Beyond the root: geometric characterization for the diagnosis of syndromic heritable thoracic aortic diseases

Pau Romero[a], Miguel Lozano[a], Lydia Dux-Santoy[b], Andrea Guala[b,c], Gisela Teixido-Tura[c,d], Rafael Sebastián[a] and Ignacio García-Fernández[a,1]

[a]CoMMLab – Computational Multiscale Simulation Lab. University of Valencia. Spain.
[b]Vall d'Hebron Institut de Recerca (VHIR), Barcelona, Spain
[c]CIBER de Enfermedades Cardiovasculares, CIBER-CV, Instituto de Salud Carlos III, Madrid, Spain
[d]Department of Cardiology, Hospital Universitari Vall d'Hebron, Barcelona, Spain

Abstract

Syndromic heritable thoracic aortic diseases (sHTAD), such as Marfan (MFS) or Loeys-Dietz (LDS) syndromes, involve high risk of life threatening aortic events. Diagnosis of syndromic features alone is difficult, and negative genetic tests do not necessarily exclude a genetic or hereditary condition. Periodic 3D imaging of the aorta is recommended in patients with aortic disease. Thus, an imaging-based approach aimed at identifying unique features of aortic geometry can be highly effective for diagnosing sHTAD and assessing risk. In this study, we present a method that can help identify the manifestations of sHTAD by focusing on the entire geometry of the thoracic aorta, rather than only using measurements of dilation of the aortic root. We analyze the geometric phenotype of 97 patients with genetically confirmed sHTAD (79 MF and 18 LDS) and of 45 healthy volunteers, using 3D aorta meshes obtained from phase contrast-enhanced magnetic resonance angiograms computed from 4D flow cardiac magnetic resonance. We build a geometric encoding of the aorta, based on a vessel coordinate system, and use several mathematical models to discriminate between controls and patients with sHTAD: a baseline scenario, based on aortic root dimensions only, a descriptor typically used in sHTAD patients; a low dimensional scenario, with a reduce encoding using principal component analysis; and a high-dimensional scenario, which included the full coefficient representation for geometry encoding, aiming to capture finer geometric details. The results indicate that considering the anatomy of the whole thoracic aorta can improve predictive ability. We achieve precision and sensitivity values over 0.8, with a specificity of over 70% in all the models used, while a single value classifiers (based only on aortic root diameter) demonstrated a trade-off between sensitivity and specificity. Using the mathematical properties of the vessel coordinate system representation, feature importance is mapped onto a set of anatomical traits that are used by the models to do the classification, thus providing interpretability of the results. This analysis indicates that in addition to the diameter of the aortic root, aortic elongation and a narrowing of the descending thoracic aorta may be markers of positive sHTAD.

Keywords: Aorta; Syndromic heritable thoracic aortic diseases; Genetic aortopathies; Manifestation of disease; Aortic dilatation; Imaging based diagnosis.

[1]Corresponding Author <ignacio.garcia@uv.es>.
Tel. +34 96 354 3064. Fax. +34 96 354 47 68.
Journal article published in Computers in Biology and Medicine: https://doi.org/10.1016/j.compbiomed.2024.109176

1 Introduction

Syndromic heritable thoracic aortic diseases (sHTADs) are a subset of thoracic aortic aneurysms present in patients with a genetic or heritable condition involving multisystemic syndromic manifestations (Loeys et al., 2010). Among sHTAD, Marfan (MFS) and Loeys-Dietz (LDS) syndromes are connective tissue disorders associated with different gene mutations (FBN1 or certain TGFB pathway-related genes, respectively) (Isselbacher et al., 2022; Schepers et al., 2018). Patients affected by sHTAD are at high risk of life-threatening adverse aortic events, such as aortic dissection and rupture. Unfortunately, the capacity to predict these major adverse events is very limited, underscoring the need for further research.

When syndromic features are identified, genetic testing is recommended to potentially identify pathogenic variants that confirm the specific aortic syndrome and assess related risks. However, genetic testing presents a number of difficulties, including the complexity in the interpretation of results. When performed, a negative genetic panel in patients presenting syndromic features does not necessarily exclude a genetic or hereditary condition. A positive genetic testing outcome should trigger cascade testing for at-risk family members; while in cases where the genetic panel yields negative results, or detects variants of unknown significance, a recommended approach involves implementing imaging-based screening for aortic dilation or aneurysms in at-risk relatives (Isselbacher et al., 2022).

Periodic assessment of aortic size using imaging techniques is also recommended for clinical follow-up of patients with sHTAD. When the maximum aortic root or ascending aorta is greater than a certain threshold (50mm for MFS, 45mm for LDS), elective substitution of the aorta is recommended to prevent the risk of aortic dissection or rupture (Isselbacher et al., 2022). Nonetheless, this surgical intervention has some inherent risk, and thus the timing of surgery should be tailored according to patient-specific risk factors (genetic variant, aor-

tic growth rate, familial history,...). Given that imaging is routinely obtained for the diagnosis and clinical follow-up of sHTAD, an imaging-based approach to identify high-risk patients with sHTAD may be of significant interest.

This study aims to identify the manifestations of sHTAD on thoracic aorta geometry beyond aortic root dilation and to test whether a mathematical representation of thoracic aorta geometry (i.e. a shape representation) could be useful in the diagnosis of sHTAD. Additionally, we offer a geometric explanation of the decision-making process employed by the mathematical models used.

2 Material and methods

2.1 Material

2.1.1 Cohort

Retrospective imaging data of prospectively enrolled patients obtained in the context of previous research was used. A total of 97 patients with a genetically-confirmed sHTAD (79 MFS and 18 LDS) and 45 healthy volunteers were included. Inclusion criteria were age ≥ 18 years, no significant aortic valvular disease by echocardiography (aortic regurgitation $\leq II/III$, maximum aortic velocity ≤ 3 m/s) or other cardiovascular disorders, tricuspid aortic valve, and no history of surgical repair or aortic valve replacements. The study was approved by the local ethics committee, and informed consent was obtained from all participants. Table 1 reports demographic and clinical characteristics of the patients.

Table 1: Demographical description of the cohort used for the study. BSA : Body Surface Area, AoRD : Aortic Root Diameter, Z-score is described in Section 2.2.2.

