WILEY

Generation of synthetic aortic valve stenosis geometries for in silico trials

Sabine Verstraeten[1] | Martijn Hoeijmakers[2] | Pim Tonino[3] |
Jan Brüning[4] | Claudio Capelli[5] | Frans van de Vosse[1] |
Wouter Huberts[1,6]

[1]Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

[2]ANSYS, Zoetermeer, The Netherlands

[3]Department of Cardiology, Catharina Hospital, Eindhoven, The Netherlands

[4]Institute of Computer-assisted Cardiovascular Medicine, Charite Universitaetsmedizin, Berlin, Germany

[5]Institute of Cardiovascular Science, University College London, London, UK

[6]Department of Biomedical Engineering, CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht, The Netherlands

Correspondence
Sabine Verstraeten, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.
Email: s.c.f.p.m.verstraeten@tue.nl

Abstract

In silico trials are a promising way to increase the efficiency of the development, and the time to market of cardiovascular implantable devices. The development of transcatheter aortic valve implantation (TAVI) devices, could benefit from in silico trials to overcome frequently occurring complications such as paravalvular leakage and conduction problems. To be able to perform in silico TAVI trials virtual cohorts of TAVI patients are required. In a virtual cohort, individual patients are represented by computer models that usually require patient-specific aortic valve geometries. This study aimed to develop a virtual cohort generator that generates anatomically plausible, synthetic aortic valve stenosis geometries for in silico TAVI trials and allows for the selection of specific anatomical features that influence the occurrence of complications. To build the generator, a combination of non-parametrical statistical shape modeling and sampling from a copula distribution was used. The developed virtual cohort generator successfully generated synthetic aortic valve stenosis geometries that are comparable with a real cohort, and therefore, are considered as being anatomically plausible. Furthermore, we were able to select specific anatomical features with a sensitivity of around 90%. The virtual cohort generator has the potential to be used by TAVI manufacturers to test their devices. Future work will involve including calcifications to the synthetic geometries, and applying high-fidelity fluid–structure-interaction models to perform in silico trials.

KEYWORDS
aortic valve stenosis, In silico trials, statistical shape modeling, synthetic geometries, transcatheter aortic valve implantation, virtual cohort

1 | INTRODUCTION

The process from development to the clinical introduction of cardiovascular implantable devices is very long and expensive due to the strict regulatory requirements regarding device safety and efficacy. Depending on the type of device, the

validation chain to achieve regulatory approval consists of several pre-clinical steps: in vitro bench testing, ex vivo testing, and in vivo animal testing. Then, multi-phase clinical trials, that is, tests on humans must be completed.[1,2] However, this validation chain faces some problems. Firstly, the animal model is not always a reliable representation of the human body. As a result, devices may be rejected after the phase of human trials, despite significant financial investments during the animal testing phase. Moreover, animal testing involves ethical issues regarding animal welfare. Thirdly, medical devices are typically tested at single physiological states, without taking into consideration varying physiological conditions that can occur in the human body (e.g., varying heart rate, cardiac output, and blood pressure). Iterative validation and refinement of the devices are also challenging since it requires new randomized clinical trials every time the device is adapted. Lastly, the sample size in animal and clinical trials is relatively small, as a result of which the heterogeneity of the population is not fully covered. These problems could possibly be solved by incorporating in silico trials into the validation chain prior to the animal tests and clinical trials.[1,2]

During in silico trials devices are tested on groups of virtual patients, that is, virtual cohorts, in which individual patients are represented by computer models. These models can be either based on quantitative parameters measured on an individual (subject-specific models) or sampled from population distributions of those parameters (population-specific models).[3] These computer models can be used to simulate physiological behavior and the response after device implantation. There are many potential advantages of in silico trials. They can easily be repeated multiple times, without the need to set up new animal trials or to organize new patient inclusions for clinical trials. Additionally, virtual cohorts can be much larger than real patient cohorts, and rare cases that are challenging to include in real patient populations can be incorporated more easily. Lastly, in silico simulations can help to improve the design of real clinical trials, for example by gaining insight into the optimal number and selection of patients necessary to address the research question effectively. Thus, by performing in silico trials, the number of animal tests and clinical trials could be reduced. As a result animal, and patient burdens, as well as costs and time-to-market, can be reduced.[1,2]

An example of a cardiovascular implantable device that would benefit from in silico trials during its development is the transcatheter aortic valve implantation (TAVI) device. TAVI is a minimally invasive treatment for patients with aortic valve stenosis where the aortic valve is replaced with a prosthetic valve. Despite TAVI has been successfully employed over the past two decades, many complications have been reported such as conduction problems, leading to pacemaker dependency after treatment, and aortic regurgitation (AR).[4–6] TAVI devices, that is, the prosthetic valves, are subject to continuous development to improve current devices and overcome these complications. In silico trials could possibly speed up the clinical introduction of improved TAVI devices. Computational fluid dynamics (CFD) and fluid–structure interaction (FSI) models are typically used to simulate aortic valve dynamics before,[7,8] and after treatment.[9–13] Furthermore, finite element methods can be used to simulate contact forces between the aortic root and the TAVI valve during deployment.[14–16] Performing in silico trials using these models requires patient-specific aortic valve geometries and boundary conditions. These geometries, represented by three-dimensional meshes, could be obtained from medical image data, such as computed tomography (CT) images. However, reconstructing the geometries from those images is time-consuming and suffers from observer variability.[17] Furthermore, the available patient data is scarce and often incomplete. Additionally, sharing of patient data usually involves many ethical and data privacy issues. A good alternative would be to generate synthetic geometries. Synthetic geometries must be realistic and should mimic real patient populations, while simultaneously not representing any specific patients.

To generate synthetic aortic valve geometries, the geometries must be parameterized first. Several approaches have been used in previous studies to parameterize aortic valve geometries. For example, Vergara et al.,[18] parameterized bicuspid aortic valves by two parameters: valve opening and the opening's angle of rotation about the centerline of the aorta. Hosny et al.[19] proposed a method to parameterize the leaflets with seven points: the point of coaptation of the leaflets, the three attachment points at the level of the sinotubular junction (STJ) and the three basal attachment points. Another method to parameterize aortic valve geometries is statistical shape modeling (SSM),[20] where the parameters describing a geometry, that is, shape coefficients, are based on the shape variations within a data set of geometries. These shape coefficients are generally obtained using principal component analysis (PCA).[21] SSM has been used in multiple studies to generate synthetic geometries. In Romero et al.[17] the concept was used to generate synthetic aorta geometries for a virtual patient cohort, whereas Hoeijmakers et al.[8] used it to generate synthetic aortic valve geometries for training a surrogate model of a CFD simulation model. The methods used in those studies are based on parametric methods for which iso-topology between the meshes is required. However, especially when dealing with complex, amorphous geometries it is challenging to preserve this iso-topology. Moreover, when geometries in the training data set are reconstructed or post-processed differently, which is the case in this work (Section 2.2), iso-topology is also not maintained. Therefore, Mansi et al.[22–24] introduced the non-parametric SSM framework, developed by Durrleman et al.,[25] to the cardiovascular field.

This method relies on deformations of a template geometry to patient-specific geometries. An SSM is then built on these deformations, instead of on the geometries themselves. The advantages of this framework are that it does not require iso-topological geometries, and that it is applicable to sets of geometries with large variability. Several other studies successfully made use of this framework to obtain either an SSM to extract 3D shape biomarkers from aortic shapes after repair of aortic coarctation,[26] or to generate a set of synthetic geometries to train deep neural networks for fast acquisition of aortic 3D pressure and velocity flow fields.[27] After setting up an SSM based on a real patient population, synthetic geometries can be generated by sampling new combinations of shape coefficients.

In Niederer et al.[28] two different sampling strategies for obtaining a virtual cohort were proposed. The first strategy is sampling parameter sets, representing virtual patients, from an inferred distribution, which is a statistical distribution fitted to the parameters of a real patient cohort. The second strategy is random sampling a parameter set and using acceptance criteria to decide if these result in anatomically or physiologically plausible models. Romero et al.[17] introduced two types of acceptance criteria: data-driven, and clinically-driven acceptance criteria. The first type requires that the virtual cohort is compatible with a certain reference cohort, for example, a real patient cohort. An example of a data-driven acceptance criterion is that the parameters must be within the minimum and maximum of the parameters observed in the real patient cohort. For the second type, the members of the virtual cohort have to meet certain clinical requirements, such as specific anatomical features.

Clinically-driven acceptance criteria are relevant for in silico trials of cardiovascular implantable devices, as some specific anatomical features might affect the success of device implantation and performance. For example, specific anatomical features of aortic valves are associated with the occurrence of complications after TAVI. In the study of Ishizu et al.[4] it was found that patients with a converging left ventricular outflow tract (LVOT), that is, LVOT smaller than annulus, have a higher chance of suffering from conduction problems that can lead to pacemaker dependency. Furthermore, Condado et al.[5] found that non-tubular, that is, converging and diverging LVOTs are associated with mild to severe paravalvular leakage. Sherif et al. found a positive relationship between the chance of significant AR, and the angle between the LVOT and the ascending aorta.[6]

To be effective, in silico trials need to include virtual cohorts with anatomical features associated with the complications that are being addressed in these trials. For example, if a TAVI device is improved to address the conduction problems, virtual patients with a converging LVOT should be included in the in silico trials. By using in silico trials, one can easily scale up the number of study samples in which this feature is present. Therefore, the aim of this research is to develop a framework that is (1) able to generate anatomically plausible aortic valve stenosis geometries (data-driven approach), and (2) allows for the selection of clinically relevant anatomical features; in our case two: the shape of the LVOT (converging or diverging), and the angle between the LVOT and the ascending aorta (clinically-driven approach). The framework, required for this aim, makes use of a novel combination of techniques: non-parametric statistical shape modeling to generate synthetic geometries and clinically-driven acceptance criteria for the selection of the anatomical features of interest.

2 | METHODS

The presented work is based on a data set op 97 aortic valve geometries with a severe stenosis (Section 2.2). These geometries were all post-processed differently, because of which, iso-topology was not maintained. To tackle this non-iso-topology, we used a non-parametric SSM framework, proposed by Durrleman et al.[25] to generate synthetic aortic valve stenosis geometries. In this section, we will first give an elaborate explanation of the theory behind the framework, aiming to ensure a shared conceptual understanding among all readers (Section 2.1). Then, the data that was used to build the SSM is described (Section 2.2) in more detail and the required steps to build the model are discussed (Section 2.5). Once the SSM is built, data-driven synthetic geometries are generated with a sampling technique discussed in Section 2.6. After validating the synthetic geometries, clinically-driven geometries are generated by applying several filtering methods (Section 2.7).

