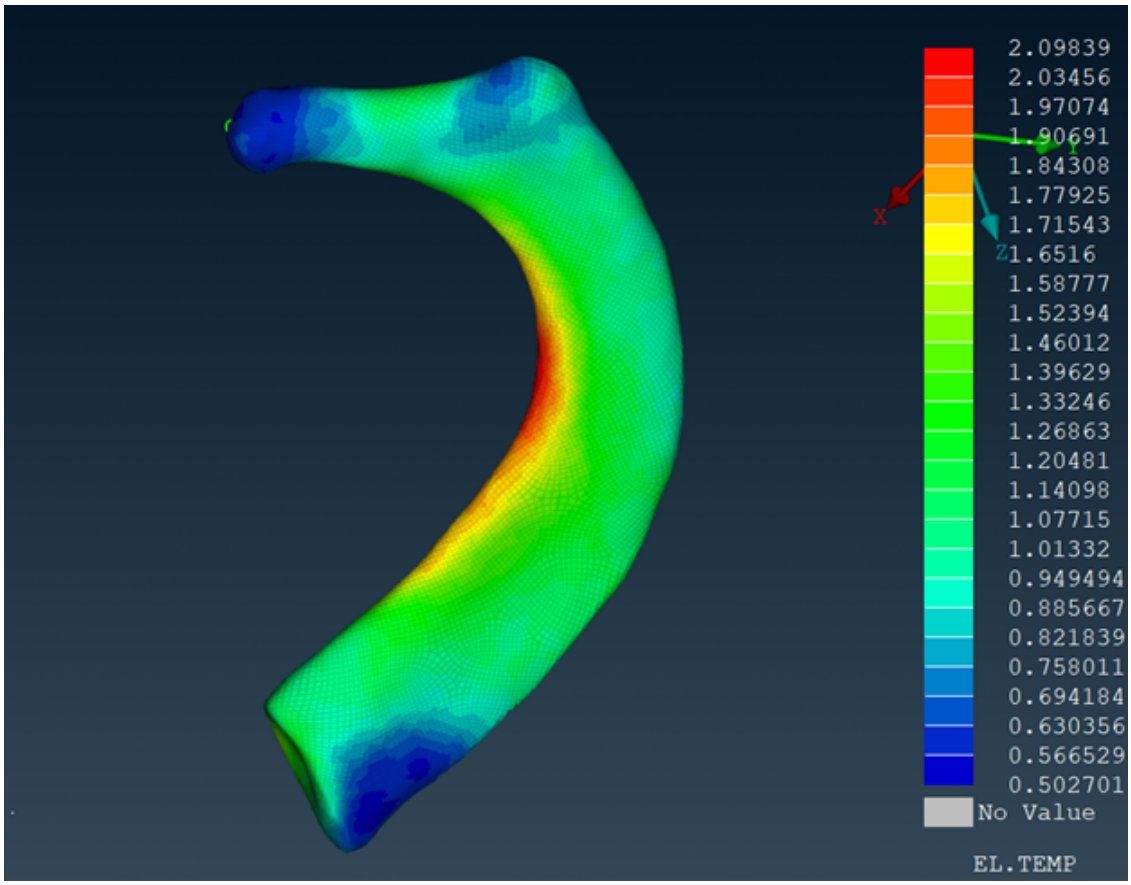
CHALMERS
UNIVERSITY OF TECHNOLOGY



Statistical Shape Modeling of the First Human Rib

Modeling Using CT Data and Demographic Predictors

Master's thesis in MSc. Biomedical Engineering

SHANKHARAN SRINIVASAN AND JOHAN BERG

DEPARTMENT OF MECHANICS AND MARITIME SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Statistical Shape Modeling of the First Human Rib

Modeling Using CT Data and Demographic Predictors

SHANKHARAN SRINIVASAN
JOHAN BERG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology
Gothenburg, Sweden 2025

Statistical Shape Modeling of the First Human Rib
Modeling Using CT Data and Demographic Predictors
SHANKHARAN SRINIVASAN, JOHAN BERG

Supervisors :Johan Iraeus, Department of Mechanics and Maritime Sciences.
Karl-Johan Larsson, Autoliv AB
Examiner : Johan Davidsson, Department of Mechanics and Maritime Sciences.

Master's Thesis 2025
Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Average shape of the first rib with cortical thickness mapped onto the surface after morphing and alignment.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2025

Statistical Shape Modeling of the First Human Rib
Modeling Using CT Data and Demographic Predictors
Shankharan Srinivasan, Johan Berg
Department of Mechanics and Maritime Sciences
Chalmers University of Technology

Abstract

The first human rib plays a critical role in thoracic biomechanics and is increasingly relevant for improving simulations in human body models used in vehicle and traffic safety research. This thesis aims to develop a statistical shape model (SSM) of the first rib that captures morphological variability in terms of sex, age, height, and body mass index (BMI), based on computed tomography (CT) data.

A total of 50 anonymized CT scans, sourced from a trauma hospital through the University of Michigan, were used from an available dataset of 104 samples. The workflow included segmentation in 3D-Slicer, landmarking in ANSA, and cortical bone thickness mapping using Stradview. To enable statistical analysis, a template rib surface was morphed to match each segmented rib using the Infepy Python library, so that all surfaces shared the same number and arrangement of nodes. Rib thickness values were also transferred onto these common nodes to combine shape and cortical thickness data. Generalized Procrustes Analysis (GPA) and Principal Component Analysis (PCA) were performed to align the ribs and identify dominant modes of variation, respectively. Linear regression was employed to examine relationships between principal component scores and demographic variables, and to predict shape and thickness parameters, which were then used to reconstruct corresponding 3D meshes.

Four principal components showed statistically significant correlations ($p \leq 0.05$) with demographic variables (age, sex, height, and BMI) in regression analysis. These components primarily encoded variation in rib size, cortical thickness, and subtle shape changes. However, age, sex, height, and BMI alone could capture only 15% of the total variability, indicating that these parameters were insufficient to fully predict rib morphology.

Despite these limitations, the resulting model represents a valuable step toward anatomically realistic rib shape modeling and offers potential for enhancing personalized simulations in safety engineering. Additionally, PCA proved useful beyond model building by enabling quantification and visualization of how rib morphology varies across individuals.

Keywords: First rib, statistical shape model (SSM), cortical bone mapping, Generalized Procrustes Analysis (GPA), Principal Component Analysis (PCA), regression

Acknowledgements

We would like to express our sincere appreciation to all those who contributed to the successful completion of this thesis.

We are particularly grateful to our supervisors, Johan Iraeus and Karl-Johan Larsson (Autoliv AB), for their expert guidance, constructive feedback, and continuous support throughout the course of this work. Their insights were invaluable in shaping the quality and direction of the research.

We would also like to acknowledge Autoliv AB for initiating and supporting this thesis project, and for providing a valuable industrial context in which to conduct our research.

We are grateful to Sven Holcombe (University of Michigan) for providing the dataset used in this work.

We extend our thanks to Chiara Fichera for her technical assistance, especially in providing data processing methods and for her thoughtful input at critical stages of the project.

We also wish to thank Oscar Hallberg for his helpful guidance and foundational input at the outset of the thesis work, which greatly facilitated our initial progress. We would like to thank our examiner Johan Davidsson.

Our gratitude further goes to Professor Graham Treece (University of Cambridge) for his timely support and expertise in cortical bone mapping, which proved instrumental in addressing key challenges.

Finally, we would like to acknowledge the unwavering support of our families and friends, whose encouragement and understanding have been a constant source of motivation.

To all who contributed, directly or indirectly to this work, we extend our heartfelt thanks.

Shankharan Srinivasan
Johan Berg
Gothenburg, June 2025

List of Acronyms

Below is the list of acronyms and abbreviation that have been used throughout this thesis.

AIC	Akaike Information Criterion
ATD	Anthropometric Test Device
BMI	Body Mass Index
CBM	Cortical Bone Mapping
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
FWHM	Full Width at Half Maximum
GPA	Generalized Procrustes Analysis
HBM	Human Body Model
MRI	Magnetic Resonance Imaging
PCA	Principal Component Analysis
PC	Principal Component
PMHS	Post Mortem Human Subject
RBF	Radial Basis Function
ROI	Region of Interest
RSS	Residual Sum of Squares
SSM	Statistical Shape Model
Tn	Thoracic Vertebra n
TPS	Thin Plate Spline
TSS	Total Sum of Squares

Contents

List of Figures

List of Tables

1

Introduction

Motor vehicle crashes are the second leading cause of severe or fatal thoracic injuries [1] and the leading cause of rib fractures [2]. However, the risk is not uniform across the population [1]. Factors such as age, obesity, and sex influence injury susceptibility, with women exhibiting a higher risk than men. To enhance vehicle safety for a diverse population, accurate simulations, such as finite element analysis of vehicle restraint performance and injury risk, are essential.

Crash test dummies, also known as anthropomorphic test devices (ATDs), are traditionally used in crash testing to estimate injury risk and ensure regulatory compliance [3]. For frontal impact tests, the commonly used Hybrid III ATD exists in three versions: a 5th-percentile female and 50th- and 95th-percentile males [3]. However, these models do not adequately capture anatomical and demographic diversity in the population and may therefore fail to accurately evaluate injury risks for all groups.

To complement ATDs, finite element human body models (HBMs) have been developed. A key advantage of HBMs is their ability to be scaled and morphed to account for variations in human anatomy [4]. An essential step in improving these models is the development of Statistical Shape Models (SSMs), also known as morphometric models, which capture population-level anatomical variations. These models help in the estimation of the bone's morphology using attributes such as stature, sex, and age from individual bones.

Importantly, the medical imaging field is advancing rapidly, aided by increasing accessibility to modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) [5]. These technologies provide valuable insights into anatomical variation in bones, supporting applications that extend beyond clinical practice into domains such as automotive safety. They also facilitate the development of more precise SSMs by using imaging techniques like CT scans to capture highly detailed bone structures.

Anatomically, the first rib plays a pivotal role in thoracic biomechanics by maintaining the structural stability of the rib cage. Due to its proximity to critical nerves, soft tissues, and blood vessels, the first rib influences the vulnerability of these structures to injury. Fractures of the first rib often occur near its articulation with the spine, where forces from sudden neck movements or strong muscle contractions concentrate stress. In vehicle crashes, these forces can be

transmitted through the clavicle, scapula, or spine, causing fractures at these sites [6]. Therefore, accurately capturing the shape and mechanical properties of the first rib is essential for simulating thoracic injuries and assessing its interaction with restraint systems.

Despite the availability of SSMs for various body parts, no such model currently exists for the first rib. A recent study [7] developed statistical shape models for ribs but excluded the first and twelfth ribs due to their distinct morphology. To address this gap, a SSM for the first rib will be developed using CT images. This model will enable realistic morphing of HBMs that incorporates accurate, population-specific anatomical features of the first human rib.

1.1 Aim

The aim of this thesis is to develop a statistical model for the first rib that accounts for variations in sex, height, BMI, and age, based on a dataset of CT images.

1.2 Specification of the issue being investigated

We aim to answer the following questions in this thesis:

- How does the shape and cortical thickness of the first human rib vary across individuals?
- What key morphological patterns can be identified through principal component analysis?
- To what extent do demographic factors such as age, sex, height, and BMI influence first rib morphology?
- How do the findings of this model compare to existing data from prior studies?

1.3 Limitations

The below limitations were recognized in this thesis -

- **Sample size:** The dataset comprised about 100 ribs, which may limit statistical power and generalizability.
- **Population characteristics:** All participants were from a midwestern U.S. adult population, restricting demographic diversity.
- **Imaging resolution:** Only standard clinical CT data was available; higher-resolution imaging such as micro-CT was not accessible.

- **Anthropometric data:** Only basic anthropometric data (age, sex, height, BMI) were available.
- **Time constraints:** The fixed thesis timeline limited dataset expansion and deeper analysis.

2
Theory

2.1 Rib Anatomy

The skeleton serves several important functions in the human body [8]. It provides a protective framework for vital organs, supports movement, act as a reservoir for essential minerals, and house bone marrow. Each bone consists of a dense outer layer called cortical bone and an internal porous structure known as trabecular bone, which contains the marrow.

The human thorax includes a skeletal structure called the ribcage, comprising 12 pairs of ribs, each posteriorly connected to the thoracic vertebrae T1–T12, though the exact structure of the connection varies between different ribs [8]. The seven uppermost pairs are called true ribs as they anteriorly connect directly to the sternum via costal cartilage. Ribs 8–12 are instead classified as false ribs. Of these, ribs 11 and 12 do not connect to the sternum and are therefore called floating ribs, while the cartilage of ribs 8–10 connects indirectly to the sternum through the cartilage of the ribs above them. The sternum also connects with the clavicles.

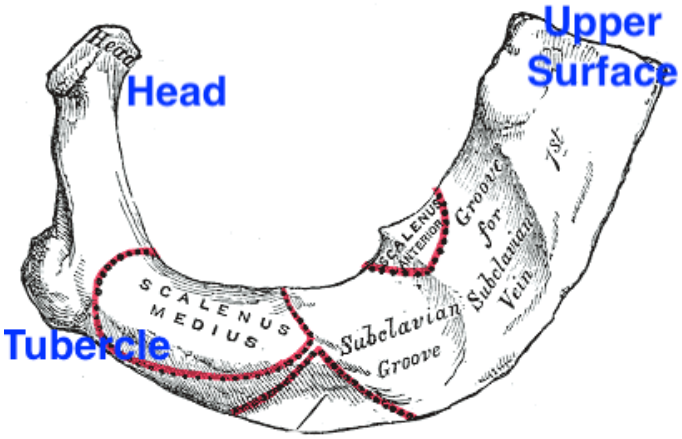


Figure 2.1: Visualization of the first rib with the three most important parts for this thesis; the head, tubercle and upper surface [9].

Rib 1, visualized in Figure 2.1, is a mostly flat bone with a cross-section that is approximately rectangular. On its posterior end, it has an expanded head that articulates with the vertebra T1, and a tubercle that forms an articulation with T1s transverse process [8]. The region between the head and the tubercle is called the neck. On its anterior end, opposite the shaft, the upper surface connects to the cartilage. The shaft is sharper along the medial edge compared to the lateral edge and features two grooves for the passage of blood vessels. Approximately halfway between the neck and the anterior surface on the sharp medial edge, the scalene tubercle, connecting to the anterior scalene muscle [10], forms a separation between the two grooves.

2.2 Statistical Shape Models

A statistical shape model is a mathematical representation of a physical object, consisting of both an average geometry and statistical information about its variations [11]. One common use of statistical shape models is to represent human anatomy beyond static average geometries. These models can also incorporate additional information such as material properties, appearances, functions, aging, or the progression of diseases. Such information can be derived from medical imaging modalities like CT or MRI, or through systematic observation of visual appearance.

Statistical shape models serve as a formalized description of anatomical variation and can be used in conjunction with biometric data, such as sex, height, medical conditions, or lifestyle habits, to identify and model correlations [11]. These models have a wide range of potential applications, including medical diagnosis, forensic analysis, and biomechanical modeling, which is the focus of this thesis.

2.3 CT imaging

Computed tomography (CT) is a medical imaging technique based on the fundamental principle of measuring X-ray attenuation within the body [12]. CT imaging captures X-ray projections from multiple angles, which can be combined to reconstruct cross-sectional slices. These slices are then assembled to produce a 3D image. The degree of X-ray attenuation depends on the properties of the tissue being imaged, with bone exhibiting higher attenuation than soft tissues. The intensity value of each voxel is determined by the average attenuation of the tissues and materials present within that voxel. These differences in attenuation generate the contrast seen in a CT image.

The relative attenuation is quantified using the Hounsfield scale, where different materials are assigned specific values [12]. Air, which has very low attenuation, is assigned a Hounsfield units of -1000, water is set at 0, and dense structures such as bone may reach values around +1000. These Hounsfield units are mapped to grayscale levels, producing a visual representation of the anatomical structures.

2.4 Segmentation

Image segmentation refers to the task of identifying and labeling specific regions or structures within an image. Based on the extent of human input required, segmentation is generally grouped into three categories - through manual effort, aided techniques, or fully automated algorithms [13]. In medical imaging, segmentation plays a critical role, especially in converting volumetric data from modalities such as CT and MRI into usable 3D models. These models are valuable for clinical research, surgical planning, and industrial applications.

Manual Segmentation

Manual segmentation is the most basic and user-driven form of segmentation [13]. In this approach, a user manually traces the region of interest (ROI) on each image slice, typically using a mouse or similar device.

Although this method allows for high control, it is labor-intensive and prone to variability, making it less reliable for reproducibility. However, for structures with complex or irregular geometries, such as certain bones, manual segmentation remains a widely accepted choice [13].

Semi-Automatic Segmentation

Semi-automatic segmentation represents a compromise between manual effort and full automation [13]. In this technique, only selected slices (e.g., every second or third slice) are outlined manually. The software then interpolates the missing contours between these slices.

This approach significantly reduces the time required while retaining a reasonable level of accuracy. The reliability of the result, however, depends on how frequently manual slices are marked. The more frequent the annotations, the more accurate the interpolated result.

Automatic Segmentation (Thresholding)

Automatic segmentation, especially using thresholding techniques, involves assigning pixel or voxel labels based on predefined intensity values [13]. This process transforms the image into a binary format, isolating regions of interest by filtering based on intensity.

In CT imaging, this is particularly useful for detecting bone structures due to their distinct Hounsfield unit range. Nevertheless, this method may struggle with very thin bones or areas where tissues of different types blend within a voxel. This partial volume effect can lead to imprecise boundaries, often requiring subsequent manual or semi-automatic corrections [13].

As deep learning methods continue to evolve, neural network-based techniques, particularly convolutional neural networks (CNNs), are increasingly adopted for CT image segmentation. Conventional segmentation approaches can support these methods by supplying annotated data required for model training [14].

2.5 Cortical Bone Mapping

Cortical Bone Mapping (CBM) refers to the technique of quantifying the thickness of both cortical and endocortical layers along the surface of a bone using CT imaging data. This typically involves evaluating properties along lines that extend perpendicularly into the bone surface. These lines are distributed across the bone, and their orientation is guided by an initial approximation of the bone's surface obtained through segmentation, as outlined in Section 2.4 [15].

Several approaches exist for estimating cortical thickness and density. One of the most basic methods is based on intensity thresholding. In this approach, thickness is calculated as the distance between points where the CT intensity, measured in Hounsfield units, crosses a specific threshold. The average Hounsfield value between these points serves as an estimate of cortical density. A variant of this is the Full Width at Half Maximum (FWHM) method, in which the threshold is dynamically set at half the maximum intensity and nearby baseline values. Here, the thickness corresponds to the width at half maximum, and cortical density is defined by the maximum Hounsfield value along the line [15].

For greater accuracy, particularly in regions with thin cortices, model-based CBM techniques are used. These methods begin by estimating features such as the position of cortical boundaries and the density of the cortex. Models are generated using these initial parameters and compared to the actual imaging data. The model is iteratively refined until the simulated data closely matches the observed CT measurements. This approach offers improved performance over simpler techniques like FWHM [15].

After the local measurements are completed, the final stage involves projecting the computed thickness values onto the 3D surface of the bone. These results are typically visualized as a color map on the bone mesh, allowing for an intuitive understanding of spatial variation in cortical thickness. Moreover, the mapped data can be exported to a template mesh, enabling statistical analysis across individuals or groups [15].