	Healthy	sHTAD
Age (years)	38.5 ± 12.3	35.7 ± 12.7
Sex (F|M)	34|66	56|44
Weight (kg)	71.6 ± 11.0	73.1 ± 15.7
Height (cm)	171.7 ± 7.4	178.8 ± 11.4
BSA (m^2)	1.83 ± 0.15	1.90 ± 0.23
AoRD (mm)	30.4 ± 3.6	37.2 ± 6.2
AoRD Z-score	-0.3 ± 1.1	2.0 ± 1.7

2.1.2 Magnetic resonance imaging

CMR studies were performed without intravenous contrast administration on a GE 1.5-T Signa scanner (GE Healthcare, Waukesha, Wisconsin, USA). The research protocol included 2D cine CMR images and a 4-dimensional (4D) phase-contrast CMR (4D flow CMR) acquisition with retrospective electrocardiographic gating during free breathing. A radially undersampled acquisition (PC VIPR) with 5-point balanced velocity encoding (Johnson et al., 2008) was used for 4D flow imaging of the entire thoracic aorta in approximately 10 min of scan time. Data were acquired using the following parameters: velocity encoding 200 cm/s, field of view 400 x 400 x 400 mm, acquisition matrix 160 x 160 x 160, voxel size 2.5 x 2.5 x 2.5 mm, flip angle 8, repetition time 4.2 to 6.4 ms, and echo time 1.9 to 3.7 ms. Reconstructions were performed offline with corrections for background phase from concomitant gradients, eddy currents, and trajectory errors of the 3-dimensional radial acquired k-space (Johnson et al., 2008).

2.2 Methodology

In the following, the methods used to describe the geometry of the aorta and to discriminate between healthy volunteers and sHTAD patients are described.

2.2.1 Aorta anatomy representation

Phase-contrast enhanced magnetic resonance angiograms were built (Johnson et al., 2008), and used to segment the aorta. Cine CMR images of the aortic root were used to measure the diameters of the aortic roots. From the aorta segmentation, 3D triangle meshes were obtained using the marching cubes algorithm (Cline and Lorensen, 1988). Due to the low resolution of the imaging modality, we applied Laplacian smoothing to the mesh, using the Python library PyVista (Sullivan and Kaszynski, 2019). All the geometries were transformed using the same parameter configuration that imposed minimal smoothing, ensuring that only high-resolution artifacts were removed while preserving the overall shape and volume. Moreover, the impact of this smoothing on the subsequent encoding is minimal, due to the least squares optimization involved in the encoding process. Then, the section spanning from the the sinotubular junction to the diaphragm (an example input mesh is shown in Figure 1, left) was extracted. The obtained triangle meshes were encoded to make the representation of all aortas consistent and adequate for training ML models. The encoding used was a patient-specific vascular model based on the definition of a vessel coordinate system (VCS) (Romero et al., 2023), as described below.

Using the VCS, an arbitrary point of the wall can be assigned an univocal set of coordinates, corresponding to its longitudinal location along the vessel, τ, and its polar coordinates on the corresponding cross section of the aorta, (θ, ρ). The origin for polar coordinates is taken at the centerline of the vessel, in the corresponding cross section. In this reference system, the wall is defined by the surface, $\rho = \rho_w(\tau, \theta)$, Then, the geometry of the thoracic aorta can be entirely represented by a curve, $\mathbf{c}(\tau)$, that approximates the vessel centerline, and the radius function $\rho_w(\theta, \tau)$, that expresses the distance from the wall to the curve $\mathbf{c}(\tau)$. Figure 1 shows an scheme of this representation for one of the aortas of the cohort. Romero et al. (2023) build their model approximating both $\mathbf{c}$ and ρ_w with uniform cubic B-splines, and define the feature vector as the vector of coefficients on a certain basis. In this way, the representation for every patient has the same dimension. Moreover, the coefficients have a geometrical interpretation that relates the shape in a particular region of the vessel with a section of the feature vector. This last property will enable the interpretability of the resulting ML models.

Using this approach, the data set was built through the following steps: we computed the VCS for every aorta; using the centerlines, we applied generalised Procrustes alignment (Heimann and Meinzer, 2009), and we approximated the wall and the centerline using uniform B-splines. To capture enough detail we used 45 coefficients for the centerline approximation and 609 coefficients for the bivariate spline wall

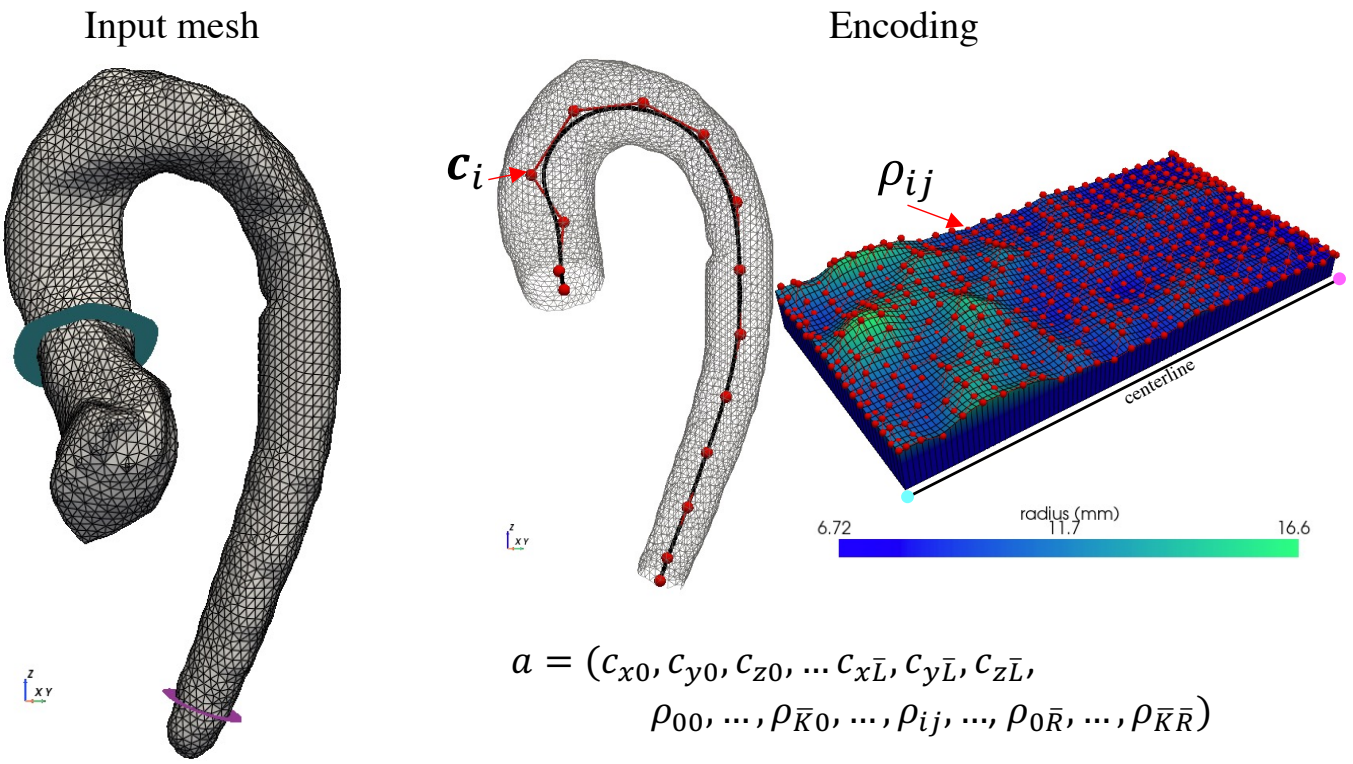

Figure 1: Computational encoding of an aorta as described by Romero et al. (2023). On the left, the input mesh, with the cross-section at the sinotubular junction (cyan) and at the diaphragm (magenta). On the right, the centerline and radius mapping with the control polygons of its B-spline based approximation. The control points (red) form the encoding vector a.

approximation, leading to feature vectors of length 654. To reduce the dimensionality of the representation, we applied a principal component analysis (PCA) to the geometrical encoding, obtaining a representation with less than 30 modes of deformation that reflect more than 98% of cumulative variance. The feature vector of every patient was labeled as H for healthy volunteers and G for patients with positive genetic test.