2.1 | Theory behind the statistical shape modeling framework

The non-parametric SSM framework proposed by Durrleman et al.,[25] does not require iso-topology between the geometries on which the SSM is built. This framework relies on deformations of a template geometry to approximate patient-

specific geometries, that is, target geometries. Each target geometry can be described by a unique set of deformation vectors (momenta) that induce this deformation. These momenta are located at grid points, that is, control points, that are identical for all geometries. Any point $\vec{x}(0) = \vec{x}_0$, belonging to the template geometry, can be deformed into a corresponding point $\vec{x}(t)$ in the deformed template using a time-varying velocity field $\dot{\vec{x}}(t)$ with pseudo-time $t \in [0,1]$. The deformed position of each point at time instant t is then given by[25]:

$$\vec{x}(t) = \vec{x}_0 + \int_0^t \dot{\vec{x}}(t)dt. \tag{1}$$

We are interested in the final deformed template $\vec{x}_{def} = \vec{x}(1)$. The velocity field depends on time-varying control points $\vec{c}_p(t)$ and momenta $\vec{\beta}_p(t)$:

$$\dot{\vec{x}}(t) = \sum_{p=1}^{p=N_{cp}} K_V\left(\vec{x}(t), \vec{c}_p(t)\right)\vec{\beta}_p(t), \tag{2}$$

with $\vec{c}_p(t)$ the pth control point, $\vec{\beta}_p(t)$ the momentum vector located at the pth control point, and N_{cp} the number of control points. K_V is a Gaussian kernel with width λ_V that has any two points $\vec{x}_1$ and $\vec{x}_2$ in the 3D space as an input:

$$K_V\left(\vec{x}_1, \vec{x}_2\right) = e^{\frac{-\|\vec{x}_1 - \vec{x}_2\|^2}{\lambda_V^2}}. \tag{3}$$

In other words, each template point $\vec{x}_0$ follows the integral curve of Equation 1 to deform to $\vec{x}_{def}$. The contribution of any momentum vector $\vec{\beta}_p$, located at control point $\vec{c}_p$ to the path of $\vec{x}_0$ is weighted by the distance between $\vec{x}$ and $\vec{c}_p$, by means of K_V. Equation (2) also describes the motion of the kth control point $\vec{c}_k(t)$:

$$\dot{\vec{c}}_k(t) = \sum_{p=1}^{p=N_{cp}} K_V\left(\vec{c}_k(t), \vec{c}_p(t)\right)\vec{\beta}_p(t). \tag{4}$$

The time-varying momenta that deform $\vec{x}_0$ into $\vec{x}_{def}$ are not unique. Therefore, the momenta that minimize the integral of the "kinetic" energy along the path are chosen, which satisfies the following differential equation[25]:

$$\dot{\vec{\beta}}_k(t) = -\sum_{p=1}^{p=N_{cp}} \vec{\beta}_k(t) \cdot \vec{\beta}_p(t) \left[\nabla_{\vec{x}_1} K_V\left(\vec{x}_1, \vec{x}_2\right)\right]_{\vec{x}_1 = \vec{c}_k(t), \vec{x}_2 = \vec{c}_p(t)}, \tag{5}$$

where $\vec{\beta}_k$ is the momentum vector at the kth control point, and $\nabla_{\vec{x}_1}$ is the gradient with respect to $\vec{x}_1$. For a detailed proof of this equation, we refer to the work of Durrleman et al.[25] In summary, to determine the deformation of the template shape towards the target shape the following set of ordinary differential equations should be solved:

$$\dot{\mathbf{S}} = \begin{cases} \dot{\mathbf{X}}(t) & \mathbf{X}(0) = \mathbf{X_0} \\ \dot{\mathbf{C}}(t) & \mathbf{C}(0) = \mathbf{C_0} \\ \dot{\mathbf{B}}(t) & \mathbf{B}(0) = \mathbf{B_0} \end{cases}, \tag{6}$$

where $\mathbf{C}$, and $\mathbf{B}$ are the $N_{cp} \times 3$ matrices of control points and momenta respectively, and $\mathbf{X}$ the $N_v \times 3$ matrix with the N_v template vertices. The three columns in those matrices store the $x-$, $y-$, and $z-$ coordinates of the control points, momenta and template points. Thus, to approximate each target geometry a template geometry $\mathbf{X_0}$, a set of initial control points $\mathbf{C_0}$, and a set of initial momenta $\mathbf{B_0}$ is required. The template and control points are equal for all target geometries, but the initial momenta are different for each geometry. Therefore, each target geometry of the data set can be parameterized by a unique set of initial momenta. Figure 1 visualizes the deformation of a template by means of momenta to approximate an aortic valve geometry. Now, we understand how the initial momenta deform the template
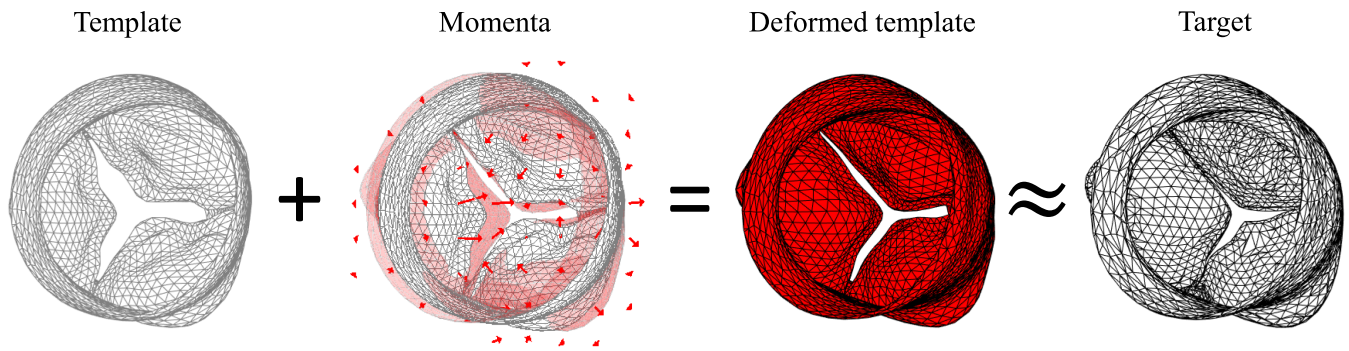
FIGURE 1 The deformation of the template geometry (gray) into the approximation (red) of the target geometry (black) with help of the initial momenta (red) belonging to the target geometry.

shape to approximate the target shape. In the next paragraph, it will be explained how these initial momenta and the template are obtained.

Suppose that Φ_i is the deformation, induced by initial momenta $\mathbf{B_{0,i}}$, that maps the template geometry $\mathbf{T}$ to the deformed template, approximating the ith target geometry $\mathbf{S_i}$[29]:

$$\mathbf{S_i} \approx \Phi_i(\mathbf{B_{0,i}})\mathbf{T}. \tag{7}$$

The template and deformations are estimated by alternately minimizing a cost function C with respect to the template and to the deformations:

$$C = \sum_{i=1}^{i=N_S} d_W(\Phi_i(\mathbf{B_{i,0}})\mathbf{T}, \mathbf{S_i})^2/\sigma^2 + \sum_{i=1}^{i=N_S} \boldsymbol{B_{0,i}}^\top \mathbf{K}_W\left(\vec{c}, \vec{c}\right)\boldsymbol{B_{0,i}}, \tag{8}$$

with N_S the number of target geometries in the data set. In the upcoming paragraph, this cost function will be explained step by step. The first term of the cost function contains the sum of the distances d_W between all target geometries $\mathbf{S_i}$ and their approximations $\Phi_i(\mathbf{B_{0,i}})\mathbf{T}$. This distance is the Varifold distance, a similarity metric that does not require point-to-point correspondence between geometries. In the Varifold framework, two geometries, that is, meshes, are embedded into a Hilbert space (W^*), where the inner-product between two meshes $\mathbf{S}$ and $\mathbf{S}'$ is defined as[25]:

$$(\mathbf{S}, \mathbf{S}')_{W^*} = \sum_{p=1}^{p=N_f} \sum_{q=1}^{p=N_f'} K_W\left(\vec{c}_p, \vec{c}_q{}'\right)\frac{\left(\vec{n}_p \cdot \vec{n}_q{}'\right)}{|\vec{n}_p||\vec{n}_q{}'|}, \tag{9}$$

with N_f and N_f' the number of faces, $\vec{c}_p$ and $\vec{c}_q{}'$ the centers of the faces and $\vec{n}_p$ and $\vec{n}_q{}'$ the face normals of the meshes $\mathbf{S}$ and $\mathbf{S}'$ respectively. K_W is a Gaussian kernel function equal to Equation (3). The distance between $\mathbf{S}$ and $\mathbf{S}'$ is then:

$$d_W(\mathbf{S}, \mathbf{S}')^2 = \|\mathbf{S} - \mathbf{S}'\|_{W^*}^2 = (\mathbf{S}, \mathbf{S})_{W^*} + (\mathbf{S}', \mathbf{S}')_{W^*} - 2(\mathbf{S}, \mathbf{S}')_{W^*}. \tag{10}$$

Thus, Equation (10) is the definition of the distance d_W that is used in the first term of the cost function (Equation (8)). The second term of the cost function is a penalization term to avoid deformations with high "kinetic" energy,[30] where $\boldsymbol{B_{0,i}}$ is the $N_{cp} \times 3$ matrix of momenta belonging to target geometry $\mathbf{S_i}$. $\mathbf{K}_W(\vec{c}, \vec{c})$ is the matrix that is obtained by computing K_W for all combinations of control points $\vec{c}$. The relative importance of the two terms in the cost function (Equation (8)) is determined by σ.

To minimize the cost function in Equation (8) two steps are alternated.[29] The first step is to keep the template geometry fixed and minimize the cost function with respect to the initial momenta $\mathbf{B_{0,i}}$. This step consists of
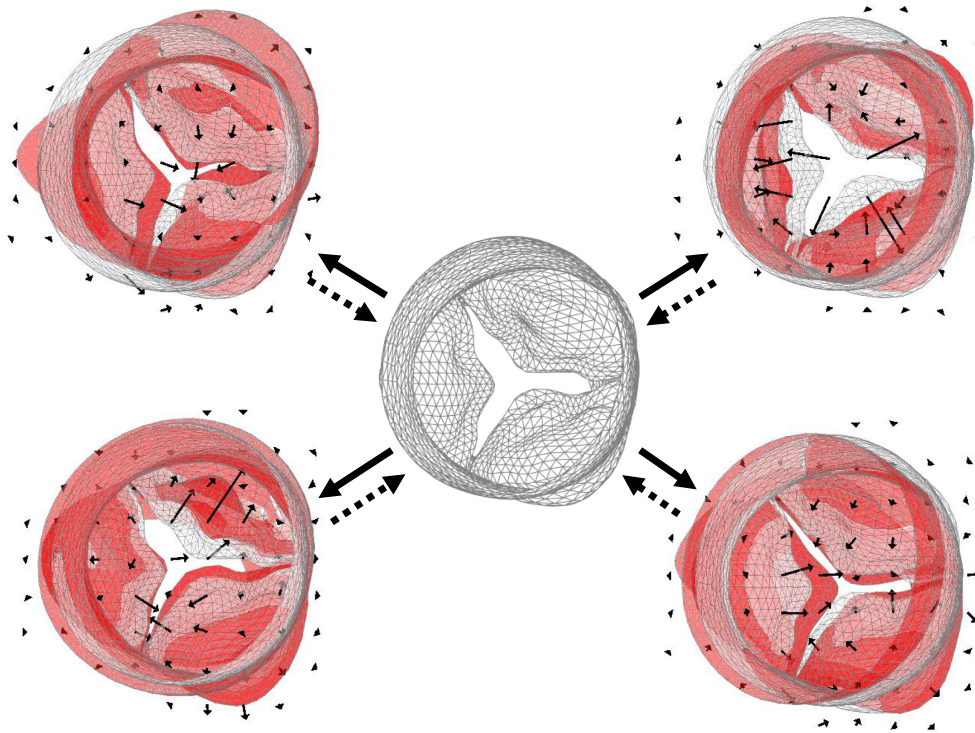
FIGURE 2 Visualization of the two-step minimization approach to estimate the momenta (black vectors) and the template geometry (middle). The first step (black, solid arrows) is minimizing the cost function (Equation (8)) with respect to the momenta, that is, registering the template to the target geometries (red). The second step (black, dashed arrows) is minimizing the cost function with respect to the template.

