Jan Oldenburg*, Finja Borowski, Laura Supp, Matthias Leuchter, Alper Öner, Klaus-Peter Schmitz and Michael Stiehm

Cluster Analysis of Volumetric Calcification Distributions in the Aortic Valve Region of TAVR Patients

https://doi.org/10.1515/cdbme-2025-0133

Abstract: Minimally invasive transcatheter aortic valve replacement (TAVR) has become the preferred procedure for patients with aortic valve stenosis or insufficiency who are at high risk for conventional open-heart surgery. The favorable clinical outcomes observed in high-risk patients have led to an expansion of the eligible cohort, now including intermediate- and low-risk patients. A critical aspect of advancing TAVR procedures lies in preoperative simulation, which in future can integrate patient-specific *in-silico* deployment simulations and post-deployment fluid mechanics assessments. In light of the aforementioned context, this study investigates the volumetric distribution of calcifications in the aortic cusps of a TAVR patient population and explores whether naturally occurring clusters should be considered in *in-silico* simulations. We analyze clustering results based on different methods for feature extraction, using circular measured volumetric calcification distributions. The clustering method applied is hierarchical clustering. Our findings identify distinct calcification clusters across different feature extraction methods. The Framework can be incorporated into *in-silico* trials and clinical studies assessing the impact of calcification patterns on clinical outcomes and TAVR device optimization.

Keywords: TAVR, calcification volume, cluster, aortic stenosis

1 Introduction

The minimally invasive implantation of transcatheter aortic valve replacement (TAVR) has become the standard procedure for patients with severe aortic valve stenosis or insufficiency who are considered high-risk candidates for open surgery. Given its promising clinical outcomes, there is ongoing debate about expanding TAVR eligibility to include patients with intermediate or even low surgical risk. [1]. To further enhance clinical outcomes, preoperative TAVR planning approaches—including patient-specific *in-silico* deployment simulations and post-deployment flow analyses—are being actively explored. Numerical simulation models for TAVR deployment in patient-specific geometries have already been developed and applied, as demonstrated in studies such as [2]. Certain research groups are investigating the development of virtual cohorts using shape modeling techniques to adopt a more comprehensive, population-based approach, allowing for the generation of synthetic patient geometries. For instance, Verstraeten et al. developed a shape model of the aortic root of TAVR patients [3].

In this context, gaining deeper insights into the volumetric calcification patterns of real clinical TAVR cohorts is essential. More explicitly, ensuring a robust modeling approach requires accurately capturing naturally occurring calcification patterns, so they can be appropriately incorporated into numerical modeling techniques.

In this study, we investigate whether distinct, naturally occurring calcification patterns exist in TAVR patients and whether these patterns can be grouped into meaningful clusters. To achieve this, we present a cluster analysis that explores different feature extraction methods, clusters circular measured calcification volumes based on the extracted features, and evaluate the resulting clusters using internal cluster validation metrics and visual inspection. Among these methods, we include Wasserstein-PCA for circular measures, recently proposed by [4], which appears to be particularly well-suited for this type of analysis.

Based on this calcification clustering framework and results, we aim to establish a robust baseline for future analyses focused on modeling realistic calcification patterns in virtual cohorts. These cohorts can be integrated into *in-silico* modeling frameworks to evaluate the impact of calcification patterns on the TAVR procedure and subsequent clinical outcomes.

***Corresponding author: Jan Oldenburg:** Institute for ImplantTechnology and Biomaterials e.V., Friedrich Barnewitz Str. 4, Rostock-Warnemünde, Germany, e-mail: jan.oldenburg@uni-rostock.de
Finja Borowski, Laura Supp, Matthias Leuchter, Klaus-Peter Schmitz, Michael Stiehm: Institute for ImplantTechnology and Biomaterials e.V.
Alper Öner: Heart Center/Department of Cardiology, Rostock University Medical Center

2 Material and methods

2.1 Data processing of calcified TAVR populations

The cluster analysis is based on 96 post-operative TAVR CT patient datasets provided by the Department of Cardiology at the Rostock University Medical Center. The patient data are uniform distributed between female and male patients, with an average age of 86 ± 17 years. The reconstruction of the aortic root and the calcifications was previously described in [5]. Variations in the aortic root within the population were modelled using a shape model based on principal component analysis (PCA). By combining this approach with a mapping routine between a mean aortic root geometry and individual patient anatomies, as described in [5], calcification volumes were normalized (CVn) onto the mean aortic root. In this study, we focus exclusively on calcifications in the region of the aortic valve and root, specifically between the annulus and the sinotubular junction (STJ). We propose investigating the calcifications as a pseudo-circular volumetric distribution spanning the aortic root in a circular manner, with the axis aligned from proximal to distal locations, following the aortic valve orientation. We refer to it as a "pseudo-circular distribution" because the segmentation is not based on equidistant points along the entire circle. Instead, the circle is first divided into three segments corresponding to the three cusps: the left coronary cusp (LCC), the right coronary cusp (RCC), and the non-coronary cusp (NCC). Within each of these three segments, equidistant points are placed along the curve to further segment the calcifications and calculate their volumes within these subdivisions, as illustrated in Figure 1.