2.6 Landmarking

Landmarking is a method to compare anatomical features between individuals by specifying points of interest [16]. Homologous landmarks are anatomical landmarks that correspond to the same anatomical location making them meaningful references for comparison across individuals. Identifying these points is the first step of the

landmarking process. However, the number of homologous points on the ribs is limited.

One solution to this challenge of limited homologous points is to place pseudo-landmarks between them, for example, by distributing them equidistantly along a curve [16]. This ensures a standardized and dense representation of the mesh geometry while limiting the loss of consistency across individuals.

2.7 Morphing

A synonym for "transforming", morphing aims to apply a similar mesh structure from the template mesh (originating from the source bone) to the target mesh. A mesh is a discrete representation of an object, created by dividing the domain into a finite number of elements.

Thus, if n landmarks are identified in the template mesh, the same n landmarks will be present in the target mesh after morphing. The trade-off in mesh morphing involves balancing time and accuracy. Following morphing, the target mesh will exactly match the template mesh at the locations of the landmarks, using interpolation to fill in the 'gaps' between them. However, this interpolation may not always result in the required accuracy. In general, a greater number of landmarks improves accuracy, as there are more points where the target mesh will exactly match the template mesh. However, landmarking is time-consuming, which calls for a compromise between precision and efficiency.

The exact method of morphing used in this thesis is the Radial Basis Function (RBF) mesh morphing method [17]. The "Thin Plate Spline" interpolation, which is commonly employed for smooth deformations, has been used.

The RBF is generally defined as

$$s(\mathbf{x}) = p(\mathbf{x}) + \sum_{i=1}^{n} \lambda_i \, \varphi(\|\mathbf{x} - \mathbf{x}_i\|) \tag{2.1}$$

where $p(\mathbf{x})$ is a low-order polynomial, λ_i is the weighting coefficient, φ is the basis function, and $\|\mathbf{x} - \mathbf{x}_i\|$ is the Euclidean distance between $\mathbf{x}$ and $\mathbf{x}_i$.

The thin plate spline function is defined as

$$\phi(r) = r^2 \log(r) \tag{2.2}$$

where r is the Euclidean distance.

For this thesis, corresponding landmarks are identified on the template and target meshes. The displacement of each landmark is computed from its position on the template to its position on the target. These displacements are used in the RBF equation (Equation 2.1) to interpolate smooth deformations across the entire

rib mesh.

To improve the stability of the RBF interpolation, a smoothing parameter can also be included. This parameter slightly relaxes the requirement for the interpolated surface to pass exactly through all the landmarks. Instead, it allows a small amount of flexibility, which helps reduce the impact of any noise or small errors in the landmark positions and results in a smoother overall deformation.

2.8 Generalized Procrustes Analysis

The location of the subjects are based on the individual CT coordinate systems and are visibly misaligned in both location and orientation. Generalized Procrustes Analysis (GPA), introduced by Gower [18] and later popularized by Rohlf and Slice [19], addresses this issue by translating and rotating each specimen to minimize the summed squared distances, known as the Procrustes distance, between corresponding landmarks of each configuration and an iteratively computed mean configuration. This process aligns all specimens into a common coordinate system, such that differences in landmark coordinates reflect differences in shape and size.

2.9 Principal Component Analysis

Principal Component Analysis (PCA) is a dimensionality reduction technique designed to simplify complex datasets that contain many correlated features. It transforms the original variables into a new coordinate system composed of linearly uncorrelated variables known as principal components (PCs). The components are arranged so that those capturing the most variation in the data appear first [20].

Consider a data vector $\mathbf{x} = (x_1, x_2, \ldots, x_p)^\top$, representing p variables with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. PCA seeks directions $\mathbf{a}_k$ that capture the greatest variance when the data is projected onto them.

The first principal component z_1 is obtained by projecting the data onto the direction $\mathbf{a}_1$ that explains the greatest variance:

$$z_1 = \mathbf{a}_1^\top \mathbf{x}.$$

$\mathbf{a}_1$ is a vector that indicates the direction in which the data exhibits the highest variance.

Subsequent principal components, z_k, are found by projecting the data onto directions $\mathbf{a}_k$ that capture the next highest variance, subject to being orthogonal to all previously found directions:

$$z_k = \mathbf{a}_k^\top \mathbf{x}.$$

Each $\mathbf{a}_k$ represents a unique direction in the data space, ensuring that the principal components are uncorrelated and capture different aspects of the data's variability [21].

Each principal component accounts for a portion of the total variance in the data. To quantify how much of the original variability is retained, the proportion of variance explained by the first m components is calculated as:

$$\text{Proportion Explained} = \frac{\sum_{k=1}^{m} \lambda_k}{\sum_{k=1}^{p} \lambda_k},$$

where λ_k represents the variance captured by the k-th principal component. This ratio is a key criterion for selecting the number of components to retain, providing a balance between dimensionality reduction and information preservation [21].

When PCA is applied to some data, the original data points can be approximately reconstructed using a limited number of principal components. For the i-th observation, the approximation using the first m components is given by:

$$\tilde{\mathbf{x}}_i^{(m)} = \boldsymbol{\mu} + \sum_{k=1}^{m} z_{ik}\mathbf{a}_k \tag{2.3}$$

In this expression, $z_{ik} = \mathbf{a}_k^{\top}(\mathbf{x}_i - \boldsymbol{\mu})$ represents the score of the i-th observation on the k-th principal component. This reconstruction approximates the original data in the initial coordinate space using only the most informative components [21].

2.10 Regression

This step identifies which demographic data points, if any, have statistically significant predictive power for the PC score, based on a linear regression model. Using this information, a linear model of the PC score can be constructed.

The linear model is defined in Equation 2.4:

$$y = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n + \epsilon \tag{2.4}$$

where y is the dependent variable (a PC score), w represents the weights, x the independent variables (biometric data points), and ϵ is a noise term accounting for the difference between the actual and predicted PC score.

To compute the weights, an error function is selected, commonly the sum of squared errors, also known as the residual sum of squares (RSS), as defined in Equation 2.5:

$$E(\mathbf{w}) = \sum_{i=1}^{N}(\mathbf{w}^T\mathbf{x}_i - y_{i,\text{true}})^2 \tag{2.5}$$

Here, E denotes the RSS, $\boldsymbol{w}^T$ the transposed weight vector, $\boldsymbol{x}_i$ is the demographic data for sample i, and $y_{i,\text{true}}$ the actual PC score (ground truth) for that sample.

Assuming the noise term follows a Gaussian distribution, minimizing the RSS yields the maximum likelihood estimate of the weights [22].

To quantify how well the regression model predicts the PC scores, the coefficient of determination, R^2, is used, as defined in Equation 2.6:

$$R^2 = 1 - \frac{RSS}{TSS} \tag{2.6}$$

where TSS is the total sum of squares.

To compare the how good the fit is when comparing models with different number of independent variables, the adjusted R^2 value, defined by Equation 2.7:

$$R^2_{\text{adj}} = 1 - \left(\frac{(1 - R^2)(n - 1)}{1 - p - n} \right) \tag{2.7}$$

where n is the total number of datapoints and p refers to the number of independent variables used in the regression model. The adjusted R^2 therefore takes the number of variables into account to get more meaningful values in multivariate regressions [23].

Variable selection

Selecting the relevant variables the multivariate regression helps to improve the models and prevents overfitting the regression model [24]. It also serves to simplified the model by excluding irrelevant parameters. There are different systematic methods for this process that makes the selection based on a objective criteria such as p-value, R^2 or Akaike information criterion (AIC). The method used in this thesis is *backward elimination* based on p-value.

Backward elimination is done by starting with a multivariate linear regression that includes all demographic parameters [24]. If there are one or more parameters with a p-value over a predefined threshold (0.05) the highest p-value variable is removed and a new regression calculated. This process is iterated until all variables have a value equal to or below the threshold, which at this point includes the significant variables in the model. An other possibility is that the iteration continues until no variable are left, in this case no variables are considered statistically significant and the specified regression model cannot be used to make a prediction.

3
Methods

The method described in this section consists of several steps from the data preparation to the predicted SSM as well as a literature review. The flowchart in Figure 3.1 shows the most important steps described in the sections below and how they are connected to each other.

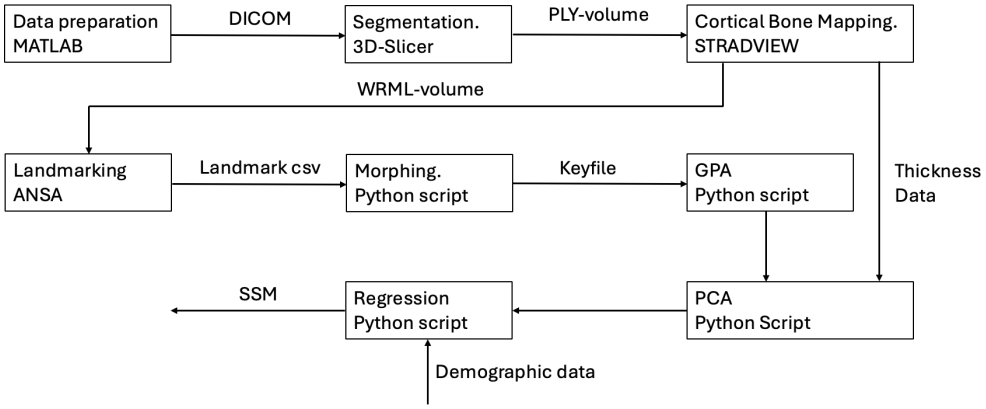


Figure 3.1: Flowchart connecting the parts described in the following sections.

3.1 Dataset

For this thesis, the data set originates from the University of Michigan and was obtained from a trauma hospital (approval HUM00041441) [25]. It contains 104 cropped and anonymized CT volumes including the first rib and the demographic parameters. The CT images were collected for reasons unrelated to this thesis.

3.1.1 Data preparation

The CT data was provided by the supervisors as MATLAB MAT-files with the Hounsfield values as spatial information for each voxel, from which voxel size could calculated from coordinate tables and reconstructed in as described in section 3.2. A MATLAB script was used for repackaging the CT data into DICOM files for work further work in 3D-Slicer.

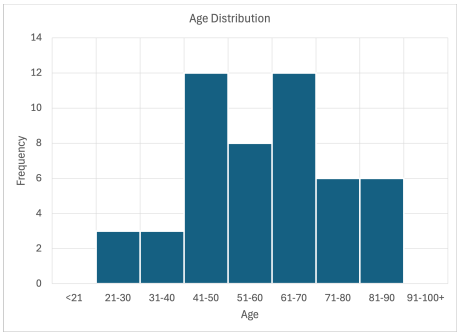
3.1.2 Exclusion criteria

Table 3.1 presents the exclusion criteria along with the corresponding number of CT images excluded for each category.

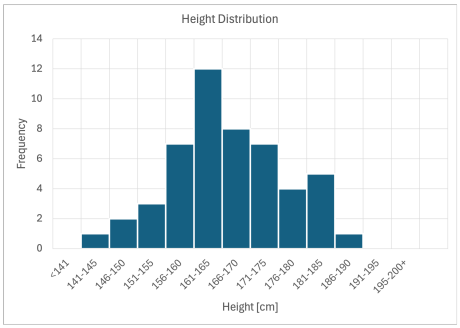
Exclusion Criteria	Number of Exclusions
Missing demographic data	8
Highly cropped/zoomed in, poor image quality, CT artifacts, broken ribs, or other uncertainties in the segmentation process	30
Failed cortical bone mapping	15
Demographic data outlier (BMI = 62.5)	1
Total	**54**

Table 3.1: Exclusion criteria and number of CT images excluded

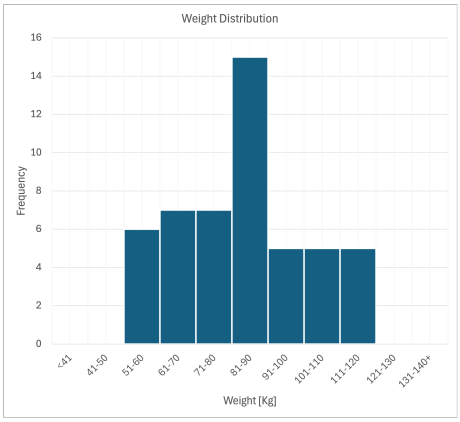
3.1.3 Demographic data distribution



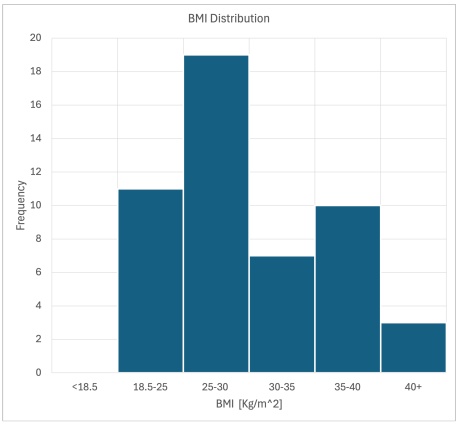
(a) Age distribution



(b) Height distribution



(c) Weight distribution



(d) BMI distribution

Figure 3.2: Distribution histogram of the dataset: (a) Age, (b) Height, (c) Weight, and (d) BMI.

Table 3.2: Mean and standard deviation of the demographic variables in the dataset

Variable	Mean	Standard Deviation
Age (years)	59	16
Height (cm)	167	10
Weight (kg)	83	17
BMI (kg/m^2)	30	6

Out of the 104 samples, 50 was used and the demographic data distributions are visualized in Figure 3.2a, 3.2b, 3.2c and 3.2d. Out of the 50 data samples, 29 was from female subjects and 21 was from male subjects. Table 3.2 shows the mean and standard deviations of the demographic variables in the dataset.

3.2 Segmentation

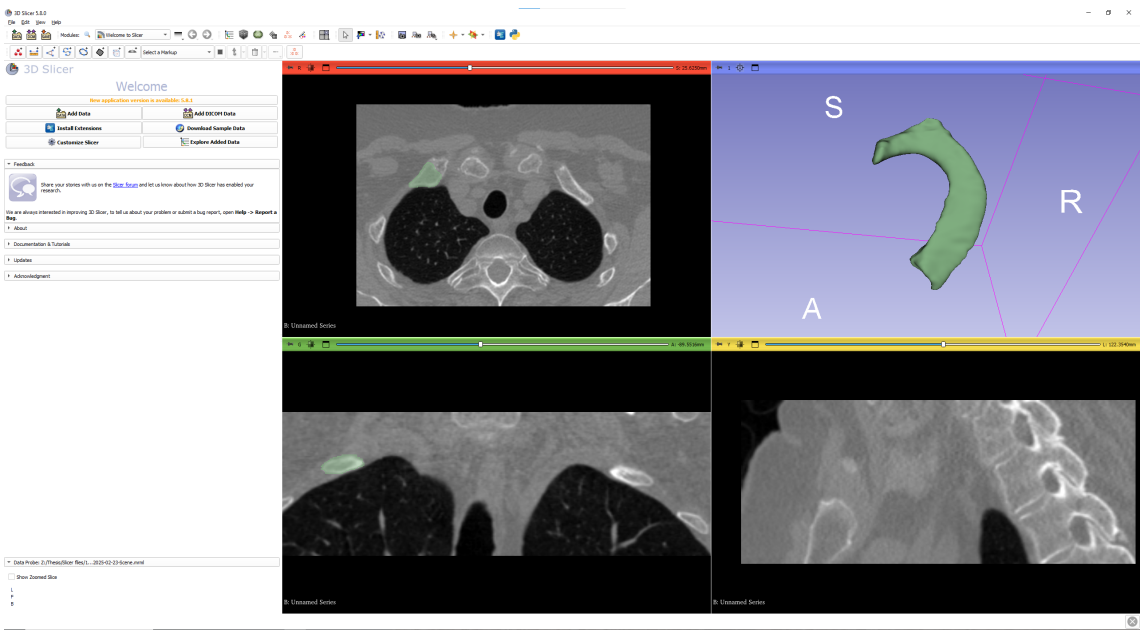


Figure 3.3: Segmenting the first rib in 3D-Slicer

Segmentation was performed in a semi-automatic manner using 3D-Slicer [26]. The DICOM images were loaded into the software, and the first human rib was manually segmented using the pencil and brush tools on every second slice. The slices in between were then interpolated using the "Fill between slices" option. The segmentation shape was inspected for irregularities, which were corrected manually in the corresponding slices using the eraser tool.

The DICOM images produced by the MATLAB script (see Section 3.1) do not contain the necessary metadata for pixel spacing and frame spacing. Therefore, as an additional step, each segmentation was resized in 3D-Slicer to ensure the

correct dimensions. This was achieved by first exporting the segmentation as a labelmap. The image and labelmap volumes were then resized in the "Volumes" module. Finally, the resized labelmap was re-imported as a segmentation, completing the resizing process.

The segmented surfaces were exported as .ply files, as this format is compatible with Stradview, used for cortical bone mapping in a subsequent step.

3.3 Cortical Bone Mapping

Cortical bone mapping was performed using Stradview [27] to estimate local cortical thickness across the first rib surface.

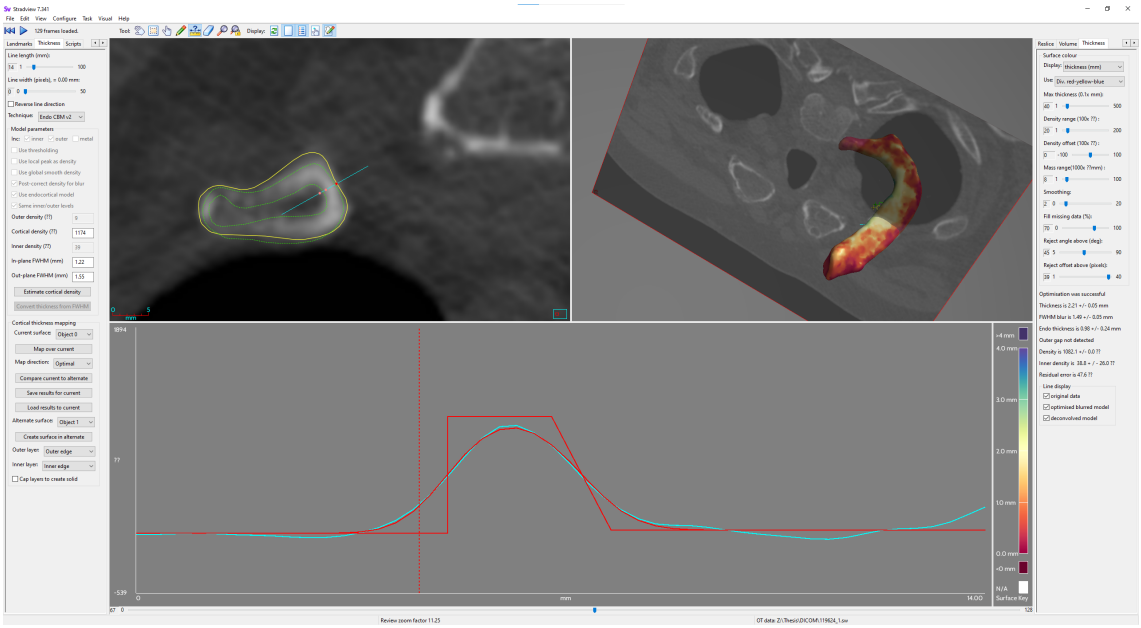


Figure 3.4: Mapping the cortical thickness of the rib in Stradview. The top-left pane displays the CT slices, allowing inspection of the detected outer and inner cortical surfaces. The top-right pane shows the 3D rib surface with cortical thickness mapped onto it. The bottom pane presents plots of Hounsfield units measured along the normals to the segmented surface

The segmented rib surfaces, obtained from 3D-Slicer, were imported into Stradview along with the corresponding CT image volumes. The surface resolution was set to "Very High", and "High" smoothing was applied to mitigate surface irregularities introduced during segmentation. Cortical thickness was estimated using the Endo CBM v2 technique [28], which detects both the outer and inner cortical boundaries along projected normals from the rib surface.