N_S registrations of the template $\mathbf{T}$ to each target geometry $\mathbf{S_i}$. The second step is to fix the deformations and minimize the cost function with respect to the template, which leads to minimizing the first term of Equation (8). For this minimization, a gradient descent scheme is used. These two steps are alternated until convergence. Figure 2 visualizes the two-step minimization approach. For detailed information about the minimization process, we refer to Durrleman et al.[29]

After estimating the template and initial momenta, each geometry can be represented by a unique set of initial momenta, all located at the same grid of control points. Then, the initial p momenta for each geometry $\mathbf{S_i}$ are concatenated in a row vector:

$$\underset{\sim}{v}_i^\top = \left[\beta_{x,1},...,\beta_{x,p},\beta_{y,1},...,\beta_{y,p},\beta_{z,1},...,\beta_{z,p}\right], \tag{11}$$

with β_x, β_y, and β_z, the x, y, and z component of each momentum vector respectively. The average of N_S sets of momenta was calculated as follows:

$$\underset{\sim}{\bar{v}} = \frac{1}{N_S}\sum_{i=1}^{i=N_S}\underset{\sim}{v}_i. \tag{12}$$

The row vectors $\underset{\sim}{v}_i^\top$ were collected in a matrix of size $[N_S, 3p]$, in which each row contained the $3p$-dimensional coordinates, representing one geometry. Principal component analysis (PCA)[21] was performed to reduce the dimensionality of this matrix. The resulting principal components, that is, shape modes, describe the directions in which the data varies the most. The shape modes ϕ_m are vectors with length $3p$, and they are ordered from high to low explained variance. Each set of momenta $\underset{\sim}{v}_i$ can be approximated $\left(\widehat{\underset{\sim}{v}}_i\right)$ by adding a weighted combination of N_m shape modes to the average set of momenta:

$$\widehat{v}_i = \overline{v} + \sum_{m=1}^{m=N_m} \alpha_{i,m}\phi_m \quad i \in \{1,2,...,N_S\}, \tag{13}$$

with ϕ_m the mth shape mode, and $\alpha_{i,m}$ the shape coefficient for the mth shape mode, belonging to the ith geometry. Shape coefficients are the projections of the vectors v_i onto the directions of the shape modes. Each target geometry $\mathbf{S_i}$ can be approximated by deforming the template with momenta $\widehat{v}_i$ according to Equation (1). The larger the number of shape modes that are taken into account, the more accurate the approximation is.

2.2 | Data acquisition

To set up the statistical shape model, a data set of 97 peak systolic aortic valve geometries in STL format was used, originating from a population of patients with aortic stenosis. The meshes included the LVOT, the sinuses of Valsalva, the ascending aorta, and the aortic arch. The retrospective data set was collected by the Institute of Computer-assisted Cardiovascular Medicine, Charité Universitaetsmedizin Berlin. The data were collected in the scope of the H2020 project SIMCor. Data were obtained from clinical routine. The collection of informed consent was waived by the institutional review board which approved the study. The meshes were reconstructed from pre-operative CT images of patients who underwent TAVI treatment. Reconstruction of the patient-specific anatomy of the left ventricular outflow tract, the aortic root including the aortic valve leaflets, as well as the ascending aorta was performed semi-automatically using a combination of a shape-constrained deformable model for the whole heart[31] and one for the tricuspid aortic valve.[32] A detailed description of the definition of the parametric aortic valve model and the methods for fitting the model to the CT image data is provided in Weese et al.[32] However, a manual correction of these automated reconstructions was necessary to account for the complex asymmetry in the stenosed valves. To facilitate this, a new coordinate system was introduced that allowed to slice through the CT image data in an orientation parallel to the aortic annulus. Using this orientation, the contours of the valve leaflet model were manipulated via drag and drop. The contours of each leaflet were moved to match the Hounsfield contrasts and contours visible in the CT image data. This procedure was performed slice by slice from the annulus towards the slice in which no leaflet tip was discernible anymore. This procedure is also described in more detail in Franke et al.[33] After reconstruction, the meshes were meshed, which led to non-iso-topological meshes. Therefore the non-parametric SSM framework described in Section 2.1 was used.

2.3 | Pre-processing

For the non-parametric SSM, a registration of the geometries was required. Therefore, all target geometries were translated, rotated and scaled to align with the template geometry (Section 2.5). For this purpose, the ascending aorta was temporarily cut off at 1–2 cm from the sinotubular junction (STJ). This step was performed to be able to align the sinuses of the geometries and to avoid the alignment being influenced by the ascending aorta. The first step in aligning two geometries was to identify the geometry with the smallest number of vertices. Then, for each point in this geometry, the closest point in the other geometry was selected, resulting in a number of point pairs. In the second step, these point pairs were aligned by a Procrustes analysis[34] which determined the optimal translation, rotation and scaling by minimizing the sum of squared errors between the point pairs. The resulting transformation was applied to the target geometry to align it with the template. Steps one and two were repeated until convergence. After alignment, the centerlines of the geometries were extracted using VMTK.[35] The geometries were cut proximally 0.5 cm below the valve annulus and distally, 3 cm above the STJ, perpendicular to the centerline. For 37 geometries, the LVOT was too short. For those geometries, the LVOT was extruded by a straight tube. Lastly, the meshes were made coarser (4300 vertices) to reduce computational time during the computation of the template and initial momenta.

2.4 | Extract geometrical features

For purposes that will be discussed in Sections 2.6 and 2.7, eight anatomical features were automatically extracted from the aortic valve geometries: the diameters of the LVOT, annulus, sinuses Valsalva, and STJ, the aortic valve area
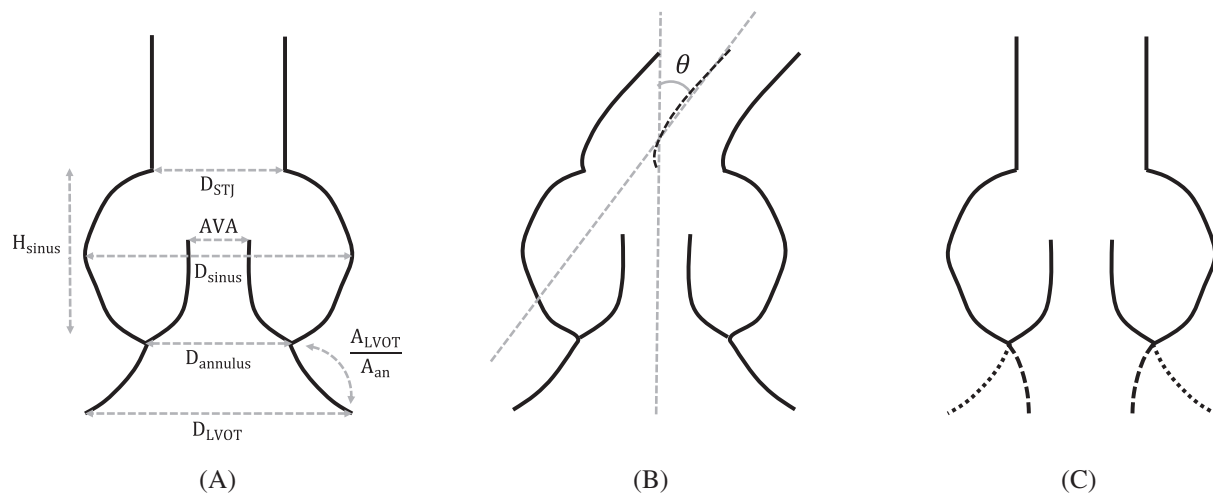
FIGURE 3 Schematic overview of the eight anatomical features: (A) STJ diameter (D_{STJ}), AVA, sinus of Valsalva diameter (D_{sinus}), annulus diameter ($D_{annulus}$), LVOT diameter (D_{LVOT}), sinus of Valsalva height (H_{sinus}), and the ratio between LVOT and annulus area (A_{LVOT}/A_{an}). (B) The angle between the LVOT and the ascending aorta (θ). (C) The shape of the LVOT, either converging (dashed line) or diverging (dotted line).

(AVA), the ratio between the LVOT and annulus area, the sinuses of Valsalva height (Figure 3A), and the angle between the LVOT and the ascending aorta (Figure 3B). To determine the LVOT, annulus and STJ diameter respectively, the geometry was divided into 0.5 mm slices by planes parallel to the LVOT cutting plane. The cross-section of the first plane with the geometry was the LVOT area. The annulus area was measured just before the plane of insertion of the aortic cusps. The insertion plane could be easily detected automatically since it is the first plane that resulted in a cross-section consisting of more than one piece. To compute the STJ area, the cross-sections with all the planes after the leaflets were determined. The cross-section with the smallest area corresponds to the STJ area. To determine the area of the sinus of Valsalva all vertices between the annulus and STJ plane were projected onto the annulus plane. The outline of the resulting point cloud defined the sinus area. The diameters were extracted from the measured areas by assuming circular cross-sectional areas. To determine the AVA, random points were generated within the outline surrounding the sinus of Valsalva. Imaginary beams, perpendicular to the LVOT plane, were sent from these points through the geometry. The beams passing through the geometry without intersecting formed the AVA. The angle between the LVOT and the ascending aorta was calculated by extracting the centerlines of the LVOT (from the annulus to the proximal end of the geometry) and the ascending aorta (from the STJ to the distal end of the geometry), fitting a straight line through the centerlines and computing the angle between those lines using the inner product.

2.5 | Application of the statistical shape modeling framework

In Section 2.1 the non-parametric SSM framework was explained. In this section, it will be discussed how it was applied to the 97 aortic valve geometries. First, the template and the patient-specific initial momenta were computed in Deformetrica.[30] To be able to do the computation, three preparation steps had to be completed first.

The first preparation step was to obtain an initial template geometry that was used in the very first iteration of the two-step minimization approach to estimate the initial momenta and the final template geometry. As an initial template, the mean geometry of another data set of aortic valve geometries, that was used in the study of Hoeijmakers et al.,[8] was used. The original LVOT and ascending aorta were extended with a straight and a slightly curved tube, respectively (Figure 4).

The second step was to find optimal values for the two parameters mentioned in Section 2.1: λ_V, the kernel width of K_V, and λ_W, the kernel width of K_W. Kernel width λ_V, that is, the stiffness parameter, influences the stiffness of the deformations. For high values of λ_V the deformations are more global ("stiffer"), whereas for low values the deformations are more local ("less stiff"). Kernel width λ_W, that is, the resolution parameter, influences the level of detail that is
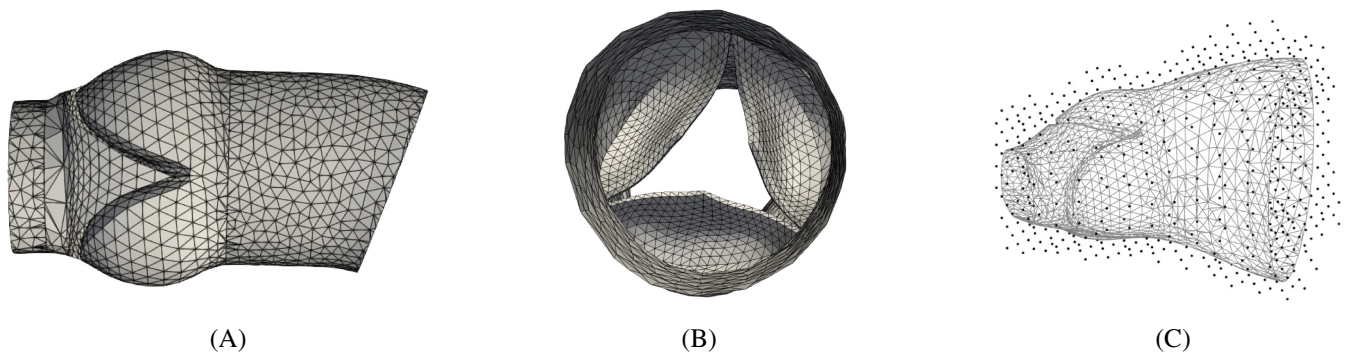
FIGURE 4 Initial template geometry from side view (A) and top view (B), and the initial control points surrounding one of the target geometries (C).

captured in the Varifold distance between two geometries. The higher λ_W, the more detail is captured. To find suitable settings for this application, parameter optimization was performed. The template and initial momenta were computed with different combinations of λ_V and λ_W, varying between 5 and 15 mm, and between 3 and 10 mm respectively. The optimal values of λ_V and λ_W were assessed by using only two aortic valve geometries of the data set: the one with the smallest and the one with the largest aortic valve area. These two geometries were chosen because it is important that the template can deform to these two extremes regarding valve opening. Only two geometries were used to increase the speed of the optimization. The quality of the resulting deformations was expressed in the maximum surface distance between the target geometry and the deformed template geometry. The combination that resulted in the smallest distance was $\lambda_V = 5$ mm, and $\lambda_W = 8$ mm, and was therefore used to calculate the momenta for all geometries and the final template geometry.