2.2.2 *Patient stratification methodology*

We will consider three scenarios, with different levels of complexity in geometrical encoding. The first scenario, termed *baseline*, aims to emulate clinical practice: In the clinical context, patient stratification typically involves Z-score standardization of the aortic root diameter, adjusted for patient age, sex, and body size (Campens et al., 2014). In our baseline scenario, geometry encoding will only consider the diameter of the aortic root. In the second scenario, termed *low-dimensional* scenario, we will use the coordinates in the PCA space for geometry encoding. In the third scenario, termed *high-dimensional* scenario, we will adopt the full representation of the B-spline coefficient for geometry encoding, aiming to capture finer geometric details. For each scenario, we will assess the performance of several predictive models when classifying the aortas among patients with sHTAD and healthy volunteers, under the same training workflow. Since the diameter of the aortic root is a very relevant biomarker in clinical practice, by adding this measurement to the two geometric feature vectors, we can give all the available information to the models. However, we have found that the results are very similar regardless of the presence of the diameter of the aortic root, and thus they are not presented here. The performance for this scenario is presented in Section 1 of the supplementary material.

For the assessment of each scenario we will use three linear models, logistic regression (LR), linear discriminant analysis (LDA) and linear-kernel support vector machine (SVM); and a non-linear model, the random forest (RF). The predictive performance of the models will be analysed using their precision, recall (sensitivity), area under the receiver operating characteristic curve (AUC), and specificity. In addition, we will compute the f1 score, which uses a combination of the precision and the recall. The results will be presented for 5-fold stratified cross-validation, averaging the results of the 5 folds. In the case of the SVM and the RF, nested cross validation (Cawley and Talbot, 2010), with 3 folds on the training set, will be performed for hyperparameter adjustment, taking the set of parameters with best f1 score. Due to the class imbalance, all the models will be trained using weights inversely proportional to class frequencies. To prevent bias in the low-dimensional scenario, at each fold the PCA will be computed exclusively with the training set in its high-dimensional representation. Then, the PCA transformation will be applied to the train and test sets.

2.2.3 *Interpretability of the results*

A common drawback alleged against complex ML models is that they can lead to poor understanding on what are the reasons for their decisions once trained (Rudin, 2019). We will benefit from the properties of our geometry encoding to provide an interpretation of the results, that will inform about the parts of the anatomical characteristics that have been more relevant for patient stratification.

In the anatomy encoding of the VCS, the aorta centerline and

Table 2: Performance metrics of the different models for each of the described scenarios. The first row corresponds to stratification based on Z-score of the root diameter (Loeys et al., 2010) applied to our cohort. The following rows contain the measurement of the metrics (per column) of each model for the three scenarios considered. Each cell contains the mean and the variance for a stratified 5-fold cross validation.

Scenario	Model	Precision	Recall (sens)	f1-score	AUC	Specificity
Z-score>2	-	1.00	0.47	0.64	-	1.00
Baseline	LR	0.78 ± 0.05	0.89 ± 0.09	0.83 ± 0.05	0.81 ± 0.11	0.39 ± 0.15
	LDA	0.78 ± 0.05	0.88 ± 0.08	0.83 ± 0.06	0.81 ± 0.11	0.38 ± 0.17
	SVM	0.90 ± 0.07	0.67 ± 0.08	0.76 ± 0.06	0.81 ± 0.11	0.80 ± 0.17
	RF	0.83 ± 0.08	0.65 ± 0.08	0.72 ± 0.05	0.72 ± 0.09	0.64 ± 0.20
Low-dimensional	LR	0.90 ± 0.05	0.82 ± 0.08	0.85 ± 0.03	0.87 ± 0.03	0.80 ± 0.14
	LDA	0.88 ± 0.04	0.88 ± 0.08	0.88 ± 0.05	0.89 ± 0.04	0.73 ± 0.10
	SVM	0.89 ± 0.05	0.80 ± 0.07	0.84 ± 0.04	0.87 ± 0.04	0.78 ± 0.14
	RF	0.84 ± 0.05	0.87 ± 0.10	0.85 ± 0.07	0.83 ± 0.08	0.64 ± 0.09
High-dimensional	LR	0.87 ± 0.07	0.87 ± 0.07	0.87 ± 0.04	0.89 ± 0.044	0.71 ± 0.19
	LDA	0.84 ± 0.06	0.80 ± 0.09	0.82 ± 0.07	0.83 ± 0.104	0.67 ± 0.14
	SVM	0.89 ± 0.08	0.79 ± 0.13	0.83 ± 0.07	0.89 ± 0.084	0.78 ± 0.16
	RF	0.85 ± 0.04	0.89 ± 0.11	0.87 ± 0.06	0.89 ± 0.054	0.67 ± 0.11

wall are represented by spline coefficients that are directly related to their geometry. For example, an increase in the value of a particular wall coefficient is directly related to the diameter and shape of the aorta in a very localised region of the wall. To understand what anatomical traits are more determinant for the stratification of patients, we will compute feature importance for the different models and identify the regions that correspond to the most relevant features. Moreover, B-Splines form a vector space, which allows interpreting differences in the feature vectors as shape deformations (Romero et al., 2023). This deformation indicates the geometric traits that are used by the models to distinguish sHTAD patients from healthy individuals, and will allow to generate aorta shapes that are representative of each class. Furthermore, thanks to the VCS, we will be able to split the importance in the contribution of the wall deformation (variation in radius) and the centerline warping. For this analysis, we will use the whole cohort to train each model, thus achieving the most informed model with the available data.

We compute feature importance as follows. Fitting of linear models, LR, SVM, and LDA, results in a separating hyperplane in the space of features. The direction orthogonal to this plane defines a weight vector that indicates what are the features that best separate the two classes. Note that the separating direction in the high-dimensional scenario does not account for correlations on the features defining the physiological limits. Thus, we will interpret the importance of a feature as an indication of the local relevance of the associated region of the vessel to classify patients. We will visualise the mapping of importance on the different regions of the vessel wall with a colour scale to show the parts of the anatomy that have been more relevant for classification. We will show this mapping represented on the mean aorta shape, computed as the average of all the feature vectors in the dataset.

For the RF algorithm, feature importance is typically computed by averaging the cumulative impurity decrease of the decision trees in the forest. Unfortunately, interpretation of RF features in a way similar to that for the linear models does not have physiological meaning and can be misleading. Thus, we will not include the results here. Further discussion on this point can be found in Section 2 of the supplementary material.