A normal line length of approximately 16 mm was initially used, following standard recommendations to exceed three times the expected cortical thickness [29]. Cortical density was estimated first to assess the reliability of the parameter

settings. The line length was adjusted, either increased or decreased, based on the resulting uncertainty in cortical density, and this process was repeated until a low uncertainty was achieved.

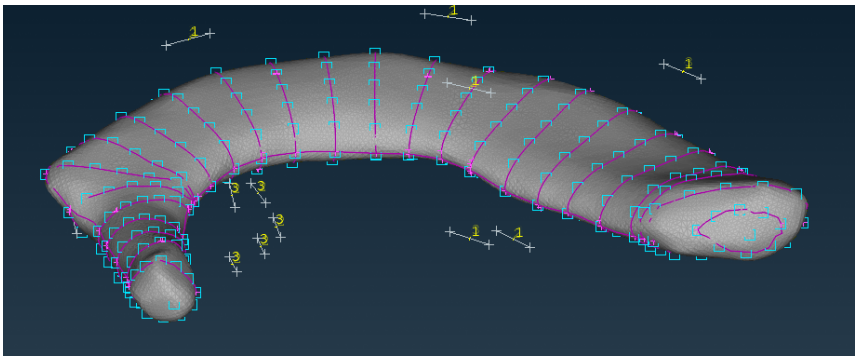
Once the uncertainty in the cortical density estimation was sufficiently low, the corresponding cortical thickness was mapped onto the imported surface. Outlier rejection parameters, such as "Reject Angle Above" and "Fill missing data", were manually tuned through trial and error, with adjustments guided by visual inspection of the outer and inner cortical surfaces to achieve accurate and continuous mapping. Smoothing was then applied to fill in regions where the algorithm failed to map thickness. The mapped outer surface, which represents the corrected "true" rib surface, was extracted and saved as a new geometry. Cortical thickness values were exported as binary point files, which were subsequently converted to CSV format via a custom Python script for integration into the shape modeling and statistical analysis pipeline.

3.4 Landmarking

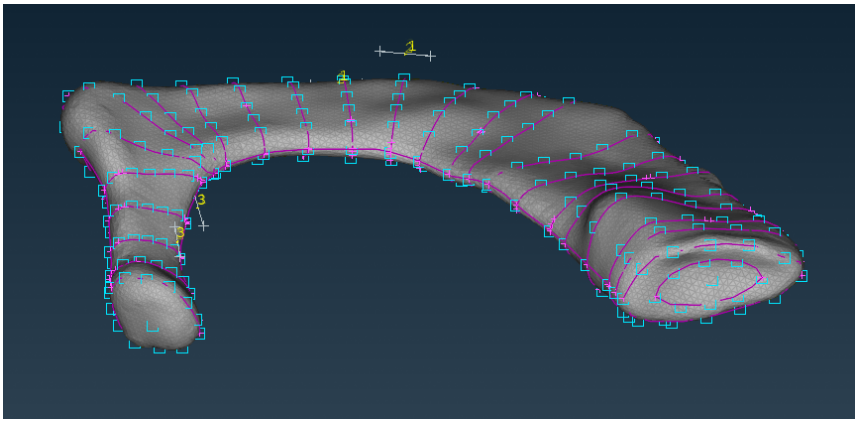
The first step in the landmarking process involved developing a scheme and then systematically applying it to both the template and all samples. The template used was a model of an average female first rib [30]. As this thesis used both left and right ribs, mirroring around the sagittal plane was used to get consistency between all samples and the template. The landmarking process then began with identifying anatomical points that could serve as consistent references across all rib samples. Three anatomically significant features were selected: the head, the tubercle, and the upper surface.

In addition to these anatomical landmarks, pseudo-landmarks were placed based on relative distances or angles. A combination of automatic placement in ANSA [31] and manual estimation was employed, with automatic placement preferred whenever it could enhance reproducibility. In cases where the output surface was only slightly cropped in the tubercle region, landmarks were manually adjusted. One example of such manual adjustment is shown in Appendix B.

Throughout this iterative process, the morphed template was compared with each landmarked rib to evaluate how well the landmarking captured the shape. Both qualitative and quantitative analyses were performed. The qualitative analysis primarily focused on ensuring that no morphing artifacts appeared and that the general shape was accurately represented. Quantitatively, morphing accuracy was measured in ANSA by comparing the maximum distance between the morphed surface and the corresponding landmarked surface. Figure 3.5 shows the landmarking scheme and a detailed procedure is described in Appendix A.



(a) The template mesh landmarked with the landmarking scheme.



(b) Example of a sample landmarked with the landmarking scheme.

Figure 3.5: Examples of landmarking using the scheme: (a) the landmarked template mesh, and (b) a landmarked sample mesh.

3.5 Morphing

The morphing was implemented using the Infepy Python library [32], which supports various radial basis function (RBF) methods for non-rigid transformations. In this study, the thin-plate spline (TPS) method was employed due to its suitability for smooth, biologically realistic deformations. A smoothing factor of 1 was applied to control the regularity of the transformation and to reduce potential overfitting to landmark noise.

Each subject's rib shape was defined by a set of manually placed anatomical landmarks. These landmarks were read from CSV files and used to compute the transformation from a reference (template) shape to the target geometry. The computed transformation was then applied to the nodes of the template mesh, producing a morphed mesh that retained the topological structure of the template while conforming to the target shape.

The output meshes were saved in the LS-DYNA keyword format (.key) format for

consistency with the python scripts for GPA, and visualization with ANSA. It was also essential to map the rib thickness onto the morphed rib, as a consistent number of nodal thickness values is imperative for comparison and statistical analysis. This mapping was performed in ANSA. Some examples of this mapping are shown in Appendix C.

3.6 Generalized Procrustes Analysis

To enable shape comparison across individuals, Generalized Procrustes Analysis (GPA) was performed on the morphed surface meshes of the first rib. The goal of this step was to bring all samples into a common coordinate system, ensuring that observed variation reflected true morphological differences.

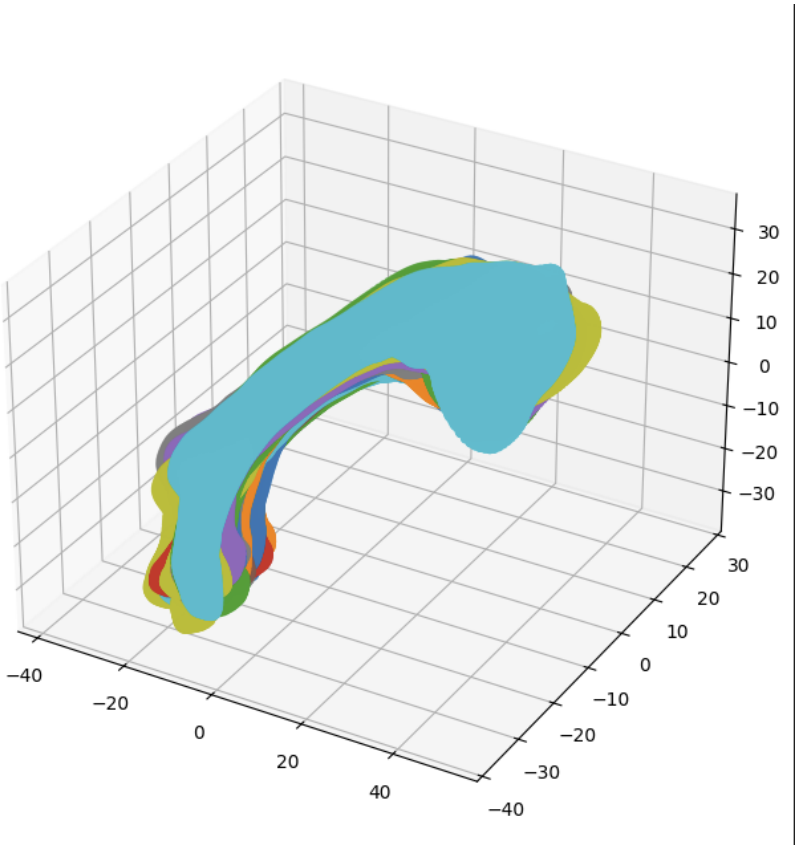


Figure 3.6: Surfaces after GPA

For each subject, anatomical landmarks were read from CSV files and aligned to a common template using a custom Python implementation of GPA. Since all ribs were mirrored to the right side during landmarking, reflection was not part of the alignment procedure. The algorithm computed optimal rigid transformations, translation and rotation, for each sample to minimize the root mean squared distance between corresponding landmarks from the template.

The resulting transformations were then applied to the full surface mesh of

each rib to register all geometries in a shared coordinate space. Following alignment, corresponding surface nodes across the dataset enabled point wise comparison of shape and associated cortical thickness values. Mean coordinates and mean cortical thickness were computed to generate an average rib geometry and thickness distribution.

Aligned meshes and mapped scalar values were exported in CSV format for further statistical analysis and visualization. 3D plots were used to qualitatively assess the alignment and distribution of anatomical features.

3.7 Principal Component Analysis

To investigate the dominant modes of morphological variation in the first rib, a Principal Component Analysis (PCA) was performed on the combined shape and cortical thickness data. For each subject, the 3D mesh coordinates were concatenated into a single vector of length $3N$, with N representing the number of surface nodes and the factor 3 accounting for the x, y, and z dimensions. Corresponding cortical thickness values were obtained from precomputed mapping results, sorted by node ID, and organized into vectors of length N. The thickness data were mean-centered across the sample to isolate subject-specific deviations while preserving the absolute magnitude of thickness variation. No further standardization was applied to either the shape or the thickness data to maintain the anatomical scale and to allow direct interpretation of reconstructed geometries.

Each subject's data were then concatenated into a single vector of length $4N$, combining both shape and mean-centered thickness. This resulted in a data matrix of size $M \times 4N$, where M is the number of subjects. PCA was conducted on this matrix using the scikit-learn [33] implementation. All principal components were initially computed, and the number of retained components was chosen to explain at least 95% of the total variance. This produced a low-dimensional representation of the data (PCA scores), the loading vectors for each principal component, and the corresponding explained variance ratios.

To visualize and interpret the principal modes, rib geometries was reconstructed corresponding to ± 2 standard deviations along each component. For each principal component, both the shape and thickness deviations were added to the mean configuration to generate hypothetical extreme geometries. These reconstructions were saved in both CSV and .key formats. Additionally, animated 3D scatter plots that transitioned smoothly between -2σ and $+2\sigma$ deviations were generated. These animations were rendered with node color maps corresponding to cortical thickness values, providing an intuitive representation of how shape and thickness vary jointly along each mode.

The cumulative variance explained by the components was plotted to assess the dimensionality of variation. PCA score scatterplots (e.g., PC1 vs. PC2) were also generated to inspect potential patterns or clustering within the sample.

3.8 Regression

This step used the Statsmodels toolkit [34] to calculate linear regression models. The sex parameter was defined as 0 for female and 1 for male. The other demographic variables (age, height, BMI) were normalized by dividing each by its average to simplify model evaluation.

After preparing the demographic data, the next step involved analyzing each principal component (PC) individually using backward elimination, as described in section 2.10, to identify the statistically significant demographic independent variables. All regression models included a intercept term. Whether or not the intercept term was statistically significant was not considered during the backward elimination process.

For each demographic variable, a scatter plot of the PC scores along with the corresponding linear regression line was generated. This was done to visually assess whether the absence of statistical significance could be attributed to poor model fit of linear regression.

This procedure was repeated for each PC individually. In cases where statistically significant demographic variables were identified, all the corresponding weights were saved for use in predicting and reconstructing new SSMs, as per Equation 2.3. No statistical model was made for PCs without any statistically significant biometric variables. These PCs were not used in the reconstruction process.

4

Results

4.1 Average Shape and Thickness

Following GPA, node-wise coordinates and cortical thickness values were averaged to generate a mean rib geometry with corresponding thickness distribution. The resulting visualization is presented in Figure 4.1.

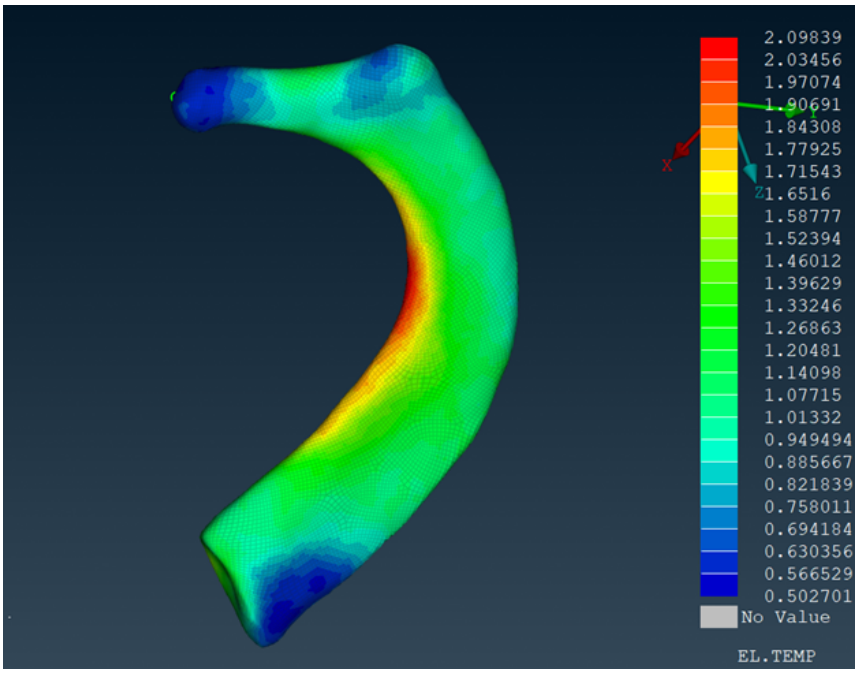


Figure 4.1: Average Shape and Thickness

4.2 Principal Component Analysis

PCA was applied to the combined shape and cortical thickness data of the first human rib to identify dominant modes of morphological variation. A total of **21** components were extracted, which explained 95% of total variance.

4.2.1 Explained Variance

Figure 4.2 illustrates the cumulative variance explained by successive components.

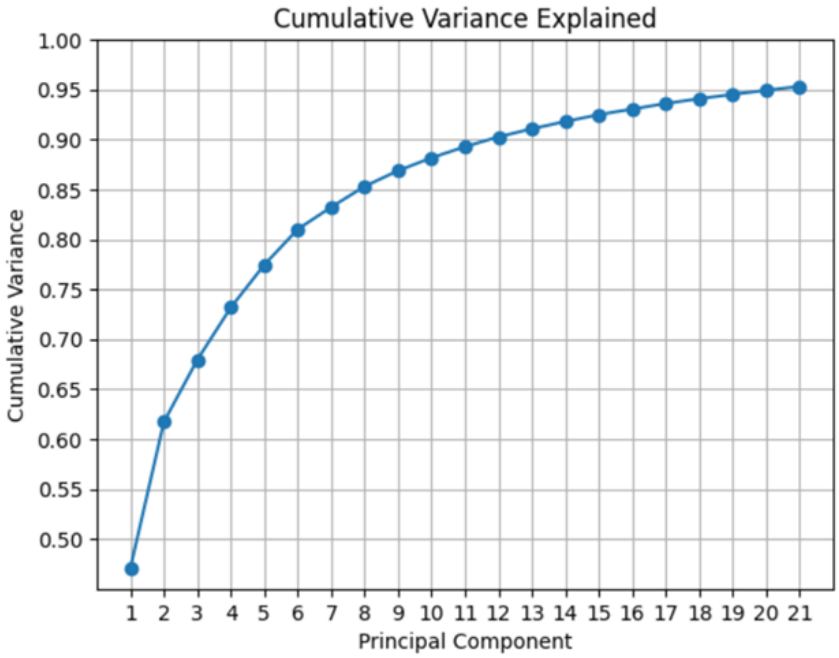


Figure 4.2: Cumulative variance explained by PCA of rib shape and thickness.

4.2.2 Modes of Variation captured by PCA

Visualizations of the principal components with statistically significant demographic variables at ± 2 standard deviations show how rib shape and thickness varies. The PCs without statistically significant demographic variables are visualized in Appendix D.