The last preparation step was to initialize the control points. To this end, a box filled with points at a distance of λ_V from each other surrounded all the geometries. Then, the points that were at a distance greater than 10 mm from any geometry were removed. The resulting initial control points are visualized in Figure 4C.

The initial control points, the initial template, and the selected values for λ_V and λ_W were used to compute the final template geometry and the patient-specific initial momenta in Deformetrica. To prepare the template for CFD and FSI simulations, small holes and distorted elements in the proximity of the STJ and the annulus were removed. Subsequently, the template was remeshed to obtain suitable surface mesh density for simulations. The mesh consisted of 2934 vertices and 5685 triangles, and the edge size ranged from 0.2 to 1.9 mm. Then, the 97 target geometries were approximated by applying the resulting momenta to the template. Next, the shape modes and shape coefficients of the resulting momenta were computed in Matlab 2021a. To approximate the patient-specific momenta sets, according to Equation (13), a finite number of shape modes N_m was used. Since the geometries were all scaled for alignment (Section 2.3), the vertices of the approximated geometries had to be multiplied by the corresponding scaling factors s to obtain the original size of the geometries. Thus, each target geometry was described by $N_m + 1$ parameters, that is, N_m shape coefficients, and one scaling factor s.

2.6 | Data-driven generation of synthetic geometries

To generate synthetic aortic valve geometries, new combinations of shape coefficients $\alpha_1 ... \alpha_{N_m}$ and scaling factors (s) were obtained by random sampling from an inferred distribution.[28] This distribution was obtained from the shape coefficients and scaling factors of the real patient data set. Each patient-specific geometry was represented by a point in the $N_m + 1$-dimensional space (N_m shape coefficients and one scaling factor). The distribution was mapped to a uniform distribution between 0 and 1 with a cumulative distribution function. To ensure that the correlations between the parameters were accurately captured, a Gaussian copula[36] was fitted to the data, which allowed for modeling the relationship between the parameters while being independent of their marginal distributions. This approach was necessary since the individual parameters were not all normally distributed. New points were randomly sampled from this copula distribution. Lastly, these new points, that is, the synthetic data, were mapped back to the original distribution with an inverse cumulative distribution. The synthetic momenta were reconstructed using Equation (13), and the synthetic geometries were created by deforming the final template (Figure 6) with these momenta

(Equation (1)). In this way, 500 synthetic aortic valve geometries were generated. To validate whether the generated virtual cohort of synthetic aortic valve geometries was realistic, it was compared to the real patient cohort, based on the eight anatomical features described in Section 2.4. The Spearman correlation coefficients between all features were determined and the correlations within the two cohorts were compared. Thereafter, a non-parametric multivariate ANOVA test was done to assess if the 8-dimensional distribution of the synthetic data differed significantly from the distribution of the real data. This test was done using the R statistical software package "non-parametric comparison of multivariate samples (npmv)".[37] Moreover, we performed a literature study to be able to compare the distributions (mean $\pm$ standard deviation) for the eight features we found in our real and virtual cohorts, with distributions found in the literature.

2.7 | Clinically-driven generation of synthetic geometries

In addition to the data-driven approach, a clinically-driven approach[17] was used to generate aortic valve geometries with certain clinically relevant anatomical requirements. We focused on the two anatomical features that strongly influence the occurrence of procedural complications after TAVI (Section 1). The first feature is the LVOT shape, either converging or diverging (Figure 3C), thereby increasing the chance of conduction problems, and paravalvular leakage respectively. The second feature is the angle between LVOT and the ascending aorta (Figure 3B), which is also correlated with the degree of paravalvular leakage. Therefore, an LVOT shape filter and an angle filter were designed and implemented in the workflow to generate synthetic geometries.

To obtain the LVOT shape filter, the annulus and LVOT diameter were extracted from the 500 synthetic geometries as explained in Section 2.4. The difference between the two diameters was calculated. The LVOTs were considered to be converging for a positive difference, and diverging for a negative difference (Figure 3C). Then, a logistic regression model was fitted to the data, with $\alpha_1 ... \alpha_{N_m}$, s as predictors, and the LVOT shape as the dependent variable. After sampling a new point from the copula distribution, the probability to obtain a converging geometry was calculated by substituting the sampled parameters in the model. To generate converging geometries, the set of parameters was accepted if the probability was 0.9, otherwise, it was discarded. To generate diverging geometries, the set of predictors was accepted if the probability was below 0.1.

The angle filter was obtained by extracting the angle between LVOT and the ascending aorta, as explained in Section 2.4, from the 500 synthetic geometries. A linear regression model with linear and quadratic terms was fitted to the data with $\alpha_1 ... \alpha_{N_m}$, s the predictors, and the angle between the LVOT and the ascending aorta the dependent variable. Using this model, it was possible to estimate the angle and its corresponding 5% confidence interval for a newly sampled point. If the confidence interval's lower and upper bounds fell within the desired angle range, the point was accepted; otherwise, it was rejected.

The data-driven and clinically-driven approaches to generate synthetic aortic valve geometries were combined in one virtual cohort generator, implemented in Matlab 2021a code. The user of the virtual cohort generator can set the desired number of synthetic geometries, the preferred LVOT shape, and the desired angle range. A flowchart summarizing the process of generating synthetic geometries within the virtual cohort generator is visible in Figure 5. The virtual cohort generator was used to create several virtual cohorts, including 500 purely data-driven synthetic geometries (Section 2.6), 500 geometries with converging LVOT, 500 geometries with diverging LVOT, and 500 geometries for each of the following angle ranges: a small angle range (0°–20°), medium angle range (20°–40°), and a large angle range (>40°).
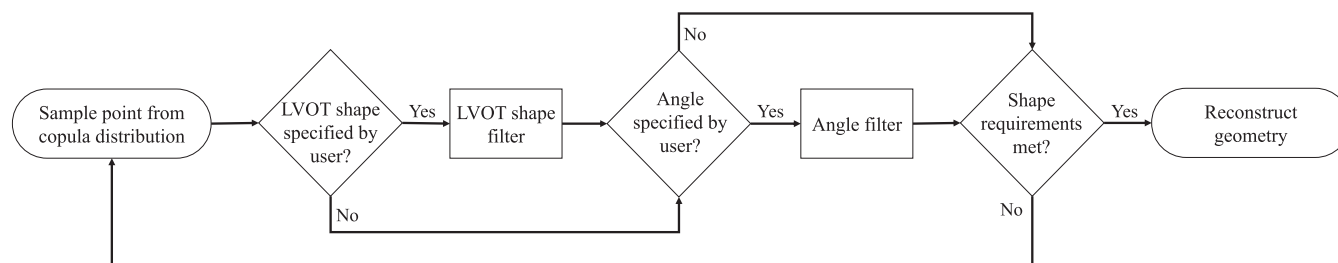

FIGURE 5 Flowchart showing how one synthetic geometry is generated within the virtual cohort generator.

3 | RESULTS

In the non-parametric SSM framework, the 97 aortic valve geometries were approximated in two steps. Therefore, the accuracy of the approximations is evaluated in Section 3.1. Then, the plausibility of the generated data-driven and clinically-driven virtual cohorts of synthetic geometries is assessed in Sections 3.2 and 3.3.

3.1 | Application of the statistical shape modeling framework

A set of momenta for each of the 97 real aortic valve geometries, together with a final template were obtained with Deformetrica. The final template is shown in Figure 6. The 97 geometries were approximated by applying the resulting momenta to the template. The accuracy of the approximations was evaluated by computing the distance from each vertex of the approximated geometry to the target geometry. We considered a distance as being acceptable when it was below 0.47 mm, which is the accuracy of the algorithm with which the geometries were reconstructed from CT images.[32] For three geometries the surface distances, normalized with respect to 0.47 mm, are visualized in Figure 8 (approximation 1). This group of three geometries has been selected for visualization because it contains the geometry with the smallest AVA (Figure 8A), the geometry with the largest AVA (Figure 8B), and the geometry with the largest sum of surface distances (Figure 8C). The distances at all vertices of the approximated geometry are summarized in the box plots. Overall, the target geometries and their approximations showed high similarity with errors reported as median (first quartile-third quartile) of 0.09 mm (0.04–0.18 mm), 0.11 mm (0.05–0.22 mm), and 0.12 mm (0.05–0.26 mm) for the cases in Figure 8A–C respectively. The highest errors were found at the valve leaflets and equaled 1.50, 1.43, and 1.87 mm for the cases in Figure 8A–C, respectively. However, even in the approximation with the largest sum of absolute errors (Figure 8C), 88% of all surface distances were still below 0.47 mm. Therefore, we consider the approximations to be sufficiently accurate to represent the target geometries.

Principal component analysis was applied to the 97 sets of momenta. The momenta sets were approximated with an increasing number of shape modes and applied to the template to approximate the target geometries. The eight anatomical features were computed for all approximations and compared to the anatomical features of the target geometries. Moreover, the surface distance, averaged over all vertices and all geometries was computed for an increasing number of shape modes. Figure 7 shows how the errors between approximated and target geometries decrease with the increasing number of shape modes used for the approximation. Note, that the sharp drops in the sinus height error plot (Figure 7) are due to the discrete nature of the measurement method, which uses 0.5 mm slices to measure the distance between the annulus and the STJ, that is, the sinus height (Section 2.4). For most of the anatomical features, the mean error seems to be stabilized after 32 shape modes. Furthermore, the mean surface distance is below 0.47 mm after 32 modes. These 32 shape modes capture 91% of the shape variation within the data set. The normalized surface distance between
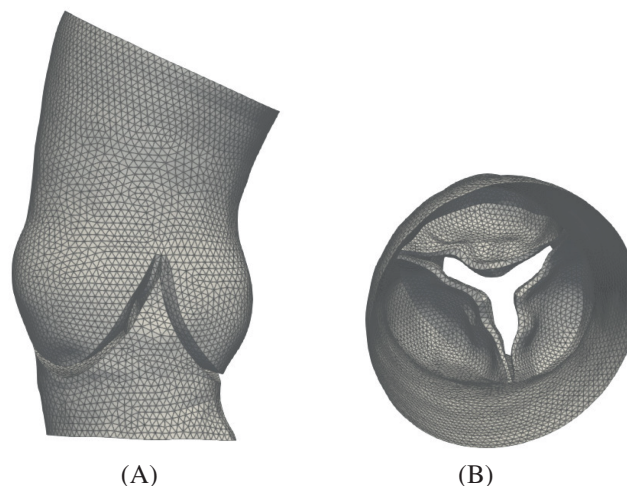

(A)   (B)