3 Results

The performance metrics of the different models trained in this study are presented in Table 2. To provide a reference for a criterion that is used in clinical practice, the table also presents the performance of the aortic root Z-score on the cohort. For each model, the results presented are the average $\pm$ standard deviation over the 5 folds of the 5-fold cross validation methodology. Furthermore, Figure 2 provides further information on the baseline scenario, as well as on the outcomes of the Z-score threshold when applied to the available dataset.

We see that the baseline scenario (classification based solely on the diameter of the aortic root) shows the poorest performance for all the models. Comparing with the Z-score>2 classification, this criterion yields in favour of precision and specificity with a very poor recall for our cohort. In contrast, the trained models, that optimise for f1-score, present more balanced results for the different metrics. An explanation of the poor results in this scenario can be seen in Figure 2; many positive genetic patients share the value of aortic root diameter with healthy volunteers, making this biomarker insufficient for discrimination. Note that, in this scenario, the linear models simply define a threshold trading off between precision versus recall. As a consequence, they share the same ROC curve and, hence, the same AUC values in Table 2.

When accounting for the whole anatomy, all the models reach an average precision above 84% and a recall above 80%. The sensitivity worsens in all scenarios, with the linear models in the low-dimensional scenario the most robust ones, reaching an 80% for LR. The main performance difference between the results in low- and high-dimensional scenarios seems to be the standard deviation. Higher values of standard deviation are usually related to a high dependence of the results on the fold

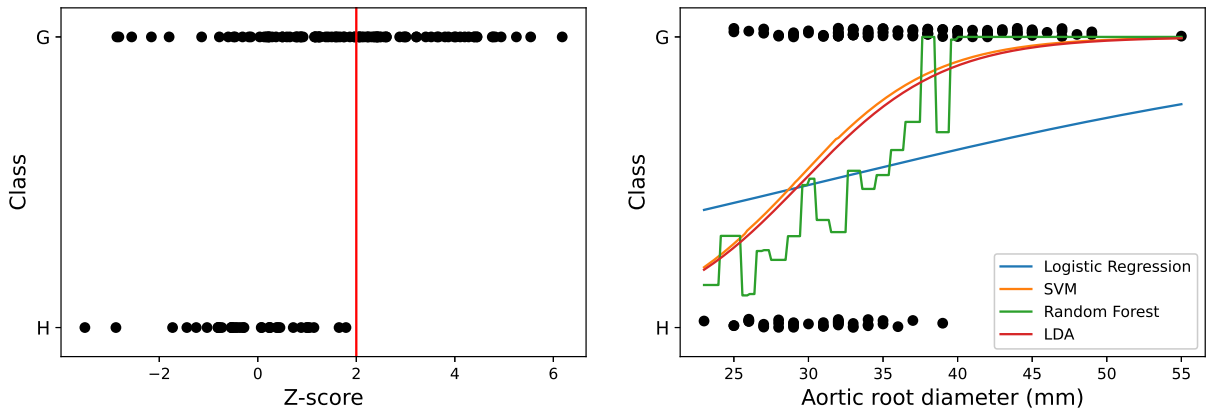

Figure 2: Distribution of Z-score value (left) and aortic root diameter (right), in mm, for sHTAD patients (G, for genetic) and healthy volunteers (H), together with information of the classifiers. A red vertical line at Z-score$= 2$ has been added to the left plot, while on the right plot the probability functions of each model are displayed. The models used were trained on the whole dataset with the best parameter configuration found during the 5-fold testing.

partition. In this sense, low-dimensional scenarios seem to be more robust than high-dimensional ones, particularly regarding the sensitivity.

Explainability analysis of the models is presented in Figures 3 and 4. They provide insight about the decision-making process in low- and high-dimensional scenarios, respectively, as described in Section 2.2.3. Note that the interpretation of importance needs a different approach between the low-dimensional scenario, which is based on PCA dimension reduction, and the high-dimensional scenario, which utilises the complete VCS encoding of the anatomy.

For the low-dimensional scenario, Figure 3 shows the deformation of the mean geometry using two standard deviations in the negative (healthy) direction on the left and the positive (sHATD) direction on the right. In both directions the deformation has been split into radius deformation (mapped on the surface colour) and centerline warping. The phantom of the mean geometry has been added to provide a reference.

In the high-dimensional scenario, the normal of the separating hyperplane cannot be used as a deformation mode, since this vector does not account for the physiological correlations in the coefficient representation. Nonetheless, this normal direction encodes the importance of each coefficient in the weight of each of the vector components. Hence, we used these weights to map the importance onto the centerline and the wall. Figure 4 shows the relevant regions for each of the linear models in the high-dimensional scenario. There, the mean geometry has been used to map the importance split in radius and centerline. Note that while radius importance can be computed with sign, allowing us to distinguish between positive and negative predictive values, centerline importance is only mapped in positive values, and it must be interpreted as predictive power.

4 Discussion

In this study, we explore the feasibility of employing three-dimensional aortic geometry as a biomarker to discern genetically-triggered sHTAD from healthy volunteers, beyond the prevailing approach of relying on the Z-score at different aortic levels (aortic root, mid-ascending aorta...) (Loeys et al., 2010) in patients with sHTAD. For three different representations of the geometry, we evaluate the performance of four classifiers: three linear classifiers, namely LR, LDA, and SVM; and one non-linear classifier, RF.

The baseline scenario (using the aortic root diameter alone) served to illustrate the limitations of using a uni-dimensional representation of the geometry, which leads to a trade off between specificity versus sensitivity. With the second and third scenarios, we assess the ability of a multidimensional, geometric encoding of the whole thoracic aorta phenotype to identify the characteristic traits of each class. The evaluation of model performance in these two scenarios showcases the ability to achieve high precision and recall while maintaining moderate to high levels of specificity, addressing a primary limitation observed in the baseline scenario. Both scenarios exhibit comparable metrics, with the low-dimensional scenario demonstrating slightly greater robustness. Our findings are in concordance with recent findings that also propose expanding beyond the root the biomarkers that can help understand heritable aortic diseases (Lovato et al., 2024).