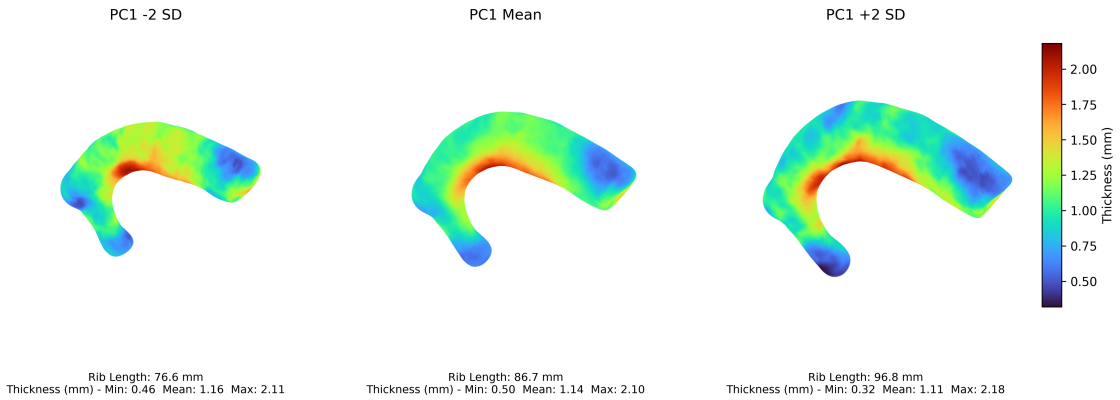


Figure 4.3: PC1 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

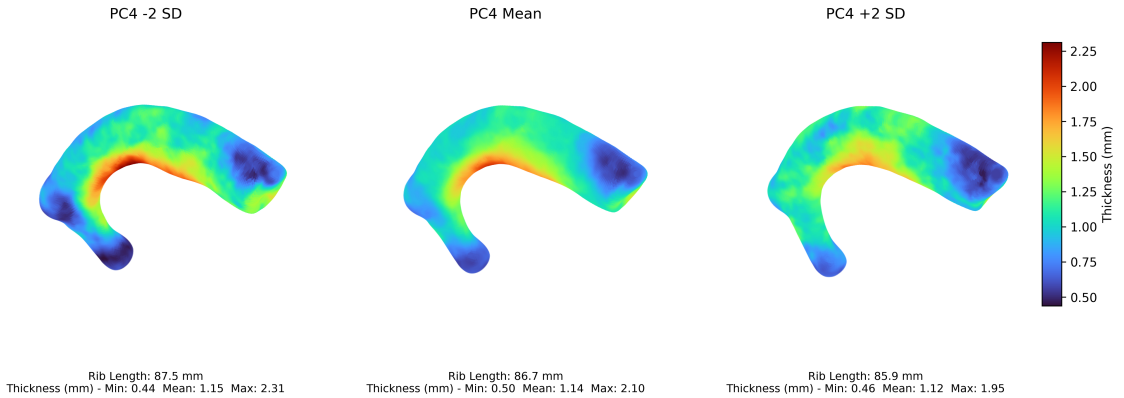


Figure 4.4: PC4 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

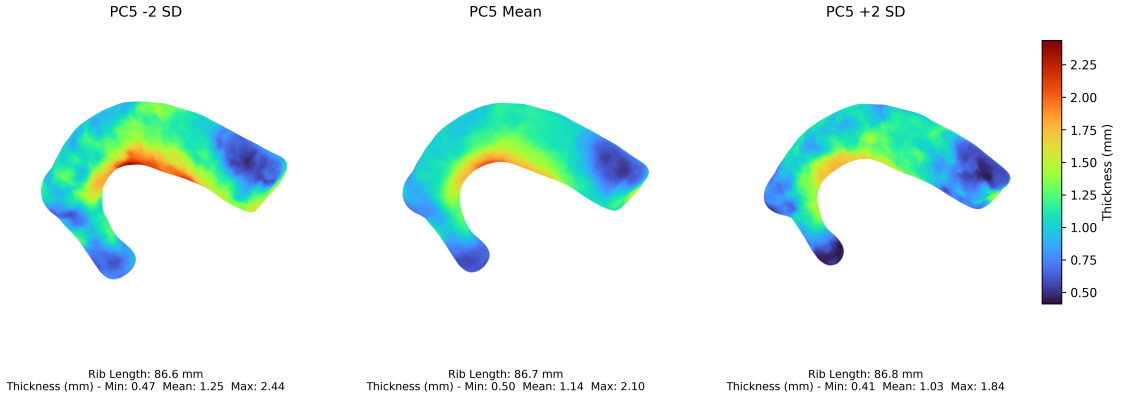


Figure 4.5: PC5 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

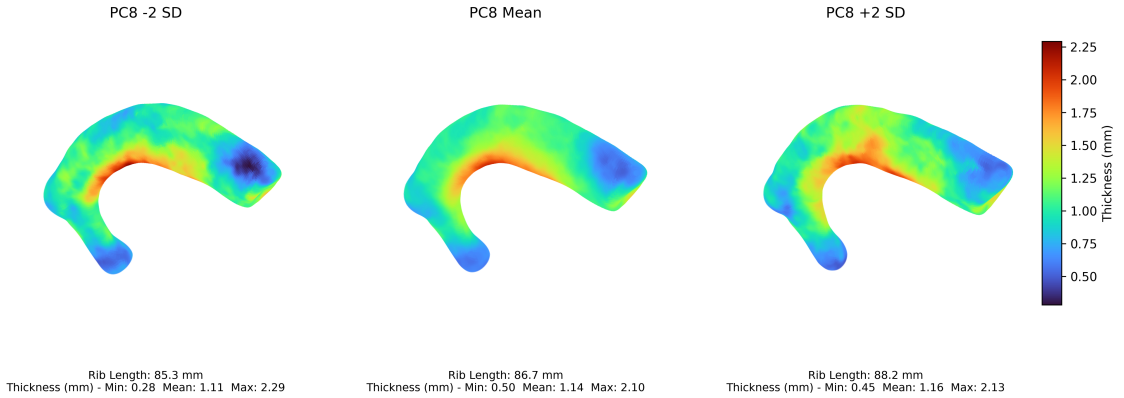


Figure 4.6: PC8 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

4.2.3 Score Distribution and Anatomical Trends

The scatter plot (Figure 4.7) of the first two principal components (PC1 and PC2) shows a continuous distribution of specimens without clear clustering, indicating that variation in rib shape occurs along a spectrum rather than discrete groups. A small number of specimens lie at the extremes of these components, representing notable deviations from the average rib morphology.

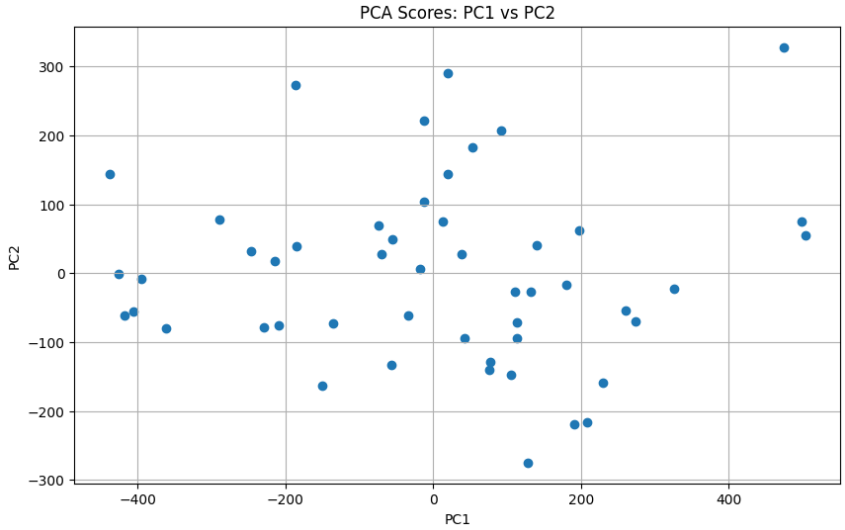


Figure 4.7: Scatter plot of PCA scores for first rib across PC1 and PC2.

4.3 Regression

Table 4.1: Regression coefficients, p-values, and adjusted R^2 for each principal component (PC).

PC		Age		Sex		Height		BMI		R^2_{adj}
	w_0	w_1	p	w_2	p	w_3	p	w_4	p	
1	-2477.5	349.4	0.002	–	–	2128.0	0.000	–	–	0.314
2	–	–	–	–	–	–	–	–	–	–
3	–	–	–	–	–	–	–	–	–	–
4	17.8	–	–	-44.4	0.0049	–	–	–	–	0.059
5	867.4	–	–	–	–	-791.1	0.000	-76.3	0.049	0.398
6	–	–	–	–	–	–	–	–	–	–
7	–	–	–	–	–	–	–	–	–	–
8	63.8	–	–	–	–	–	–	-63.8	0.050	0.058
9+	–	–	–	–	–	–	–	–	–	–

Table 4.1 shows the regression coefficients (w_1, w_2, w_3 and w_4) and their respective p-values for statistically significant and normalized independent variables. It also includes the intercept (w_0) for each model. The R^2_{adj} values for each linear regression model are stated to quantify the predictive power for each PC.

When this predictive power is considered alongside the explained variance in Figure 4.2, it provides insight into how much of the total shape variance is associated with the independent variables. Together, the statistically significant PCs explain about 15% of the total variance.

PC1 and PC5 have the highest predictive power based on R^2_{adj}. PC1 is positively correlated with both age and height, with the coefficient for height being approximately six times larger. However, it should be noted that normalized age has greater variance than height. PC5 is negatively correlated with height and is therefore positively associated with higher rib thickness, as illustrated in Figure 4.5.

The remaining principal components with statistically significant independent variables, PC4 and PC8, have much lower predictive power, with $R^2_{\text{adj}} \approx 0.06$. Being male is associated with a lower PC4 score, while a higher BMI is associated with a lower PC8 score.

These results suggest height followed by age are the most important variables while BMI and sex have statistically significant predictive power, but are overall much weaker predictors.

4.4 Predictions

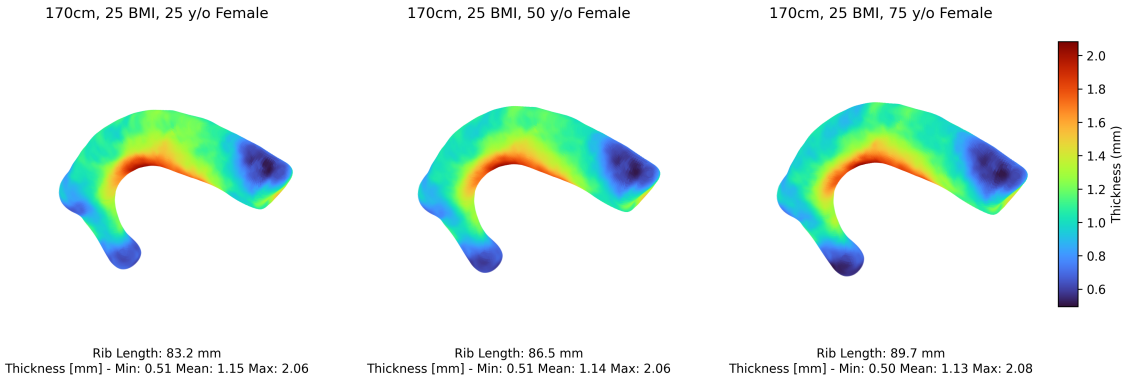


Figure 4.8: Age effect: 25 years (left), 50 years (center), 75 years (right), with cortical thickness mapped

In Figure 4.8, predictions are made for individuals aged 25, 50, and 75 years, with height, BMI and sex held constant (170 cm, BMI 25, female). Rib length increases from 83.2 mm to 89.7 mm (approximately 8%), while thickness changes only marginally.

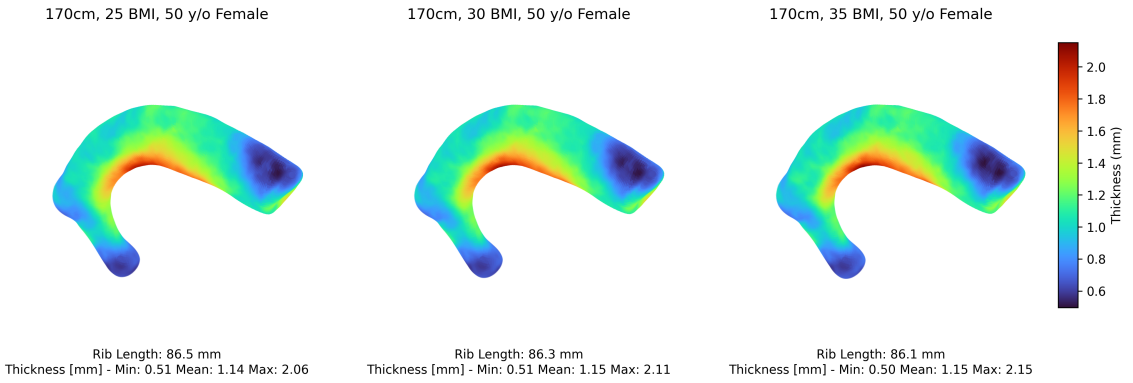


Figure 4.9: BMI effect: 25 BMI (left), 30 BMI (center), 35 BMI (right), with cortical thickness mapped

Figure 4.9 shows predictions for BMIs of 25, 30, and 35, with height, age, and sex held constant (170 cm, 50 years old, female). The rib length and average thickness show only marginal changes. However, maximum thickness increases from 2.06 mm to 2.15 mm (approximately 4%).

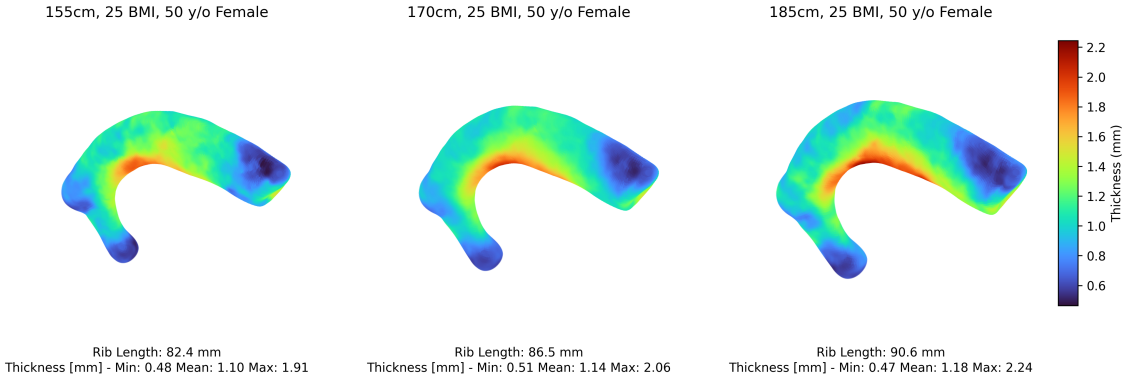


Figure 4.10: Height effect: 155 cm (left), 170 cm (center), 185 cm (right), with cortical thickness mapped

Figure 4.10 presents predictions for statures of 155 cm, 170 cm, and 185 cm, with age, BMI and sex held constant (50 years old, BMI 25, female). Rib length increases from 82.4 mm to 90.6 mm (approximately 10%). The average thickness increases from 1.10 mm to 1.18 mm (approximately 7%), and the maximum thickness from 1.91 mm to 2.12 mm (approximately 11%).

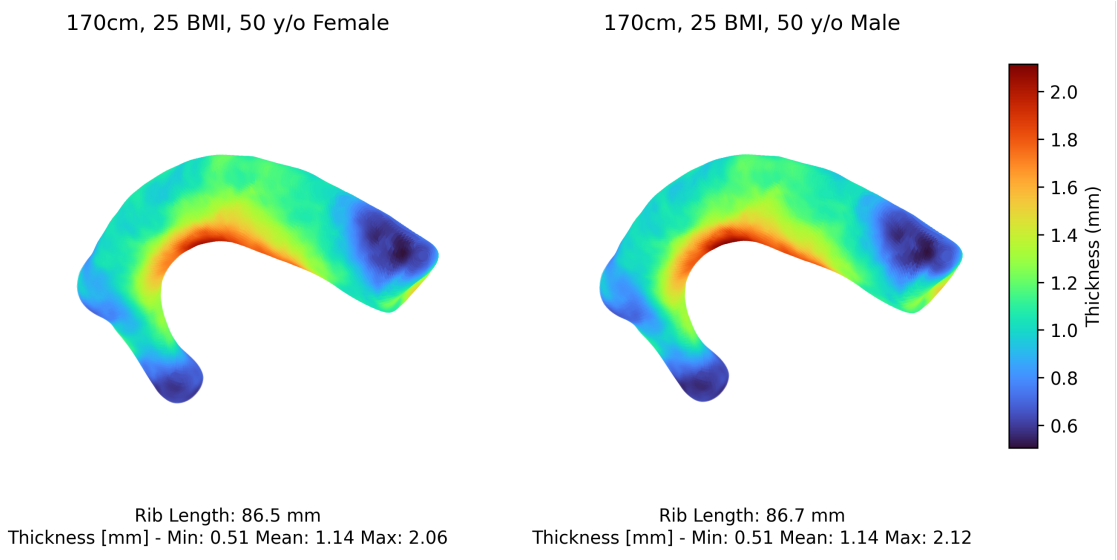


Figure 4.11: Sex effect: Female (left), Male (right), with cortical thickness mapped

Finally, Figure 4.11 compares predictions for a male and a female subject, with height, BMI, and age held constant (170 cm, 25 BMI, 50 years old). The differences in rib length, minimum thickness, and average thickness are marginal. However, maximum thickness increases from 2.06 mm for a female to 2.14 mm for a male (approximately 4%).

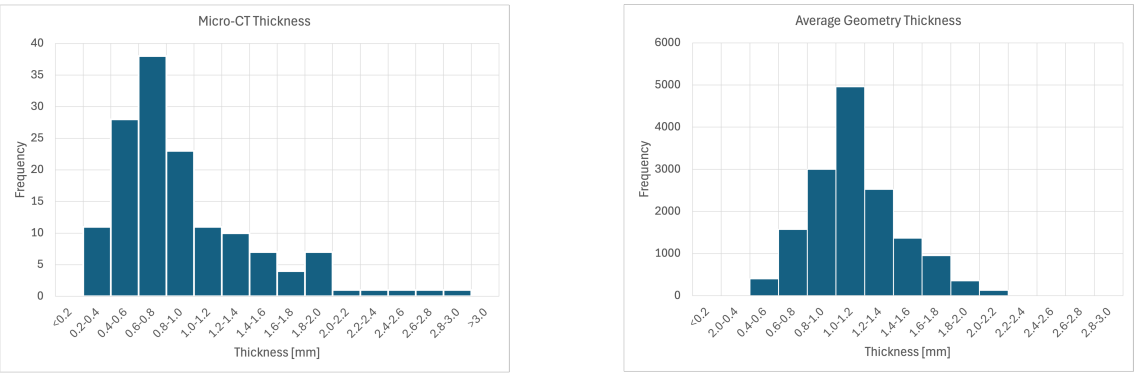
5

Discussion

5.1 Overview

The statistical shape model of the first rib revealed that overall size and cortical thickness variation dominate the morphological variability, with more subtle deformations (curvature, tubercle prominence and torsion) represented by higher-order PCs which were not statistically significant. Specifically, PC1; capturing global scaling of rib length and cortex distribution; accounted for the largest variance. The takeaway of this analysis was that taller individuals have longer, generally thicker ribs. Older individuals also have larger ribs, but lacks the thickness increase seen with height. Instead, older individuals have a marginal decrease in cortical thickness. Notably, PC1 and PC5 were most strongly linked to body size (height) and age in our sample (adjusted $R^2 \approx 0.31$–0.40), whereas PCs that encode curvature, tubercle size, and twist showed weaker demographic associations. This pattern is consistent with literature: taller, heavier individuals develop larger bones [7], and males in general have larger and thicker ribs [7][35].

However, it is worth noting that the average thickness in our model is higher than in the micro-CT data reported in a previous study based on seven post mortem human subjects (PMHS) [36]. The variance is also more compressed, with a higher minimum thickness and a lower maximum thickness.



(a) Micro CT thickness.

(b) Average geometry thickness.

Figure 5.1: Thickness distribution from micro-CT data [36] compared to the thickness distribution from the average geometry.

Table 5.1: Comparison of maximum, average, and minimum rib thickness, and demographic variables between micro-CT and average geometry.

Metric	Micro-CT	Average geometry
Maximum thickness (mm)	2.82	2.10
Average thickness (mm)	0.92	1.14
Minimum thickness (mm)	0.22	0.50
Average age (years)	71	59
Average height (cm)	51	83

Figure 5.1 and Table 5.1 show a comparison between the demographic and thickness data derived from the average geometry (see Section 4.1) and micro-CT data [36].

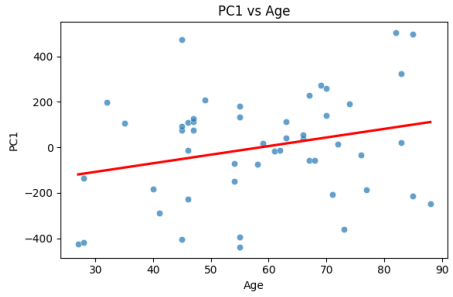
The micro-CT data is based on nine cross-sections, each containing 16 measurement points. Each value represents the average thickness at the corresponding point, calculated across several rib samples. In contrast, the thickness analysis from the average geometry is based on the thickness value at each individual node.

This compression of variance is likely explained by our method of node-wise thickness averaging, as discussed in Section 5.4. Notably, the results visualized prior to averaging in Appendix C show variance more comparable to the micro-CT findings, further suggesting that the averaging process contributes to the reduced variance.

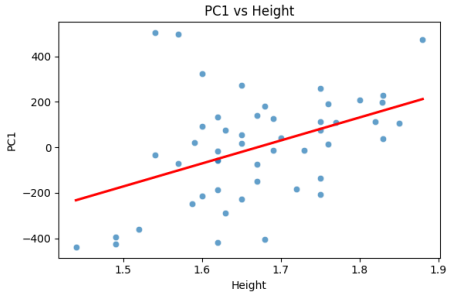
It also remains unclear why the average thickness differs between the micro-CT data and our thickness model. The consistent difference across all data points indicates a potential systematic bias, either in one or both datasets, in our method or exclusion criteria, or due to the limited coverage of the micro-CT data, which is based on only nine cross-sections.