FIGURE 6   Final template geometry obtained through atlas computation in Deformetrica from side view (A) and top view (B).
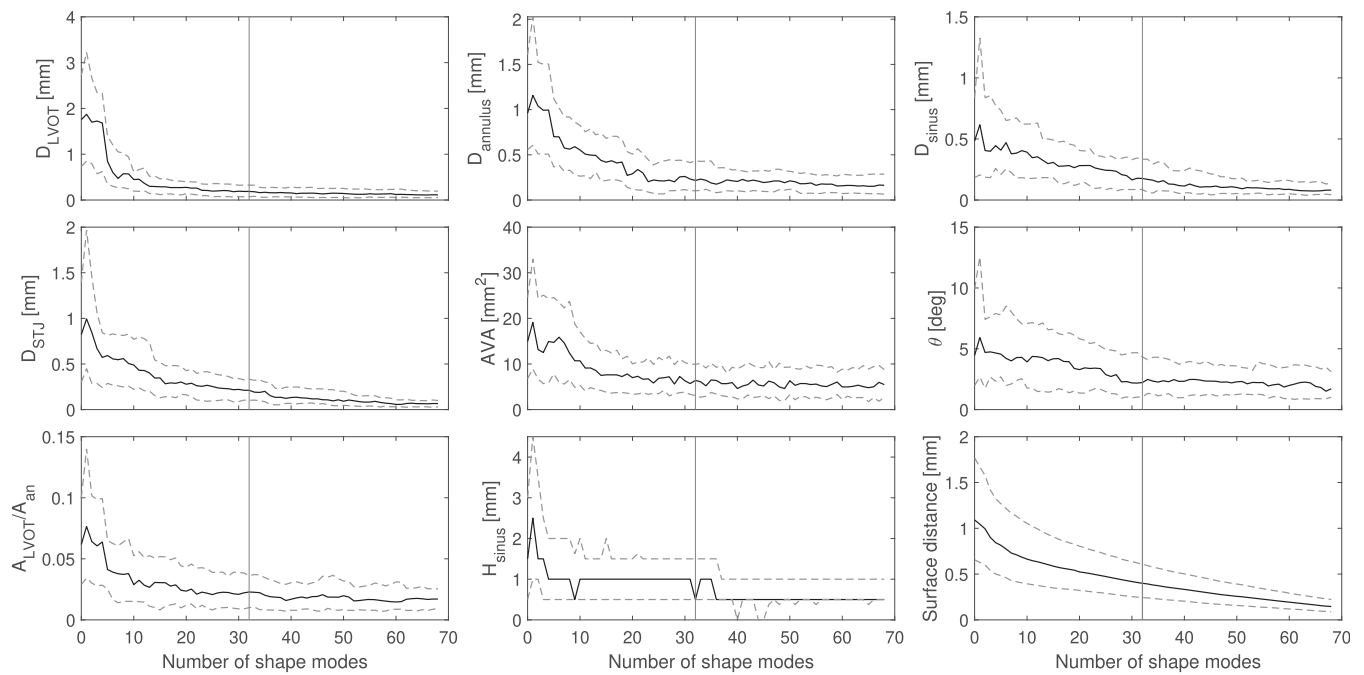
FIGURE 7 Median (black solid line), first and third quartile (gray dashed lines) of error in eight geometric features, and the surface distances as a function of the number of shape modes included in the approximation of the geometry. The largest error reduction is reached within 32 shape modes (indicated with black vertical line). The abbreviations D_{LVOT}, D_{an}, D_{sinus}, D_{STJ}, AVA, θ, A_{LVOT}/A_{an}, and H_{sinus} stand for the diameters of the LVOT, the annulus, the sinuses of Valsalva, and the STJ, the aortic valve area, the angle between the LVOT and the ascending aorta, the ratio between LVOT and annulus area, and the sinus height respectively.

the geometry reconstructed with 32 shape modes and the target geometry is visualized in Figure 8 (approximation 2). It is visible that the overall characteristics of the target geometries are still captured in the reconstruction with 32 modes. The errors, reported as median (first quartile-third quartile) were 0.26 mm (0.13–0.45 mm), 0.28 mm (0.13–0.46 mm), and 0.26 mm (0.12–0.49) and the maximum distances were 1.63, 1.92, and 1.88 mm for the cases in Figure 8A–C, respectively. Approximately 74% of the surface distances were still below the reconstruction accuracy of 0.47 mm. Therefore, it was decided to generate synthetic geometries with 32 shape modes.

3.2 | Data-driven generation of synthetic geometries

The first eight geometries from the virtual cohort, are shown in Figure 9. The individual geometries appear realistic and physiologically plausible.

To evaluate the compatibility of the virtual cohort with the real cohort, the eight-dimensional distributions of features of both cohorts were compared by means of a non-parametric multivariate ANOVA test. The test revealed that the 500 synthetic geometries do not differ significantly from the real geometries ($p = 0.86 > 0.05$). To visualize the similarity between the features of both cohorts, scatter plots of all combinations of features are depicted in Figure 10. It shows that the trends and distributions of anatomical features in both cohorts were similar. The similarity is also confirmed by the Spearman correlation coefficient matrices in Figure 11, with a mean difference in correlation of 0.03. Thus, based on the correlations and the ANOVA test, we consider the synthetic cohort as being physiologically plausible and realistic.

Finally, the anatomical feature values of both the synthetic and real cohorts were compared with values reported in the literature (Table 1). The diameters of the LVOT, annulus, sinus and STJ were compared with area-derived diameters from Buellesfeld et al.,[38] which were determined based on CT images. Additionally, we compared AVA values with the results from the same study. Both the diameters and the AVA values were in good agreement. However, the values in Buellesfeld et al. were approximately 1 mm lower on average. To assess the angles between the LVOT and ascending aorta, we referred to the values presented by Sherif et al.,[6] who derived their measurements from 2D angiography images. In contrast to their findings, our study's angles were on average seven degrees higher. The LVOT over annulus
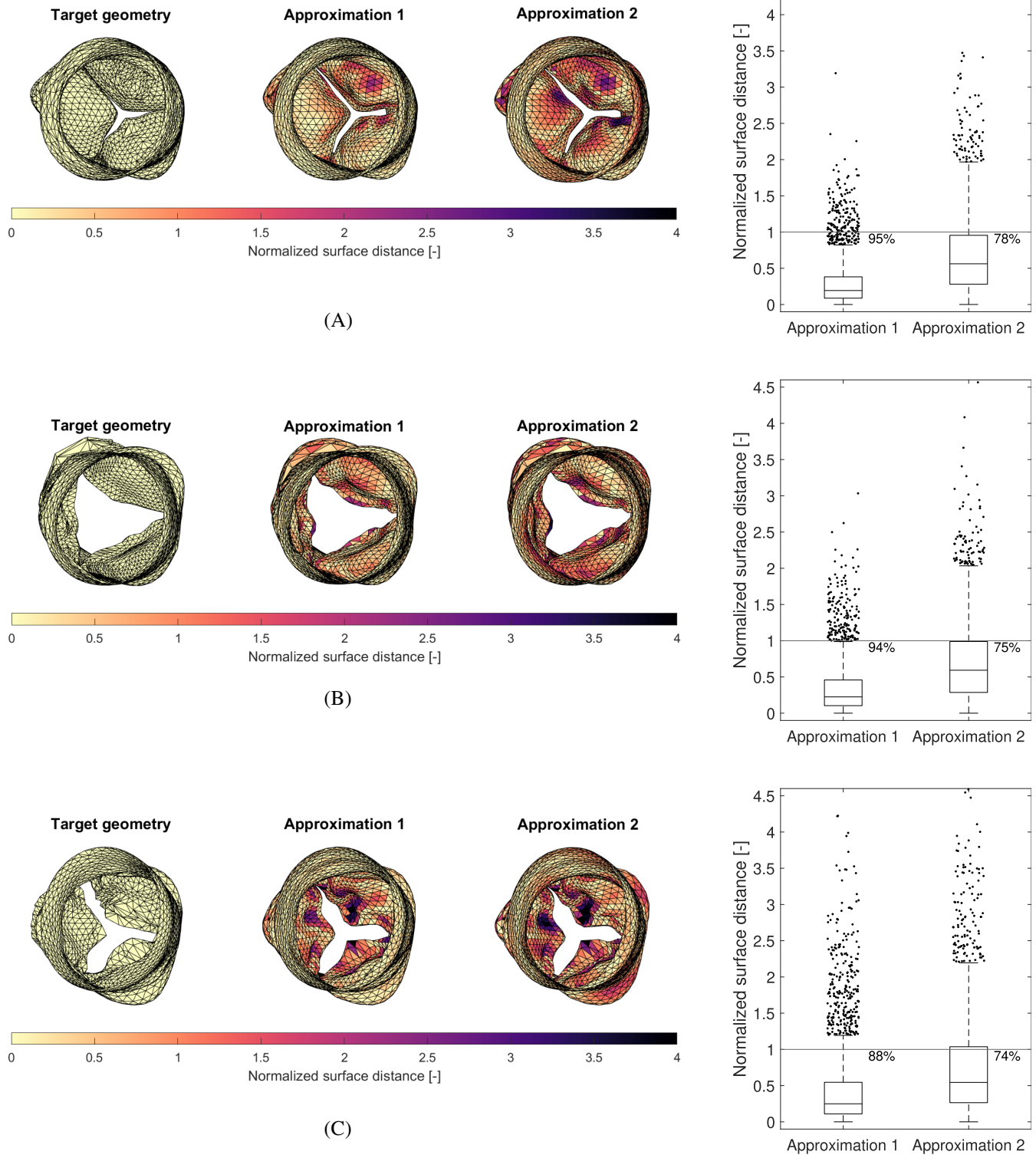
FIGURE 8 Target geometry (first column), approximated geometry after applying momenta to the template (approximation 1, second column), and geometry approximated with 32 shape modes (approximation 2, third column) for the geometry with the smallest AVA (A), the geometry with the largest AVA (B), and the geometry with the largest sum of surface distances (C). The colors on the geometries indicate the distance at each vertex to the target geometry, normalized with respect to the reference value, that is, the reconstruction accuracy of 0.47 mm.[32] Box plots showing all surface distances are shown in the fourth column, with the percentile of distances below 0.47 mm indicated.
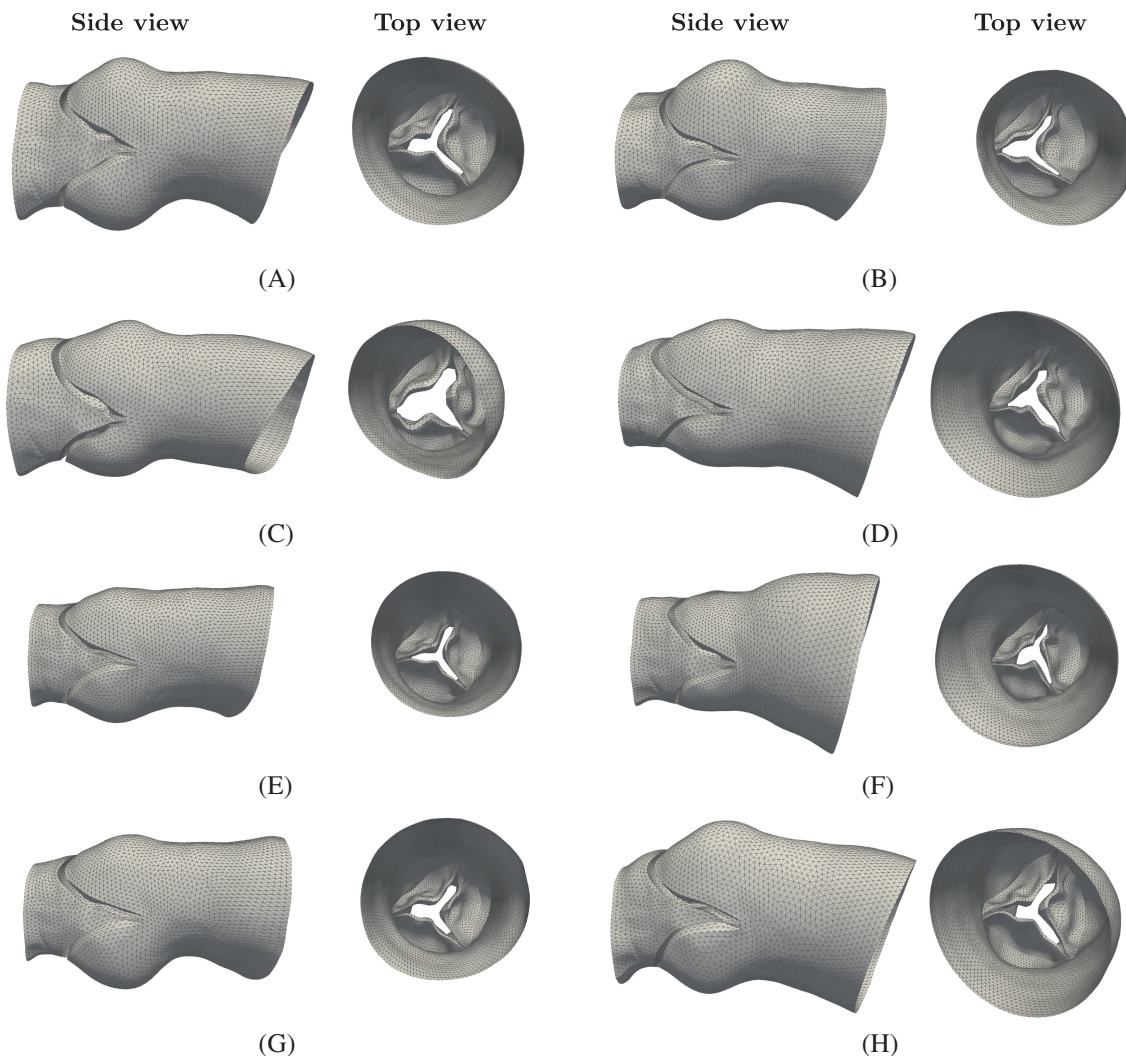
FIGURE 9 The first eight geometries ((A)–(H)) of the data-driven virtual cohort from the side view (left) and top view (right).

