Review report

This manuscript presents a machine learning framework that develops a surrogate model for predicting patient-specific aortic wall mechanics in ascending thoracic aortic aneurysms (ATAAs), together with a data generation pipeline for ATAA surfaces. The problem addressed is relevant, and the numerical pipeline is technically sound. However, the methodological novelty is limited, and the analysis of the results remains largely descriptive, providing insufficient insight into the observed performance. In particular, the predictive accuracy for patient-specific geometries is only moderately convincing, and the presentation of several figures lacks clarity and interpretability. Substantial improvements in result analysis, figure presentation, and clarification of methodological assumptions are required for the work to meet the standards of the journal.

For these reasons, I recommend a decision of major revision.

Major Comments:

In Section 2.4, Young's modulus is sampled from a uniform distribution $U(1,4)$MPa. What is the motivation for this choice? Is this sampling intended primarily to increase the size and coverage of the training dataset, or is it meant to represent physiologically realistic variability? Given the wide uniform range assumed, how practically relevant is this random sampling strategy?

R: The uniform distribution $U(1,4)$ MPa was selected based on the experimental characterization of ascending thoracic aortic aneurysm tissue reported by Duprey et al. [35], who observed both physiological and upper-bound elastic moduli within this range. The adopted sampling strategy serves a dual purpose: (i) to represent the documented inter-patient variability in aortic wall stiffness, and (ii) to ensure sufficient coverage of the input space for robust training of the surrogate model. We acknowledge, however, that uniform sampling does not reflect the true population distribution of material properties and may therefore over-represent extreme values. This limitation is now explicitly discussed in the revised Discussion section, "Also, uniform distributions were adopted for the sampling of Young's modulus and wall thickness. This choice does not reflect the true population distribution of aortic wall properties, which is likely non-uniform".

Figure 3 presents a summary of the NMAE and R values for all test samples; however, the figure is vague and difficult to interpret. There is no legend or explanation for the labels A, B, C, and D. Moreover, the distribution of errors across the dataset is not clearly conveyed. The figure should be revised to improve clarity and to more effectively communicate the model's performance.

R: We agree that the figure did not sufficiently explain the meaning of the labels A, B, C, and D. To address these issues, Figure 3 has been revised to include a clear legend explicitly explaining that labels A, B, C, and D correspond to representative test cases exhibiting very high, high, moderate, and low levels of agreement, respectively, which are further analysed

in Figures 4 and 5. In addition, we have added Table 2, which summarises the distribution of the performance metrics across all test samples using descriptive statistics.

In Section 3.2.1, the authors state that “the y-axis corresponds to the circumferential direction, with BC and SC denoting the big and small curvatures, respectively”. What exactly is meant by “big” and “small” curvature in this context? Are these standard or commonly accepted concepts in the biomechanics literature, or are they defined specifically for this study?

R: The terms "big curvature" (BC) and "small curvature" (SC) refer to the outer (convex) and inner (concave) aspects of the ascending aorta, respectively. These correspond to the "greater curvature" and "lesser curvature" in standard anatomical nomenclature. We have revised the manuscript to adopt the more widely accepted terminology and added a clarifying statement with appropriate anatomical references.

Regarding the numerical simulation part (section 2.4), was any mesh convergence or mesh sensitivity study performed?

R: We acknowledge that a formal mesh convergence or mesh sensitivity study was not conducted as part of the original manuscript. The mesh density of the template mesh adopted in this work ($n = 1354$ surface elements) was selected through a trial-and-error process aimed at balancing geometric accuracy and robustness of the mesh-morphing procedure across the full dataset. Template meshes with approximately 700, 2500, and 5000 surface elements were also evaluated. While finer meshes improved geometric resolution, they led to increased element distortion and occasional mesh degradation following the morphing process, whereas coarser meshes compromised geometric fidelity. The selected mesh density was therefore identified as the most reliable compromise.

With respect to the numerical simulations themselves, the volumetric meshes were generated using element sizes consistent with those adopted in our previous FSI studies of the aortic wall [10.3390/biomechanics2020016, 10.1016/j.cmpb.2024.108475], where similar discretisation levels were shown to provide stable and accurate results.

We recognise that the absence of a formal mesh convergence analysis constitutes a limitation of the present study, which is now explicitly acknowledged in the revised manuscript. Accordingly, the following statements have been added to Section 2.4 and to the Discussion section, respectively:

- *“These simulations were performed using meshes with element sizes varying between 0.8–1 mm. These values were inspired by previous work from our group.”*
- *“and a systematic mesh convergence analysis to quantify the sensitivity of the numerical results to spatial discretisation and to further assess the impact of mesh resolution on surrogate model accuracy.”*

Regarding the evaluation of model performance, the manuscript does not provide any aggregated or global assessment over the entire dataset. The results are presented mainly through scatter plots of NMAE and R values for individual data point, e.g., in Figure 1, 6, and 7. These plots do not allow the reader to form a clear overall understanding of the model’s performance. Summary statistics, or a more comprehensive quantitative comparison

between the target and prediction values over the full dataset would be necessary to support the conclusions drawn.

R: We agree that, in the original manuscript, model performance was primarily communicated through per-sample scatter plots of NMAE and R, which can make it difficult to obtain a clear global assessment over the full test dataset. To address this limitation, we have added Table 2, which provides an aggregated quantitative summary of model performance across all test samples. Specifically, this Table reports descriptive statistics (mean, median, interquartile range and min-max range) for both NMAE and R, enabling the reader to directly assess the overall accuracy and variability of the predictions across the dataset. The Results section has been updated accordingly to explicitly reference this table when discussing global model performance.

Together, the added summary statistics and the existing scatter plots provide both a global assessment of the model's performance and insight into inter-sample variability.

Section 3.3 discussed the correlation between error and inputs. However, this interpretation is not sufficiently supported or clearly explained. The observed trend in high errors can be a consequence of imbalanced training datasets. The manuscript does not provide any analysis on the distribution of training samples across the input space. As such, the current analysis does not convincingly demonstrate any causal relationship between input parameters and model performance.

R: We agree that the analysis presented in Section 3.3 does not provide sufficiently strong evidence to support a causal interpretation between the input parameters and the observed model errors, particularly in the absence of a detailed analysis of the training data distribution across the input space. In light of this, and to avoid potentially misleading interpretations, we have decided to remove this analysis from the revised manuscript. The corresponding section has been deleted, and the manuscript has been updated accordingly.

Minor Comments:

Typos

Page 5, line 4: "In tis case" should be corrected to "In this case".

Page 5, line 7: "a acceptable" should be corrected to "an acceptable".

Page 5, paragraph 3: "Section 3 present" should be corrected to "Section 3 presents".

Page 6, Section 2.2, paragraph 2: "Source Pointss (SPs)" should be corrected to "Source Points (SPs)", and "These SP" should be corrected to "These SPs".

Page 11, Section 3.1: "The Second and third" should be corrected to "The second and third".

In the Introduction, the authors state that the model predicts the *Second Piola–Kirchhoff stress* and the *deformation gradient tensor*, whereas the Abstract and the rest of the manuscript report results for the *Second Piola–Kirchhoff stress* and the *Right Cauchy–Green*

strain tensor. Since these are different quantities, this inconsistency should be clarified and corrected.

R: All listed corrections have been made in the revised manuscript. Some additional notes are:

- “In tis case” was simply removed. This sentence is now “In the realm of cardiovascular biomechanics analysis, surrogate models…”
- Regarding the last aspect, the surrogate model predicts the Right Cauchy-Green deformation tensor (C), not the deformation gradient tensor (F). The Introduction has been corrected to consistently refer to the Second Piola-Kirchhoff stress tensor and the Right Cauchy-Green deformation tensor throughout the manuscript.