The demographic variables was to a large degree similar between the PHMS in the micro-CT and the demographics used in this thesis. The average weight was 50.9 kg for the PMHS, compared to 83 kg in this study. Although dehydration may reduce body weight relative to that of living individuals, this alone is unlikely to account for the observed discrepancy. However, the results of this thesis suggest that BMI have only a marginal effect and the height is similar. Therefore, it seems unlikely that the differences can be explained solely by these demographic variables.

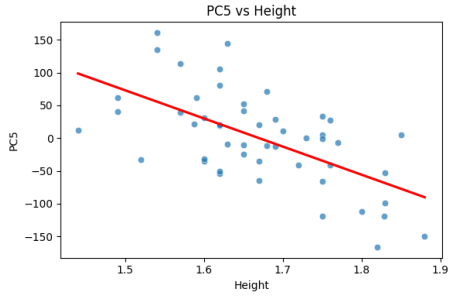
5.2 Model Results



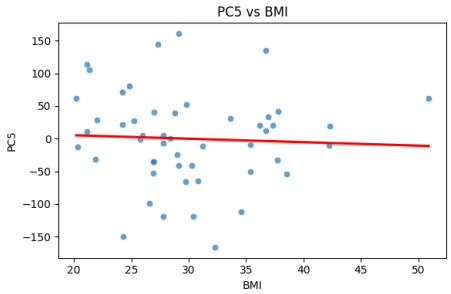
(a) PC1 scores vs Age with regression fit



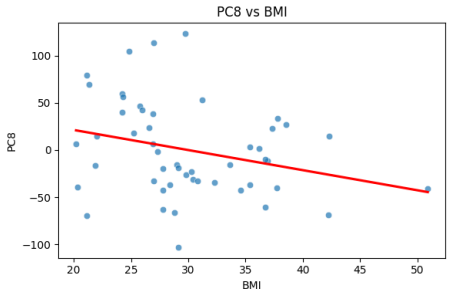
(b) PC1 scores vs Height with regression fit



(c) PC5 scores vs Height with regression fit



(d) PC5 scores vs BMI with regression fit



(e) PC8 scores vs BMI with regression fit

Figure 5.2: Regression fits for selected principal components (PCs) with significant predictors: (a) PC1 vs Age, (b) PC1 vs Height, (c) PC5 vs Height, (d) PC5 vs BMI, (e) PC8 vs BMI.

The regression findings further suggest that first rib length and cortical thickness tend to increase with height, BMI, and male sex. These trends mirror population-level analyses of ribs 2–11: for example, Holcombe et al. showed that male ribs have larger cross-sectional areas (14–29% larger total area and up to 18% larger cortical area) than female ribs of equivalent weight and stature [7].

In our sample, older age correlated with higher PC1 (indicating a slightly

larger but thinner first rib). This leads to a marginal cortical thinning with increased age, but not as large as other rib studies report (on the order of 0.02 to 0.04 mm per decade [37], and a decrease in the cortical area in advanced age [7]. The adjusted R^2 values for age and height ($\approx$ 0.31–0.40) indicate that much of the shape variation is not simply explained by these linear trends. Other factors; genetic, developmental or lifestyle; likely contribute to the remaining variability. The PC's linked to shape (for example, PC2 and PC3), show only weak statistical links to demographic variables. This suggests the finer details primarily associated to rib shape vary considerably between individuals regardless of age, sex or size.

5.3 Methodological Considerations

This section discusses key methodological choices and their rationale.

Segmentation was performed using a semi-automatic approach in 3D-Slicer to balance accuracy and efficiency. Fully automatic segmentation was not feasible due to the presence of adjacent bony structures with similar densities, which would require manual intervention, and manual segmentation alone was too time-consuming for the dataset size. Therefore, manual annotation on every second slice combined with interpolation was used, followed by manual corrections to ensure quality. The segmented volumes were resized to correct spatial dimensions before export.

Regarding CBM, since there was no consistent range or trend of outlier rejection parameters that worked reliably across all samples, trial and error guided by visual inspection was employed to achieve accurate and continuous mapping.

Among the exclusion criteria (Section 3.1.2), the removal of 15 samples due to cortical bone mapping failures requires further clarification.

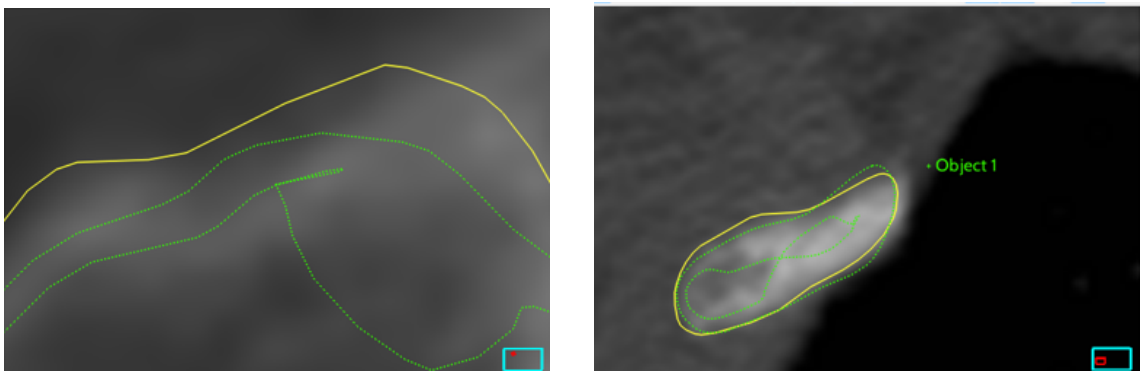


Figure 5.3: Examples of cortical bone mapping artifacts that led to exclusion from the dataset.

Specifically, these were cases where the inner or outer boundary estimations were folded or intersecting, regardless of how the outlier rejection parameters were tuned.

These errors were likely related to insufficient image resolution and the software's inability to detect a clear boundary. These mapping artifacts, such as spikes or missing regions, compromised surface quality and could not be corrected, and the affected samples were therefore excluded to preserve model integrity. Examples of such cases are shown in Figure 5.3.

Before applying PCA, both coordinate and thickness values were normalized with respect to scale and variance. This ensured that the relatively small thickness values (approximately 1 mm) did not become proportionally insignificant compared to the larger coordinate values (typically tens of millimeters).

In the regression model, the independent variables, excluding sex, were normalized in regards to magnitude to enable more intuitive comparisons of coefficients. This normalization allows the magnitude of each coefficient to reflect how much a change in the predictor affects the PC. For example, it helps compare the influence of a 10% increase in BMI versus a 10% increase in height on a particular PC.

In addition to improving interpretability, normalization also reduces the disproportionate influence of variables with large numerical scales, such as age, on the optimization of the RSS. This leads to a more balanced estimation of regression coefficients and ensures that predictors with different units or ranges are more equitably weighted.

Finally, the regression fits (Figure 5.2) showed no evidence of nonlinear relationships, supporting the use of linear models for the demographic variables analyzed.

5.4 Limitations

The results presented in this study align broadly with previous research on rib morphology. The observed relationship between rib size and factors such as height and body mass is consistent with expectations and supports trends seen in recent CT-based analyses [7], [35]. Nevertheless, several assumptions and methodological choices must be acknowledged.

One simplification made in this thesis is the assumption of bilateral symmetry between left and right ribs. This approximation was not empirically validated within the scope of the current work.

In contrast to the detailed cortical mapping performed by Holcombe et al., which demonstrated region-specific thinning, particularly among middle-aged females, this study relies on node-wise averaging of cortical thickness [37]. As a result, our principal component models reflect broader trends in thickness distribution but are not capable of identifying localized changes across the rib surface. This limitation is especially apparent in the principal components derived from PCA, where fine structural variation becomes smoothed during dimensionality reduction.

Another factor contributing to the loss of regional detail stems from the use of CBM processing in Stradview. During cortical thickness mapping, unmapped ("gray") regions often appeared as a result of the chosen settings for outlier rejection parameters. To compensate, smoothing algorithms interpolate values from adjacent regions, which can obscure subtle but potentially important variations in thickness. No consistent pattern was observed in the location of these unmapped regions.

Regarding statistical analysis, a linear regression model was selected to maintain interpretability and to reduce the risk of overfitting, given the small sample size and high variability. However, this choice excludes the possibility of capturing more complex, non-linear relationships, as no alternative modeling approaches were tested.

Given that this thesis suggests age is the second most important demographic variable (after height) in relation to rib shape variance, it is important to note the lack of younger individuals in the dataset. This limitation could result in a biased model. Of the 50 subjects in the study, only ten were between 20 and 40 years old, and none were under the age of 20. These groups, including younger adults and children, need to be analyzed separately.

It is well established that the mechanical properties of the chest cage and associated injury mechanisms change with age. In younger individuals, the ribcage is highly flexible, but it becomes stiffer as people age [38]. In younger subjects, the chest can be compressed all the way to the spine, increasing the risk of internal organ injuries. In contrast, in older individuals, the ribs may fracture before such a compressed state is reached, leading to a loss of structural integrity in the ribcage.

While these fractures may result from factors beyond rib shape and thickness, such as changes in bone properties and cartilage stiffness [38], the shape variability captured in this study may still have a meaningful impact on biomechanical modeling.

Generalizing the model to children and adolescents introduces additional uncertainties. Growth may introduce age-related effects that cause the model to behave non-linearly, unlike the gradual, more linear changes observed across adult age groups. The body proportions are also different in children compared to the adult population [39]. This may further contribute to non-linearities in the influence of height. Although only adults were included in this study, these considerations provide evidence against the effective generalization of the model to pediatric populations.

Lastly, while cortical thickness values were incorporated into the model as scalar values assigned to mesh nodes, it remains unclear whether this representation sufficiently reflects the mechanical behavior of the ribs or is robust enough to support applications in future biomechanical or clinical contexts.

5.5 Future Work

The SSM for the first rib, parameterized by age, height, BMI, and sex, can be enhanced to improve its accuracy by including additional parameters like lifestyle. Extending the model to include children and the elderly is critical, as rib geometry changes significantly across life stages. Incorporating CT data from pediatric and geriatric populations, as exemplified by Holcombe et al.'s pediatric rib model [40], would ensure the SSM better represents the full range of anatomical variation across age groups.

The results produced in this thesis are based on a single dataset with subjects from a midwestern US population. Holcombe et al. [7], stated that future work should be done to compare the results to other adults populations, the same goes for this thesis. This thesis, specifically would also benefit from a larger sample size. A total of 50 ribs was used in this thesis based on our exclusion criteria (Section 3.1.2), while Holcombe et al. [7] used 240 subjects.

The thickness related results would also benefit from future investigations due observed discrepancies compared to the previous studies. A new method without systematic errors introduced by the node-wise averaging may be needed.

Integrating the SSM into human body models (HBMs) like VIVA+ or THUMS could enhance crash simulations. Morphing HBM ribs with SSM parameters could improve injury predictions [41] supporting safer vehicle designs.

In addition to using regression models to predict rib shape from demographic variables, future work could explore sampling directly from the distribution of PC scores. This approach would allow for the generation of synthetic rib geometries that span a broader and more diverse range of anatomical variation.

5.6 Ethical Considerations

In this thesis, two main areas of ethical consideration have been analyzed. The first concerns the dataset and the subjects involved, while the second addresses potential misuse of the results produced in this work.

5.6.1 Datasets

The dataset used in this thesis originates from the University of Michigan and has been employed in previous studies (approval HUM00041441) [25]. Before it was provided, the scans were cropped to include only the region around the first rib, and no facial features were included.

In addition to cropping the volume, all personally identifiable metadata, such as names or hospital records, was removed to ensure compliance with the GDPR [42]. The remaining demographic variables, age, sex, height, weight, and BMI,were

retained and processed to enable the development of the regression models outlined in the thesis aims.

5.6.2 Misuse of Results

The SSM developed in this study is intended to reduce reliance on anthropomorphic test devices (ATDs) and to represent population-level variability. The aim is to generalize across a wide range of individuals. However, it is important to note that the model may still not provide accurate predictions at the individual level. This raises ethical concerns, particularly in safety-critical applications, where over reliance on such models could inadvertently lead to harm for outliers.

The removal of metadata beyond the included demographic variables was necessary to protect patient privacy and comply with regulations such as GDPR. However, this also introduces challenges when evaluating whether the model might contain unintended biases. This is a compromise between patient privacy and the possibility of conducting bias analysis. Nevertheless, the retained demographic variables are widely recognized in the literature (for example [43][7][1]) as predictors is SSMs.

Beyond their intended use in improving safety, SSMs may also have applications in other domains, such as forensic science. For example, they could be used in the opposite direction and predict demographic variables from skeletal data [44]. Although such use may appear benign, it carries significant ethical risks, particularly if profiling is based on biased datasets, which could lead to inaccurate or unfair conclusions.

There are also a potential dual-use of SSMs. Although this thesis focuses on reducing harm in motor vehicle crashes, SSMs could, with minor modifications, be repurposed for harmful applications. This dual-use dilemma is a well-documented concern in life sciences and related fields [45]. For instance, SSMs could hypothetically be used to design non-lethal weapons that exploit injury simulations to maximize incapacitation. While such consequences fall outside the scope of this thesis, their possibility highlights the importance of maintaining ethical awareness.

Finally, there is a historical dimension to research involving human anatomy. In the 19th and early 20th centuries, misused anthropometric studies were employed to support pseudoscientific theories of race and eugenics [46]. While these theories have since been discredited, their legacy underscores the importance of applying thoughtful methodology and ethical practices in contemporary anatomical research.

5.7 Contributions

We received scripts for converting .mat files to .dcm from Oscar Hallberg. Scripts for morphing and GPA were provided by Chiara Fichera [32].

5.8 Declaration of Authorship

We hereby declare that this Master's thesis was composed by us as a group of two authors, and that the work presented is our own except where explicitly stated otherwise. This thesis has not been submitted for any other degree or professional qualification, nor has it been published. All sources of information, data, or ideas from others; whether from printed materials, online sources, or elsewhere; have been properly acknowledged and referenced.

During the writing process, we used ChatGPT (GPT-4) [47] to assist with improving the language, flow, and readability of the text. All content generated with the help of this tool was critically reviewed and edited by us to ensure accuracy and alignment with our research. We take full responsibility for the final content of the thesis in its entirety.

6
Conclusion

This thesis set out to develop a SSM of the human first rib using CT scan data, aiming to better understand how rib morphology varies across individuals. By applying PCA, we found that the most significant differences in rib shape are related to overall size and cortical thickness. These variations, especially those captured by the first principal component, were strongly linked to demographic factors like height and age. On the other hand, more subtle shape features, such as curvature and torsion, were captured in later components and showed weaker correlations with the available anthropometric data.

The findings are largely in line with existing literature, which has shown that larger individuals tend to have larger bones, and that males generally have thicker ribs. However, the model does not capture local variations in thickness as precisely as some previous studies that used high-resolution micro-CT or detailed cortical mapping. This is likely due to the node-wise averaging and smoothing used in our processing pipeline, which may have come at the cost of capturing finer details.

There are also some important limitations to acknowledge. Our dataset, while representative in many respects, did not include children or very elderly individuals, and only a limited set of anthropometric parameters was available. These factors likely explain some of the unexplained variance in the model and point to the need for broader data in future work.

That said, this SSM offers a practical starting point for improving biomechanical models, particularly human body models (HBMs) used in crash simulations. Being able to adapt rib shape based on demographic inputs makes these simulations more realistic and inclusive. With further refinement, such as incorporating more varied data or combining this model with finite element analysis, there is real potential to improve safety systems by tailoring them to the anatomical differences seen across the population.

In summary, this work contributes a new perspective on first rib anatomy and provides a flexible tool for further research and application. As data-driven models continue to grow in importance across engineering and biomedical fields, efforts like this one help ensure that they are grounded in real-world variation.

Bibliography

[1] Y. Wang, L. Cao, Z. Bai, *et al.*, "A parametric ribcage geometry model accounting for variations among the adult population," *Journal of Biomechanics*, vol. 49, no. 13, pp. 2791–2798, Sep. 2016, ISSN: 0021-9290. DOI: 10.1016/j.jbiomech.2016.06.020. [Online]. Available: http://dx.doi.org/10.1016/j.jbiomech.2016.06.020.

[2] J. Peek, Y. Ochen, N. Saillant, *et al.*, "Traumatic rib fractures: A marker of severe injury. a nationwide study using the national trauma data bank," *Trauma Surgery amp; Acute Care Open*, vol. 5, no. 1, e000441, Jun. 2020, ISSN: 2397-5776. DOI: 10.1136/tsaco-2020-000441. [Online]. Available: http://dx.doi.org/10.1136/tsaco-2020-000441.

[3] T. Xu, X. Sheng, T. Zhang, H. Liu, X. Liang, and A. Ding, "Development and validation of dummies and human models used in crash test," *Applied Bionics and Biomechanics*, vol. 2018, pp. 1–12, Nov. 2018, ISSN: 1754-2103. DOI: 10.1155/2018/3832850. [Online]. Available: http://dx.doi.org/10.1155/2018/3832850.

[4] E. Haug, H.-Y. Choi, S. Robin, and M. Beaugonin, "Human models for crash and impact simulation," in *Computational Models for the Human Body*. Elsevier, 2004, pp. 231–452, ISBN: 9780444515667. DOI: 10.1016/s1570-8659(03)12004-2. [Online]. Available: http://dx.doi.org/10.1016/S1570-8659(03)12004-2.

[5] E. A. Audenaert, C. Pattyn, G. Steenackers, J. De Roeck, D. Vandermeulen, and P. Claes, "Statistical shape modeling of skeletal anatomy for sex discrimination: Their training size, sexual dimorphism, and asymmetry," *Frontiers in Bioengineering and Biotechnology*, vol. Volume 7 - 2019, 2019, ISSN: 2296-4185. DOI: 10.3389/fbioe.2019.00302. [Online]. Available: https://www.frontiersin.org/journals/bioengineering-and-biotechnology/articles/10.3389/fbioe.2019.00302.

[6] E. S. Yee, A. N. Thomas, and P. C. Goodman, "Isolated first rib fracture: Clinical significance after blunt chest trauma," *Annals of Thoracic Surgery*, vol. 49, no. 6, pp. 887–891, 1990. DOI: 10.1016/0003-4975(90)90115-L.

[7] S. A. Holcombe, Y. Huang, and B. A. Derstine, "Population trends in human rib cross-sectional shapes," *Journal of Anatomy*, vol. 244, no. 5, pp. 792–802, Jan. 2024, ISSN: 1469-7580. DOI: 10.1111/joa.13999. [Online]. Available: http://dx.doi.org/10.1111/joa.13999.

[8] R. Drake, A. W. Vogl, and A. Mitchell, *Gray's anatomy for students* (Gray's Anatomy), en, 3rd ed. Churchill Livingstone, Jan. 2014.

[9] Henry Gray. "First rib - gray's anatomy illustration." Public domain image from Gray's Anatomy via GetArchive.net. (1918), [Online]. Available: https://itoldya420.getarchive.net/amp/media/first-rib-gray-67ee3b (visited on 05/29/2025).

[10] Physiopedia contributors, *First rib*, https://www.physio-pedia.com/First_Rib, Accessed: 2025-06-14, 2024.

[11] F. Ambellan, H. Lamecker, C. von Tycowicz, and S. Zachow, "Statistical shape models: Understanding and mastering variation in anatomy," in *Biomedical Visualisation*. Springer International Publishing, 2019, pp. 67–84, ISBN: 9783030193850. DOI: 10.1007/978-3-030-19385-0_5. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-19385-0_5.