area values were in good agreement with the values found in Ishizu et al.[4] Lastly, the sinus heights found by Delgado et al.[39] were on average 3 mm lower than the sinus heights in our study. Overall, there are no extreme differences between the anatomical feature values in our study and those reported in the literature, which reinforces the plausibility of our synthetic cohort.

3.3 | Clinically-driven generation of synthetic geometries

The LVOT shape filter and angle between LVOT and ascending aorta filter were derived from the data set of 500 synthetic geometries and used to generate five virtual cohorts with 500 geometries each, with the following clinical requirements: converging LVOT, diverging LVOT, angles between LVOT and ascending aorta in range $0°–20°$, $20°–40°$, and greater than $40°$. The effectiveness of these filters was assessed and summarized in Table 2. The sensitivity of the filters was evaluated by calculating the ratio between the number of geometries that met the desired clinical requirements and the total number of geometries in that virtual cohort. The sensitivity of the filters was found to be high, with values exceeding 0.89. The sampling efficiency of the filters was also evaluated by calculating the ratio between the number of samples accepted by the filter and the number of discarded samples. The sampling efficiencies were relatively low and strongly dependent on the percentage of geometries with the corresponding clinical requirement in the data-driven virtual cohort of 500 synthetic geometries. Despite the low sampling efficiencies, the computational times required to generate 500 synthetic geometries with these clinical requirements remained within reasonable limits, taking no longer than 70 min. Figure 12 shows an example geometry for each of the five clinically-driven cohorts. For the cohorts with a
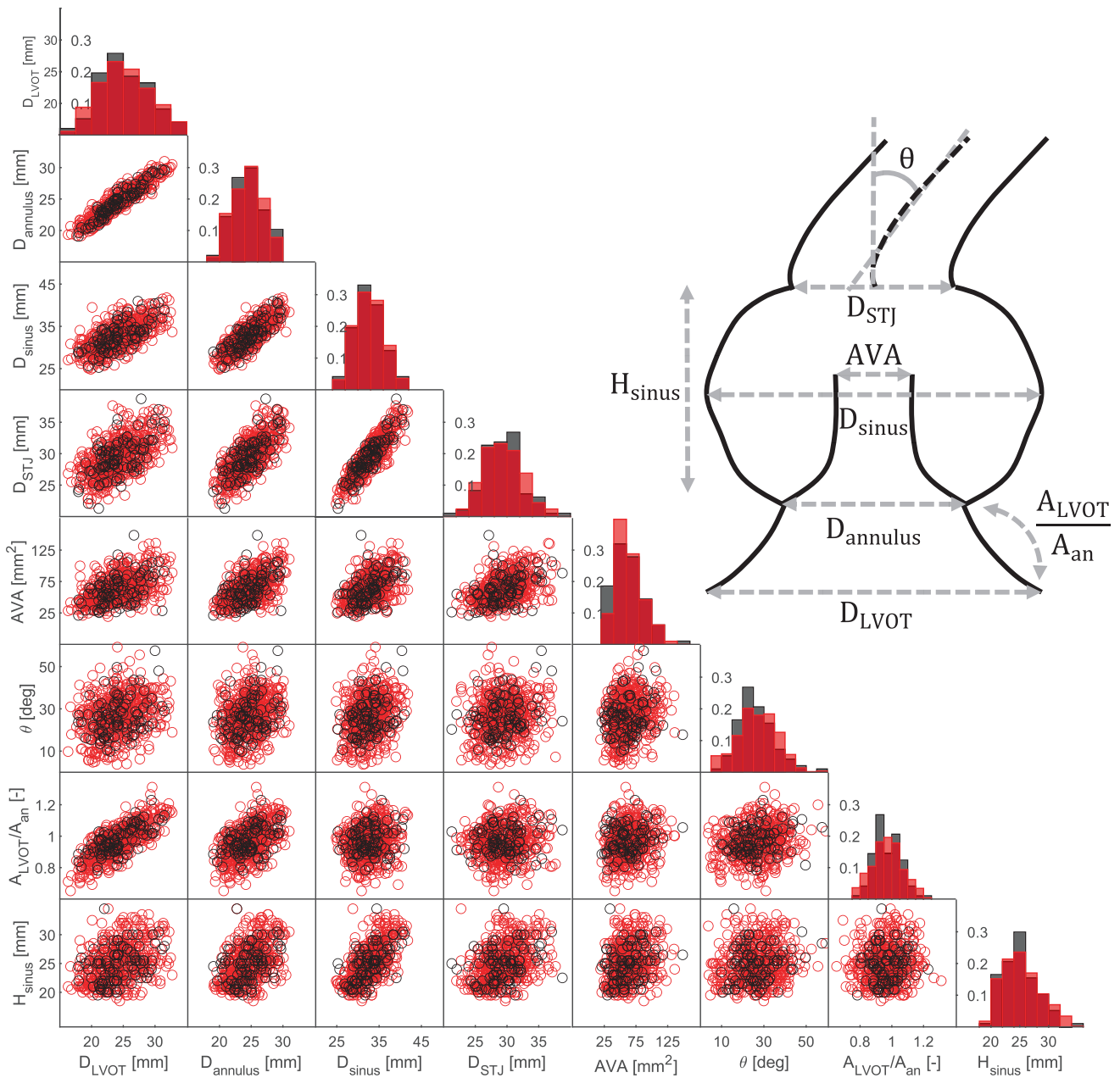
FIGURE 10 Scatter plots of all combinations of anatomical features for both the real (black) and synthetic geometries (red). On the diagonal, the distributions of the individual features are shown in histograms. The abbreviations D_{LVOT}, $D_{annulus}$, D_{sinus}, D_{STJ}, AVA, θ, A_{LVOT}/A_{an}, and H_{sinus} stand for the diameters of the LVOT, the annulus, the sinuses of Valsalva, and the STJ, the aortic valve area, the angle between the LVOT and the ascending aorta, the ratio between LVOT and annulus area, and the sinus height, respectively. The features are indicated in the schematic in the upper right corner.

specific angle range as a clinical requirement, the geometry at the median of the resulting angles is shown. For the cohorts with a converging and diverging LVOT shape as clinical requirements, the geometry with the largest and the geometry with the smallest LVOT over the annulus area are shown respectively. These geometries were chosen to provide a clear visualization of the difference between a converging and diverging LVOT.

4 | DISCUSSION

The goal of this study was to develop a framework to generate synthetic aortic valve stenosis geometries. The framework had to satisfy two main requirements. First, the framework had to be capable of generating synthetic geometries
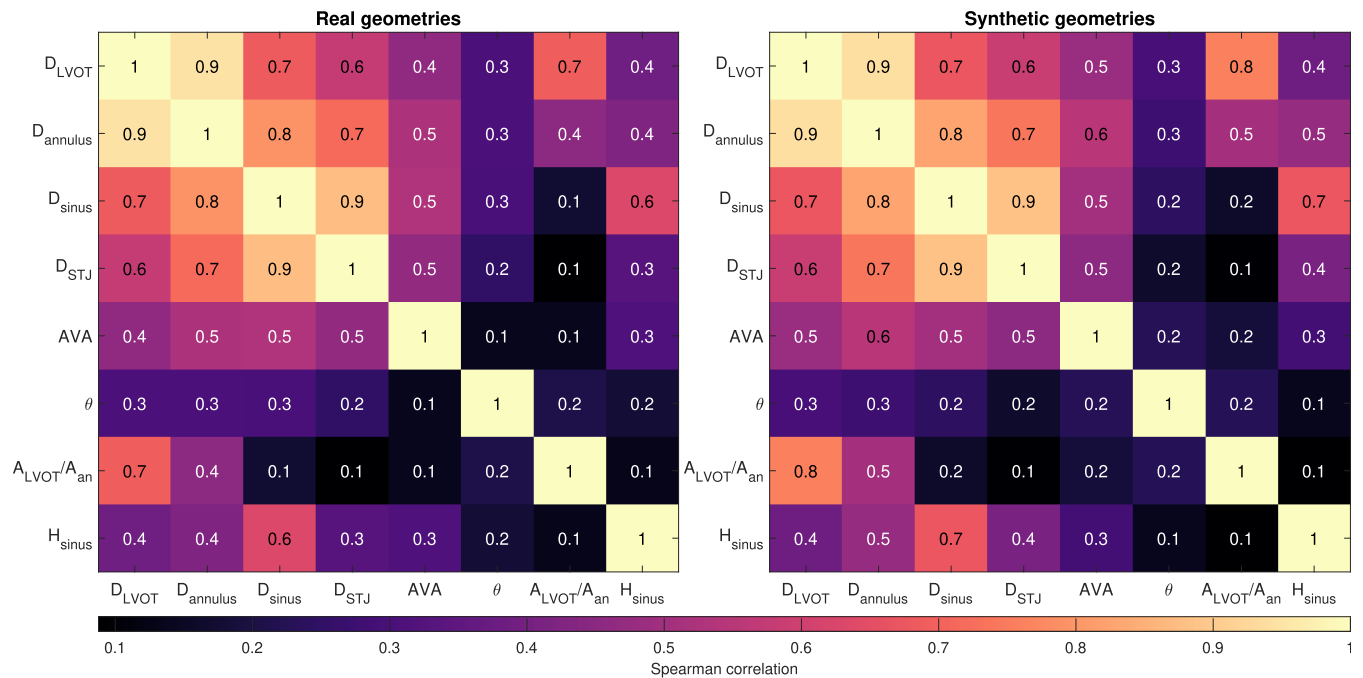
FIGURE 11 Spearman correlations describing the relations between the anatomical features shown in Figure 10 for the real geometries (left) and the synthetic geometries (right).

