Heart J.* **2021**(42), 3599–726. https://doi.org/10.1093/eurheartj/ehab368 (2021).
4. Marwick, T. H. Ejection fraction pros and cons: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **72**, 2360–79. https://doi.org/10.1016/j.jacc.2018.08.2162 (2018).
5. Meng, Q. et al. MulViMotion: Shape-aware 3D myocardial motion tracking from multi-view cardiac MRI. *IEEE Trans. Med. Imaging* **41**, 1961–74. https://doi.org/10.1109/TMI.2022.3154599 (2022).
6. Meng, Q., Bai, W., O'Regan, D. P. & Rueckert, D. DeepMesh: Mesh-based cardiac motion tracking using deep learning. *IEEE Trans. Med. Imaging* **43**, 1489–500. https://doi.org/10.1109/TMI.2023.3340118 (2024).
7. Biffi, C. et al. Learning interpretable anatomical features through deep generative models: Application to cardiac remodeling. *Med. Image Comput. Comput. Assist. Intervent. - MICCAI* **2018**(11071), 464–71. https://doi.org/10.1007/978-3-030-00934-2_52 (2018).
8. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–9. https://doi.org/10.1038/s41586-018-0579-z (2018).
9. Bai, W. et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J. Cardiovasc. Magn. Reson.* **20**, 65. https://doi.org/10.1186/s12968-018-0471-x (2018).
10. Bai, W. et al. A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion. *Med. Image Anal.* **26**, 133–45. https://doi.org/10.1016/j.media.2015.08.009 (2015).
11. Duan, J. et al. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi- task deep learning approach. *IEEE Trans. Med. Imaging* **38**, 2151–64. https://doi.org/10.1109/TMI.2019.2894322 (2019).
12. Govil, S. et al. A deep learning approach for fully automated cardiac shape modeling in tetralogy of Fallot. *J. Cardiovasc. Magn. Reson.* **25**, 15. https://doi.org/10.1186/s12968-023-00924-1 (2023).
13. Ramirez, P. Z. et al. Deep learning on object-centric 3D neural fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **46**, 9940–56. https://doi.org/10.1109/Tpami.2024.3430101 (2024).
14. De Luigi, L. et al. Deep learning on implicit neural representations of shapes. In *11th International Conference on Learning Representations, ICLR 2023* (2023).
15. McGurk, K. A. et al. Genetic and phenotypic architecture of human myocardial trabeculation. *Nat. Cardiovasc. Res.* **3**, 1503–15. https://doi.org/10.1038/s44161-024-00564-3 (2024).
16. Tadros, R. et al. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nat. Genet.* **53**, 128–34. https://doi.org/10.1038/s41588-020-00762-2 (2021).
17. Rand, W. M. Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **66**, 846–50. https://doi.org/10.1080/01621459.1971.10482356 (1971).
18. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform manifold approximation and projection. *J. Open Source Softw.* **3**. https://doi.org/10.21105/joss.00861 (2018).
19. Heidenreich, P. A. 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* **145**, e895–e1032. https://doi.org/10.1161/cir.0000000000001063 (2022).
20. Savarese, G., Stolfo, D., Sinagra, G. & Lund, L. H. Heart failure with mid-range or mildly reduced ejection fraction. *Nat. Rev. Cardiol.* **19**, 100–16. https://doi.org/10.1038/s41569-021-00605-5 (2022).
21. Redfield, M. M. & Borlaug, B. A. Heart failure with preserved ejection fraction: A review. *JAMA* **329**, 827–38. https://doi.org/10.1001/jama.2023.2020 (2023).
22. Chadalavada, S. et al. Myocardial strain measured by cardiac magnetic resonance predicts cardiovascular morbidity and death. *J. Am. Coll. Cardiol.* **84**, 648–59. https://doi.org/10.1016/j.jacc.2024.05.050 (2024).
23. Zerhouni, E. A., Parish, D. M., Rogers, W. J., Yang, A. & Shapiro, E. P. Human heart: Tagging with MR imaging-a method for noninvasive assessment of myocardial motion. *Radiology* **169**, 59–63. https://doi.org/10.1148/radiology.169.1.3420283 (1988).
24. Smiseth, O. A. et al. Myocardial strain imaging: Theory, current practice, and the future. *JACC Cardiovasc. Imaging* **18**, 340–81. https://doi.org/10.1016/j.jcmg.2024.07.011 (2025).
25. Bonazzola, R. et al. Unsupervised ensemble-based phenotyping enhances discoverability of genes related to left-ventricular morphology. *Nat. Mach. Intell.* **6**, 291–306. https://doi.org/10.1038/s42256-024-00801-1 (2024).
26. Burns, R. et al. Genetic basis of right and left ventricular heart shape. *Nat. Commun.* **15**, 9437. https://doi.org/10.1038/s41467-024-53594-7 (2024).
27. Tayal, U. et al. Precision phenotyping of dilated cardiomyopathy using multidimensional data. *J. Am. Coll. Cardiol.* **79**, 2219–32. https://doi.org/10.1016/j.jacc.2022.03.375 (2022).
28. Ishikita, A. et al. Machine learning for prediction of adverse cardiovascular events in adults with repaired tetralogy of Fallot using clinical and cardiovascular magnetic resonance imaging variables. *Circ. Cardiovasc. Imaging* **16**, e015205. https://doi.org/10.1161/CIRCIMAGING.122.015205 (2023).

29. Curran, L. et al. Genotype-phenotype taxonomy of hypertrophic cardiomyopathy. *Circ. Genom. Precis. Med.* **16**, e004200. https://doi.org/10.1161/CIRCGEN.123.004200 (2023).
30. Thanaj, M. et al. Genetic and environmental determinants of diastolic heart function. *Nat. Cardiovasc. Res.* **1**, 361–71. https://doi.org/10.1038/s44161-022-00048-2 (2022).
31. Arbelo, E. et al. ESC Scientific Document Group. ESC Guidelines for the management of cardiomyopathies. *Eur. Heart J.* **2023**(44), 3503–626. https://doi.org/10.1093/eurheartj/ehad194 (2023).
32. Bergan, N. et al. Systematic review, meta-analysis, and population study to determine the biologic sex ratio in dilated cardiomyopathy. *Circulation* **151**, 442–59. https://doi.org/10.1161/circulationaha.124.070872 (2025).
33. Tan, Y. et al. Mechanisms of diabetic cardiomyopathy and potential therapeutic strategies: Preclinical and clinical evidence. *Nat. Rev. Cardiol.* **17**, 585–607. https://doi.org/10.1038/s41569-020-0339-2 (2020).
34. Littlejohns, T. J. et al. The UK Biobank imaging enhancement of 100,000 participants: Rationale, data collection, management and future directions. *Nat. Commun.* **11**, 2624. https://doi.org/10.1038/s41467-020-15948-9 (2020).
35. Zheng, S. L. et al. Evaluation of polygenic scores for hypertrophic cardiomyopathy in the general population and across clinical settings. *Nat. Genet.* **57**, 563–71. https://doi.org/10.1038/s41588-025-02094-5 (2025).
36. Tadros, R. et al. Large-scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. *Nat. Genet.* **57**, 530–8. https://doi.org/10.1038/s41588-025-02087-4 (2025).
37. Henry A. et al. Mapping the aetiological foundations of the heart failure spectrum using human genetics. medRxiv. 2023:2023.10.01.23296379. https://doi.org/10.1101/2023.10.01.23296379 (2023).
38. Thompson, D.J. et al. A systematic evaluation of the performance and properties of the UK Biobank Polygenic Risk Score (PRS) release. *PLoS One***19**, e0307270. https://doi.org/10.1371/journal.pone.0307270 (2024).
39. Petersen, S. E. et al. UK Biobank's cardiovascular magnetic resonance protocol. *J. Cardiovasc. Magn. Reson.* **18**, 8. https://doi.org/10.1186/s12968-016-0227-4 (2016).
40. Petersen, S. E. et al. Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. *J. Cardiovasc. Magn. Reson.* **19**, 18. https://doi.org/10.1186/s12968-017-0327-9 (2017).
41. Russo, C. et al. Comparison of echocardiographic single-plane versus biplane method in the assessment of left atrial volume and validation by real time three-dimensional echocardiography. *J. Am. Soc. Echocardiogr.* **23**, 954–60. https://doi.org/10.1016/j.echo.2010.06.010 (2010).
42. Bai, W. et al. Recurrent neural networks for aortic image sequence segmentation with sparse annotations. *Med. Image Comput. Comput. Assist. Intervent. MICCAI* **2018**, 586–94. https://doi.org/10.1007/978-3-030-00937-3_67 (2018).
43. Bai, W. et al. A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat. Med.* **26**, 1654–62. https://doi.org/10.1038/s41591-020-1009-y (2020).
44. Kingma, D.P. & Welling, M. Auto-Encoding Variational Bayes. arXiv:1312.6114 (2022).
45. Higgins, I., Matthey, L., Pal, A., Burgess, C.P., Glorot, X., Botvinick, M.M., Mohamed, S. & Lerchner, A. beta-VAE: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*. https://api.semanticscholar.org/CorpusID:46798026 (2016).
46. Beetz, M., Banerjee, A. & Grau, V. Multi-domain variational autoencoders for combined modeling of MRI-based biventricular anatomy and ECG-based cardiac electrophysiology. *Front. Physiol.* **13**, 886723. https://doi.org/10.3389/fphys.2022.886723 (2022).
47. Asperti, A. & Trentin, M. Balancing reconstruction error and Kullback-Leibler divergence in variational autoencoders. *IEEE Access* **8**, 199440–8. https://doi.org/10.1109/Access.2020.3034828 (2020).
48. Barrow, H.G., Tenenbaum, J.M., Bolles, R.C. & Wolf, H.C. Parametric correspondence and chamfer matching: Two new techniques for image matching. In *Proceedings: Image Understanding Workshop*. 21–27 (Science Applications, Inc., 1977).
49. Tavenard, R. et al. Tslearn, a machine learning toolkit for time series data. *J. Mach. Learn. Res.* **21**, 1–6 (2020).
50. Serantoni, C. et al. Integrating dynamic time warping and K-means clustering for enhanced cardiovascular fitness assessment. *Biomed. Signal Process. Control* **97**. https://doi.org/10.1016/j.bspc.2024.106677 (2024).
51. Vera, J. F. & Angulo, J. M. An MDS-based unifying approach to time series K-means clustering: Application in the dynamic time warping framework. *Stoch. Environ. Res. Risk Assess.* **37**, 4555–66. https://doi.org/10.1007/s00477-023-02470-9 (2023).
52. Duo, A., Robinson, M. D. & Soneson, C. A systematic performance evaluation of clustering methods for single-cell RNA-seq data. *F1000Res.* **7**, 1141. https://doi.org/10.12688/f1000research.15666.3 (2018).

Author contributions

Conceptualisation: P-R.S. and D.P.O.; Methodology: P-R.S., S.K., J.Z.; Software: P-R.S.; Validation: P-R.S.; Formal analysis: P-R.S.; Investigation: P-R.S.; Resources: D.P.O., P.I., K.A.M., S.L.Z., W.B.,; data curation: P-R.S., K.A.M., S.L.Z.; Writing - original draft: P-R.S. and D.P.O.; Writing - review & editing: all authors; Visualisation: P-R.S., D.P.O.; Supervision: J.S.W. and D.P.O.; Project administration: D.P.O.; Funding acquisition: D.P.O.

Funding

The study was supported by the Medical Research Council (MC_UP_1605/13); the Engineering & Physical Sciences Research Council (EP/W01842X/1); the British Heart Foundation (BBC/F/21/220106, RG/F/24/110138, RE/24/130023, CH/F/24/90015, FS/IPBSRF/22/27059, NH/F/23/70013); and the National Institute for Health Research (NIHR) Imperial College Biomedical Research Centre. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any author accepted manuscript version arising.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-56151-y.

Correspondence and requests for materials should be addressed to D.P.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026