For a better understanding of how anatomy is used by the models during stratification, we analyze what are the regions of the aorta that are most relevant to distinguish between the two classes. In Figure 3, corresponding to low-dimension scenario, the regions of interest inferred by the three linear models agree on the primary anatomical traits that distinguish between healthy and sHTAD anatomies. In terms of radius, the most notable variation is observed at proximal ascending aorta, where pathological shapes tend to exhibit an increase in aortic diameter, probably linked to dilation of the aortic root, a prominent manifestation of sHTAD (Loeys et al., 2010; Ruiz-Muñoz et al., 2022). Regarding centerline warping, a difference in length is observed, with pathological shapes tending to display longer thoracic aorta. Length as a geometrical descriptor of the aorta has been studied previously, with reports suggesting that sHTAD patients have longer proximal aorta (Lovato et al., 2024; Franken et al., 2015; Schepers et al., 2018), which may be related to the cyclic longitudinal strain imposed by the left-ventricle contraction (Guala et al., 2019). Notably, excessive length was related to higher risk of aortic adverse events in patients with ascending aorta aneurysms (Wu et al., 2019). The SVM provides a clearer illustration of this phenomenon. This elongation is likely to be linked to the slight narrowing at mid ascending aorta, a distinction particularly evident in the

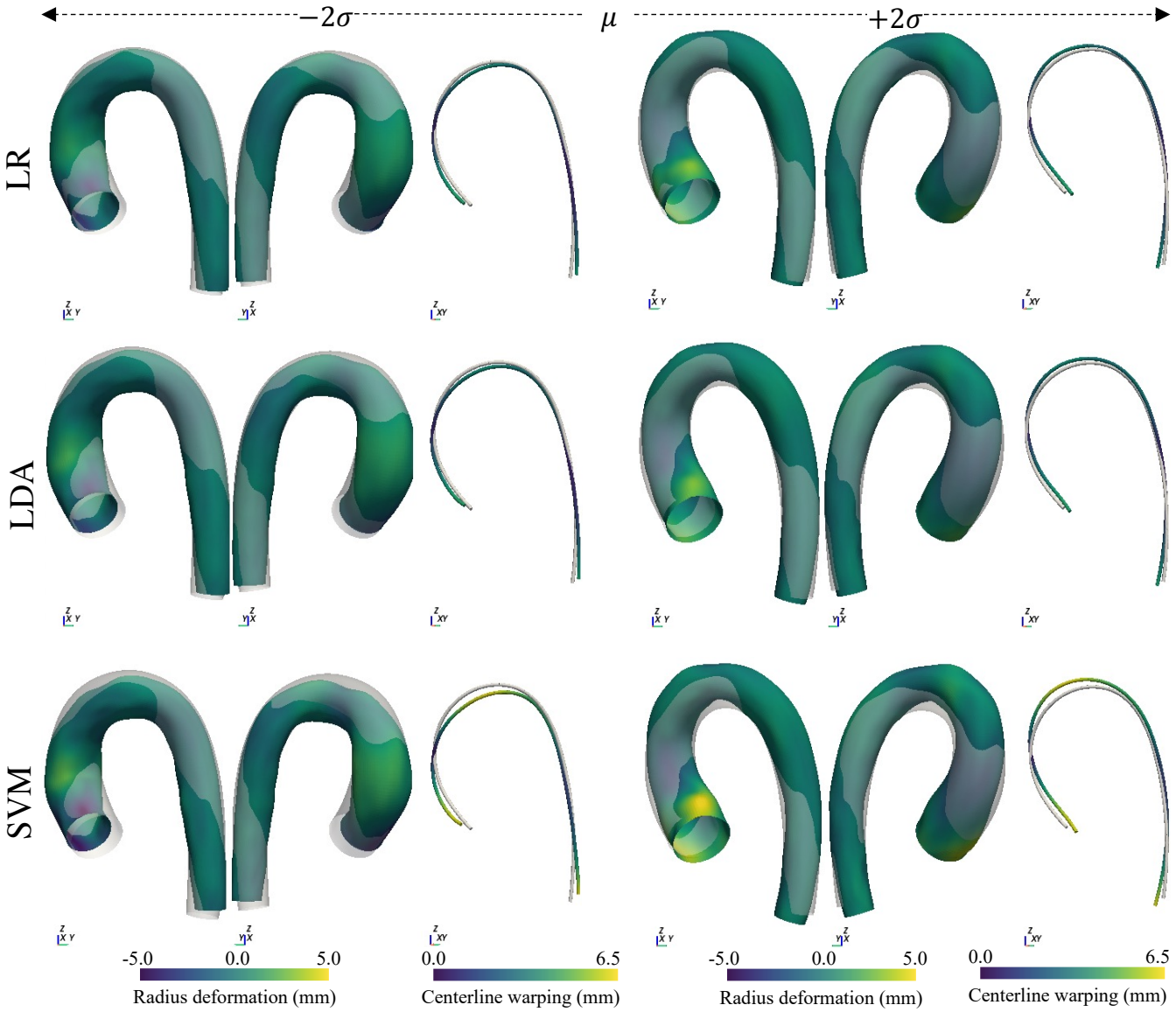

Figure 3: Variation in radius and centerline displacement along the axis, defined by normal of the decision boundary hyperplane in the low-dimension scenario. From top to bottom, LR, LDA and SVM. The range color for the radius variation is [-5, 5] and for the centerline warping is [0, 6], both expressed in milimeters.

yellow hues of healthy geometries, again especially notable in the SVM at the figure's bottom row. A relevant manifestation of progressive aortic lengthening may be aortic tortuosity (Jackson et al., 2005). Indeed, considering that the aorta is likely bounded to a fixed position where it is anatomically fixed by aortic branches, aortic elongation may force the vessel to curve and become tortuous (Jackson et al., 2005). Notably, tortuosity was related to adverse aortic events during follow-up (Franken et al., 2015) and may thus be a relevant metric for risk-stratification.

In contrast to the low-dimensional scenario, Figure 4 shows a disparity between the regions of relevance. LR and SVM appear to explain the aortic dilation proximal to the sinotubular junction and the mid-ascending aorta narrowing, as low-dimensional models did. In addition, some patches at the descending aorta level seem to coincide in positive predictive value, once again, being much more intense in the SVM. The LDA however, gives little importance to the ascending aorta, focusing primarily on the descending segment. Radius importance exhibits a more localised behaviour in the high-dimensional scenario than in the low-dimensional one. With regard to the centerline in the last two columns of Figure 4, LR seems to assign the higher importance to the extrema, which may be related to length variation, as was also noted in the low-dimensional scenario, while LDA and SVM seem to assign the importance to the arch twisting as shown by red hues in the middle of the colour importance band at the last column of the figure.