TABLE 1 Anatomical feature values measured in the synthetic and real geometries and those reported in the literature.

	Synthetic geometries	Real geometries	Literature
LVOT diameter (mm)	24.2 ± 3.3	24.2 ± 3.0	22.9 ± 3.2[38]
Annulus diameter (mm)	24.6 ± 2.5	24.5 ± 2.4	23.3 ± 2.2[38]
Sinus diameter (mm)	32.7 ± 3.4	32.6 ± 3.3	31.4 ± 3.0[38]
STJ diameter (mm)	29.4 ± 3.0	29.4 ± 3.0	28.0 ± 3.0[38]
AVA (cm^2)	0.64 ± 0.2	0.62 ± 0.2	0.60 ± 0.2[38]
θ (°)	26.7 ± 9.6	26.4 ± 8.9	19.7 ± 7.9[6]
LVOT area/annulus area (−)	0.97 ± 0.10	0.97 ± 0.09	0.96 ± 0.06[4]
Sinus height (mm)	25.1 ± 3.3	24.7 ± 3.0	22.2 ± 3.2[39]

Note: The values are reported as mean $\pm$ standard deviation.

TABLE 2 Effectiveness of the filters in generating geometries with five clinical requirements, expressed in filter sensitivity, sampling efficiency, and required time to generate 500 geometries. The occurrence of geometries with a certain clinical requirement in the unfiltered virtual cohort is shown in the fifth column.

Filter	Clinical requirement	Filter sensitivity	Sampling efficiency	Occurrence in unfiltered virtual cohort	Time generating one geometry (ms)
LVOT shape	Converging	0.98	0.59	62%	36
	Diverging	0.90	0.41	38%	60
Angle between LVOT and ascending aorta	0°–20°	0.91	0.15	23%	52
	20°–40°	0.93	0.52	69%	56
	>40°	0.89	0.04	8%	140

(A) Angle between 0° and 20°

(B) Angle between 20° and 40°

(C) Angle > 40°

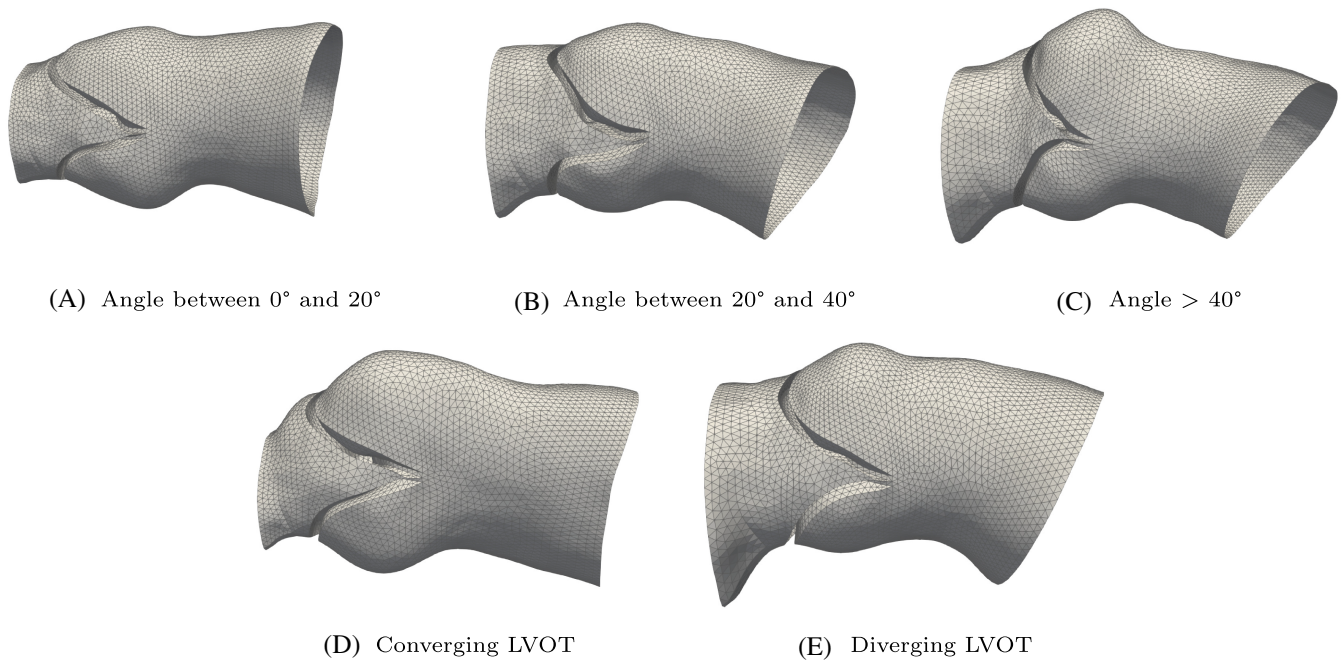(D) Converging LVOT

(E) Diverging LVOT

FIGURE 12 Synthetic geometries from clinically-driven cohorts. For the cohorts with specific ranges of angles between the LVOT and the ascending aorta, the median angle geometry is shown (A–C). For the converging LVOT cohort, the geometry with the smallest ratio between LVOT and annulus area is shown (D) and for the diverging LVOT cohort, the geometry with the largest ratio between LVOT and annulus area (C) is shown.

that are anatomically plausible using a data-driven approach. Second, it had to enable the selection of specific, clinically relevant anatomical features using a clinically-driven approach. We focused on two anatomical features: the LVOT shape (either converging or diverging), and specific ranges of angles between the LVOT and the ascending aorta.

With the use of non-parametric SSM, we managed to generate synthetic aortic valve geometries that are comparable with a real-patient cohort and therefore, are considered as being physiologically plausible. The plausibility of our synthetic geometries was reinforced by comparison with literature. In general, the values of the anatomical features were in good agreement with those reported in the literature. The small differences between the values found in the literature and our values can be attributed to differences in measurement strategies. Furthermore, by using clinically-driven acceptance criteria, we were able to select specific anatomical features: converging and diverging LVOTs and three different ranges of angles, with a sensitivity of around 90%. This novel combination of non-parametric SSM and clinically-driven acceptance criteria enables us to create virtual cohorts of aortic valve stenosis geometries with distinct anatomical features, such as a cohort comprising exclusively converging geometries. We only focused on two features as an example. However, the proposed framework could also be used to select other features, such as aortic valve area.

Furthermore, the template used to generate the synthetic geometries is very suitable for finite element, CFD or FSI simulations, since it does not contain any holes or distorted elements. Moreover, its refinement level aligns with, or even surpasses the refinement reported by Hoeijmakers et al.[7] As a result, the synthetic geometries are well-suited for simulations and require minimal pre-processing steps before integration into a simulation pipeline.

Therefore, the virtual cohort generator that was developed has the potential to be used by TAVI manufacturers to test their new or improved devices. They can generate any number of geometries with the desired anatomical feature and perform device effect simulations via for example finite element, CFD or FSI simulations on the geometries. For example, the synthetic geometries could be used to investigate the occurrence of paravalvular leakage or conduction problems in new devices, using simulation strategies similar to the studies of Jaegere et al.[40] and Rocatello et al.[16] respectively. Synthetic geometries can also be helpful to train machine learning surrogates for high-fidelity models such as CFD and FSI models.[27] To train machine learning models, very large data sets are required, especially for high-fidelity model surrogates with a large number of parameters. For this purpose, real patient data sets are often too small. Therefore, generating synthetic geometries with the generator developed in this study could be a good alternative. In

the next two paragraphs, the differences and improvements of our work with respect to other studies in this field are discussed.

The current study is building on the previous work of Hoeijmakers et al.[8] in which, synthetic aortic valve geometries were generated as well. We made two main improvements compared to the previous study. The first improvement is that we used non-parametric instead of conventional SSM. A non-parametric SSM framework does not require iso-topology between input geometries to build the SSM. The use of this framework makes our framework wider applicable, especially for complex shapes where obtaining iso-topologic meshes is challenging. For example, our framework could also be applied to generate synthetic pulmonary artery geometries for in silico trials of pulmonary artery pressure sensors. The second improvement is that we introduced a sampling strategy to make sure that the distribution of parameters in the virtual cohort was equal to that of the real cohort.

Compared to other studies focusing on the generation of synthetic aorta geometries, such as those of Romero et al.[17] and Pajaziti et al.,[27] our study distinguishes itself by the sampling strategy that is used. In this study, sampling from a copula distribution was used. Copulas disconnect marginal distributions from the dependence structure between variables. When using copulas the correlation between variables is modeled without the need to assume specific distributions of each variable.[36] Therefore, we believe that the copula distribution is a sufficiently accurate estimation of the real cohort's distribution. Points sampled from this distribution are by definition compatible with the real cohort and therefore it was not necessary to use acceptance criteria as was done in the study of Romero et al.[17]

Our approach also has some limitations. First of all, the downside of non-parametric SSM is that geometries are approximated in two steps instead of in one step. First, the geometries are approximated by deformation vectors that deform the template to the target geometries, introducing a certain level of error. Then, the momenta are approximated with PCA, adding another source of error. In contrast, conventional SSM approximates geometries solely with PCA, without introducing the error from the first step. Despite the two approximation steps, the biggest part (around 75%) of the surface distances between approximated and target geometries were below the reconstruction accuracy of 0.47 mm that was reported in Reference 32. Furthermore, for the purpose of generating synthetic data, the approximations are sufficiently accurate, since we do not aim to reproduce each patient in the real cohort. To validate if the introduced approximation errors influence physical parameters such as pressure drops, flows and wall shear stresses, CFD or FSI simulations should be performed on both the original and the approximated geometries.

The second limitation is that a relatively short segment of the ascending aorta was included in the geometries, compared to the studies of Romero et al.[17] and Pajaziti et al.[27] There is much variation in the shape of the aortic arch and descending aorta among different patients. Including a longer part of the aorta in the geometries would lead to shape modes describing this variation dominating shape modes describing variation in the valve itself. Therefore we chose to focus on a shorter segment of the aorta. Nonetheless, the segment we included is still sufficiently long to enable device effect simulations, and we do not believe that this relatively shorter segment poses a significant problem for our framework's target application.

Furthermore, the sample efficiencies for the specific anatomical features during clinically-driven geometry generation were low compared to the sample efficiencies found in Romero et al.[17] This difference is due to the fact that we always sampled from the same distribution, despite the desired anatomical features, whereas Romero et al.[17] sampled from different distributions for different anatomical features. The disadvantage of a low sample efficiency is that it takes longer to generate the desired number of geometries. The clinical requirement with the lowest sample efficiency, and thus the highest computational time was the angle greater than $40°$. However, when compared to the time it takes to acquire one CT scan to obtain a patient-specific geometry, the computational time of 144 ms is still very reasonable.

Additionally, some clinically relevant anatomical features that are crucial for TAVI were absent in this study. First of all, we did not include calcifications in the aortic valve geometries. According to the Euro Heart Survey on Valvular Heart Disease, calcified leaflets accounted for approximately 82% of the patients with aortic stenosis included in the study.[41] Calcifications significantly increase the stiffness of the leaflets, which strongly influences device implantation. Thus, it is important to take calcifications into account. Therefore, a future extension to our work is adding calcifications to the synthetic geometries. Moreover, we did not include the bicuspid aortic valve, which is present in 25% of the population of patients aged 80 years or older who undergo valve replacement.[42] A bicuspid aortic valve only has two leaflets instead of three and often leads to stenosis. Bicuspid valves were not included in the data set used in our study but should be considered in future studies by using another data set that includes this valve type. Eventually, the position of the coronary arteries could be included in the synthetic geometries as it influences the occurrence of another TAVI complication: coronary artery occlusion.[43] During coronary artery occlusion, the TAVI valve blocks the coronary artery, resulting in a lack of blood flow to the heart muscle. However, this complication is rare (incidence <1%), and therefore was not prioritized in this study.