Using this approach, we gain a more intuitive understanding of the circular position of calcifications while accounting for the different sizes of the cusp regions. The segmentation and cluster analysis were performed using PyVista library within a custom Python script [6]. In this analysis, eleven segments were used per cusp.

In the following, the CVn distributions are represented on a unit circle with equidistant segments described by the angle φ. This should be kept in mind when interpreting the results (see Figure 1 for an example). The pseudo-circular measured CVn will hereafter be referred to simply as CVn.

2.2 Cluster analysis of calcified TAVR populations

Based on CVn, we implemented a clustering routine, testing five different feature extraction mechanisms, which will be described in the following section. Using the extracted features, we applied a clustering algorithm to identify distinct calcification patterns.

2.2.1 Feature extraction on the calc volume distribution

As one of the feature extraction methods we implemented a PCA on the raw data consisting of CVn ordered according to φ as input. The second feature extraction method is the Elliptical Fourier Descriptor (EFD), introduced by [7]. This method describes the closed curve formed by the CVn distribution using Fourier sine and cosine functions. The level of detail captured in the shape representation is controlled by the parameter h, which determines the amount of higher-frequency components included. In this study $h = 5$, 10, and 20 were analyzed. The python library spatial-efd was used [8]. The third method is a wavelet transform-based feature extraction technique, in which CVn is decomposed into a series of approximation coefficients using basis functions [9]. We employed the Haar wavelet transform, by using PyWavelets [10], and used 100 coefficients to describe CVn distribution and applied PCA analysis on this features (WT-PCA). Furthermore, we tested a recently published method, the so called Wasserstein-PCA (WA-PCA) on circular domains [4]. This approach operates in the Wasserstein space of probability measures supported on the

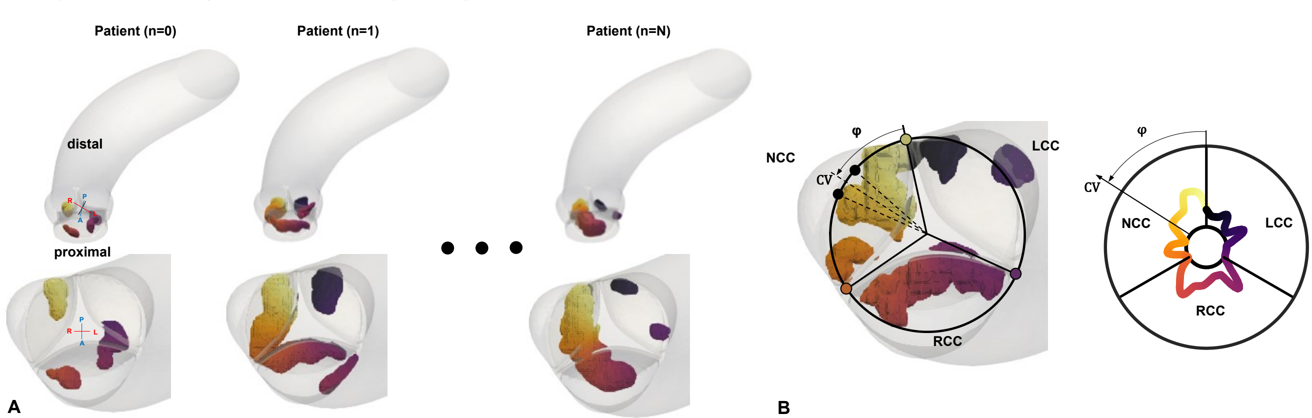

Figure 1: (A) Exemplary calcified aortic roots of TAVR patients, colored by radial angle φ, (B) and data processing routine extracting the circular measured pseudo-circular calcification volume distribution.

unit circle, combined with PCA. For all PCA analyses, we retained components that explain > 95 % of the variance in the input data.