These discrepancies in importance ranking between models could be explained by the different level of locality in the meaning of each coefficient. Although using PCA dimension reduction is a common approach in the literature (Bruse et al., 2016; Liang et al., 2017; Balaban et al., 2021; Hermida et al., 2023), deformation modes usually exert a global influence on all encoding variables, affecting, to some extent, every coefficient in high-dimensional geometry codification Heimann and Meinzer (2009). In contrast, using the high-dimension representation, enables identifying the importance of very local regions, corre-

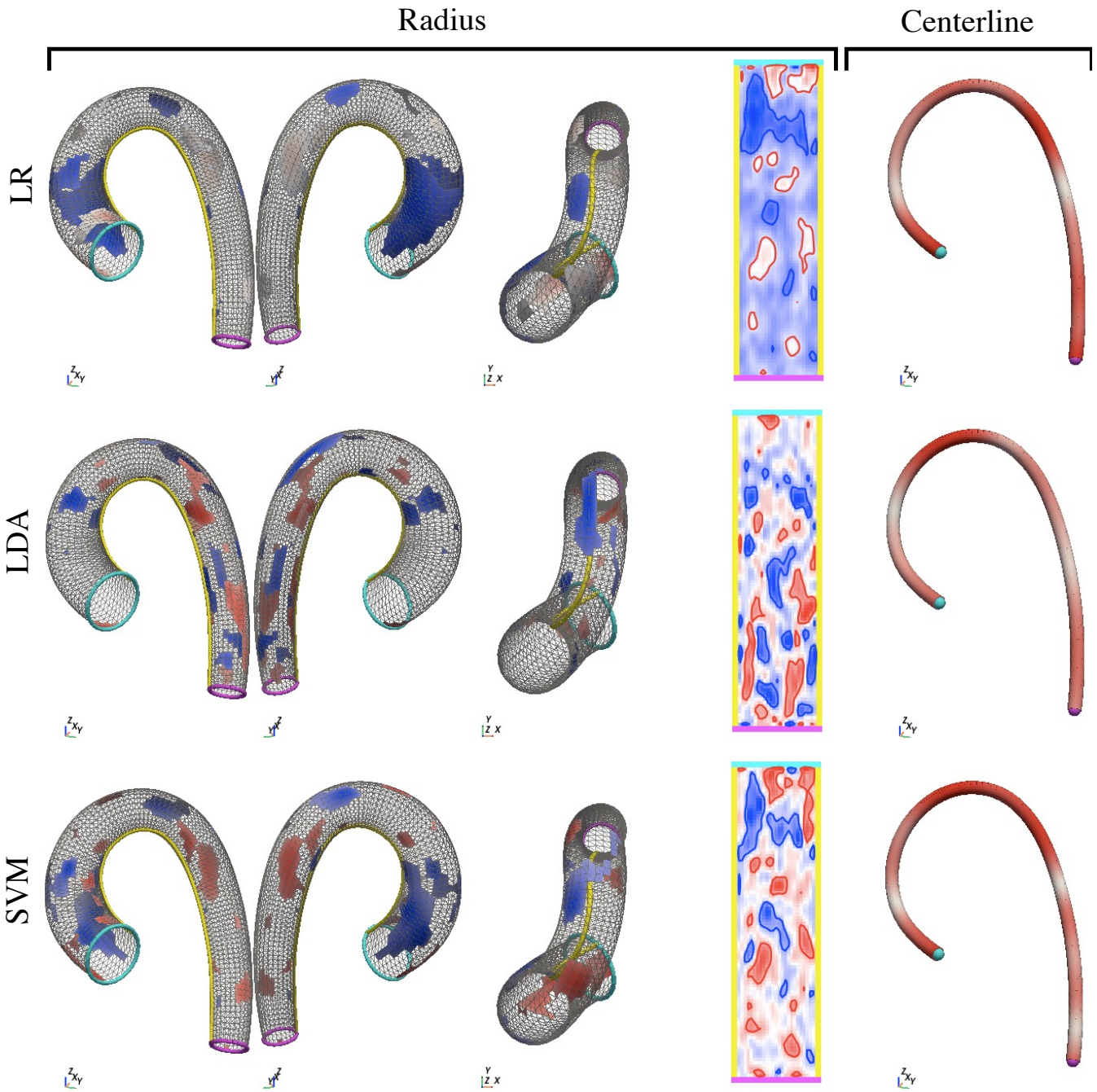

Figure 4: Decomposition of the predictive importance stored in the radius variation and the centerline warping. From left to right the saggital (front and back) and axial views with the importance distribution thresholded to show the most relevant regions. The fourth column represents the importance of the unfolded wall, with the seam highlighted in yellow, the proximal end in cyan, and the distal end in magenta. The fifth column is the centerline importance. The colour corresponds to the relevance attributed by the normal vector of the hyperplane that determines the decision boundary. Consequently, an intense blue hue denotes a negative predictive value (in favour of healthy), while a deep red hue signifies a positive predictive value (in favor of genetic-positive). In contrast, lighter shades suggest diminished relevance for classification. The centerline warping is strictly positive, and thus it is only related to predictive power regardless the sign.

sponding to individual coefficients of the feature vector. As a consequence, this scenario may be more sensitive to spurious differences between the two classes on certain dimensions.

The normal of the separating hyperplane of the linear models in the low-dimensional scenario, has been considered in the literature as the optimal discriminant axis of anatomical variation (Hermida et al., 2023). With this intuition, an anatomical biomarker that accounts for the entire three-dimensional complexity of the thoracic aorta can be defined. Similarly to how the aortic root dilation Z-score is used, a risk score can be defined using the Z-score of the projection of the encoded geometries onto this axis (Balaban et al., 2021; Hermida et al., 2023). The implementation of such a risk score in clinical practice would be highly feasible, due to the noninvasive nature of CMR imaging and to the significant level of automation in our approach, potentially improving sensitivity in clinical assessment of sHTAD.

4.1 Discussion on model selection

When delving into machine learning, a critical decision lies in choosing the pattern representation. The analysis of magnetic resonance images (MRI) is a common preprocessing problem in this area, where typically the input data (images) must be translated into patterns (feature vectors) that will feed the classical ML models. In our case, the combination of the shape encoding used within the linear ML models let us map the feature importance in a 3D space, showing the main deformation direction of the classes involved. This approach provides a ML solution that offers the possibility to interpret their results from a clinical perspective. On the contrary, convolutional models introduce a more complex features engineering, challenging the explainability of their decisions and and thus posing a challenge for their use in clinical domains. Moreover, using CNNs involve a considerably higher number of parameters/weights to fit (e.g. ResNet50 requires 0.5M of parameters approximately for a simple net with 3 convolutional layers (3x3x3) and a dense layer with 100 neurons). For these reasons, we decided to directly reconstruct 3D meshes using the marching cubes algorithm (Cline and Lorensen, 1988) and establish our own shape encoding (feature vectors).

Among all the possible approaches, we have made our model selection based on two criteria, namely, the explainability of the classifier and the number of parameters inferred by the model. The first criterion follows the rising concern of the medical community of making clinical decisions based on "black-box" algorithms from which it is hardly realistic to derive a physiological or medical explanation of the decisions taken by the algorithm (Price, 2018; Rieg et al., 2020; Rudin, 2019). The second criterion is based on preventing overfitting. When handling with data sets of hundreds of samples instead of thousands or millions, special care should be taken on the amount of parameters the model derives from the data. With these considerations in mind, we predominantly opted for linear models that hinge on the inferred separating hyperplane for classification. Consequently, the normal vectors of these planes offer valuable geometric insights into the decision-making process. Additionally, we introduced the RF model to contrast the performance of a non-linear model with the linear counterparts. The choice of RF was deliberate, as the algorithm enables the computation of feature importance by averaging the accumulation of impurity decrease across the decision trees within the RF. However, it is important to note that this metric might be misleading for continuous features, as in our case, and we encountered challenges in finding a clear method to interpret the provided information.

4.2 Limitations and future work

The geometric encoding used (Romero et al., 2023) ignores the different anatomical regions present in the thoracic aorta, namely, the ascending aorta, the aortic arch and the descending aorta. Although this is not a major limitation in this work, since we are working with the thoracic aorta as a whole, it may lead to erroneous measurements at specific landmarks close to the transition between regions. It is also to note that the meshes were derived from 4D flow CMR acquisitions and have limited spatial resolution. However, all the methods described above could be reproduced with the thoracic aorta geometry acquired from 3D imaging modalities with higher spatial resolution (i.e. computed tomography or magnetic resonance angiography).

To further assess the presented methodology in a differential diagnosis scenario, a more general study, with a cohort containing individuals with overlapping clinical manifestations such as aortic aneurysm but negative to sHTAD, would be required, as suggested by Loeys et al. (2010). In addition, analysis of differences between MFS and LDS thoracic geometries would be of high value, due to their different risk profiles. However, due to the limited size of the available cohort, such investigations are beyond the reach of the available cohort for study.

5 Conclusion

We have performed a comprehensive analysis to assess whether three-dimensional geometrical characterization of the aorta differentiates patients with syndromic heritable thoracic aortic diseases from healthy subjects. Our findings present promising results on potential shape biomarkers that could significantly contribute to the intricate processes of diagnosis and treatment planning of these patients.

6 Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7 Acknowledgements

The authors would like to express their sincere gratitude to Jordi Mill and Oscar Camara for making this work possible. This work was funded by Generalitat Valenciana Grant AICO/2021/318 (Consolidables 2021). This research is part of Project PID2020-114291RB-I00 funded by MCIN/10.13039/501100011033/ FEDER, UE. A. Guala received funding from the La Caixa Foundation (LCF/BQ/PR22/11920008). None of the sponsors were involved in any step of the execution of the study or the preparation of the manuscript.

References

Gabriel Balaban, Brian P Halliday, Daniel Hammersley, Christopher A Rinaldi, Sanjay K Prasad, Martin J Bishop, and Pablo Lamata. Left ventricular shape predicts arrhythmic risk in fibrotic dilated cardiomyopathy. *Europace*, 24 (7):1137–1147, December 2021. ISSN 1099-5129. doi: 10.1093/europace/euab306. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9301973/.

Jan L. Bruse, Kristin McLeod, Giovanni Biglino, Hopewell N. Ntsinjana, Claudio Capelli, Tain-Yen Hsia, Maxime Sermesant, Xavier Pennec, Andrew M. Taylor, Silvia Schievano, Andrew Taylor, Alessandro Giardini, Sachin Khambadkone, Marc de Leval, T. Y. Hsia, Edward Bove, Adam Dorfman, G. Hamilton Baker, Anthony Hlavacek, Francesco Migliavacca, Giancarlo Pennati, Gabriele Dubini, Alison Marsden, Jeffrey Feinstein, Irene Vignon-Clementel, Richard Figliola, John McGregor, and for the Modeling of Congenital

Hearts Alliance (MOCHA) Collaborative Group. A statistical shape modelling framework to extract 3d shape biomarkers from medical imaging data: assessing arch morphology of repaired coarctation of the aorta. *BMC Med Imaging*, 16 (1):40, 05 2016. ISSN 1471-2342.

Laurence Campens, Laurent Demulier, Katya De Groote, Kristof Vandekerckhove, Daniël De Wolf, Mary J. Roman, Richard B. Devereux, Anne De Paepe, and Julie De Backer. Reference values for echocardiographic assessment of the diameter of the aortic root and ascending aorta spanning all age categories. *American Journal of Cardiology*, 114(6):914–920, Sep 2014. ISSN 0002-9149. doi: 10.1016/j.amjcard.2014.06.024.

Gavin C. Cawley and Nicola L. C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11(70):2079–2107, 2010. URL http://jmlr.org/papers/v11/cawley10a.html.

Harvey E Cline and William E Lorensen. 4710876 System and method for the display of surface structures contained within the interior region of a solid body. *Magnetic Resonance Imaging*, 6(5):IV–V, September 1988. ISSN 0730725X. doi: 10.1016/0730-725X(88)90153-1. URL https://linkinghub.elsevier.com/retrieve/pii/0730725X88901531.

Romy Franken, Abdelali El Morabit, Vivian De Waard, Janneke Timmermans, Arthur J. Scholte, Maarten P. Van Den Berg, Henk Marquering, Nils R.N. Planken, Aeilko H. Zwinderman, Barbara J.M. Mulder, and Maarten Groenink. Increased aortic tortuosity indicates a more severe aortic phenotype in adults with Marfan syndrome. *International Journal of Cardiology*, 194:7–12, September 2015. ISSN 01675273. doi: 10.1016/j.ijcard.2015.05.072. URL https://linkinghub.elsevier.com/retrieve/pii/S0167527315011298.

Andrea Guala, Gisela Teixidó-Tura, Jose Rodríguez-Palomares, Aroa Ruiz-Muñoz, Lydia Dux-Santoy, Nicolas Villalva, Chiara Granato, Laura Galian, Laura Gutiérrez, Teresa González-Alujas, Violeta Sanchez, Alberto Forteza, David García-Dorado, and Artur Evangelista. Proximal aorta longitudinal strain predicts aortic root dilation rate and aortic events in Marfan syndrome. *European Heart Journal*, 40(25):2047–2055, July 2019. ISSN 0195-668X, 1522-9645. doi: 10.1093/eurheartj/ehz191. URL https://academic.oup.com/eurheartj/article/40/25/2047/5448881.

Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis*, 13(4):543–563, 2009. ISSN 1361-8415. doi: https://doi.org/10.1016/j.media.2009.05.004. URL https://www.sciencedirect.com/science/article/pii/S1361841509000425.

Uxio Hermida, Milou P. M. van Poppel, David F. A. Lloyd, Johannes K. Steinweg, Trisha V. Vigneswaran, John M. Simpson, Reza Razavi, Adelaide De Vecchi, Kuberan Pushparajah, and Pablo Lamata. Learning the Hidden Signature of Fetal Arch Anatomy: a Three-Dimensional Shape Analysis in Suspected Coarctation of the Aorta. *Journal of Cardiovascular Translational Research*, 16(3):738–747, June 2023. ISSN 1937-5395. doi: 10.1007/s12265-022-10335-9. URL https://doi.org/10.1007/s12265-022-10335-9.

Eric M. Isselbacher, Ourania Preventza, James Hamilton Black III, John G. Augoustides, Adam W. Beck, Michael A. Bolen, Alan C. Braverman, Bruce E. Bray, Maya M. Brown-Zimmerman, Edward P. Chen, Tyrone J. Collins, Abe DeAnda, Christina L. Fanola, Leonard N. Girardi, Caitlin W. Hicks, Dawn S. Hui, William Schuyler Jones, Vidyasagar Kalahasti, Karen M. Kim, Dianna M. Milewicz, Gustavo S. Oderich, Laura Ogbechie, Susan B. Promes, Elsie Gyang Ross, Marc L. Schermerhorn, Sabrina Singleton Times, Elaine E. Tseng, Grace J. Wang, and Y. Joseph Woo. 2022 acc/aha guideline for the diagnosis and management of aortic disease. *Journal of the American College of Cardiology*, 80(24):e223–e393, 2022. doi: 10.1016/j.jacc.2022.08.004. URL https://www.jacc.org/doi/abs/10.1016/j.jacc.2022.08.004.

Zane S. Jackson, Dorota Dajnowiec, Avrum I. Gotlieb, and B. Lowell Langille. Partial Off-Loading of Longitudinal Tension Induces Arterial Tortuosity. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 25(5):957–962, May 2005. ISSN 1079-5642, 1524-4636. doi: 10.1161/01.ATV.0000161277.46464.11. URL https://www.ahajournals.org/doi/10.1161/01.ATV.0000161277.46464.11.

Kevin M. Johnson, Darren P. Lum, Patrick A. Turski, Walter F. Block, Charles A. Mistretta, and Oliver Wieben. Improved 3d phase contrast mri with off-resonance corrected dual echo vipr. *Magnetic Resonance in Medicine*, 60(6):1329–1336, 2008. doi: https://doi.org/10.1002/mrm.21763. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.21763.

Liang Liang, Minliang Liu, Caitlin Martin, John A. Elefteriades, and Wei Sun. A machine learning approach to investigate the relationship between shape features and numerically predicted risk of ascending aortic aneurysm. *Biomechanics and Modeling in Mechanobiology*, 16(5):1519–1533, Oct 2017. ISSN 1617-7940.

B. L. Loeys, H. C. Dietz, A. C. Braverman, B. L. Callewaert, J. De Backer, R. B. Devereux, Y. Hilhorst-Hofstee, G. Jondeau, L. Faivre, D. M. Milewicz, R. E. Pyeritz, P. D. Sponseller, P. Wordsworth, and A. M. De Paepe. The revised Ghent nosology for the Marfan syndrome. *Journal of Medical Genetics*, 47(7):476–485, June 2010. ISSN 1468-6244. doi: 10.1136/jmg.2009.072785.

Luigi Lovato, Mariano Cefarelli, Luca Di Marco, Daniel Arcioni, Giada Tortora, Ada Dormi, Nicolò Schicchi, Elisabetta Mariucci, Marco Di Eusanio, Davide Pacini, and Rossella Fattori. Marfan and Loeys-Dietz aortic phenotype: A potential tool for diagnosis and management. *JTCVS Open*, 19:223–240, June 2024. ISSN 26662736. doi: 10.1016/j.xjon.2024.03.015. URL https://linkinghub.elsevier.com/retrieve/pii/S2666273624000950.

W. Nicholson Price. Big data and black-box medical algorithms. *Science Translational Medicine*, 10 (471):eaao5333, 2018. doi: 10.1126/scitranslmed.aao5333. URL https://www.science.org/doi/abs/10.1126/scitranslmed.aao5333.

Thilo Rieg, Janek Frick, Hermann Baumgartl, and Ricardo Buettner. Demonstration of the potential of white-box machine learning approaches to gain insights from cardiovascular disease electrocardiograms. *PLOS ONE*, 15(12):1–20, 12 2020. doi: 10.1371/journal.pone.0243615. URL https://doi.org/10.1371/journal.pone.0243615.

Pau Romero, Abel Pedrós, Rafael Sebastian, Miguel Lozano, and Ignacio García-Fernández. Robust inter-patient comparison and analysis of blood vessels through the univocal definition of point coordinates, 2023. URL https://doi.org/10.48550/arXiv.2311.00888.

Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, May 2019. ISSN 2522-5839. doi: 10.1038/s42256-019-0048-x. URL https://www.nature.com/articles/s42256-019-0048-x.

Aroa Ruiz-Muñoz, Andrea Guala, Jose Rodriguez-Palomares, Lydia Dux-Santoy, Luz Servato, Angela Lopez-Sainz, Lucia La Mura, Chiara Granato, Javier Limeres, Teresa Gonzalez-Alujas, Laura Galián-Gay, Laura Gutiérrez, Kevin Johnson, Oliver Wieben, Augusto Sao-Aviles, Ignacio Ferreira-Gonzalez, Arturo Evangelista, and Gisela Teixido-Tura. Aortic flow dynamics and stiffness in Loeys–Dietz syndrome patients: a comparison with healthy volunteers and Marfan syndrome patients. *European Heart Journal - Cardiovascular Imaging*, 23(5):641–649, April 2022. ISSN 2047-2404, 2047-2412. doi: 10.1093/ehjci/jeab069. URL https://academic.oup.com/ehjcimaging/article/23/5/641/6295278.

Dorien Schepers, Giada Tortora, Hiroko Morisaki, Gretchen MacCarrick, Mark Lindsay, David Liang, Sarju G. Mehta, Jennifer Hague, Judith Verhagen, Ingrid Van De Laar, Marja Wessels, Yvonne Detisch, Mieke Van Haelst, Annette Baas, Klaske Lichtenbelt, Kees Braun, Denise Van Der Linde, Jolien Roos-Hesselink, George McGillivray, Josephina Meester, Isabelle Maystadt, Paul Coucke, Elie El-Khoury, Sandhya Parkash, Birgitte Diness, Lotte Risom, Ingrid Scurr, Yvonne Hilhorst-Hofstee, Takayuki Morisaki, Julie Richer, Julie Désir, Marlies Kempers, Andrea L. Rideout, Gabrielle Horne, Chris Bennett, Elisa Rahikkala, Geert Vandeweyer, Maaike Alaerts, Aline Verstraeten, Hal Dietz, Lut Van Laer, and Bart Loeys. A mutation update on the LDS-associated genes *TGFB2/3* and *SMAD2/3*. *Human Mutation*, 39(5):621–634, May 2018. ISSN 10597794. doi: 10.1002/humu.23407. URL https://onlinelibrary.wiley.com/doi/10.1002/humu.23407.

Bane Sullivan and Alexander Kaszynski. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). *Journal of Open Source Software*, 4(37):1450, May 2019. doi: 10.21105/joss.01450. URL https://doi.org/10.21105/joss.01450.

Jinlin Wu, Mohammad A. Zafar, Yupeng Li, Ayman Saeyeldin, Yan Huang, Rui Zhao, Juntao Qiu, Maryam Tanweer, Mohamed Abdelbaky, Anton Gryaznov, Joelle Buntin, Bulat A. Ziganshin, Sandip K. Mukherjee, John A. Rizzo, Cuntao Yu, and John A. Elefteriades. Ascending aortic length and risk of aortic adverse events: The neglected dimension. *Journal of the American College of Cardiology*, 74(15):1883–1894, 2019. ISSN 0735-1097. doi: https://doi.org/10.1016/j.jacc.2019.07.078.