The last limitation is that the validation process only compared synthetic geometries to the real geometries used to build the SSM. It is important to note that the SSM inherently depends on the data set used to build it. Therefore, to establish the plausibility of the synthetic geometries it is crucial to demonstrate that their parameter distributions are comparable with those of an independent data set. To address this issue we plan to conduct cross-validation in the future by comparing the distributions and correlations of the eight anatomical features, extracted from the synthetic data set, with those from a data set of aortic valve geometries, originating from another hospital.

In future research, we plan to build an FSI model applied to synthetic geometries. Using this FSI model we could also filter on physiological plausibility, in order to obtain realistic combinations of geometries, boundary conditions and physiological outputs. Furthermore, using this model in silico TAVI trials could be performed to investigate the difference in paravalvular leakage between geometries with a converging and diverging LVOT, and to investigate the influence of the angle between the LVOT and the ascending aorta. Once we set up a high-fidelity model with which we can perform in silico trials we would like to develop a machine learning surrogate, to be able to perform real-time simulations.

5 | CONCLUSION

In conclusion, this study successfully developed a virtual cohort generator for the generation of synthetic aortic valve stenosis geometries, with the ability to select specific, clinically relevant anatomical features. Our results demonstrate that the generated geometries are anatomically plausible and their parameter distributions are comparable to those of a real-patient cohort, which makes them suitable for in silico TAVI trials. Furthermore, the generated synthetic geometries can be used to train machine learning surrogates for high-fidelity models such as CFD and FSI models. Although this study focused on aortic valve geometries, the proposed framework is more generally applicable, due to the use of non-parametric SSM. Overall, the presented work is a significant step towards implementing in silico trials in the validation chain for cardiovascular implantable devices, to make the whole process from development to the clinical introduction, faster, safer and more efficient.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are openly available in 4TU. Research Data at https://doi.org/10.4121/3f6a3788-96e6-4b81-b37b-f07eeec85965.

ORCID
Sabine Verstraeten https://orcid.org/0000-0002-1011-6627
Martijn Hoeijmakers https://orcid.org/0000-0001-9942-7597
Pim Tonino https://orcid.org/0000-0002-2099-5426
Jan Brüning https://orcid.org/0000-0002-0540-7972
Claudio Capelli https://orcid.org/0000-0002-2071-2596
Frans van de Vosse https://orcid.org/0000-0002-4084-2064
Wouter Huberts https://orcid.org/0000-0002-0779-4785

REFERENCES
1. Henney A, Morley-Fletcher E, Contin M, Viceconti M. In silico clinical trials: How computer simulation will transform the biomedical industry. *Int J Clin Trials*. 2015;9:37-46.
2. Morrison TM, Pathmanathan P, Adwan M, Margerrison E. Advancing regulatory science with computational modeling for medical devices at the fda's office of science and engineering laboratories. *Front Med*. 2018;5:241.
3. Viceconti M, Cobelli C, Haddad T, Himes A, Kovatchev B, Palmer M. In silico assessment of biomedical products: the conundrum of rare but not so rare events in two case studies. *Proc Inst Mech Eng*. 2017;231(5):455-466.

4. Ishizu K, Murakami N, Morinaga T, et al. Impact of tapered-shape left ventricular outflow tract on pacemaker rate after transcatheter aortic valve replacement. *Heart Vessels*. 2022;37:1-11.

5. Condado JF, Corrigan FE III, Lerakis S, et al. Anatomical risk models for paravalvular leak and landing zone complications for balloon-expandable transcatheter aortic valve replacement. *Catheter Cardiovasc Interv*. 2017;90(4):690-700.

6. Sherif MA, Abdel-Wahab M, Stöcker B, et al. Anatomic and procedural predictors of paravalvular aortic regurgitation after implantation of the medtronic corevalve bioprosthesis. *J Am Coll Cardiol*. 2010;56(20):1623-1629.

7. Hoeijmakers M, Huberts W, Rutten M, van de Vosse F. The impact of shape uncertainty on aortic-valve pressure-drop computations. *Int J Numer Method Biomed Eng*. 2021;37(10):e3518.

8. Hoeijmakers M, Waechter-Stehle I, Weese J, Van de Vosse F. Combining statistical shape modeling, cfd, and meta-modeling to approximate the patient-specific pressure-drop across the aortic valve in real-time. *Int J Numer Method Biomed Eng*. 2020;36(10):e3387.

9. Fumagalli I, Polidori R, Renzi F, et al. Fluid-structure interaction analysis of transcatheter aortic valve implantation. *Int J Numer Method Biomed Eng*. 2023;39:e3704.

10. Oks D, Samaniego C, Houzeaux G, Butakoff C, Vázquez M. Fluid–structure interaction analysis of eccentricity and leaflet rigidity on thrombosis biomarkers in bioprosthetic aortic valve replacements. *Int J Numer Method Biomed Eng*. 2022;38(12):e3649.

11. Luraghi G, Migliavacca F, García-González A, et al. On the modeling of patient-specific transcatheter aortic valve replacement: a fluid–structure interaction approach. *Cardiovasc Eng Technol*. 2019;10:437-455.

12. Basri AA, Zuber M, Basri EI, et al. Fluid structure interaction on paravalvular leakage of transcatheter aortic valve implantation related to aortic stenosis: a patient-specific case. *Comput Math Methods Med*. 2020;2020:1-22.

13. Basri AA, Zuber M, Basri EI, et al. Fluid-structure interaction in problems of patient specific transcatheter aortic valve implantation with and without paravalvular leakage complication. *Fluid Dyn Mater Process*. 2021;17(3):531-553.

14. Finotello A, Morganti S, Auricchio F. Finite element analysis of tavi: impact of native aortic root computational modeling strategies on simulation outcomes. *Med Eng Phys*. 2017;47:2-12.

15. Wang Q, Kodali S, Primiano C, Sun W. Simulations of transcatheter aortic valve implantation: implications for aortic root rupture. *Biomech Model Mechanobiol*. 2015;14:29-38.

16. Rocatello G, El Faquir N, De Santis G, et al. Patient-specific computer simulation to elucidate the role of contact pressure in the development of new conduction abnormalities after catheter-based implantation of a self-expanding aortic valve. *Circ Cardiovasc Interv*. 2018;11(2):e005344.

17. Romero P, Lozano M, Martínez-Gil F, et al. Clinically-driven virtual patient cohorts generation: an application to aorta. *Front Physiol*. 2021;12:1375.

18. Vergara C, Viscardi F, Antiga L, Luciani GB. Influence of bicuspid valve geometry on ascending aortic fluid dynamics: a parametric study. *Artif Organs*. 2012;36(4):368-378.

19. Hosny A, Dilley JD, Kelil T, et al. Pre-procedural fit-testing of tavr valves using parametric modeling and 3d printing. *J Cardiovasc Comput Tomogr*. 2019;13(1):21-30.

20. Heimann T, Meinzer H-P. Statistical shape models for 3d medical image segmentation: a review. *Med Image Anal*. 2009;13(4):543-563.

21. Jolliffe IT. *Principal Component Analysis for Special Types of Data*. Springer; 2002.

22. Mansi T, Durrleman S, Bernhardt B, et al. A statistical model of right ventricle in tetralogy of fallot for prediction of remodelling and therapy planning. *Med Image Comput Comput Assist Interv*. 2009;12:214-221.

23. Mansi T, Voigt I, Leonardi B, et al. A statistical model for quantification and prediction of cardiac remodelling: application to tetralogy of fallot. *IEEE Trans Med Imaging*. 2011;30(9):1605-1616.

24. Leonardi B, Taylor AM, Mansi T, et al. Computational modelling of the right ventricle in repaired tetralogy of fallot: can it provide insight into patient treatment? *Eur Heart J Cardiovasc Imaging*. 2013;14(4):381-386.

25. Durrleman S, Prastawa M, Charon N, et al. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. *Neuroimage*. 2014;101:35-49.

26. Bruse JL, McLeod K, Biglino G, et al. 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*. 2016;16:1-19.

27. Pajaziti E, Montalt-Tordera J, Capelli C, et al. Shape-driven deep neural networks for fast acquisition of aortic 3d pressure and velocity flow fields. *PLoS Comput Biol*. 2023;19(4):e1011055.

28. Niederer S, Aboelkassem Y, Cantwell CD, et al. Creation and application of virtual patient cohorts of heart models. *Philos Trans Royal Soc A*. 2020;378(2173):20190558.

29. Durrleman S, Pennec X, Trouvé A, Ayache N. Statistical models of sets of curves and surfaces based on currents. *Med Image Anal*. 2009;13(5):793-808.

30. Bône A, Louis M, Martin B, Durrleman S. Deformetrica 4: an open-source software for statistical shape analysis, pp. 3–13. 2018.

31. Ecabert O, Peters J, Walker MJ, et al. Segmentation of the heart and great vessels in ct images using a model-based adaptation framework. *Med Image Anal*. 2011;15(6):863-876.

32. Weese J, Lungu A, Peters J, Weber FM, Waechter-Stehle I, Hose DR. Cfd-and bernoulli-based pressure drop estimates: a comparison using patient anatomies from heart and aortic valve segmentation of ct images. *Med Phys*. 2017;44(6):2281-2292.

33. Franke B, Brüning J, Yevtushenko P, et al. Computed tomography-based assessment of transvalvular pressure gradient in aortic stenosis. *Front Cardiovasc Med*. 2021;8:706628.

34. Zelditch M. Morphometric tools for landmark data: geometry and biology. *Bioscience*. 1998;48(10):855-858.

35. Izzo R, Steinman D, Manini S, Antiga L. The vascular modeling toolkit: a python library for the analysis of tubular structures in medical images. *J Open Source Softw*. 2018;3(25):745.

36. McNeil AJ, Frey R, Embrechts P. *Quantitative Risk Management: Concepts, Techniques and Tools-Revised Edition*. Princeton University Press; 2015.

37. Burchett WW, Ellis AR, Harrar SW, Bathke AC. Nonparametric inference for multivariate data: the r package npmv. *J Stat Softw*. 2017; 76:1-18.

38. Buellesfeld L, Stortecky S, Kalesan B, et al. Aortic root dimensions among patients with severe aortic stenosis undergoing transcatheter aortic valve replacement. *J Am Coll Cardiol Intv*. 2013;6(1):72-83.

39. Delgado V, Ng AC, Schuijf JD, et al. Automated assessment of the aortic root dimensions with multidetector row computed tomography. *Ann Thorac Surg*. 2011;91(3):716-723.

40. de Jaegere P, De Santis G, Rodriguez-Olivares R, et al. Patient-specific computer modeling to predict aortic regurgitation after transcatheter aortic valve replacement. *J Am Coll Cardiol Intv*. 2016;9(5):508-512.

41. Iung B, Baron G, Butchart EG, et al. A prospective survey of patients with valvular heart disease in europe: the euro heart survey on valvular heart disease. *Eur Heart J*. 2003;24(13):1231-1243.

42. Vincent F, Ternacle J, Denimal T, et al. Transcatheter aortic valve replacement in bicuspid aortic valve stenosis. *Circulation*. 2021; 143(10):1043-1061.

43. Ribeiro HB, Nombela-Franco L, Urena M, et al. Coronary obstruction following transcatheter aortic valve implantation: a systematic review. *J Am Coll Cardiol Intv*. 2013;6(5):452-461.