[12] J. Enderle and J. Bronzino, *Introduction to Biomedical Engineering* (Biomedical Engineering), en, 3rd ed. Academic Press, Mar. 2015.

[13] B. Preim and C. Botha, "Chapter 4 - image analysis for medical visualization," in *Visual Computing for Medicine (Second Edition)*, B. Preim and C. Botha, Eds., Second Edition, Boston: Morgan Kaufmann, 2014, pp. 111–175, ISBN: 978-0-12-415873-3. DOI: https://doi.org/10.1016/B978-0-12-415873-3.00004-3. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780124158733000043.

[14] A. Garcia-Garcia, S. Orts, S. Oprea, V. Villena Martinez, and J. Rodríguez, "A review on deep learning techniques applied to semantic segmentation," Apr. 2017. DOI: 10.48550/arXiv.1704.06857.

[15] G. Treece and A. Gee, "Cortical bone mapping: Measurement and statistical analysis of localised skeletal changes," *Current Osteoporosis Reports*, vol. 16, Oct. 2018. DOI: 10.1007/s11914-018-0475-3.

[16] C. Fichera and O. Hallberg, *From ct scans to statistical shape models*, PowerPoint presentation, Jan. 2025.

[17] E. Hwang, J. Hallman, K. Klein, J. Rupp, M. Reed, and J. Hu, "Rapid development of diverse human body models for crash simulations through mesh morphing," Apr. 2016. DOI: 10.4271/2016-01-1491.

[18] J. Gower, "Generalized procrustes analysis," *Psychometrika*, vol. 40, no. 1, pp. 33–51, 1975, Cited by: 2536. DOI: 10.1007/BF02291478. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0001886818&doi=10.1007%2fBF02291478&partnerID=40&md5=c7113fae6e77ffaebc62f66ce8c103be.

[19] F. J. Rohlf and D. Slice, "Extensions of the procrustes method for the optimal superimposition of landmarks," *Systematic Zoology*, vol. 39, no. 1, pp. 40–59, 1990, Cited by: 3528. DOI: 10.2307/2992207. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84860428591&doi=10.2307%2f2992207&partnerID=40&md5=74ce7a85fe27c8da218b0b8e0fbb6d49.

[20] I. Jolliffe, *Principal component analysis*. New York: Springer Verlag, 2002.

[21] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer, 2002.

[22] C. M., *Pattern Recognition and Machine Learning* (Information Science and Statistics), en, 1st ed. New York, NY: Springer, Aug. 2006.

[23] D. Team, *Adjusted r-squared in regression analysis: A guide*, Accessed: 2025-05-29, 2023. [Online]. Available: https://www.datacamp.com/tutorial/adjusted-r-squared.

[24] A. Hebbali, *Olsrr: Tools for building ols regression models*, May 2017. DOI: 10.32614/cran.package.olsrr. [Online]. Available: http://dx.doi.org/10.32614/CRAN.package.olsrr.

[25] J. Iraeus, *Private communication via email*, Discussed details datasets and ethical considerations, Jan. 2025.

[26] 3D Slicer Community, *3D Slicer*, https://www.slicer.org/, Accessed: 2025-05-29, 2024.

[27] *Machine Intelligence Laboratory*. [Online]. Available: https://mi.eng.cam.ac.uk/Main/StradView.

[28] G. Treece and A. Gee, "Cortical bone mapping: Measurement and statistical analysis of localised skeletal changes," *Current Osteoporosis Reports*, vol. 16, Oct. 2018. DOI: 10.1007/s11914-018-0475-3.

[29] G. M. Treece, *How-to: Cortical bone mapping in stradview*, Accessed: 2025-05-25, 2025. [Online]. Available: https://mi.eng.cam.ac.uk/~gmt11/stradview/howto_cbm.htm.

[30] F. Gayzik, D. Moreno, C. Geer, S. Wuertzer, R. Martin, and J. Stitzel, "Development of a full body cad dataset for computational modeling: A multi-modality approach," *Annals of biomedical engineering*, vol. 39, pp. 2568–83, Jul. 2011. DOI: 10.1007/s10439-011-0359-5.

[31] BETA CAE Systems International AG, *ANSA Pre-Processor*, Version used: [24.1.2], 2024. [Online]. Available: https://www.beta-cae.com/ansa.htm.

[32] C. Fichera and J. John, *Infepy: Python library for the infet landmarking tool*, https://github.com/infekit/infepy, Accessed: 2025-05-22, 2024.

[33] scikit-learn developers, *Scikit-learn: Machine learning in python*, https://scikit-learn.org/stable/.

[34] S. Seabold and J. Perktold, *Statsmodels: Econometric and statistical modeling with python*, Austin, TX, 2010.

[35] S. Holcombe and Y. Huang, "Cross-sectional properties of rib geometry from an adult population," *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1 158 242, May 2023. DOI: 10.3389/fbioe.2023.1158242.

[36] H. Y. Choi and I. Lee, "B-27 thorax fe model for older population," *The Proceedings of Joint Symposium: Symposium on Sports Engineering, Symposium on Human Dynamics*, vol. 2009, no. 0, pp. 367–372, 2009, ISSN: 2433-1309. DOI: 10.1299/jsmesports.2009.0_367. [Online]. Available: http://dx.doi.org/10.1299/jsmesports.2009.0_367.

[37] S. A. Holcombe and B. A. Derstine, "Rib cortical bone thickness variation in adults by age and sex," en, *J Anat*, vol. 241, no. 6, pp. 1344–1356, Aug. 2022.

[38] K.-U. Schmitt, P. F. Niederer, M. H. Muser, and F. Walz, *Trauma Biomechanics: Accidental injury in traffic and sports*. Springer Berlin Heidelberg, 2010, ISBN: 9783642037139. DOI: 10.1007/978-3-642-03713-9. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-03713-9.

[39] I. Stockman, *Safety systems for children*, PowerPoint presentation, Presented during *Impact Biomechanics – Child Safety* seminar, May 2024.

[40] S. Holcombe, K. Matthew, Z. Peng, *et al.*, "Age-based predictive model of the pediatric ribcage," *International Journal of Automotive Engineering*, vol. 5, pp. 15–22, Mar. 2016. DOI: 10.20485/jsaeijae.5.1_15.

[41] K.-J. M. Larsson, *Human body model morphing for assessment of crash rib fracture risk for the population of car occupants*. Chalmers Tekniska Hogskola (Sweden), 2023.

[42] E. Commission, *Data protection: How the gdpr helps you protect your data*, Accessed: 2025-01-28, 2025. [Online]. Available: https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_en.htm#inline-nav-1.

[43] S. A. Holcombe, S. C. Wang, and J. B. Grotberg, "The effect of age and demographics on rib shape," *Journal of Anatomy*, vol. 231, no. 2, pp. 229–247, Jun. 2017, ISSN: 1469-7580. DOI: 10.1111/joa.12632. [Online]. Available: http://dx.doi.org/10.1111/joa.12632.

[44] L. Luo, M. Wang, Y. Tian, *et al.*, "Automatic sex determination of skulls based on a statistical shape model," *Computational and Mathematical Methods in Medicine*, vol. 2013, pp. 1–6, 2013, ISSN: 1748-6718. DOI: 10.1155/2013/251628. [Online]. Available: http://dx.doi.org/10.1155/2013/251628.

[45] L. Bezuidenhout and B. Rappert, "The ethical issues of dual-use and the life sciences," *CORE Issues in Professional and Research Ethics*, vol. 1, no. 1, 2012, Accessed: 2025-01-28, ISSN: 2160-8784. [Online]. Available: http://nationalethicscenter.org/content/article/183.

[46] P. J. Bowler and I. R. Morus, *Making modern science, second edition*, en, 2nd ed. Chicago, IL: University of Chicago Press, Aug. 2020.

[47] OpenAI, *Chatgpt: Language model (gpt-4)*, https://chat.openai.com, Accessed May 2025, 2024.

A

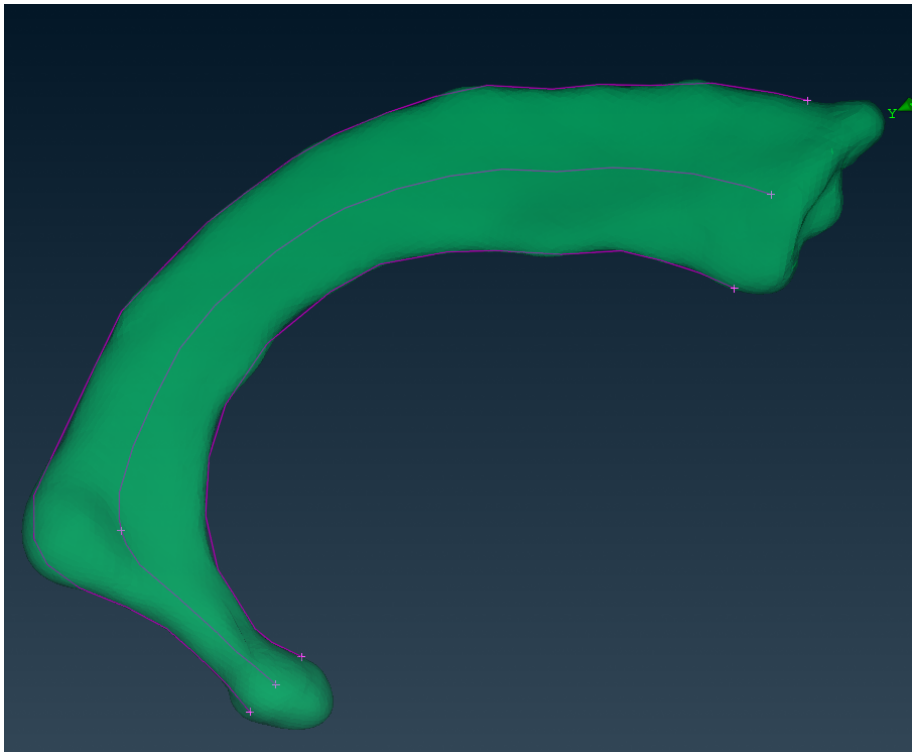
Appendix 1

Landmarking Scheme

1. If both left and right ribs have been used, make sure to mirror them around the sagittal plane for consistency with the template mesh.

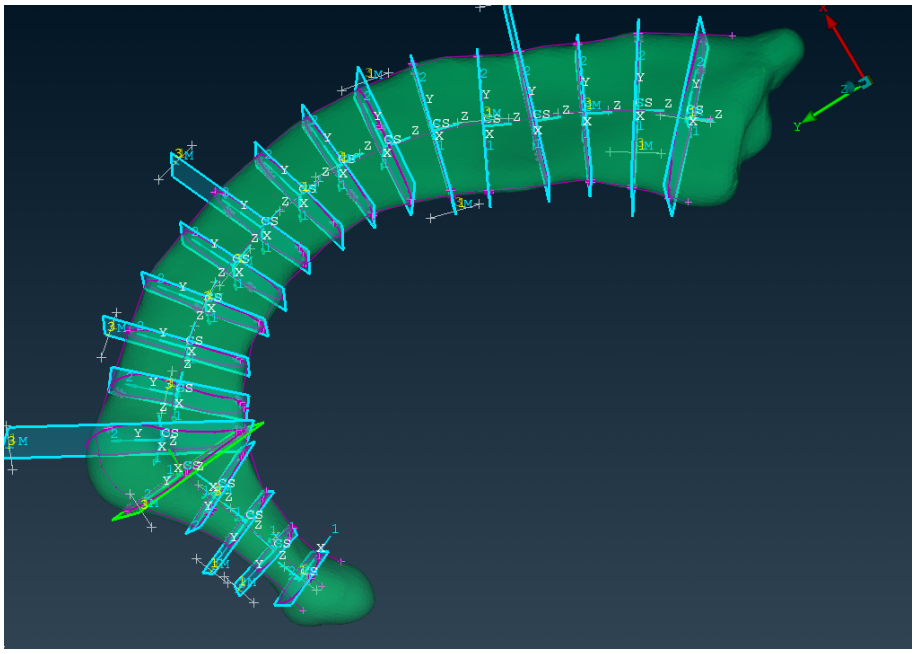
2. Draw a line from before the head starts to slightly before the upper surface starts on both sides, as shown in the figure.

3. Use the “middle curve” tool to create a new curve in between.



4. Use the “hotpoints project” tool at the tubercle to cut the middle curve into two parts.

5. Use the “multicut” tool to create 20 cross sections orthogonal to the middle line, 15 on the shaft (toward the upper surface) and 5 on the neck (toward the head). Use the “cons to curve” tool.



6. From the tubercle, draw two manual curves ending at the same point near the line created in step 2.

7. Remove the two curves closest to the line created in step 6. Create two new manual curves starting halfway between the lines from step 6 and ending at the same point as the lines in step 6.

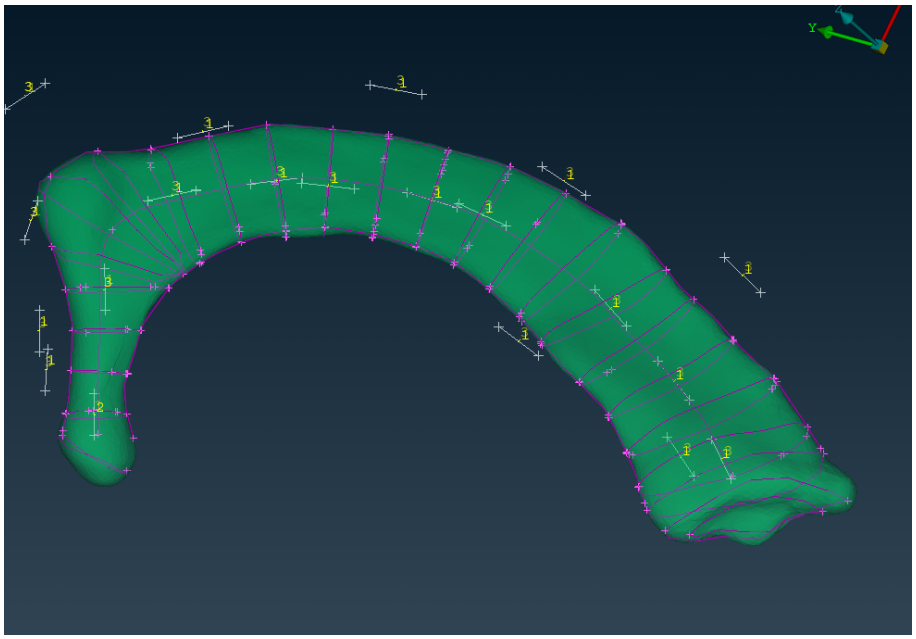
8. Use the “intersect” tool on the cross-section curves to ensure each cross-section consists of two half cross-sections that start and end at the lines created in step 2.

9. Make two curves similar to those in step 6 around the head.

10. Make two curves similar to those in step 6 around the upper surface.

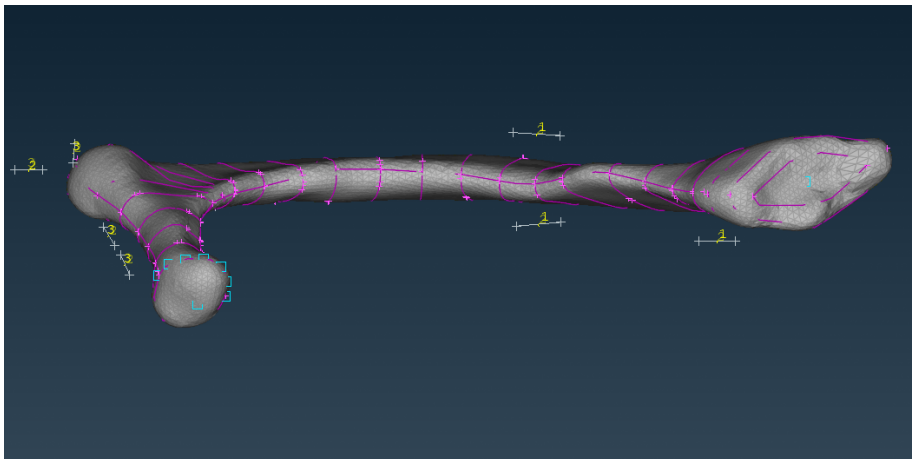
11. Create a curve similar to step 10 halfway toward the middle of the upper surface.

12. Draw a line similar to those in step 6 between the curve created in step 10 and the cross-section created in step 5.



13. Place a point on the top of the head and one point in the middle of the upper surface.

14. Use the "points on curve" tool to place 7 points on each curve (halfway around), except for the curve created in step 11.

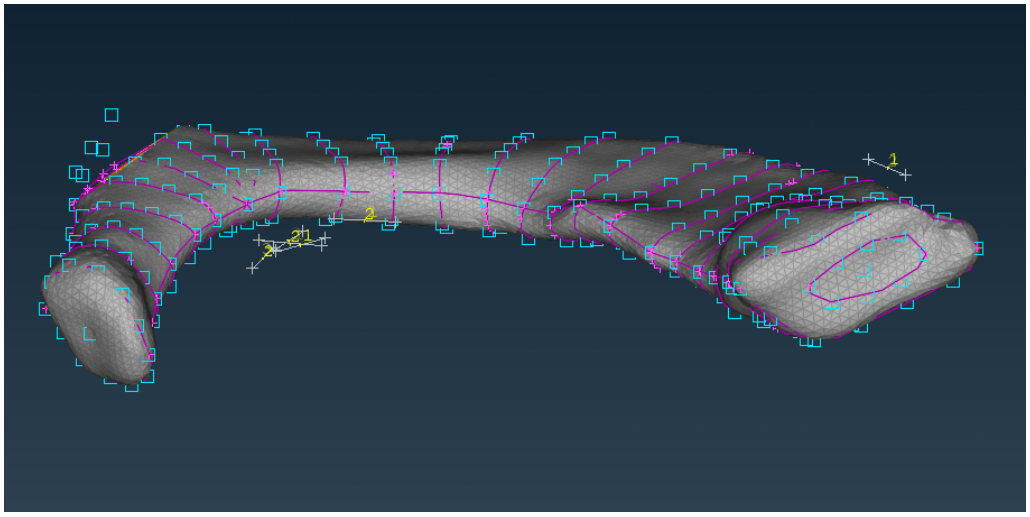


15. On the curve created in step 11, place points on the curve as in step 14, but with 5 points instead of 7.

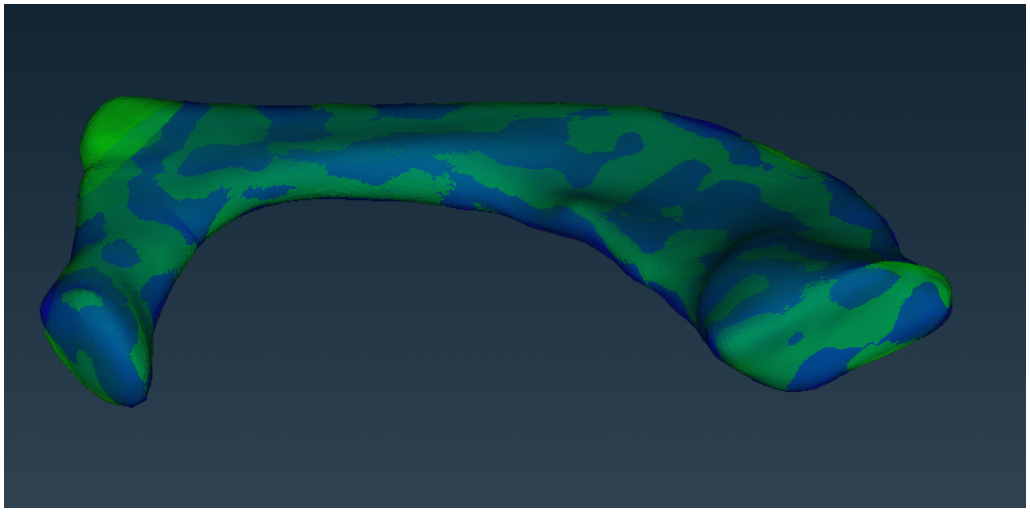
16. The order the points are created in doesn't matter as long as it is consistent between all ribs and the template. Ensure all point IDs correspond to the same positions.