2.2.2 Cluster analysis

We performed agglomerative hierarchical clustering using Ward's linkage method, by means of the Scikit-learn python module [11]. To determine the optimal number of clusters, we tested different cluster numbers and computed internal cluster evaluation metrics, including the silhouette score and Calinski-Harabasz index (CHI). To assess the robustness and generalizability of the clustering process, we randomly sampled a subset of the cohort (80 %), which was included in the clustering process. This procedure was repeated five times, and we averaged the computed internal evaluation metrics.

3 Results and Discussion

3.1 Internal evaluation of the calcification cluster

For each clustering result obtained from all feature extraction methods, the silhouette score and CHI were visualized across varying cluster numbers. These values were computed using the hierarchical clustering method and are presented with a banded standard deviation (scaled by 0.5 for improved visualization) based on a random train-test split of the CVn data. To determine the optimal number of clusters, we aimed to identify configurations with more than two clusters. A distinct and meaningful cluster number was identified for the EFD method, both with and without PCA applied to the features. The results for different harmonic choices were highly similar, all indicating an optimal cluster number of four, as suggested by both the silhouette score and CHI. Applying PCA directly to the raw data resulted in a rapid decline in both the silhouette score and CHI, with an optimal cluster number of three selected for visualization. For the WT feature extraction method, the highest silhouette score for number of cluster > 2 was observed at five cluster when combined with PCA. Clustering based on the raw WT features showed a gradual decrease in silhouette scores up to ten cluster. Finally, clustering based on Wasserstein PCA (WA-PCA) extracted features exhibited relatively small variations in silhouette scores as the cluster number increased. However, a slight inflection point was observed at nine cluster, while both CHI and the silhouette score plot indicated an optimal cluster number of three.

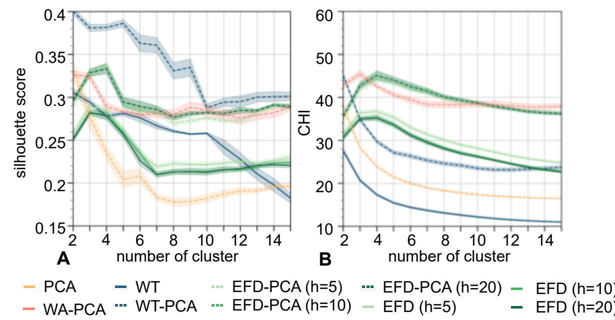

Figure 2: Internal cluster evaluation metrics computed for different cluster numbers using various feature extraction methods in clustering pseudo-circular volumetric calcification distributions. From left to right: Silhouette score and Calinski-Harabasz index (CHI).

3.2 Visual calcification cluster evaluation

The most distinct and meaningful clusters, based on the optimal cluster numbers discussed in Section 3.1, are visualized below for WA-PCA, PCA on raw data, EFD-PCA, and WT-PCA. For each cluster, the median along with the 25th and 75th percentiles are shown. The number of patients in each cluster is annotated. To facilitate interpretation, clusters with visually similar characteristics across different feature selection methods have been grouped together. The most visually identifiable cluster, C1, is characterized by a high CVn in the NCC region, accompanied by a smaller but relatively high amount of CVn in the RCC and LCC regions compared to other clusters. This finding holds for all feature selection methods. In EFD-PCA, this cluster is further subdivided, with one sub cluster exhibiting calcifications localized solely in the NCC region. WA-PCA emphasizes the identification of concentrated CVn corresponding to specific cusp regions, with C2 exhibiting higher CVn in RCC and C3 showing higher CVn in LCC. However, these patterns, while identifiable, also display a non-negligible overlap of calcifications between the cusp regions, indicating shared occurrence rather than entirely distinct distributions.

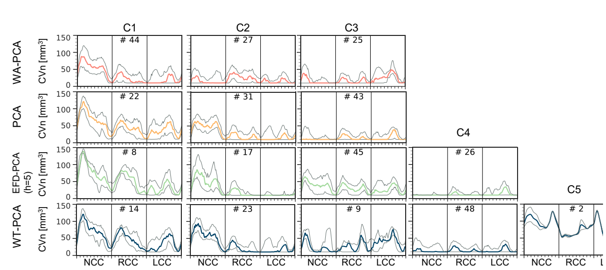

Figure 3: Exemplary visual comparison of clustering results of pseudo-circular volumetric calcification distributions using different feature extraction methods.

In Figure 4, the clustering results for WA-PCA and WT-PCA are shown for an analysis with nine clusters (number of patients: #). Clustering into nine groups reveals that the distributions can be further subdivided into clusters with distinct differences in the location of calcifications. Two clusters, C4 and C6, were identified, bridging cusp regions. C4 exhibits calcifications across both the RCC and LCC regions, while C6 shows calcifications spanning NCC and RCC. The WA-PCA method resulted in more evenly distributed patient numbers across clusters compared to the WT-PCA method, which exhibited greater variability in cluster sizes.