B

Appendix 2



(a) Landmarking for cropped surfaces.



(b) Morphed surface (green) and landmarked surface (blue) corresponding to (a).

Figure B.1: Example of landmarking on cropped surfaces: (a) raw cropped surface with landmarks, and (b) comparison of morphed and landmarked surfaces.

C
Appendix 3

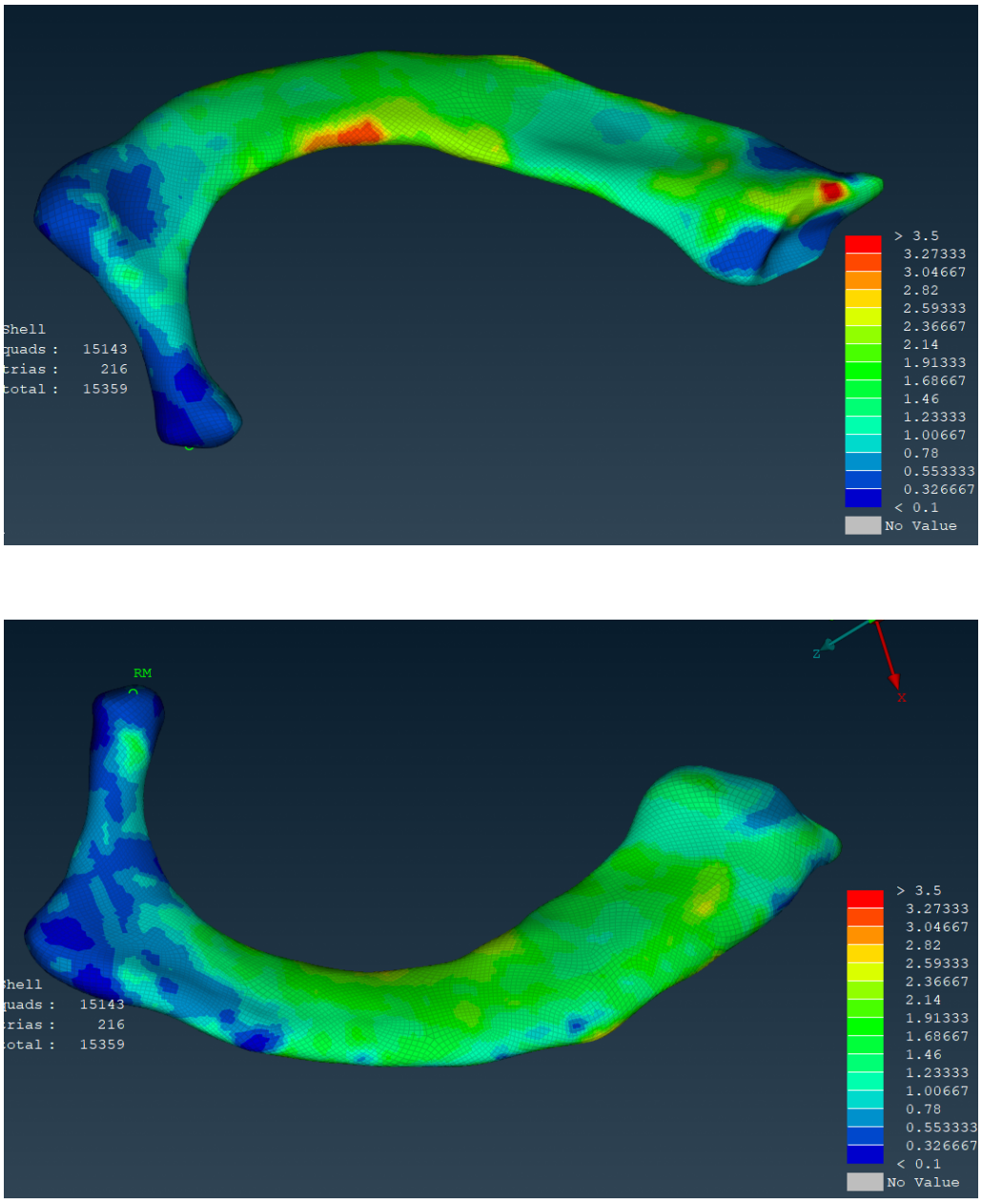


Figure C.1: Example of thickness values from Stradview mapped onto a morphed surface.

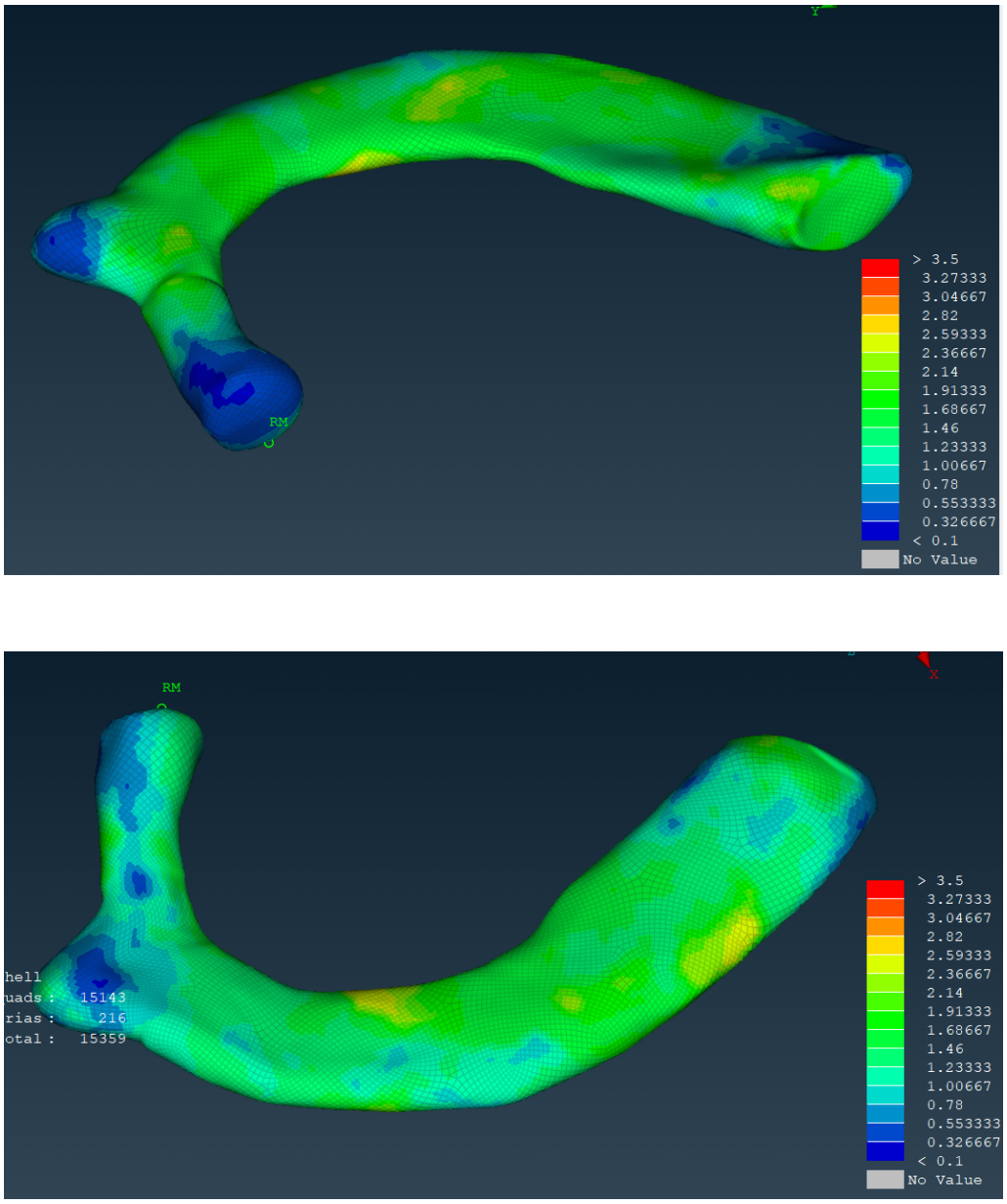


Figure C.2: Example of thickness values from Stradview mapped onto a morphed surface.

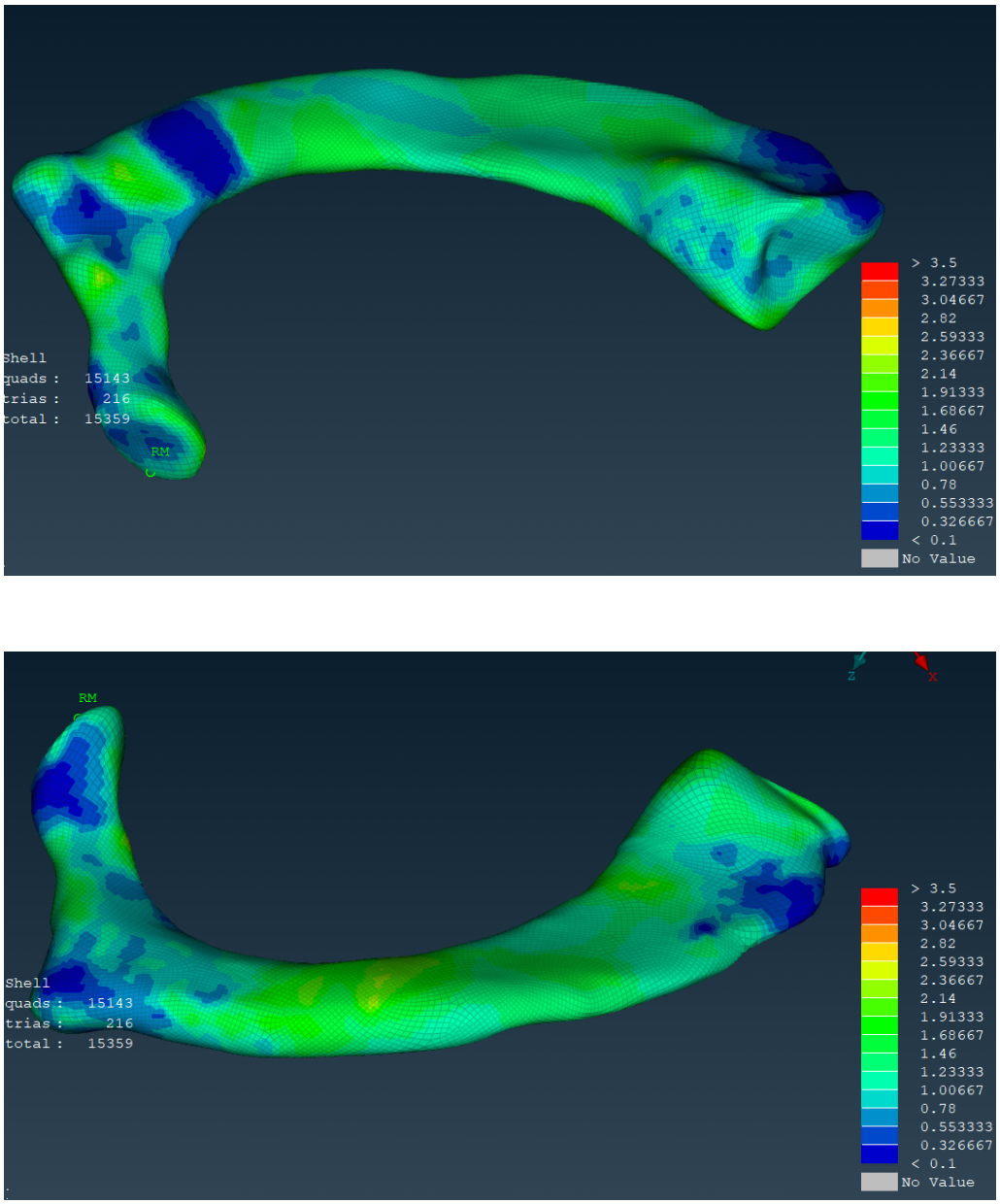


Figure C.3: Example of thickness values from Stradview mapped onto a morphed surface.