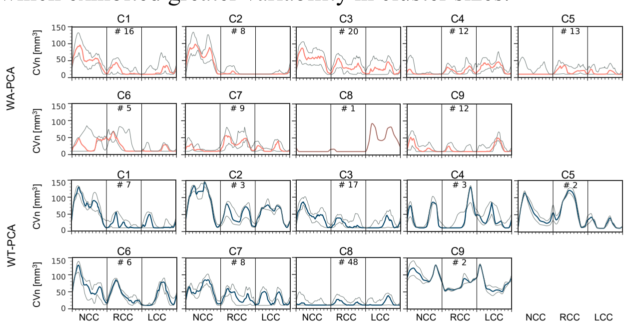

Figure 4: Exemplary visual comparison of clustering results with nine cluster based on pseudo-circular volumetric calcification distributions using WA-PCA and WT-PCA.

This study demonstrates that distinct patterns of calcifications can be grouped into clusters sharing similar CVn distributions in the region of the native aortic valve. The largest clusters identified exhibited a higher amount of calcification in the NCC, a finding consistent with previous clinical investigations [12,13]. Based on the presented clustering results, future research should investigate whether calcification clusters impact clinical outcome post TAVR, such as paravalvular leakage or thrombosis. Understanding whether these clusters provide meaningful prognostic information could contribute to stronger mechanistic insights, which remains insufficiently explored in clinical studies but is discussed [13]. We identified subgroups of calcification patterns that exhibit characteristics of functional bicuspid valves. Further investigation is warranted to assess the safety of current TAVR devices for such cases and to evaluation of bicuspid-adapted TAVR devices may be necessary. The clustering methods and results presented in this study serve as a baseline for future analyses assessing the impact of calcification patterns on clinical outcomes and TAVR device optimization. Furthermore, these results provide a reference, enabling the validation of synthetic calcification generators within the concept of virtual TAVR cohorts. Finally, the findings of this study should be cross-validated using data from other clinical centers to ensure generalizability and robustness.

Author Statement

Research funding: This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017578. The EU is not responsible for any use that may be made of the results shown in this publication or any information contained therein. This study was conducted in accordance with the ethics application no. A 2020-0250.

References

[1] Leon MB, Smith CR, Mack MJ, et al. Transcatheter or Surgical Aortic-Valve Replacement in Intermediate-Risk Patients. The New England journal of medicine 2016;374:1609–1620.
[2] Pak DH, Liu M, Kim T, et al. Patient-Specific Heart Geometry Modeling for Solid Biomechanics Using Deep Learning. IEEE transactions on medical imaging 2024;43:203–215.
[3] Verstraeten S, Hoeijmakers M, Tonino P, et al. Generation of synthetic aortic valve stenosis geometries for in silico trials. International journal for numerical methods in biomedical engineering 2024;40:e3778.
[4] Beraha M, Pegoraro M. Wasserstein Principal Component Analysis for Circular Measures; 2023 Apr 5. Available from: URL:http://arxiv.org/pdf/2304.02402.
[5] Oldenburg J, Borowski F, Supp L, Öner A, Schmitz K-P, Stiehm M. Generation of Shape Models of Calcified TAVR populations for Solid Mechanics Simulations by means of Deep Learning. Current Directions in Biomedical Engineering 2024;10:469–472.
[6] Sullivan C, Kaszynski A. PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK). JOSS 2019;4:1450.
[7] Kuhl FP, Giardina CR. Elliptic Fourier features of a closed contour. Computer Graphics and Image Processing 1982;18:236–258.
[8] W D Grieve S. spatial-efd: A spatial-aware implementation of elliptical Fourier analysis. JOSS 2017;2:189.
[9] Antoine J-P, Barachea D, Cesar RM, da Fontoura Costa L. Shape characterization with the wavelet transform. Signal Processing 1997;62:265–290.
[10] Lee G, Gommers R, Waselewski F, Wohlfahrt K, O'Leary A. PyWavelets: A Python package for wavelet analysis. JOSS 2019;4:1237.
[11] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research (.
[12] Pollari F, Dell'Aquila AM, Söhn C, et al. Risk factors for paravalvular leak after transcatheter aortic valve replacement. The Journal of thoracic and cardiovascular surgery 2019;157:1406-1415.e3.
[13] Kim W-K, Bhumimuang K, Renker M, et al. Determinants of paravalvular leakage following transcatheter aortic valve replacement in patients with bicuspid and tricuspid aortic stenosis. European heart journal. Cardiovascular Imaging 2021.