D

Appendix 4

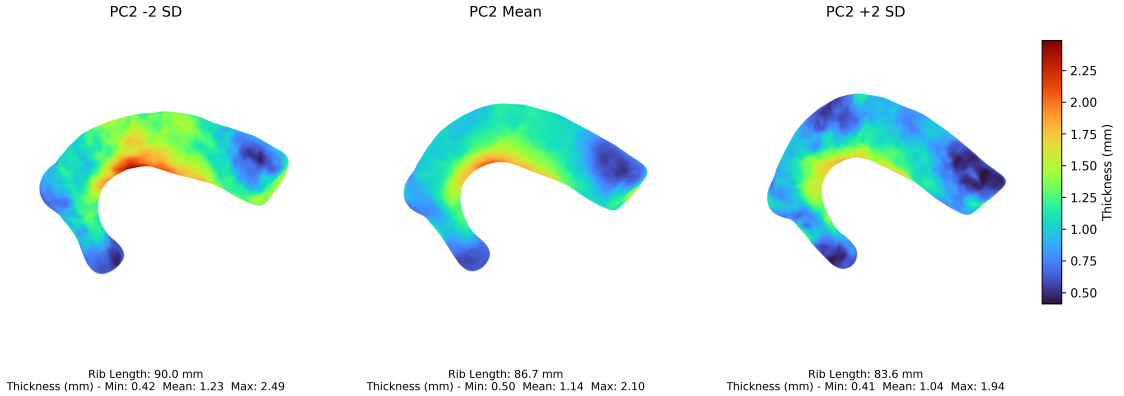


Figure D.1: PC2 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

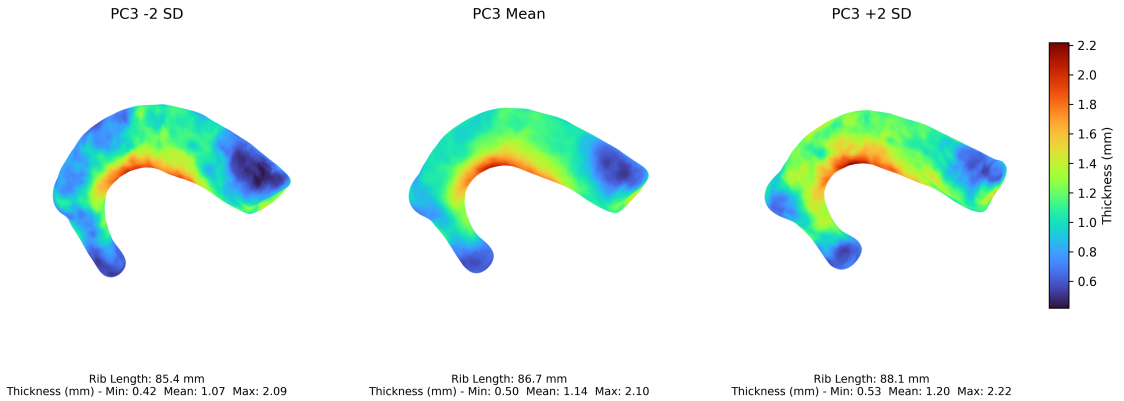


Figure D.2: PC3 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

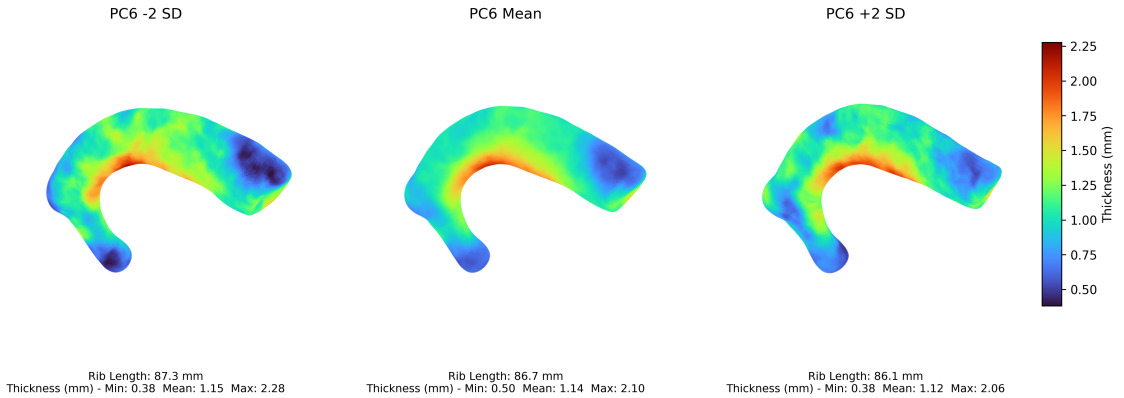


Figure D.3: PC6 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

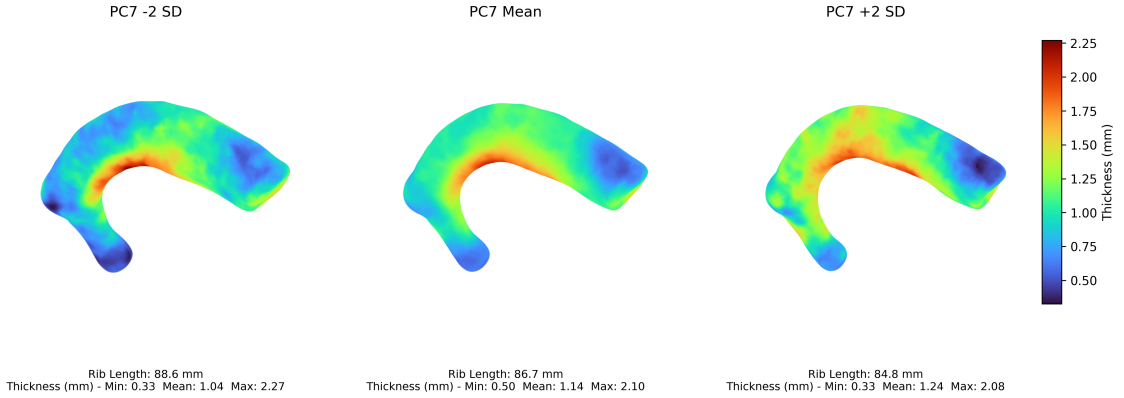


Figure D.4: PC7 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

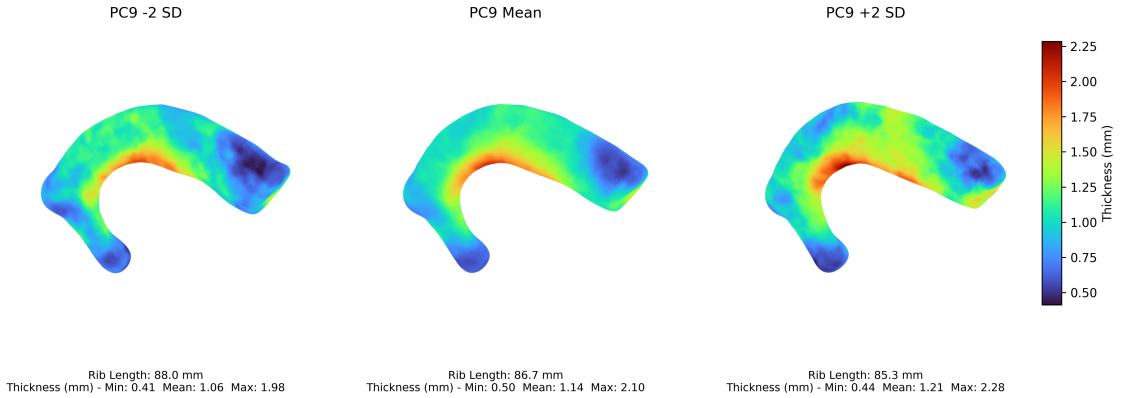


Figure D.5: PC9 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped

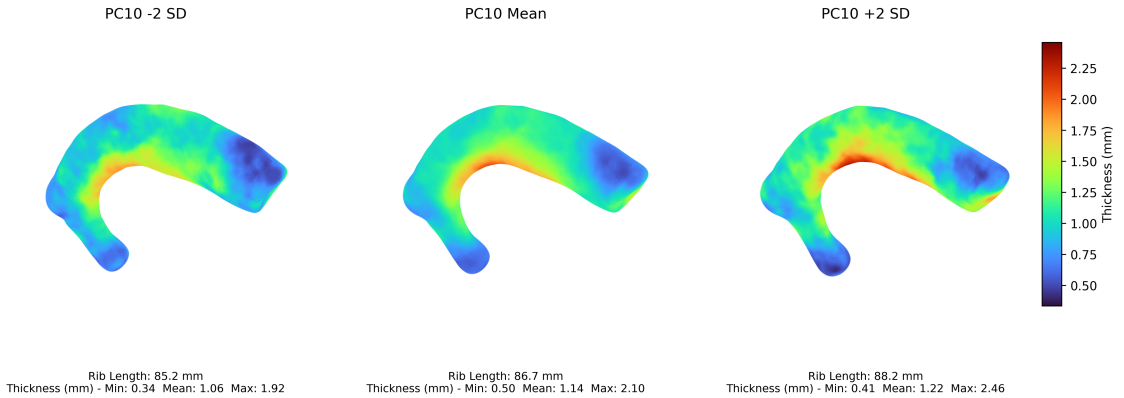


Figure D.6: PC10 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

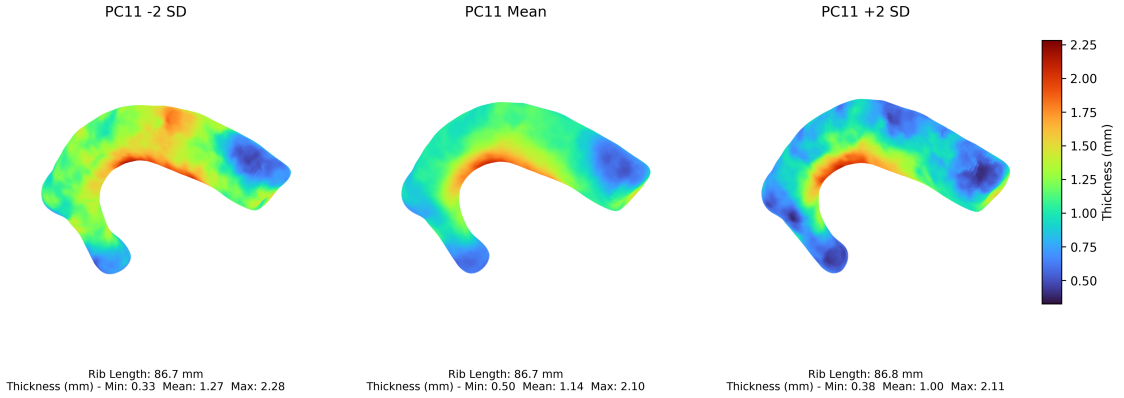


Figure D.7: PC11 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

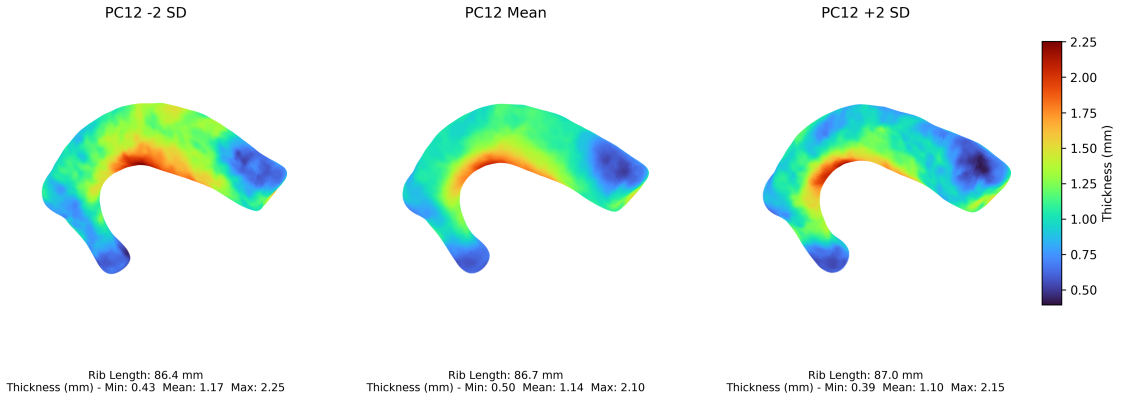


Figure D.8: PC12 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

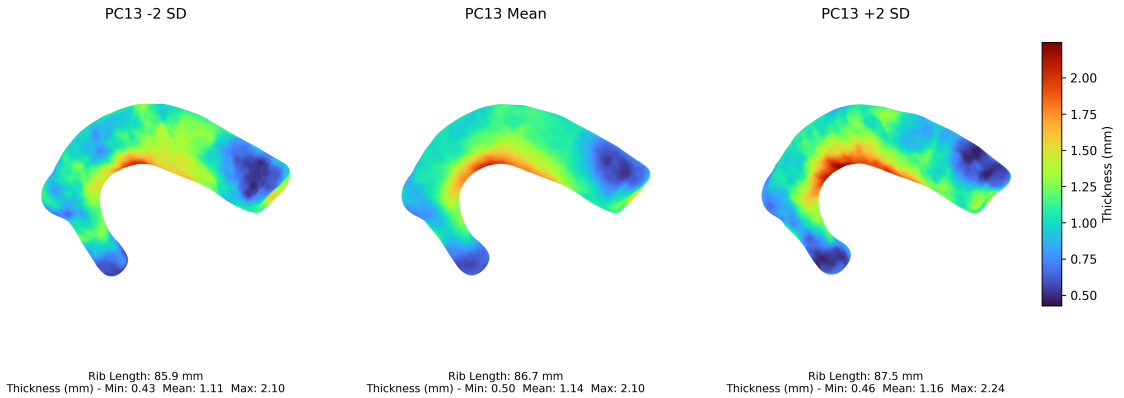


Figure D.9: PC13 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

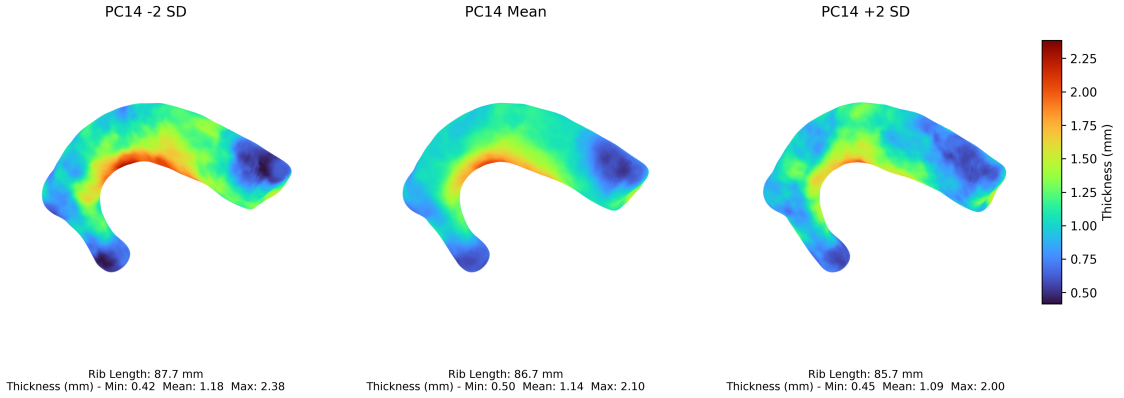


Figure D.10: PC14 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

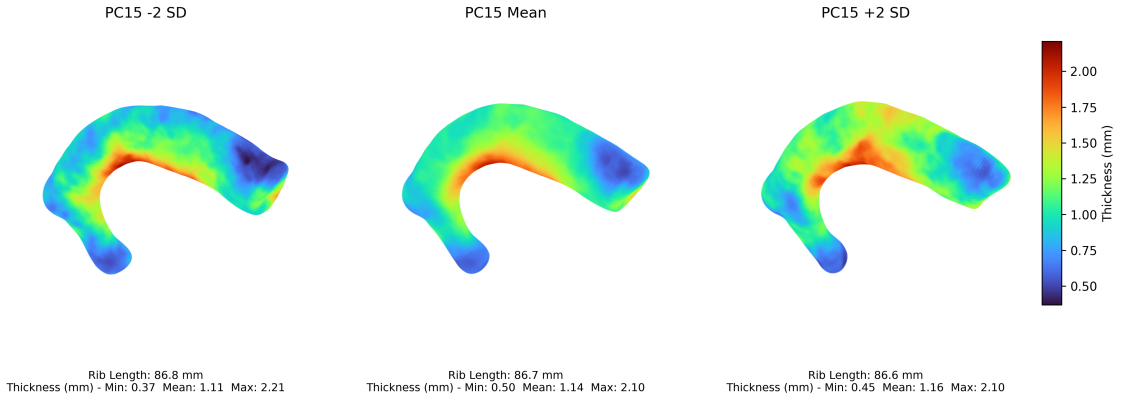


Figure D.11: PC15 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

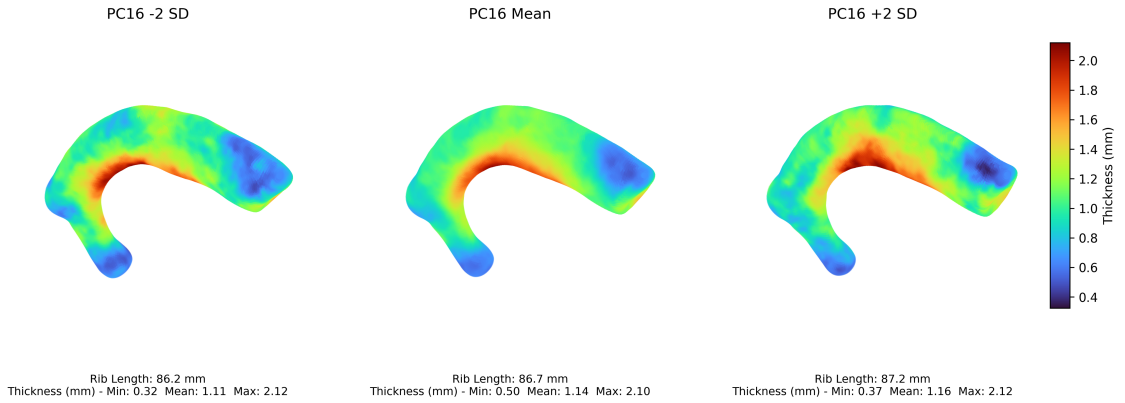


Figure D.12: PC16 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

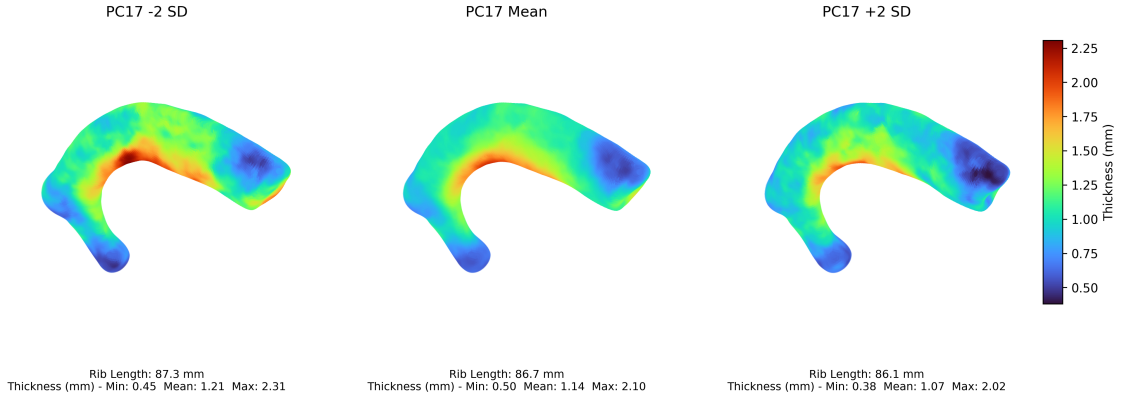


Figure D.13: PC17 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

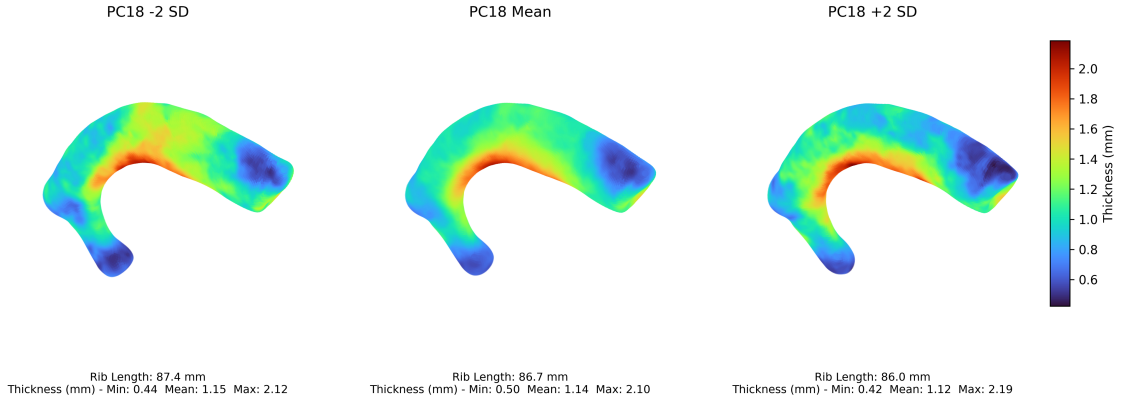


Figure D.14: PC18 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

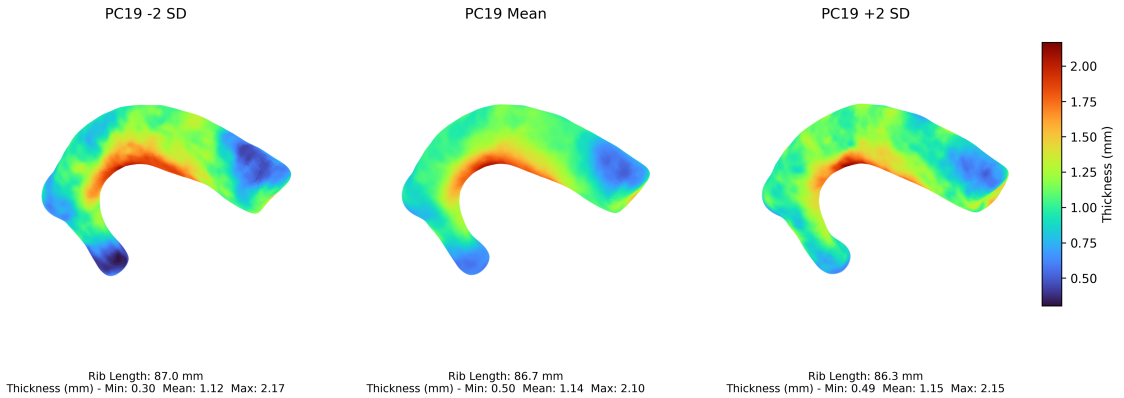


Figure D.15: PC19 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

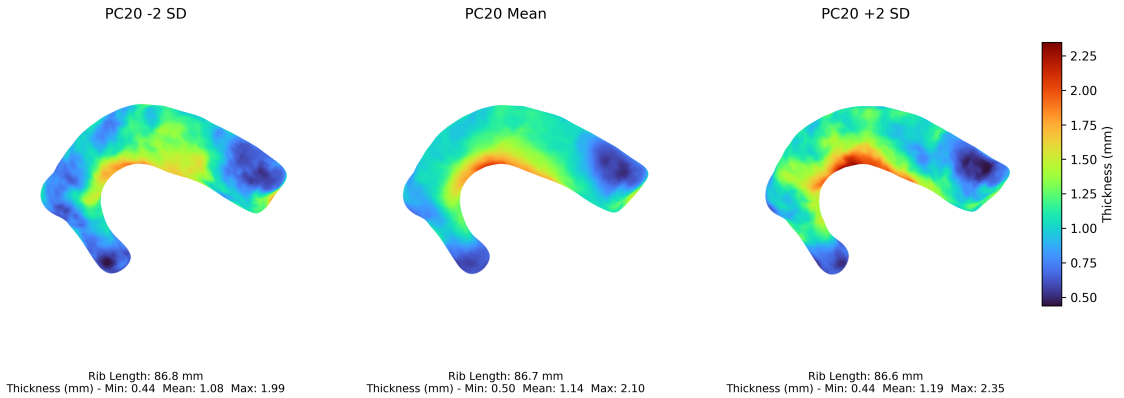


Figure D.16: PC20 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

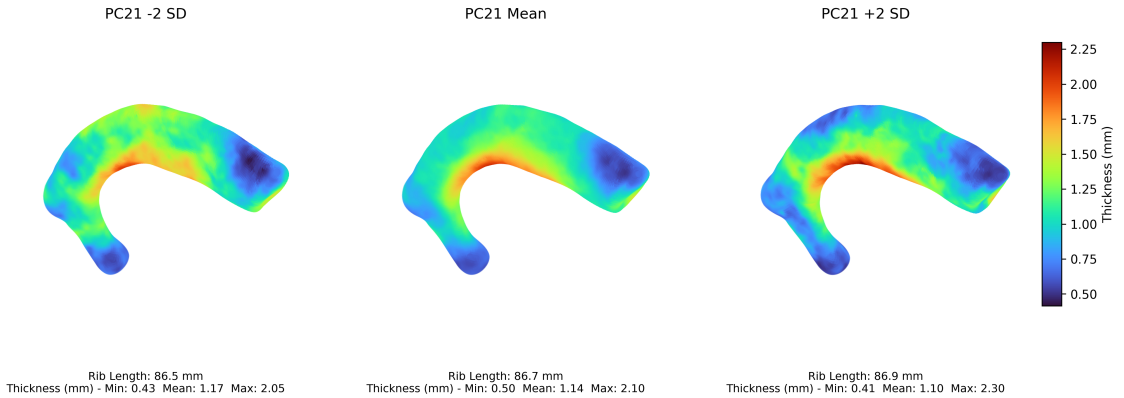


Figure D.17: PC21 mode of variation: -2 SD (left), mean (center), +2 SD (right), with cortical thickness mapped.

DEPARTMENT OF MECHANICS AND MARITIME SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY