Statistical Shape and Bone Property Models of Clinical Populations as the Foundation for Biomechanical Surgical Planning: Application to Shoulder Arthroplasty

Azita Sharif-Ahmadian[a], Aren Beagley[a], Claire Pearce[a], David Saliken[b,c], George S. Athwal[d], Joshua W. Giles[a,c,e,†]

[a] *Department of Mechanical Engineering, University of Victoria, Victoria, British Columbia, Canada*
[b] *RebalanceMD Clinic, Victoria, British Columbia, Canada*
[c] *Department of Orthopaedics, University of British Columbia, Vancouver, British Columbia, Canada*
[d]*Division of Orthopaedic Surgery, Schulich School of Medicine & Dentistry, Western University, London, Ontario, Canada*
[e]*Institute on Ageing and Lifelong Health, University of Victoria, Victoria, British Columbia, Canada*

[†]*Corresponding Author: jwgiles@uvic.ca*

ORCID #'s:
Joshua W Giles: 0000-0001-6997-8873
Aren Beagley: 0000-0003-2520-8118

Email: sharif@uvic.ca
Email: abeagley@uvic.ca
Email: claireapearce@uvic.ca
Email: saliken@ualberta.ca
Email: gathwal@uwo.ca
Email: jwgiles@uvic.ca

Abstract

This work developed, validated, and compared Statistical Shape, Statistical Intensity, and Statistical Shape and Intensity Models (SSMs, SIMs, SSIMs) of scapulae from a clinical population. SSMs efficiently describe bone shape variation while SIMs describe bone material property variation, and SSIM's combine description of both variables. This work establishes these models' efficacy and whether they can be used in surgical planning. Models were developed using shoulder arthroplasty data of patients with bone erosion, which is challenging to treat and would benefit from improved surgical planning. Models were created using previously validated non-rigid registration and material property assignment processes that were optimized for scapula characteristics. The models were assessed using standard metrics, anatomical measurements, and correlation analyses. The SSM and SIM specificity and generalization error metrics were 3.4mm and <1mm & 184 HU and 156 HU, respectively. The SSIM did not achieve the same level of performance as the SSM and SIM in this study (e.g. Shape Generalization: SSIM – 2.2mm vs SSM – <1mm). Anatomical correlation analysis showed that the SSM more effectively and efficiently described shape variation compared to the SSIM. The SSM and SIM modes of variation were not strongly correlated (e.g. $r_{max}=0.56$ for modes explaining ≤2.1% of variance). The SSIM is outperformed by the SSM and SIM and the latter two are not strongly correlated; therefore, using the SSM and SIM in conjunction will generate synthetic bone models with realistic characteristics and thus can be used for biomechanical surgical planning applications.

Keywords

statistical modeling; shoulder arthroplasty; surgical planning; morphological analysis

1. Introduction

Proper classification of bone morphology, bone quality (i.e. material properties), and disease state are critical initial steps in the orthopaedic surgical planning process. Once the disease state is established, it is necessary for surgeons to use this information to guide decisions regarding the most appropriate intervention. In the case of shoulder arthritis, patients are classified with respect to glenoid erosion type (e.g. Walch-type erosion), erosion level (e.g. levels 0, 1, 2, or 3 of the Walch Classification), and bone quality. Many studies have established classification systems and assessment protocols to determine these patient characteristics and how they should inform choices that will influence arthroplasty fixation and function [1–10]. These classification systems help surgeons understand complex patient characteristics and recently, surgical planning software have begun to assist in these classifications[11,12]. However, these systems remain limited in that they use a small number of discrete measurements to classify complex disease conditions including complex heterogeneous bone distributions, and this may oversimplify the information required to determine the optimal treatment. Therefore, the next generation of surgical planning software should not only provide classification information, but help to provide context and guidance with respect to preferred surgical treatments and potential biomechanical outcomes (e.g. fixation quality). Statistical bone modeling is an approach that can lay the foundation for such surgical planning tools through their efficient description of the full-field characteristics of a patient's bones. Some possible uses of statistical model models in future surgical planning applications include 1) disease state classification to directly inform clinicians or as an input to big data-based predictive algorithms, 2) morphological characteristic extraction for use in patient matched computational modeling such as musculoskeletal modeling or finite element analysis, or machine learning derived surrogates of these physics-based modeling methods, and 3) generation of bone instances with *a-priori* defined characteristics for use in computational virtual clinical trials.

Statistical Shape Models (SSMs) systematically describe the mutually independent ways geometry varies (i.e. modes of variation) across a given population of shapes (e.g. a cohort of patient scapulae) by applying Principal Component Analysis to a set of 3D geometric models that have corresponding point architectures [13]. By describing mutually independent modes of variation it is possible to capture the variability efficiently and to determine the significance of each mode of variation. SSMs have become widely published [14–19] with most being used as descriptive tools to understand bone shape variation. These models have primarily been built using healthy bones that describe the non-clinical population, while Lee et al. recently published an SSM analysis that determined differentiating factors between scapulae with and without rotator cuff tears[20]. Others have used SSMs built using healthy bones to predict the pre-morbid anatomy of patients with bone defects[21–26]. However, such models are not appropriate for classifying disease states and for associated next generation planning methods because they are not able to efficiently describe the patients' deformed anatomy (e.g. glenoid erosion).

Statistical Intensity Models (SIMs) (aka Statistical Density Models) describe the continuous distribution of bone material properties. The terms Intensity and Density are used because some models are constructed using Hounsfield Unit (HU) voxel intensity data from CT scans while others convert the intensity information to density data using published equations. SIMs are significantly less widely used than SSMs and in the shoulder context have only been used as a descriptive tool[27,28].

Finally, Statistical Shape and Intensity Models (SSIMs) attempt to efficiently describe shape and bone property information in a single model[29–32]. The proposed benefit of SSIMs is that correlated changes in shape and bone properties will be captured within the same mode of variation; however, these combined models are significantly more difficult to interpret due to the simultaneous changes in both variables. A previous study demonstrated that strong correlations do not exist between shape and bone properties in healthy populations [27] and thus SSIMs may not be necessary to generate representative bone instances (N.B. 'instance' refers to a 3D bone model that is synthetically created by adjusting the weights of the statistical model(s)). Rather, an SSM and SIM can be independently used to generate shape and bone property morphologies that can be combined to create a representative bone instance. However, it is unclear if this lack of correlation holds true for diseased populations.

With the current state-of-the-art in mind, the goals of this study were: 1) to develop and validate an SSM and an SIM that would be useful in surgical planning applications by building them using clinical patient scans; 2) to develop and validate a combined SSIM that integrates shape and bone property information for the same patient population; 3) to describe the variation captured by each model; and 4) to determine whether a combined SSIM is needed to describe the diseased bone states or if the independent SSM and SIM, which are simpler to understand, can be used. This will establish the efficacy of these models and whether they can be used in the development of surgical planning tools that assist with disease state classification and biomechanical analysis.

2. Methods

2.1. Patient & Data Characteristics

A database of 39 pre-operative patient CT scans (Table 1) with posterior glenoid erosions classified as Walch-type B2 & B3 were used for this study with demographic information in Table 2. Research Ethics was obtained for the use of these data (Certificate #: 18-217 & 18-226).

Table 1. Scan parameters used for imaging by GE Lightspeed CT scanner.

Beam energy	120 KV_p
Slice thickness	0.5-1.25 mm
Pixel spacing (Resolution)	~ 0.5 x 0.5 mm
Total pixels (Matrix size)	512 x 512
Convolution kernel	BonePlus
Tube current	Auto-adjusted to patient

Table 2. The demographic data of all scapulae in training set.

Sex	Quantity	Age
Female	20	72.7 ± 7.1
Male	19	66.2 ± 8.6
Left	19	67.5 ± 9.3
Right	20	71.3 ± 7.2
All	39	69.4 ± 8.5

2.2. Generating Models & Establishing Correspondence

To create an accurate scapula SSM, SIM, and SSIM, all patient bone geometries must be described using the same point architecture (e.g. same point quantity and enumeration) with matching enumerated points located on corresponding anatomical sites across all bone meshes. Mimics and 3-matic (Materialise NV) were used to segment all scans and generate a 3D scapula surface model for each with a uniform triangular mesh with an average edge length of 0.6 mm. One mesh was chosen as the 'source' model for establishing correspondence between all models by heuristically choosing the model closest to the median shape.

To create the corresponding surface meshes, three registration steps were undertaken (Fig. 1): 1) rigid registration to achieve initial alignment between the source and target models, 2) uniform scaling of the source model to match the size of the target model, which maintains the size information of the target for inclusion in the statistical model but eases the next step, and 3) mesh morphing to deform the source model to match the exact geometry of the target while maintaining the original anatomical meaning of each point (i.e. maintaining correspondence). For steps 1 and 2, registration was achieved using algorithms based on Besl et al.'s iterative closest point method [33]. For step 3, the authors optimized Amberg et al.'s correspondence preserving non-rigid iterative closest point (NR-ICP) registration algorithm for use with scapula geometries [34]. [See Supplemental Materials for details.]

The quality of fitting of the deformed source mesh to the shape of each target mesh's geometry was assessed by measuring the Root Mean Squared (RMS) distance error and the Hausdorff distance error between them. The resulting correspondence error was evaluated by calculating the Euclidean distance between manually selected ground truth points and the corresponding point locations produced by the algorithm. Specifically, the error for 12 anatomical landmarks were evaluated: four points on the glenoid fossa, three on the acromion process, two on the coracoid process, one at the medial termination of the scapular spine, one on superior angle, and one at the inferior angle. To assess the manual selection process, an inter- and intra-observer reliability analysis was performed on a set of 10 scapulae with the 12 anatomical landmarks selected by two independent observers and each observer repeated the landmark selection process two times. [See the Supplemental Materials for details.]

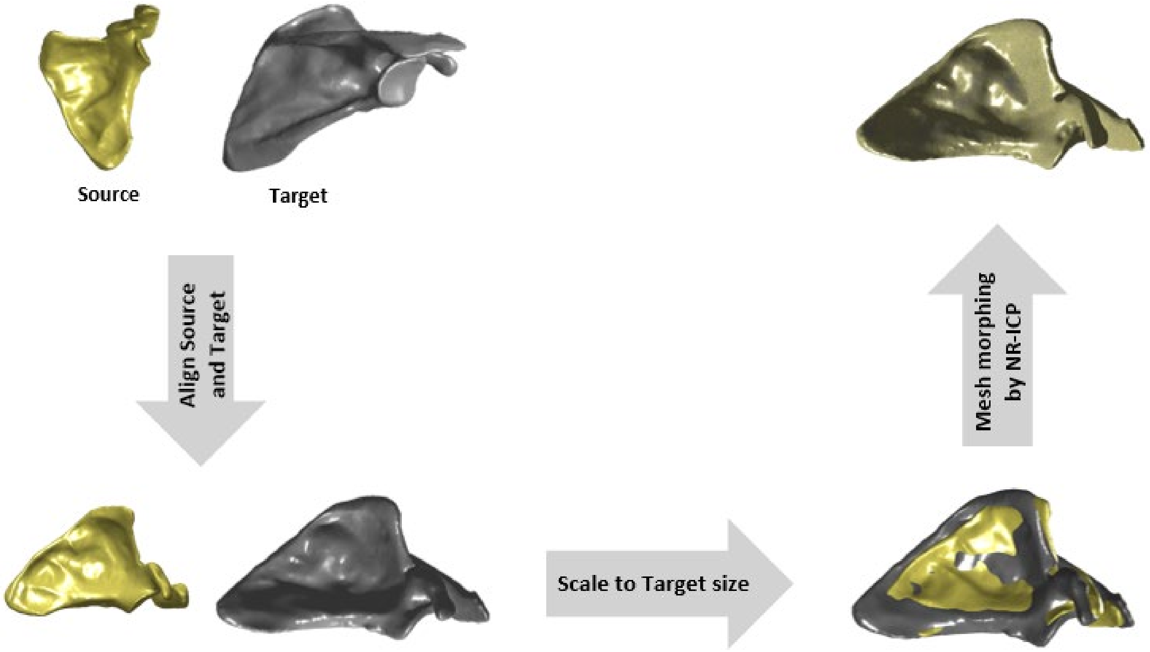


***Fig. 1** – Illustration of registration process on a target sample model using optimized NR-ICP method.*

Volumetric meshes with both surface and internal correspondence were then generated using Soltanmohammadi et al.'s method[30]: 1) using 3-matics (Materialise NV), generate a source volumetric mesh using the vertices of the source surface mesh as seed points, 2) calculate the deformation of each surface mesh point from its original source coordinates to its deformed source coordinates for a given patient, 3) generate a finite element simulation in Abaqus (Dassault Systems) using the source volumetric mesh from step 1 as the model geometry, 4) define the surface point deformations calculated in step 2 as displacement boundary conditions of the simulation, and 5) solve the system to yield the deformed volumetric mesh for that patient. [See the Supplemental Materials for details.]

With the corresponding volume meshes established, 3-matics and Mimics (Materialise NV) were used to assign Hounsfield Unit values to each volumetric mesh element based on the respective patient CT scan. Prior to material property assignment, internal intensity calibration and partial volume correction was applied [32, 33].

2.3. Calculating & Evaluating Statistical Models

SSM generation was undertaken by conducting Principal Component Analysis (PCA) on an N x N covariance matrix computed for all surface point coordinates of each patient model (N=39). This yields N-1 mutually independent modes of variation (i.e. eigenvectors or principal components) ordered from most to least variance explained (i.e. descending eigenvalues). The SIM was created using the same process but with the covariance matrix calculated based on the HU value of each volumetric element across all patient models. Finally, the SSIM was created by combining the spatial coordinates for all points that make up the nodes of the volumetric mesh along with the intensity data for each volumetric element such that the first P*3 rows represent the x, y, and z coordinates of the P volume nodes, and the remainder of the rows contain the volumetric element intensity. To account for differences in the unit magnitudes and number of data points for the shape and intensity data, a scaling factor method was adapted from the work of Cootes and Taylor [35]. To equalize the variance, after subtracting the mean shape, the shape data of each input bone was multiplied by the square root of the ratio of the total intensity variance to the total shape variance [See the Supplemental Materials for details of the scaling method]. The same covariance and PCA calculations were then conducted in the same way they were conducted for the SSM.

Each of the statistical models were assessed using three accepted metrics[36]: 1) Compactness gives the cumulative variance explained by modes from 1 to N-1; 2) Specificity is the average approximation error between object class instances generated by the model and the most similar object in the training dataset where this study averaged across 10,000 generated instances that were generated with mode weightings within the range of ±3 SDs; 3) Generalization ability is the average approximation error of the model to an unseen dataset and is assessed by performing a Leave One Out analysis across all input datasets. Standard-Generalization is typically assessed by fitting the model to the unseen dataset for varying numbers of modes of variation. An additional use is Sample-Size-Generalization where models are generated with increasing numbers of input datasets to assess at what number of input datasets (i.e. sample size) the fitting error stabilizes, indicating more data will not improve the model quality. [See the Supplemental Materials for details of these metrics including their equations.]

2.4. Characteristic Variation Captured by Principal Components of Statistical Models

It is useful to describe the characteristic variation captured by the most important principal components (i.e. modes of variation representing the largest percentage of variance) for each model. The results present a systematic description of the characteristic variation for the 2nd to 7th modes of the SSM and SIM. Mode 1 is not shown in the results as it captures uniform changes in scale and homogeneous intensity changes that are intuitive in nature; however, for completeness these modes can be seen in the Supplemental Materials. The characteristic variation captured by Modes 1, 2, and 3 of the SSIM are also described.

2.5. SSM Anatomical Measurements

Specific anatomical measurements have previously been shown to be important to shoulder surgical planning [37–41]. Therefore, the SSM and SSIM were assessed in terms of their ability to generate anatomical measurements that are in agreement with the training datasets and published values. Given the nodal correspondence of any shape instance generated by a given SSM/SSIM, measurement of a desired anatomical variable can be automated across the range of variation by specifying the nodes that form that measurement (e.g. two points defining a distance). Sixteen anatomic parameters are calculated including 9 length and 7 angular measurements (Table 3). [Definitions of the measurements and illustrative images can be found in the Supplemental Materials]. The required points were extracted from the 39 deformed source meshes in the training dataset and from the 10,000 shape instances generated for the SSM and SSIM as described for the Specificity evaluation, which generated mode weightings within ±3 SDs of the mean.

A correlation analysis was undertaken to evaluate the characteristic shape effect of each mode of variation of the SSM and SSIM, and to assess the degree to which each mode of variation affects one or more bone characteristics simultaneously. Specifically, the weight of each of the modes of variation used to generate each of the 10,000 instances generated by each model were correlated to each of the 16 measurements taken for that instance.

2.6. SSM-SIM Correlations

The final goal of this study was to assess whether bone instances would be realistic when they are generated by independently varying the principal component weights of the developed SSM and SIM. If the variations in scapula shape and intensity are weakly correlated or if moderate correlations only exist for modes of variation that explain little variance, then bone instances generated by independently varying SSM and SIM parameters would have bone shape and intensity characteristics that can be observed in the population. Therefore, a correlation analysis was undertaken similar to Burton et al. in which the developed SSM and SIM were fit to each of the 39 patient datasets, which resulted in 38 principal component weighting factors for each patient dataset for each of the SSM and SIM [27]. All permutations of the weighting factors for each of the first 30 principal components of the SSM and SIM (i.e. the 30 describing the most variance) across all 39 patients were correlated against each other (i.e. resulting in 30 SSM mode weights correlated to 30 SIM mode weights = 900 correlations). It was then determined whether/which combinations of SSM and SIM principal components vary together and their significance based on the variance each mode explains.

3. Results

3.1. Geometric Fitting and Correspondence Evaluation

Across all datasets in this study, geometric fitting RMS distance error was 0.25±0.04mm (mean±SD) while the Hausdorff error was 0.76±0.14mm. Across the 12 landmarks and 10 scapulae assessed, the correspondence error was 2.5±1.0mm. The inter-observer reliability was ±0.3mm and the intra-observer reliability was ±0.06mm.

3.2. Compactness, Specificity, and Generalization

Regarding Compactness, the SSM was markedly more compact than the SIM or SSIM, capturing 90% of data variance in the first 7 modes whereas the SIM required 30 modes to capture 90% of variance, and the SSIM's Compactness fell roughly halfway between the SSM and SIM across all numbers of modes as it was describing the variance contained in both, and required 26 modes to capture 90% of variance (Fig. 2).

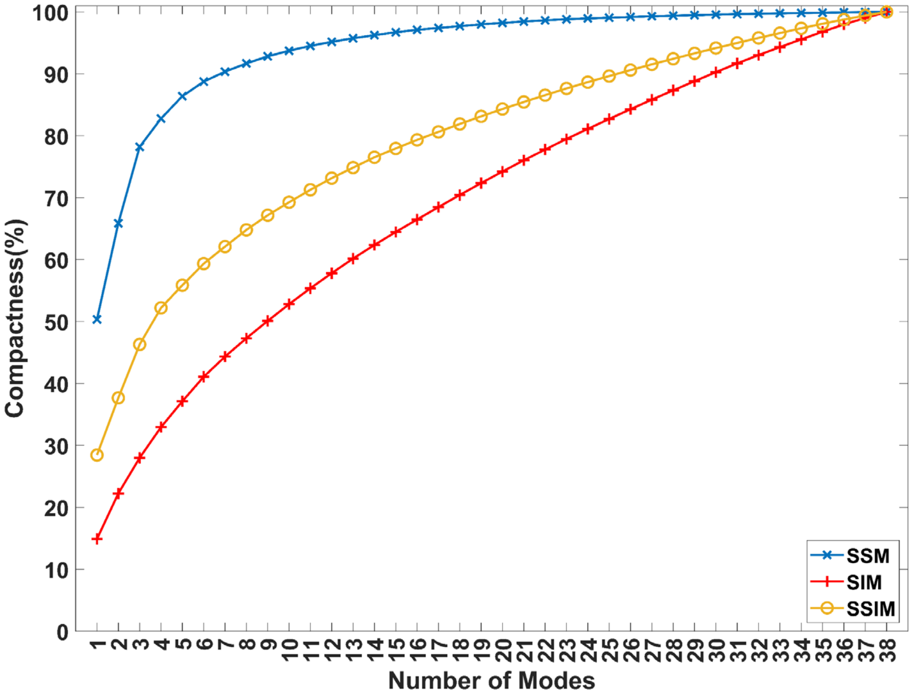


Fig. 2 *– Compactness of SSM, SIM, SSIM. Compactness describes the cumulative variance explained by 1 to m modes of variation, where m≤N-1, where N is the number of training datasets (N=39 in this study).*

The SSM, SIM, and SSIM Specificity error increases as the number of modes used to generate the instances increases (Fig. 3 & Fig. 4). Comparing the SSM and SSIM, for instances generated with sufficient modes to capture 90% of variance (7 modes for the SSM and 26 modes for the SSIM) (Fig. 3), the median Shape Specificity error was 0.39mm lower for the SSM (3.40mm vs 3.79mm, respectively). Comparing the SIM and SSIM, the median Intensity Specificity error was on average 15.7 HU lower with the SIM than the SSIM across all number of modes. For models generated with sufficient modes to capture 90% of variance (30 modes for the SIM and 26 modes for the SSIM), the Specificity was 184.3HU for the SIM and 200.8HU for the SSIM (Fig. 4). The SSIM Intensity Specificity exhibits minimal variance when using only Mode 1 indicating that this mode primarily describes shape variation, which agrees with Figure 8. However, except for when using only Mode 1, the SSIM exhibited greater variance in Intensity Specificity and more outliers with markedly higher error.

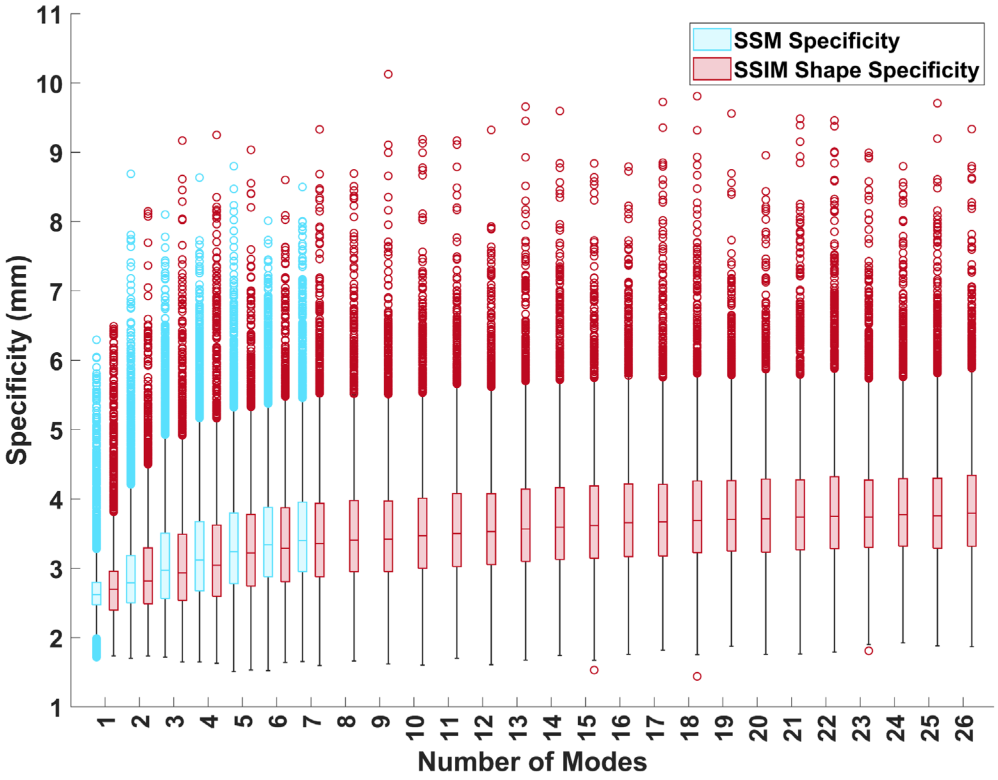


Fig. 3 *– Shape specificity of SSM & SSIM. Shape specificity indicates the amount of error between an artificially generated shape instance and the closest dataset used to create the statistical model, which varies with the number of modes of variation used to create the instance. Here specificity is calculated for instances generated using 1 to m modes, where m≤*M_{90} *where* M_{90} *is the number of modes capturing 90% of variance as determined by Compactness.* M_{90} *is 7 for the SSM and 26 for the SSIM.*

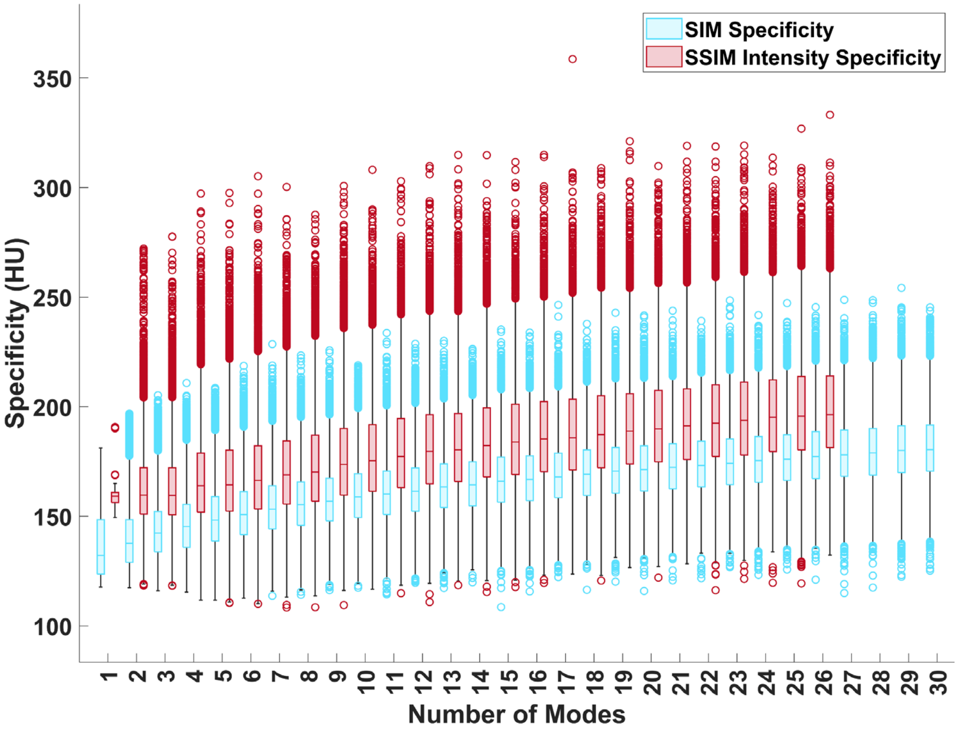


Fig. 4 *– Intensity Specificity of SIM and SSIM. Intensity specificity indicates the amount of error between an artificially generated bone material property instance and the closest dataset used to create the statistical model, which varies with the number of modes of variation used to create the instance. Here specificity is calculated for instances generated using 1 to m modes, where* $m \leq M_{90}$ *where* M_{90} *is the number of modes capturing 90% of variance as determined by Compactness.* M_{90} *is 30 for the SIM and 26 for the SSIM.*

It is generally accepted that when the Sample-Size-Generalization error changes by less than 1% after adding an additional training dataset, addition of more datasets will not add more meaningful description of bone variability, and thus is used to determine the required sample size. For the SSM, Sample-Size-Generalization error change dropped below 1% for models created with more than 23 datasets (Fig. 5a); further, the median Standard-Generalization error decreased with increasing modes used and dropped below 1mm when using more than four modes (Fig. 5a). Regarding SSIM Shape-Sample-Size-Generalization, error change dropped below 1% for models created with more than 31 datasets (Fig. 5b), and the median Shape-Standard-Generalization error dropped from 4.1mm when using only Mode 1 to 2.2mm when using the first 26 modes (Fig. 5b). For the SIM, Intensity-Sample-Size-Generalization error change dropped below 1% for models created with more than 17 of the 39 datasets (Fig. 5c). For the SIM, the median Standard-Generalization error dropped only moderately from 167.3HU when using only Mode 1 to 155.9HU when using the first 30 modes (Fig. 5c). For the SSIM, Intensity-Sample-Size-Generalization error change dropped below 1% for models created with more than 14 of the 39 datasets (Fig. 5d). For the SSIM, the median Standard-Generalization error dropped only moderately from 170.3HU when using only Mode 1 to 161.5HU when using the first 26 modes (Fig. 5d). It can also be seen that for both Shape and Intensity, the SSIM's Standard Generalization exhibits outliers with much larger magnitude, especially when using only the first few modes to fit an unseen shape.

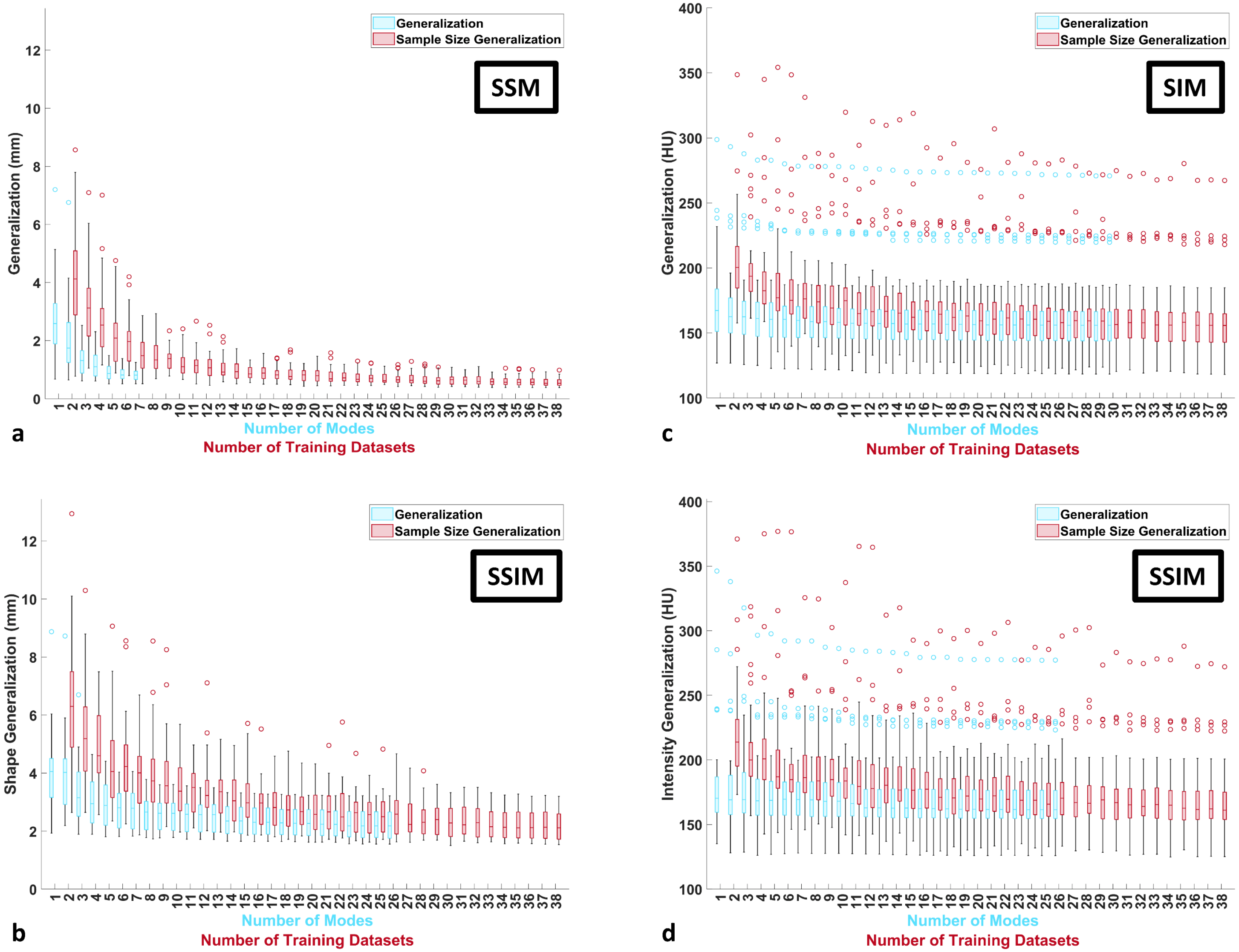


Fig. 5 – SSM, SIM, & SSIM Generalization. (a) SSM Generalization, (b) SSIM Shape Generalization, (c) SIM Generalization, (d) SSIM Intensity Generalization. Generalization is the error when a statistical model is fit to an input dataset not used to train the model. Standard-Generalization is calculated for instances generated using 1 to m modes of models created using N-1 datasets (N=39), where m≤M_{90} and M_{90} is the number of modes capturing 90% of variance as determined by Compactness. M_{90} is 7 for the SSM, 30 for the SIM, and 26 for the SSIM. Sample-Size-Generalization is calculated the same way, but the fitting is conducted using all available modes of variation of the model, which equals one less than the number of datasets used to make it, which ranges from 2 to N-1.

3.3. Characteristic Variation Captured by Principal Components of Statistical Models

The characteristic variations captured by Modes 2-to-7 of the SSM and SIM are shown in Fig. 6 & Fig. 7, respectively, and by Modes 1-3 for the SSIM in Fig. 8. Mode 1 is not shown for the SSM or SIM as it captures homogeneous size/intensity changes, which are not important in interpreting the model. [Refer to the Supplemental Materials for detailed descriptions of the characteristic variations.]

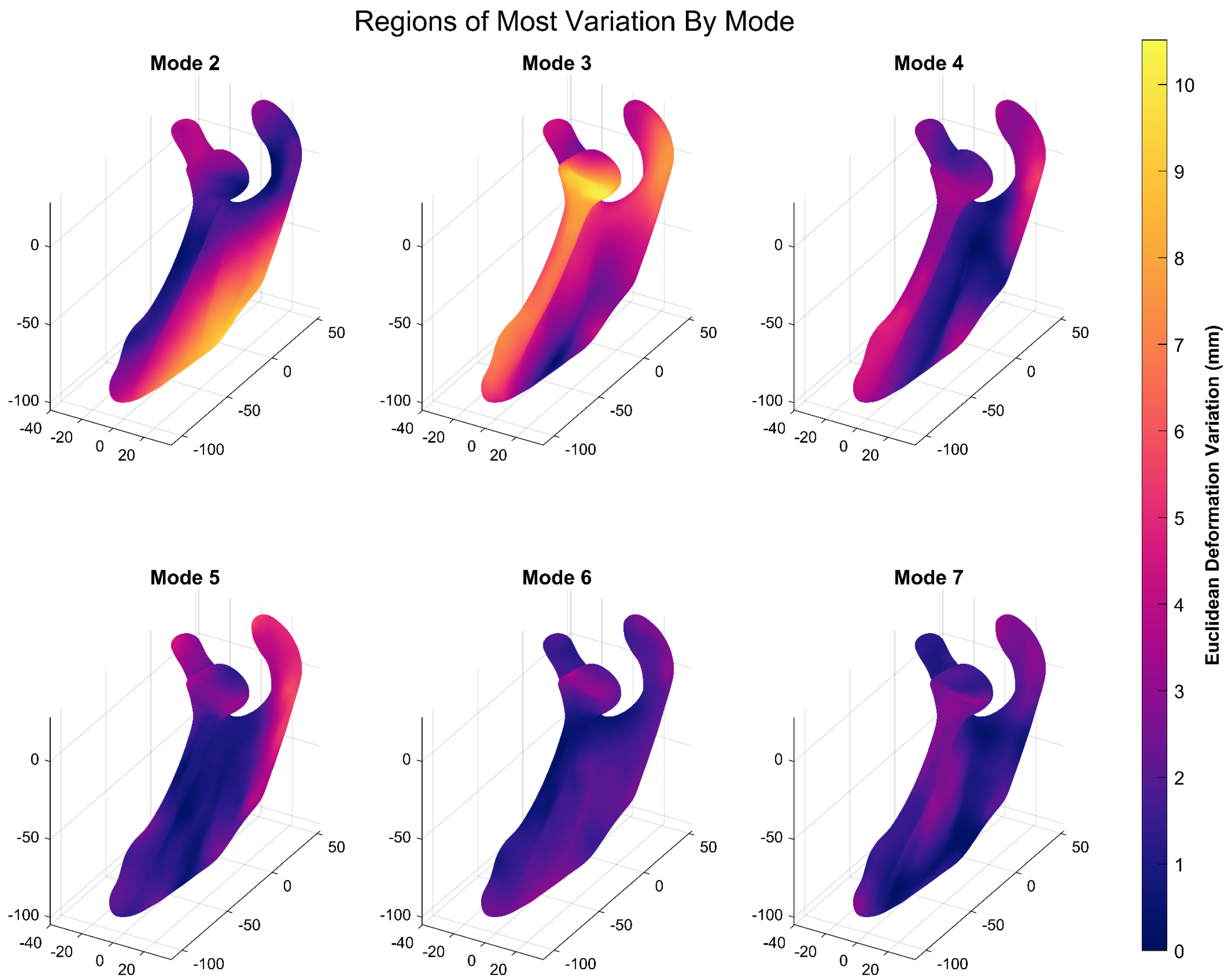


Fig. 6 *– SSM Characteristic Shape Changes. The scapula in each pane represents the average shape with changes in shape shown by the colormap variation. The colormaps represent the magnitude of spatial variation when the weighting for the given mode of variation is set to 1.96 SDs.*

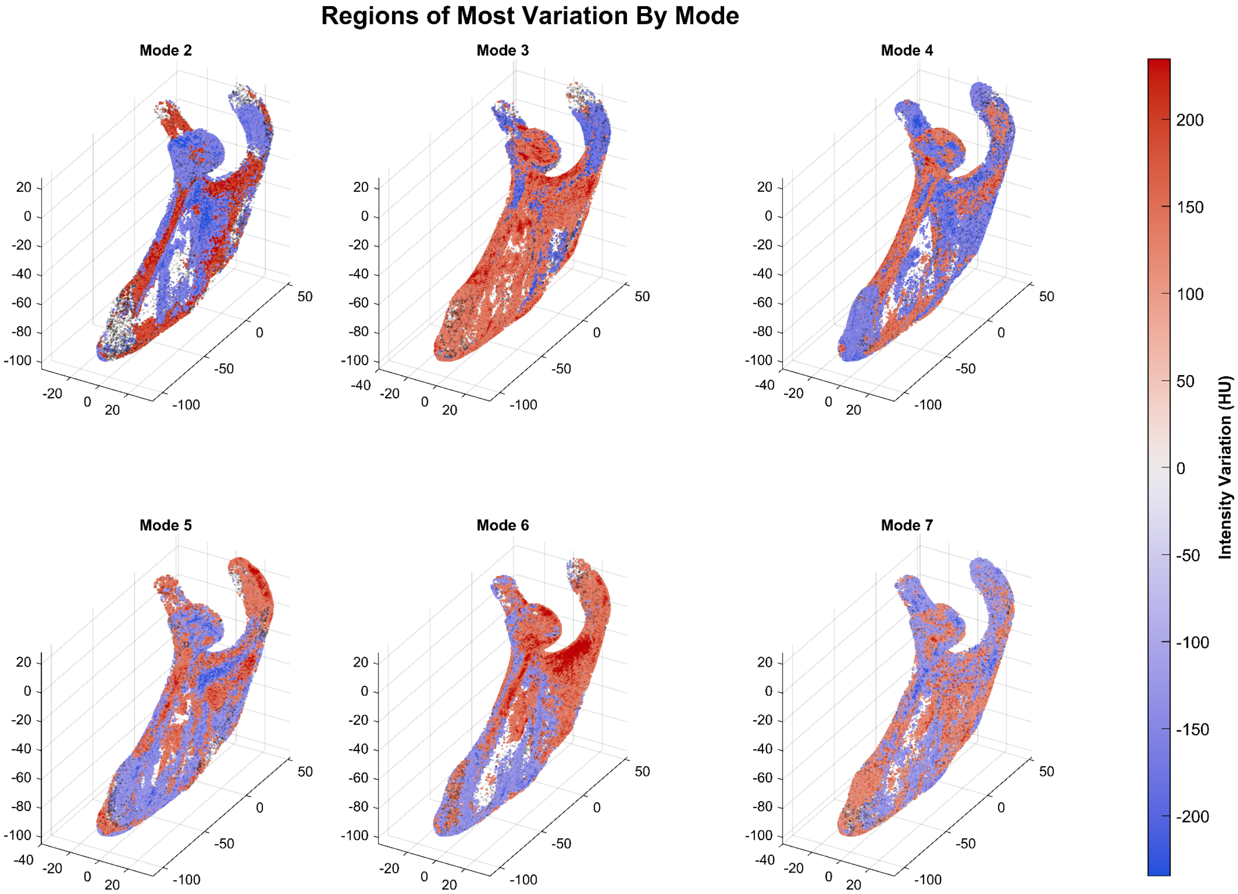


***Fig.** 7 – SIM Characteristic Bone Property Changes. The colormaps represent the change in intensity when the weighting for the given mode of variation is set to 1.96 SDs. Given that change happens throughout the bone volume, to enable improved visualization, only points with increases or decreases in intensity in the top 20% of absolute change for each mode (i.e. each figure pane) are displayed. Furthermore, to improve readability of the colorbar, any point whose increase or decrease is in the top 10% across all modes (i.e. top 10% of all changes shown in entire figure) is shown as having maximal red or blue saturation, respectively (i.e. if the change is above a given HU threshold the color saturation remains constant). Note that the sign of the weightings of each mode of variation is arbitrary and thus a + or - SD does not inherently infer increases and decreases in intensity, respectively.*

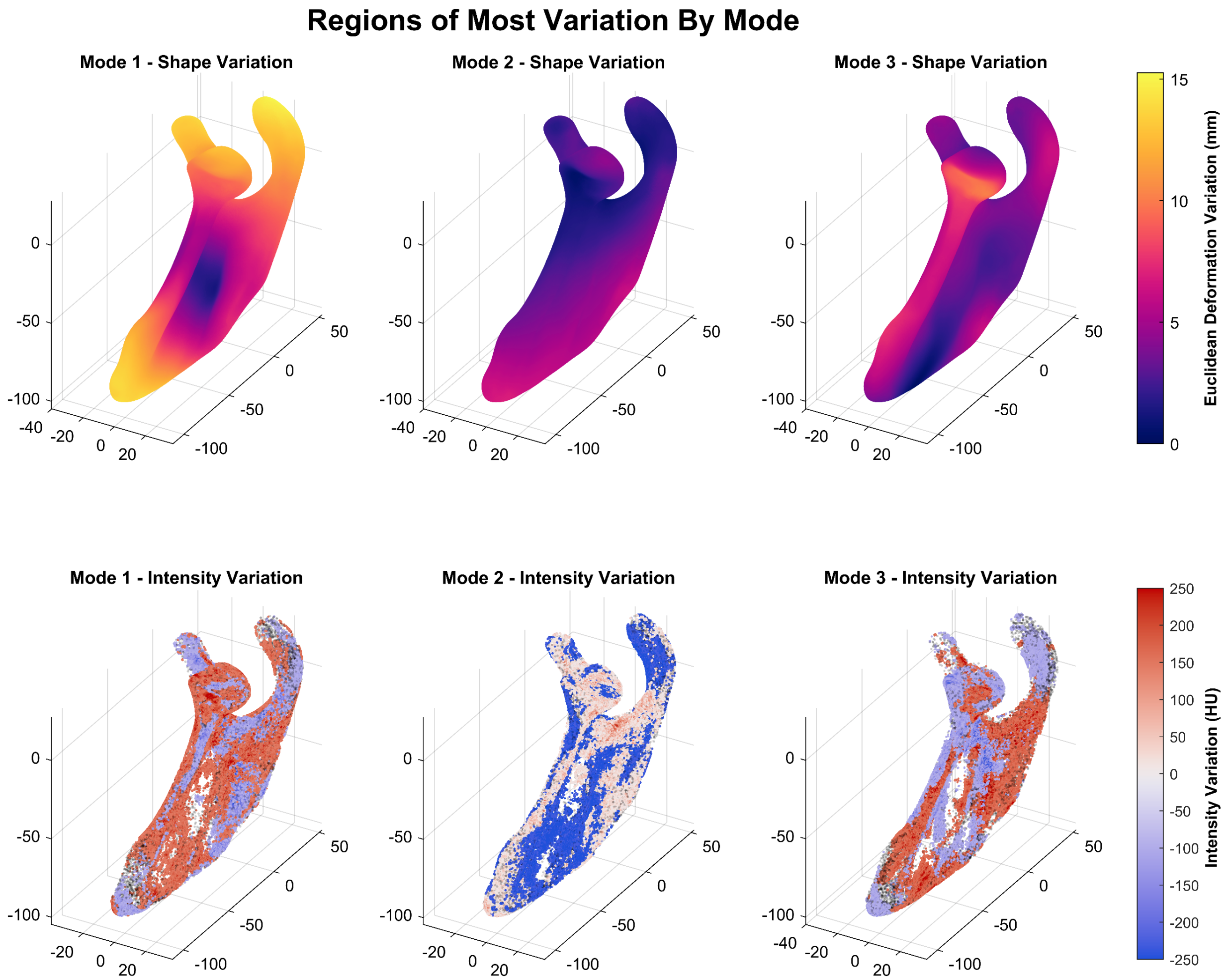


Fig. 8 – *SSIM characteristic shape and intensity variations. These shape and intensity subplots were created using the same process as for the SSM and SIM in the previous two figures.*

3.4. Anatomical Measurements

Table 3 shows the results for the 16 anatomical measurements. The average values of all measurements are comparable between the input datasets and the 10,000 generated instances for the SSM and the SSIM while the standard deviations of the measurements are consistently lower for the SSM and SSIM than for the input datasets.

Regarding the correlation between mode of variation weights and anatomical measurements, it was found that there were strong correlations for the SSM and SSIM. Specifically, for the weights and measurements across the 10,000 SSM instances generated, 1 measurement was very strongly correlated ($r>0.9$) to one mode of variation, 7 measurements were strongly correlated ($0.7<r\leq0.9$) to one mode of variation, and 8 measurements were moderately correlated ($0.5<r\leq0.7$) to one mode of variation. For the weights and measurements across the 10,000 SSIM instances generated, there were no measurements that were very strongly correlated ($r>0.9$) to any mode of variation, 2 measurements were strongly correlated to a mode of variation ($0.7<r\leq0.9$), only 9 measurements were moderately correlated ($0.5<r\leq0.7$) to a mode of variation, and the remainder were weakly correlated. Furthermore, for the SSM instances, each measurement was primarily correlated to a single mode of variation, but some had a correlation to another mode of secondary strength (e.g. Glenoid height was primarily correlated to Mode 1 ($r=0.76$) and secondarily correlated to Mode 2 ($r=0.57$)). Conversely, for the SSIM instances, with the exception of the first two measurements (e.g. scapular height and width) that were almost exclusively correlated to Mode 1, almost all measurements had noticeable correlations to multiple modes of variation (e.g. Glenoid width was correlated to Mode 1 to 3 with r values of 0.55, 0.51, 0.39, respectively).

Table 3. Scapular measurement results for 16 parameters.

Parameter name	Input Datasets	SSM Generated (7 Modes)*	SSIM Generated (26 Modes)*	Unit
Scapula Height	155.5 ± 14.2	155.3 ± 13.6	155.7 ± 14.0	mm
Scapula Width	99.1 ± 8.6	99.0 ± 8.4	99.1 ± 8.4	mm
Glenoid Height	38.2 ± 4.4	38.1 ± 3.8	38.2 ± 4.3	mm
Glenoid Width	28.3 ± 3.9	28.1 ± 3.1	28.3 ± 3.7	mm
Acromion Length	45.7 ± 4.8	45.6 ± 3.9	45.8 ± 4.4	mm
Lateral Acromion to Glenoid Center Distance	23.6 ± 6.0	23.6 ± 5.2	23.6 ± 5.9	mm
Coracoid Tip to Glenoid Center Distance	13.3 ± 4.7	13.3 ± 3.3	13.4 ± 4.5	mm
PI Acromion to Glenoid Center Distance	36.6 ± 4.7	36.6 ± 4.0	36.6 ± 4.5	mm
SA Acromion to Glenoid Center Distance	-1.2 ± 6.2	-1.2 ± 4.5	-1.2 ± 5.9	mm
Fulcrum Axis	91.9 ± 4.4	91.9 ± 3.4	92.0 ± 4.1	degrees
Inclination Angle	97.2 ± 4.6	97.3 ± 1.8	97.2 ± 4.2	degrees
Version Angle	19.5 ± 7.2	19.5 ± 3.6	19.4 ± 6.5	degrees
Acromial Tilt Angle	32.5 ± 3.7	32.4 ± 0.9	32.4 ± 3.5	degrees
Critical Shoulder Angle (CSA)	28.4 ± 5.8	28.4 ± 3.4	28.4 ± 5.3	degrees
SI Glenoid-Acromion Angle	61.1 ± 5.0	60.9 ± 3.7	61.1 ± 4.8	degrees
SI Glenoid-Acromion-Coracoid Angle	94.9 ± 4.9	94.9 ± 2.8	94.9 ± 4.5	degrees

***10K shapes generated using modes of variations with random weights ranging within ±3 SDs.**

3.5. SSM, SIM Correlations

With respect to correlations between SSM and SIM modes, no correlation exceeded r=0.56 (i.e. a moderate correlation) (Fig. 9). This strongest correlation was between Shape Mode 30 and Intensity Mode 15, which respectively accounted for only 0.1% and 2.1% of variance. The correlation accounting for the most shape variance (Mode 1: 50.4%) had only a weak, r=0.36, relationship to intensity Mode 8, which only accounts for 3.0% of variance. Conversely, the correlation accounting for the most intensity variance (Mode 1: 14.9%) had a moderate, r=0.49, correlation to shape Mode 22, which only accounts for 0.2% of variance. Finally, there were 7 correlations with a weak to moderate correlation (0.4≤r≤0.5), but the shape and intensity variance their modes explained was always <2.5% and <2.2%, respectively, except for the one correlation described above.

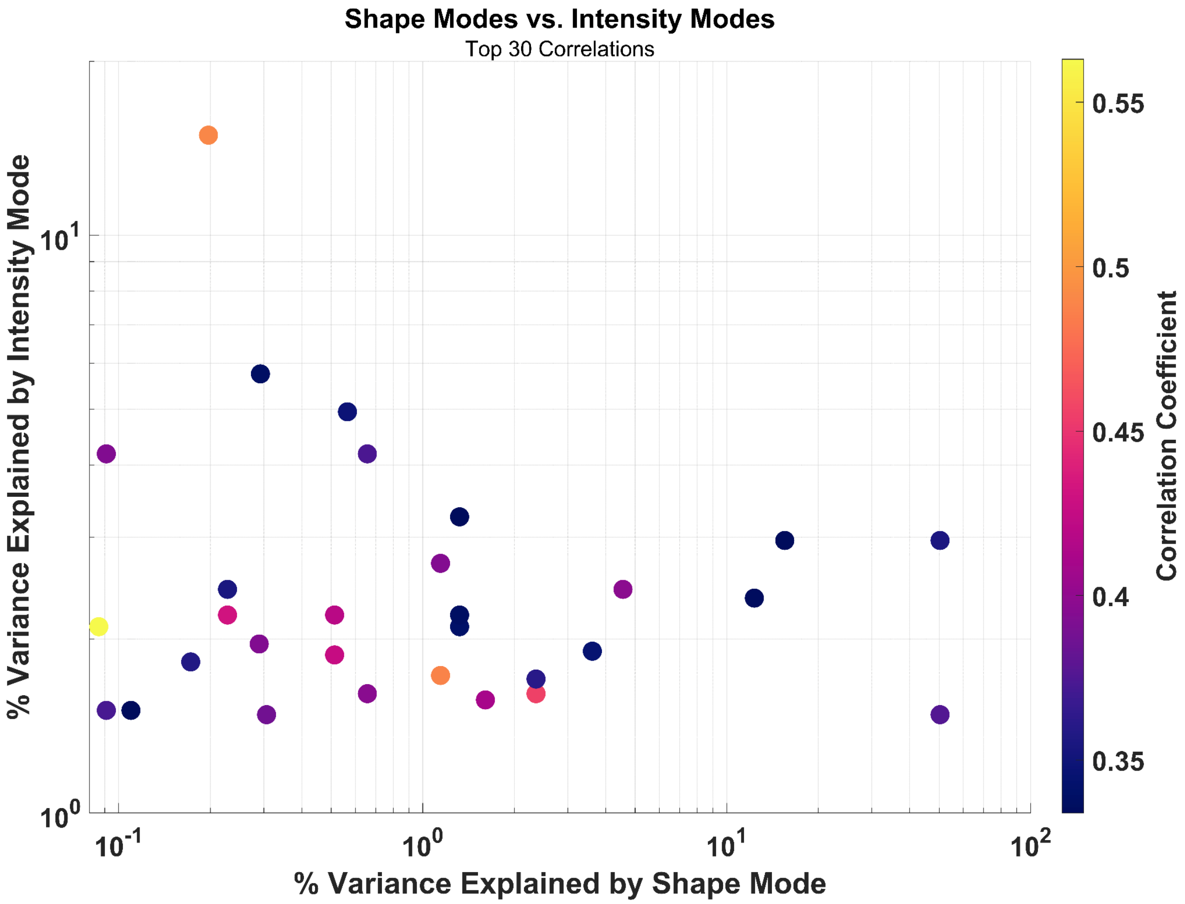


Fig. 9 *– Variance explained and strength of correlation. Correlations are particularly important if their correlation coefficient (r) was high, and they explained a large amount of variance in both modes. These three factors are visualized in this figure, where strong correlations of modes that explain a high percentage of shape and intensity variation would be represented as bright dots in the top right of the graph. Note that x and y axes are logarithmic.*

4. Discussion

To lay the foundation for future shoulder surgical planning tools, this work created three types of models, assessed their comparative efficacy, and determined that use of an independent SSM and SIM to model patient morphological variation instead of a combined SSIM is permissible and would provide more accurate, efficient, and understandable results.

4.1. SSM, SIM, and SSIM Validity

SSM, SIM, and SSIM model validity was assessed using standard methods and demonstrated strong results. Specifically, the input shape data showed high quality with very little shape error (RMSE: 0.25mm) and low correspondence error of 2.5±1.0mm, which is comparable to a published method [42]. Regarding compactness, the SSM accounted for 90% of shape variance with 7 modes of variation, which is more compact than the 9 modes required by Soltanmohammadi et al.'s scapula SSM, which may relate to the patient cohort [28]. The SIM and SSIM respectively required 30 and 26 modes to capture 90% of variance, which is fewer modes than Soltanmohammadi et al.'s scapula SIM. This difference compared to Soltanmohammadi et al.'s work may be as a result of the patient cohort in this study being more homogenous than their general subject sampling and smaller cohort used in this study. Additionally, as seen in Figure 2, the compactness curve of the SSIM falls approximately halfway between the SSM and SIM indicating that the variance normalization method applied to the SSIM input data was effective at equalizing the importance of the shape and intensity datasets despite their differing unit magnitudes and number of data points per bone.

The SSM's specificity is better than that of the SSIM's Shape Specificity as the median error is ~0.4mm lower when using sufficient modes to capture 90% of variance (i.e. 7 for the SSM and 26 for the SSIM). Furthermore, the SSM produces somewhat smaller variance and outliers compared to the SSIM. From a practical standpoint, the SSM's ability to achieve better error using significantly fewer modes is advantageous for potential surgical planning applications. Although, the SIM and SSIM produce approximately parallel trends in Specificity, the SSIM's error is consistently ~16HU higher and more importantly demonstrates markedly great standard deviations, many more outliers, and outliers with greater magnitude. There is no published Specificity data for scapula models for comparison purposes.

The SSM and SIM Sample Size Generalization error dropped below a 1% change for models created with 23 and 17 inputs, respectively, indicating that the addition of more input data will not significantly improve the models' ability to create realistic instances. For the SSIM, Intensity Sample Size Generalization error dropped below a 1% change when using 15 or more inputs; and for Shape when using 32 or more available inputs.

With respect to Standard-Generalization ability, the SSM achieved a median error below 1mm when using 5 or more modes of variation whereas the SSIM's Shape-Generalization median error demonstrated an expected marked decay as the number of modes increased, but still remained above 2mm when using 26 modes. Importantly, the difference was most pronounced when using 5 or fewer modes (e.g. differences >2mm), meaning that a sub-set of the SSIM's 26 modes could not be used when needing to achieve low shape error, while the SSM could use as few as 5 modes. The SIM and SSIM achieve similar Generalization error, with the SIM having a median value less than or equal to 157 HU when using 15 or more modes and the SSIM having a consistent 6HU higher error. However, the SSIM also exhibits outliers that are approximately 40-50HU higher than the SIM. It is important to note that although an error of 157 HU appears high, it is reasonable in the context of patients with osteoporotic and sclerotic bone, where intensity can range by up to 2000 HU [43,44]. This indicates that the SSIM is able to be accurately fit to unseen bones almost to the same level as the SIM, and in fact do so more compactly (using 26 vs 30 modes) but may be limited in that some cases may result in large outliers not seen with the SIM.

Considering these standard metrics together indicates that the developed SSM and SIM are capable of compactly capturing the variation across the patient population and accurately fitting to unseen patient data (i.e. generalization error), both of which are important characteristics for these models to be useful as part of future surgical planning tools where the model is meant to efficiently describe characteristics of patient morphology. The standard metrics indicate that the SSIM is largely as effective as the SIM at fitting to unseen intensity data and produces acceptable shape fitting accuracy but is not as accurate or compact as the SSM and produces markedly higher outliers with respect to shape and intensity in some cases. As a result, using the SSM and SIM separately for surgical planning tools that are meant to fit to patient data is likely to produce better results. Furthermore, the SSM and SIM perform better (e.g. lower specificity error and/or smaller variance/outliers) than the SSIM at generating synthetic bone instances that replicate the patient population, which make the former two models preferrable for biomechanical and surgical planning applications where synthetic bones need to be created. The SSIM's poorer performance can likely be attributed to the challenge in accurately capturing both the shape and intensity variance in one model. It may be possible to improve the efficacy of the SSIM in the future by devising an application/bone specific shape and intensity normalization methodology that balances the importance of the shape and intensity variance better than the method adopted in this study, which was derived from the work of Cootes and Taylor [35].

4.2. Anatomical Measurements and Characteristic Variations

Comparison of the measurements from the datasets used to construct the statistical models are in overall good agreement with previously published measurements [37]. Specifically, the glenoid height and width, and fulcrum axis are in close agreement to Jacobson et al., while the glenoid version angle is greater in this cohort as expected due to this cohort's Watch-type glenoid erosion. The acromion length and posterior-inferior acromion to glenoid center distance are in good agreement to Jacobson et al. with differences of only 8-15%, which can be attributed to the likely differences in selection of the posterior-inferior acromion point which has low inter-rater reliability. Lateral acromion and lateral coracoid distance to the glenoid center differs to Jacobson et al., which is explained by the different definitions of the glenoid center. Specifically, here the glenoid center is identified as the center of the inferior glenoid circle while Jacobson et al. used the glenoid centroid, which is not as relevant in reverse shoulder arthroplasty. Comparing the measurements of the input data to those extracted from 10,000 instances generated by each of the SSM and SSIM, all measurements are in very close agreement. The only noticeable difference is that the SSM measurements have lower standard deviations, however, this is likely attributable to the fact that the variance for the measurements on the input shapes would be more strongly affected by any outliers given there are only 39 samples compared to the 10,000 generated instances.

Consideration of the anatomical measurement to mode of variation weight correlation analysis clearly demonstrates that each mode of variation of the SSM is moderately to very strongly correlated to an anatomical measurement indicating that the model is very efficiently capturing meaningful anatomical features. Furthermore, the fact that each measurement has only one primary and one secondary mode that it is correlated to indicates that different morphological characteristics can be efficiently modeled by generating instances containing variations of only one or two modes. This is in agreement with the results of Burton et al., who performed a similar analysis on an SSM of a health scapula population and found that their model's modes correlated strongly with anatomical measurements [27]. Conversely, with the SSIM, many modes of variation are moderately correlated to the anatomical measurements indicating that the model is not efficiently capturing anatomical features. As well, when generating shape instances, it would be necessary to vary many SSIM modes to mimic the variability of a given morphological feature, but this would also change multiple other features, which is not ideal for multiple research and planning applications. The interactions between SSIM modes does allow for capturing variations that do not directly correspond to a single mode, as the partial effects of a mode may be canceled by varying another mode, but this also makes the SSIM less intuitive and more difficult to use in applications where the intent is to generate bone instance with specific characteristics.

With regard to the characteristic changes observed in the SSM, SIM, and SSIM, modes of variation of these models each capture important morphological changes with some commonalities between the SSM and SSIM with respect to shape. Specifically, SSM Mode 3 and SSIM Mode 3 both capture the expected posterior glenoid erosion characteristic of this patient population. Conversely, the modes of variation of the SIM and SSIM appear to describe the populations intensity variation patterns differently as similarities cannot be identified between them. This can likely be attributed to the fact that the SIM is able to describe intensity variation as efficiently as possible, while the SSIM's modes are a balance between describing shape and intensity variation.

Therefore, although the SSIM when used in its entirety can generate shapes with realistic anatomical measurements, considering the measurement to model weight correlation analysis and the characteristic shape and intensity model variations, it is clear that the SSIM does not efficiently capture and isolate variation of key anatomical shape and intensity features and thus would be more challenging to use for research and planning applications where understanding the model's mode of variation characteristics is critical.

4.3. ***SSM-SIM Correlation***

Given the SSM and SIM's superior performance compared to the SSIM with respect to specificity, generalization ability, and ease of distinguishing the effects of each of their modes of variation, it would be preferable to use these two models in research and surgical planning applications where synthetic bone instances must be generated. However, independent specification of weightings for the modes of variation in these two models will only yield bone instances with realistic shapes and bone property distributions if the modes of each model are uncorrelated and thus any combination of SSM and SIM weightings chosen by a user would yield a plausible bone. Therefore, consideration of the correlation analysis results is critical to assessing whether the SSM and SIM can be used in conjunction for such applications or if the SSIM must be used despite its poorer performance. Considering the correlation results, it is clear that none of the modes of variation of the SSM and SIM are strongly correlated. Furthermore, only a few correlations reached a moderate level ($0.4 \leq r \leq 0.56$) and these were between SSM and SIM modes accounting for little variance. This finding is in agreement with past work by Burton et al. who performed a similar analysis of SSM-SIM correlation for a healthy scapula population [27]. This lack of correlation is despite characteristic shape and bone property changes associated with the glenoid erosion present in this cohort. Therefore, in research and surgical planning applications where bone instances must be generated, independent specification of SSM and SIM mode of variation weightings (i.e. without considering correlations) are unlikely to produce implausible shape and bone property characteristic combinations. Indeed, the level of error that may be introduced by random weight specification is unlikely to be any more significant than noise-related errors introduced in traditional methods based on patient CT scans. However, for any future work that creates new SSMs and SIMs a similar correlation and variance analysis must be undertaken to confirm this result holds.

4.4. Limitations & Strengths

This study has a number of limitations. First, the models were generated using 39 patient datasets due to the difficulty of acquiring datasets with the specific type of glenoid erosion. This is fewer than in previous studies that used healthy subjects; however, the sample size generalization results presented indicate that 39 patient datasets are sufficient, such that addition of further data does not significantly improve the model's ability to accurately describe unseen bone datasets. Second, the available datasets did not provide information about the patient race, which may limit the general applicability of the model; therefore, future research should seek to build a similar model that can ensure equal representation of all racial groups. Third, sex-specific models were not created; however, this was deemed unnecessary because the SSM and SIMs had nearly equal male and female representation and their specificity and generalization metrics were deemed sufficiently accurate without making sex-specific models.

This work benefits from a number of strengths. First, this is the first set of scapula statistical models that was constructed using clinical patient datasets (i.e. of B2/B3 glenoid erosion), which is required for future research and surgical planning tool development where understanding the unique characteristics of patients is essential. Second, these models are based on approximately equal male and female cohorts (19 vs 20), which ensures that the models are not biased to one sex. Third, this is the first study to compare the efficacy of SSM and SIM models to an SSIM, which is an essential step in laying the foundation for the use of these models in future surgical planning tools.

4.5. *Conclusions*

This study successfully developed and validated an SSM, SIM, and SSIM for a patient population having Walch-type glenoid erosions, who have poor outcomes and who could benefit from improved understanding of their characteristics and through future development of statistical model-derived surgical planning tools. The work demonstrated that the SSM and SIM outperform the SSIM at describing bone shape and material property variation, and that SSM and SIM generated synthetic bone instances will produce realistic results because their respective modes of variation are not correlated. Therefore, the developed SSM and SIM can be used in conjunction for future surgical planning tools that are based on the concept of generating and analysing synthetic bone instances or fitting models to patient’s bone. Additionally, if a specific surgical planning application or patient cohort (e.g. one with strong shape-to-intensity correlations) necessitates the use of an SSIM, further work should first be undertaken to devise improved shape and intensity weighting strategies that decrease error and outliers.

5. Funding Information, Ethical Approval, & Conflict of Interest

This work was supported by a Natural Science and Engineering Research Council of Canada Discovery Grant (RGPIN-2018-03894) and a Michael Smith Health Research British Columbia Scholar Award (SCH-2021-1562).

This work received approval from the University of Victoria Research Ethics Board for the use of previously collected patient medical imaging data (Approval #s: 18-217 & 18-226).

All authors have no competing interests to declare in relation to this work except the author GSA who has previously served as an unpaid consultant for ImaScap, which produces pre-operative planning software. However, that relationship had no influence on this work.

6. Data Availability Statement

The datasets and algorithms generated and supporting the findings of this article are obtainable from the corresponding author upon request.

7. References

[1] Bercik, M. J., Kruse, K., Yalizis, M., Gauci, M.-O., Chaoui, J., and Walch, G., 2016, "A Modification to the Walch Classification of the Glenoid in Primary Glenohumeral Osteoarthritis Using Three-Dimensional Imaging," J Shoulder Elbow Surg, **25**(10), pp. 1601–1606.

[2] Walch, G., Badet, R., Boulahia, A., and Khoury, A., 1999, "Morphologic Study of the Glenoid in Primary Glenohumeral Osteoarthritis," J Arthroplasty, **14**(6), pp. 756–760.

[3] Neyton, L., Gauci, M. O., Deransart, P., Collotte, P., Walch, G., and Athwal, G. S., 2019, "Three-Dimensional Characterization of the Anteverted Glenoid (Type D) in Primary Glenohumeral Osteoarthritis," J Shoulder Elbow Surg, **28**(6), pp. 1175–1182.

[4] Rouleau, D. M., Kidder, J. F., Pons-Villanueva, J., Dynamidis, S., Defranco, M., and Walch, G., 2010, "Glenoid Version: How to Measure It? Validity of Different Methods in Two-Dimensional Computed Tomography Scans," J Shoulder Elbow Surg, **19**(8), pp. 1230–1237.

[5] Sirveaux, F., Favard, L., Oudet, D., Huquet, D., Walch, G., and Mole, D., 2004, "Grammont Inverted Total Shoulder Arthroplasty in the Treatment of Glenohumeral Osteoarthritis with Massive Rupture of the Cuff," J Bone Joint Surg Br, **86-B**(3), pp. 388–395.

[6] Chamseddine, M., Breden, S., Pietschmann, M. F., Müller, P. E., and Chevalier, Y., 2019, "Periprosthetic Bone Quality Affects the Fixation of Anatomic Glenoids in Total Shoulder Arthroplasty: In Vitro Study," J Shoulder Elbow Surg, **28**(1), pp. e18–e28.

[7] Mariaux, S., Obrist, R., Farron, A., Becce, F., and Terrier, A., "Is Preoperative Glenoid Bone Mineral Density Associated with Aseptic Glenoid Implant Loosening in Anatomic Total Shoulder Arthroplasty?"

[8] Knowles, N. K., Athwal, G. S., Keener, J. D., and Ferreira, L. M., 2015, "Regional Bone Density Variations in Osteoarthritic Glenoids: A Comparison of Symmetric to Asymmetric (Type B2) Erosion Patterns," J Shoulder Elbow Surg, **24**(3), pp. 425–432.

[9] Mahaffy, M. D., Knowles, N. K., Berkmortel, C., Abdic, S., Walch, G., Johnson, J. A., and Athwal, G. S., 2020, "Density Distribution of the Type E2 Glenoid in Cuff Tear Arthropathy," J Shoulder Elbow Surg, **29**(1), pp. 167–174.

[10] Letissier, H., Chaoui, J., Bercik, M. J., Boileau, P., Le Nen, D., Stindel, E., and Walch, G., 2020, "Glenoid Subchondral Bone Density in Osteoarthritis: A Comparative Study of Asymmetric and Symmetric Erosion Patterns," Orthopaedics & Traumatology: Surgery & Research, **106**(6), pp. 1127–1134.

[11] Walch, G., Vezeridis, P. S., Boileau, P., Deransart, P., and Chaoui, J., 2015, "Three-Dimensional Planning and Use of Patient-Specific Guides Improve Glenoid Component Position: An in Vitro Study.," Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons ... [et al.], **24**(2), pp. 302–9.

[12] Boileau, P., Cheval, D., Gauci, M.-O., Holzer, N., Chaoui, J., and Walch, G., 2018, "Automated Three-Dimensional Measurement of Glenoid Version and Inclination in Arthritic Shoulders," J Bone Joint Surg, **100**(1), pp. 57–65.

[13] Sarkalkan, N., Weinans, H., and Zadpoor, A. A., 2014, "Statistical Shape and Appearance Models of Bones.," Bone, **60**, pp. 129–40.

[14] Inyang, A. O., Fouefack, J.-R., Sivarasu, S., Roche, S., Borotikar, B., Burdin, V., and Mutsvangwa, T., 2017, "Assessment of 3D Morphological Characteristics of the Shoulder Bones Using Statistical Shape Modeling: Prospective Application to Handedness," *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, pp. 1629–1632.

[15] Tsai, T.-Y., Li, J.-S., Wang, S., Li, P., Kwon, Y.-M., and Li, G., 2015, “Principal Component Analysis in Construction of 3D Human Knee Joint Models Using a Statistical Shape Model Method.,” Comput Methods Biomech Biomed Engin, **18**(7), pp. 721–9.

[16] Hollenbeck, J. F. M., Cain, C. M., Fattor, J. A., Rullkoetter, P. J., and Laz, P. J., 2018, “Statistical Shape Modeling Characterizes Three-Dimensional Shape and Alignment Variability in the Lumbar Spine,” J Biomech, **69**, pp. 146–155.

[17] Mayya, M., Poltaretskyi, S., Hamitouche, C., and Chaoui, J., 2013, “Scapula Statistical Shape Model Construction Based on Watershed Segmentation and Elastic Registration,” *2013 IEEE 10th International Symposium on Biomedical Imaging*, IEEE, pp. 101–104.

[18] Casier, S. J., Van den Broecke, R., Van Houcke, J., Audenaert, E., De Wilde, L. F., and Van Tongel, A., 2018, “Morphologic Variations of the Scapula in 3-Dimensions: A Statistical Shape Model Approach,” J Shoulder Elbow Surg, **27**(12), pp. 2224–2231.

[19] Sintini, I., Burton, W. S., Sade, P., Chavarria, J. M., and Laz, P. J., 2018, “Investigating Gender and Ethnicity Differences in Proximal Humeral Morphology Using a Statistical Shape Model,” Journal of Orthopaedic Research, **36**(11), pp. 3043–3052.

[20] Lee, E. C. S., Roach, N. T., Clouthier, A. L., Bicknell, R. T., Bey, M. J., Young, N. M., and Rainbow, M. J., 2020, “Three-Dimensional Scapular Morphology Is Associated with Rotator Cuff Tears and Alters the Abduction Moment Arm of the Supraspinatus,” Clinical Biomechanics, **78**, p. 105091.

[21] Plessers, K., Verhaegen, F., van Dijck, C., Wirix-Speetjens, R., Debeer, P., Jonkers, I., and vander Sloten, J., 2020, “Automated Quantification of Glenoid Bone Defects Using 3-Dimensional Measurements,” J Shoulder Elbow Surg, **29**(5), pp. 1050–1058.

[22] Plessers, K., vanden Berghe, P., van Dijck, C., Wirix-Speetjens, R., Debeer, P., Jonkers, I., and vander Sloten, J., 2018, “Virtual Reconstruction of Glenoid Bone Defects Using a Statistical Shape Model,” J Shoulder Elbow Surg, **27**(1), pp. 160–166.

[23] Poltaretskyi, S., Chaoui, J., Mayya, M., Hamitouche, C., Bercik, M. J., Boileau, P., and Walch, G., 2017, “Prediction of the Pre-Morbid 3D Anatomy of the Proximal Humerus Based on Statistical Shape Modelling.,” Bone Joint J, **99-B**(7), pp. 927–933.

[24] vanden Berghe, P., Demol, J., Gelaude, F., and vander Sloten, J., 2017, “Virtual Anatomical Reconstruction of Large Acetabular Bone Defects Using a Statistical Shape Model,” Comput Methods Biomech Biomed Engin, **20**(6), pp. 577–586.

[25] Verhaegen, F., Meynen, A., Pitocchi, J., Debeer, P., and Scheys, L., 2023, “Quantitative Statistical Shape Model-Based Analysis of Humeral Head Migration, Part 2: Shoulder Osteoarthritis,” Journal of Orthopaedic Research®, **41**(1), pp. 21–31.

[26] Verhaegen, F., Meynen, A., Matthews, H., Claes, P., Debeer, P., and Scheys, L., 2021, “Determination of Pre-Arthropathy Scapular Anatomy with a Statistical Shape Model: Part I—Rotator Cuff Tear Arthropathy,” J Shoulder Elbow Surg, **30**(5), pp. 1095–1106.

[27] Burton, W. S., Sintini, I., Chavarria, J. M., Brownhill, J. R., and Laz, P. J., 2019, “Assessment of Scapular Morphology and Bone Quality with Statistical Models,” Comput Methods Biomech Biomed Engin, **22**(4), pp. 341–351.

[28] Soltanmohammadi, P., Elwell, J., Veeraraghavan, V., Athwal, G. S., and Willing, R., 2020, “Investigating the Effects of Demographics on Shoulder Morphology and Density Using Statistical Shape and Density Modeling,” J Biomech Eng, **142**(12).

[29] Nicolella, D. P., and Bredbenner, T. L., 2012, “Development of a Parametric Finite Element Model of the Proximal Femur Using Statistical Shape and Density Modeling,” Comput Methods Biomech Biomed Engin, **15**(2), p. 101.

[30] Soltanmohammadi, P., 2019, “Finite Element Analysis of Hollow-Stemmed Shoulder Implants in Different Bone Qualities Derived from a Statistical Shape and Density Model,” Electronic Thesis and Dissertation Repository.

[31] Campbell, J. Q., and Petrella, A. J., 2016, “Automated Finite Element Modeling of the Lumbar Spine: Using a Statistical Shape Model to Generate a Virtual Population of Models,” J Biomech, **49**(13), pp. 2593–2599.

[32] Bredbenner, T. L., Mason, R. L., Havill, L. M., Orwoll, E. S., and Nicolella, D. P., 2014, “Fracture Risk Predictions Based on Statistical Shape and Density Modeling of the Proximal Femur,” Journal of Bone and Mineral Research, **29**(9), pp. 2090–2100.

[33] Besl, P. J., and McKay, N. D., 1992, “Method for Registration of 3-D Shapes,” *Robotics - DL Tentative*, P.S. Schenker, ed., International Society for Optics and Photonics, pp. 586–606.

[34] Amberg, B., Romdhani, S., and Vetter, T., 2007, “Optimal Step Nonrigid ICP Algorithms for Surface Registration,” *2007 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 1–8.

[35] Cootes, T., and Taylor, C., 2004, *Statistical Models of Appearance for Computer Vision*, Manchester.

[36] Taylor, C., Twining, C., and Davies, R., 2008, “Evaluation of Statistical Models,” *Statistical Models of Shape: Optimisation and Evaluation*, Springer London, London.

[37] Jacobson, A., Gilot, G., Hamilton, M., Greene, A., Flurin, P.-H., Wright, T., Zuckerman, J., and Roche, C., 2015, “Glenohumeral Anatomic Study. A Comparison of Male and Female Shoulders with Similar Average Age and BMI - PubMed,” Bull Hosp Joint Dis, **73**(Suppl 1), pp. 68–78.

[38] Knowles, N. K., Keener, J. D., Ferreira, L. M., and Athwal, G. S., 2015, “Quantification of the Position, Orientation, and Surface Area of Bone Loss in Type B2 Glenoids,” J Shoulder Elbow Surg, **24**(4), pp. 503–510.

[39] Knowles, N. K., Ferreira, L. M., and Athwal, G. S., 2016, “Premorbid Retroversion Is Significantly Greater in Type B2 Glenoids,” J Shoulder Elbow Surg, **25**(7), pp. 1064–1068.

[40] Boileau, P., Gauci, M.-O., Wagner, E. R., Clowez, G., Chaoui, J., Chelli, M., and Walch, G., 2019, “The Reverse Shoulder Arthroplasty Angle: A New Measurement of Glenoid Inclination for Reverse Shoulder Arthroplasty,” J Shoulder Elbow Surg, **28**(7), pp. 1281–1290.

[41] Li, X., Olszewski, N., Abdul-Rassoul, H., Curry, E. J., Galvin, J. W., and Eichinger, J. K., 2018, “Relationship Between the Critical Shoulder Angle and Shoulder Disease,” JBJS Rev, **6**(8), pp. e1–e1.

[42] Danckaers, F., Huysmans, T., Lacko, D., Ledda, A., Verwulgent, S., Dongen, S. van, and Sijbers, J., 2014, “Correspondence Preserving Elastic Surface Registration with Shape Model Prior,” *2014 22nd International Conference on Pattern Recognition*, IEEE, pp. 2143–2148.

[43] Daalder, M. A., Venne, G., Sharma, V., Rainbow, M., Bryant, T., and Bicknell, R. T., 2018, “Trabecular Bone Density Distribution in the Scapula Relevant to Reverse Shoulder Arthroplasty,” JSES Open Access, **2**(3), p. 174.

[44] Fat, D. L., Kennedy, J., Galvin, R., O’Brien, F., Grath, F. M., and Mullett, H., 2012, “The Hounsfield Value for Cortical Bone Geometry in the Proximal Humerus--an in Vitro Study,” Skeletal Radiol, **41**(5), pp. 557–568.

[45] Azita Sharif Ahmadian, by, 2021, “Development of Statistical Shape and Intensity Models of Eroded Scapulae to Improve Shoulder Arthroplasty,” University of Victoria.

[46] Michalski, A. S., Besler, B. A., Michalak, G. J., and Boyd, S. K., 2020, “CT-Based Internal Density Calibration for Opportunistic Skeletal Assessment Using Abdominal CT Scans,” Med Eng Phys, **78**, pp. 55–63.

[47] Pakdel, A., Fialkov, J., and Whyne, C. M., 2016, “High Resolution Bone Material Property Assignment Yields Robust Subject Specific Finite Element Models of Complex Thin Bone Structures,” J Biomech, **49**(9), pp. 1454–1460.

Figure Captions

Fig. 1 – Registration Process. *Illustration of registration process on a target subject model using optimized NR-ICP method.*

Fig. 2 – Compactness of SSM, SIM, SSIM. *Compactness describes the cumulative variance explained by 1 to m modes of variation, where m≤N-1, where N is the number of training datasets (N=39 in this study).*

Fig. 3 – Shape Specificity of SSM & SSIM. *Shape specificity indicates the amount of error between an artificially generated shape instance and the closest dataset used to create the statistical model, which varies with the number of modes of variation used to create the instance. Here specificity is calculated for instances generated using 1 to m modes, where* $m \leq M_{90}$ *where* M_{90} *is the number of modes capturing 90% of variance as determined by Compactness.* M_{90} *is 7 for the SSM and 26 for the SSIM.*

Fig. 4 – Intensity Specificity of SIM and SSIM. *Intensity specificity indicates the amount of error between an artificially generated bone material property instance and the closest dataset used to create the statistical model, which varies with the number of modes of variation used to create the instance. Here specificity is calculated for instances generated using 1 to m modes, where* $m \leq M_{90}$ *where* M_{90} *is the number of modes capturing 90% of variance as determined by Compactness.* M_{90} *is 30 for the SIM and 26 for the SSIM.*

Fig. 5 – SSM, SIM, & SSIM Generalization. *(a) SSM Generalization, (b) SSIM Shape Generalization, (c) SIM Generalization, (d) SSIM Intensity Generalization. Generalization is the error when a statistical model is fit to an input dataset not used to train the model. Standard-Generalization is calculated for instances generated using 1 to m modes of models created using N-1 datasets (N=39), where m≤M90 and M90 is the number of modes capturing 90% of variance as determined by Compactness. M90 is 7 for the SSM, 30 for the SIM, and 26 for the SSIM. Sample-Size-Generalization is calculated the same way, but the fitting is conducted using all available modes of variation of the model, which equals one less than the number of datasets used to make it, which ranges from 2 to N-1.*

Fig. 6 – SSM Characteristic Shape Changes. *The scapula in each pane represents the average shape with changes in shape shown by the colormap variation. The colormaps represent the magnitude of spatial variation when the weighting for the given mode of variation is set to 1.96 SDs.*

Fig. 7 – SIM Characteristic Bone Property Changes. *The colormaps represent the change in intensity when the weighting for the given mode of variation is set to 1.96 SDs. Given that change happens throughout the bone volume, to enable improved visualization, only points with increases or decreases in intensity in the top 20% of absolute change for each mode (i.e. each figure pane) are displayed. Furthermore, to improve readability of the colorbar, any point whose increase or decrease is in the top 10% across all modes (i.e. top 10% of all changes shown in entire figure) is shown as having maximal red or blue saturation, respectively (i.e. if the change is above a given HU threshold the color saturation remains constant). Note that the sign of the weightings of each mode of variation is arbitrary and thus a + or - SD does not inherently infer increases and decreases in intensity, respectively.*

Fig. 8 – SSIM characteristic shape and intensity variations. *These shape and intensity subplots were created using the same process as for the SSM and SIM in the previous two figures.*

Fig. 9 – Variance explained and strength of correlation. *Correlations are particularly important if their correlation coefficient (r) was high, and they explained a large amount of variance in both modes. These three factors are visualized in this figure, where strong correlations of modes that explain a high percentage of shape and intensity variation would be represented as bright dots in the top right of the graph. Note that x and y axes are logarithmic.*

Table Captions

Table 1. *Scan parameters used for imaging by GE Lightspeed CT scanner.*

Table 2. *The demographic data of all scapulae in the training set.*

Table 3. *Scapular measurement results for 16 parameters.*

Supplemental Materials

Additional Methods Details

Generating Models & Establishing Correspondence

To create an accurate scapula SSM, SIM, and SSIM, all of the patient bone geometries must be described using the same point architecture (e.g. same number and enumeration) with matching enumerated points located on corresponding anatomical sites across all bone meshes. Surface meshes with correspondence are generated first for each bone and used to create an SSM, while volumetric meshes with correspondence are subsequently generated based on the created surfaces meshes and used to create an SIM and SSIM as described below. Mimics (Materialise NV) was used to segment all scans and generate a 3D scapula surface mesh for each. Subsequently, 3-matic (Materialise NV) was used to mirror right scapulae to appear as left and to uniformly remesh the models to have an average edge length of 0.6 mm resulting in an average of 106k nodes per model. From the 39 models, one was chosen as the 'source' model when establishing correspondence between all models. This source model was chosen by finding the model that was closest to the median shape of the 39 models.

To create the corresponding surface meshes, three registration steps were undertaken (Figure 1): 1) rigid registration to achieve initial alignment between the source and target models (N.B. 'target' refers to the patient model for which the source is trying to be registered), 2) uniform scaling of the source model to match the size of the target model, and 3) mesh morphing to deform the source model to match the exact geometry of the target while maintaining the original anatomical meaning each vertex (i.e. maintaining correspondence). The first two steps were achieved using well established algorithms based on Besl et al.'s iterative closest point method [33], while for the third step, the authors optimized Amberg et al.'s correspondence preserving non-rigid iterative closest point (NR-ICP) registration algorithm for use with scapula geometries [34,45]. In this approach, affine transformations are iteratively assigned to each vertex of the source model to progressively deform its geometry to match the target. This is achieved by using a regularisation that initially has high stiffness to maintain correspondence but is iteratively relaxed to allow progressively more localized deformation. Further details can be found in the dissertation of Sharif-Ahdmadian [32].

To construct a meaningful SSM it is critical to achieve high quality fitting of the deformed source mesh to the shape of each target mesh's geometry. This was assessed by measuring the Root Mean Squared (RMS) distance error and the Hausdorff distance error between the target and deformed source meshes. RMS distance error measures the distance from each point of deformed source surface to the closest point on the target surface. Hausdorff distance measures the greatest of all the distances from a point in one deformed source surface to the closest point in its target surface, which gives a measure of the worst fit node in the model.

To produce a meaningful SSM, it is also critical to achieve high quality nodal correspondence where the same enumerated point is accurately located on the same anatomical location across all surface meshes after the NR-ICP algorithm converges. Therefore, the resulting correspondence error was evaluated by comparing the positions of ground truth points and the corresponding point locations produced by the algorithm. This was achieved by manually selecting anatomical landmarks on the source surface mesh and recording their enumeration, as well as identifying the same anatomical landmarks on a set of target meshes and recording their spatial coordinates. After non-rigid registration, the coordinates of the previously recorded source nodes were captured based on their unchanged enumerated index values. The Euclidean distance between the landmarks manually identified on the target meshes and those algorithmically determined from the deformed meshes were then calculated for each model in our evaluation set. Specifically, the error for 12 anatomical landmarks were evaluated: four points on the glenoid fossa (superior, inferior, anterior and posterior side of the glenoid), three on the acromion process, two on the coracoid process, one at the medial termination of the scapular spine, one on superior angle, and one at the inferior angle. It is important to understand the level of variability in the manual selection process to place the registration error results in context.

Therefore, an inter- and intra-observer reliability analysis was performed on a set of 10 scapulae with the 12 anatomical landmarks selected by two independent observers and each observer repeated the landmark selection process two times per set.

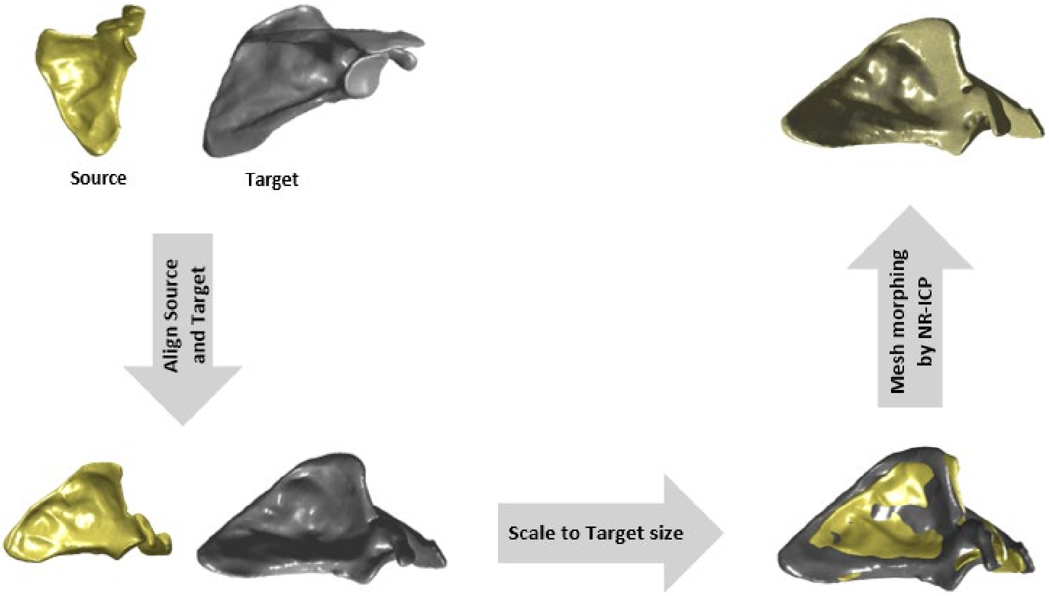


Figure 1. Illustration of registration process on a target sample model using optimized NR-ICP method.

With the corresponding surface meshes created for each of the 39 patient scans, volumetric meshes with both surface and internal correspondence can be generated. To create volumetric meshes with correspondence, a 'source' volumetric mesh is required. This is created in 3-matics (Materialise NV) by using the vertices of the source surface mesh as seed points to generate a uniform volumetric mesh with average edge length of 0.6mm, which resulted in ~1.1 million elements. Use of the source surface mesh ensures the surface nodes of the volumetric model has the same enumeration as the source surface mesh, which allows later integration of SSM and SIM results, and greatly simplifies the volumetric mesh morphing process required to deform the volumetric source to each target scapula. As a result of this common enumeration of the vertices on the surface of the volumetric source, creating corresponding volumetric meshes for each of the 39 patients is easily achieved. This was done using the method described by Soltanmohammadi et al.[30]: 1) calculating the deformation of each vertex in the mesh from their original coordinates in the source surface mesh to their coordinates in the deformed source surface mesh for a given patient, 2) generating a finite element simulation in Abaqus (Dassault Systems) with the source volume mesh as the model and the calculated surface vertex deformations as displacement boundary conditions, and 3) solving the system to yield the deformed volumetric mesh for that patient.

Once the 39 volumetric meshes with correspondence were established, using 3-matics and Mimics software (Materialise NV), they were each aligned with their respective patient CT scan and the Mimics material assignment feature was used to assign a Hounsfield Unit value to each volumetric element. Prior to material property assignment, two steps were taken to improve the accuracy of the assigned Housfield Unit value: 1) a partial volume correction algorithm derived from Pakdel et al. was applied to each patient scan to mitigate the effects of the cortical shell Hounsfield Unit values being reduced due to their inclusion of both cortical bone and soft tissues, 2) the scans were internally calibrated using the method of Michalski et al. to ensure the accuracy of the Hounsfield Unit values as clinical scans are not routinely acquired with a calibration phantom within the field of view[46,47]. With these properties assigned, the surface and volumetric meshes are ready for use in making statistical shape and intensity models.

Calculating & Evaluating Statistical Models

SSM generation was undertaken by first calculating the covariance matrix of an array with rows composed of the x, y, and z coordinates of each node of a mesh and each of N columns representing a given patient's geometry, where N = 39 patients. To yield a tractable problem, covariance was computed to yield an N x N matrix. Principal Component Analysis (PCA) was then applied to the covariance matrix yielding N-1 mutually independent modes of variation (i.e. eigenvectors or principal components) ordered from most to least variance explained (i.e. descending eigenvalue). The SIM was created using the same process but with its data values representing the HU values of each volumetric element with no reference to variation in spatial coordinates (i.e. each row was an HU value for a given volumetric element and each column represented a given subject's intensity data). Finally, the SSIM was created by combining the spatial coordinates for all points that make up the nodes of the volumetric mesh along with the intensity data for each volumetric element such that the first P*3 rows represent the x, y, and z coordinates of the P volume nodes, and the remainder of the rows contain the volumetric element intensity. To account for differences in the unit magnitudes and number of data points for the shape and intensity data, a scaling factor method was adapted from the work of Cootes et al. (2004). The scaling factor is the square root of the ratio of the total intensity variance to the total shape variance and was applied to all of the shape data to equalize the variance. The scaling adapted from Cootes is calculated as follows: 1) create the input data matrix, S, with dimensions (3P+I x 39) where P is the number of shape points, I is the number of intensity data elements, and 39 is the number of subjects; 2) calculate the average scapula S_avg with dimensions (3P+I x 1) and subtract this from each column of S to give a matrix, D, which is the variability of each scapula from the mean; 3) calculate the variance of each row of the matrix D to give a vector, V, with dimensions (3P+I x 1); 4) calculate the total shape variance by summing rows 1:3P and the total intensity variance by summing rows (3P+1):(3P+I); 5) calculate the scaling factor r^2 = total intensity variance/total shape variance; 6) multiply all shape data [D(1:3P,1:39)] by r, while leaving all intensity data [D((3P+1):(3P+I)),1:39)] unchanged. The data is now scaled in an analogous manner to the method proposed by Cootes and Taylor and the matrix, D, can be used for computation of the covariance matrix. The same covariance and PCA calculations were then conducted as with the SSM and SIM.

Each of the statistical models (i.e. the SSM, SIM, and SSIM) were evaluated using three widely accepted metrics to assess their efficacy[36]:

1. **Compactness:** The compactness is defined as the cumulative variance of the Mth mode used in the object reconstruction (e.g. shape for an SSM).

$$C(M) = \frac{\sum_{i=1}^{M} \lambda_m}{\sum \lambda}$$

Where λ_m is the eigenvalue of the mth component of the PCA.

2. **Specificity:** Specificity measures a model's ability to generate instances of the object class that are similar to those in the training set. It is measured by generating a large set of object instances using the constructed model. The approximation error is defined as the difference between the generated object instance and its most similar sample in the training data. The specificity is the average approximation error of all the generated object instances.

$$S(M) = \frac{1}{N}\sum_{i=1}^{N} \parallel S_i(M) - \acute{S}_i \parallel_2$$

Where $S_i(M)$ are object instances randomly generated using M principal modes and $\acute{S}_i$ is the deformed source model (N.B. there it has point correspondence with the generated instance) from the training dataset that is nearest in characteristics (i.e. shape or intensity) to $S_i(M)$. $S_i(M)$ is given by:

$$S(M) = \bar{S} + \sum_{m=1}^{M} P_m b_m$$

Where $\bar{S}$ is the mean shape, m is the mth principal component, M is the total number of components used to generate the object instance, P_m is the eigenvector from the PCA and b_m is a principal component weighting factor such that:

$$b_m \epsilon [-3\sqrt{\lambda_m}, 3\sqrt{\lambda_m}]$$

Where λ_m is the eigenvalue of the mth component of the PCA and its square root equals the standard deviation in the population. Calculation of average Specificity was conducted by generating 10,000 instances with randomly generated weighting factors for M = 1 to M_{max} principal components, where M_{max} is chosen based on the Compactness of the model such that it includes the number of modes need to capture 90% of the variance.

3. **Generalization ability:** Generalization measures a model's ability to represent unseen instances of the class of object. Two forms of Generalization assessment as possible and each provide useful information. First, Generalization can be assessed for an increasing number of principal components (i.e. modes of variation) to assess the influence that number has on the ability to accurately generate unseen objects. Second,

Generalization can also be used to assess the number of samples that are needed to produce a model for which adding more samples would not improve the quality of generalization; this can be termed "Sample Size Generalization". For the former, the Generalization ability of each model is measured from the training set using leave-one-out (LOO) cross validation. A model is built using all but one member of the training set (i.e. N-1) and the resulting model is fit to the excluded sample using increasing principal components from 1 to N-2. The accuracy to which the model can describe the unseen sample is measured and the process is repeated until all samples are tested. The approximation error (i.e. Generalization ability) is given by:

$$G(M) = \frac{1}{N}\sum_{i=1}^{N} \parallel S_i - \acute{S}_i(M) \parallel_2$$

Where $G(M)$ is the Generalization when the excluded sample is fit using M principal components, S_i is the excluding model, N is the number of LOO repetitions (NB: this equals the total training dataset size), and $\acute{S}_i(M)$ is the fitted object defined by:

$$\acute{S}(M) = \bar{S} + \sum_{m=1}^{M} P_m b_m$$

Where the variables on the right have the same meaning as for Specificity.

For the latter, Sample Size Generalization is assessed by excluding one sample and generating models using an increasing number of randomly selected samples drawn from the remaining training set from 2 to (N-1), and, for each model, fitting it to the excluded sample and calculating the approximation error. This is repeated with each sample being excluded and then the approximation error is averaged across all excluded samples for each model made with a given number of samples. The approximation error (i.e. Sample Size Generalization ability) is given by:

$$G(n) = \frac{1}{N}\sum_{i=1}^{N} \parallel S_i - \acute{S}_i(M) \parallel_2$$

Where $G(n)$ is the Sample Size Generalization for models made using n randomly selected samples and fit to the excluded sample using all available principal components (n-1). S_i is the excluding sample, N is the number of LOO repetitions (NB: this equals the total training dataset size), and $\acute{S}_i(M)$ is the fitted object defined by:

$$\acute{S}(M) = \bar{S} + \sum_{m=1}^{M} P_m b_m$$

Where M = n-1 and all other variables on the right have the same meeting as for Specificity.

Anatomical Measurements

Although the above described Specificity and Generalization metrics are effective at assessing the quality of SSMs in replicating overall shape characteristics, specific anatomical measurements have previously been shown to be important to shoulder surgical planning [37–41]. Therefore, the SSM was also assessed in terms of its ability to generate shape instances with specific anatomical measurements that are in agreement with the training dataset used and with previously published values.

Given the nodal correspondence of any shape instance generated by a given SSM, measurement of a desired anatomical variable can be automated across the range of variation by simply defining the node enumerations that yield a desired anatomical measurement and applying that to a set of generated instances. The reliability of this automated measurement method is dependent on the correspondence quality of the data used to create the SSM, the evaluation of which was described above and results of which are presented in the Results section.

To calculate the desired set of anatomical measurements, fourteen points needed to be identified on the source mesh. Once identified, these points are extracted from the deformed source meshes of the 39 samples used to create the SSM, and from two sets of 10,000 shape instances each generated as described for the Specificity evaluation: 1) the first set using the SSM, 2) the second set using the SSIM. Sixteen anatomic parameters are calculated in different anatomical regions of the scapula including 9 length-related and 7 angle-related measurements which are presented in Table 3.

Prior to undertaking anatomical measurements, the scapular geometries were transformed to an anatomical coordinates system as some measurements are defined with respect to anatomical axes. The points used to define the coordinate system are given in Table 1 and the result is shown in Figure 2. The 14 scapular points used to calculate each of the measurements are defined in Table 2 and shown in Figure 3. The definitions of the 16 measurements and their equations are given in Table 3 & 4 and shown in Figure 4, 5, 6, & 7.

Table 1. Description of three characteristic points of scapula.

Point	**Point Name**	**Point Description**
A	Glenoid center	The center of the best-fitting inferior circle on the glenoid cavity while facing the glenoid from lateral to medial.
B	Trigonum spinae	The point on the most medial part of scapular spine located at the intersection of the scapular spine and medial border of the scapular body while facing directly from anterior to posterior.
C	Angulus inferior	The most inferior point on the scapula body while facing directly from posterior to anterior.

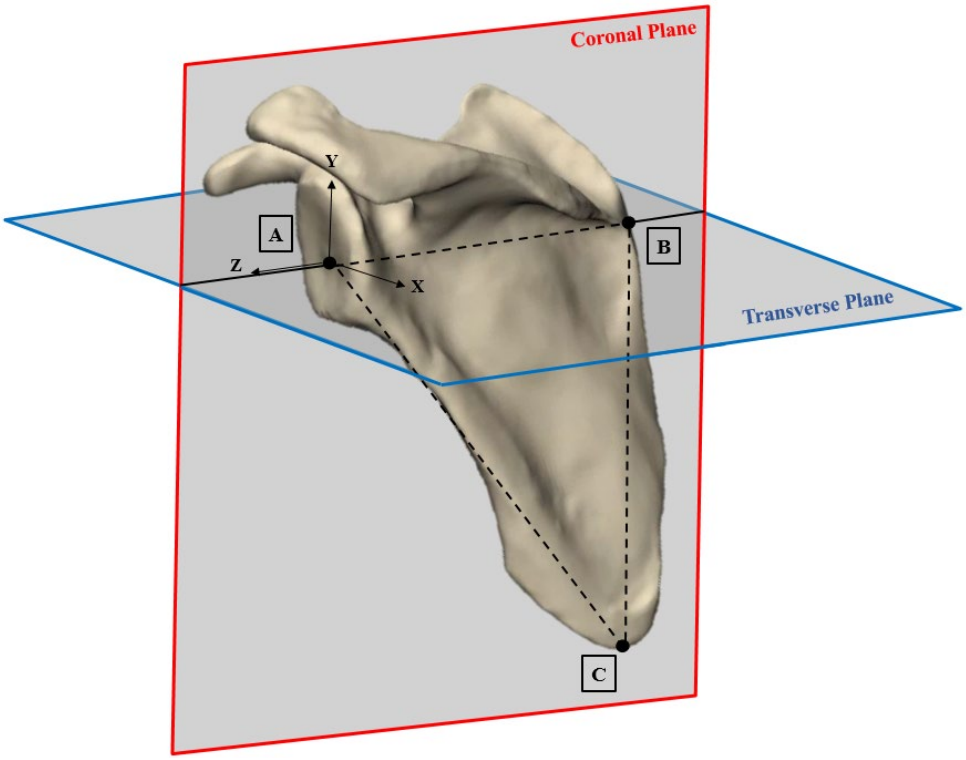
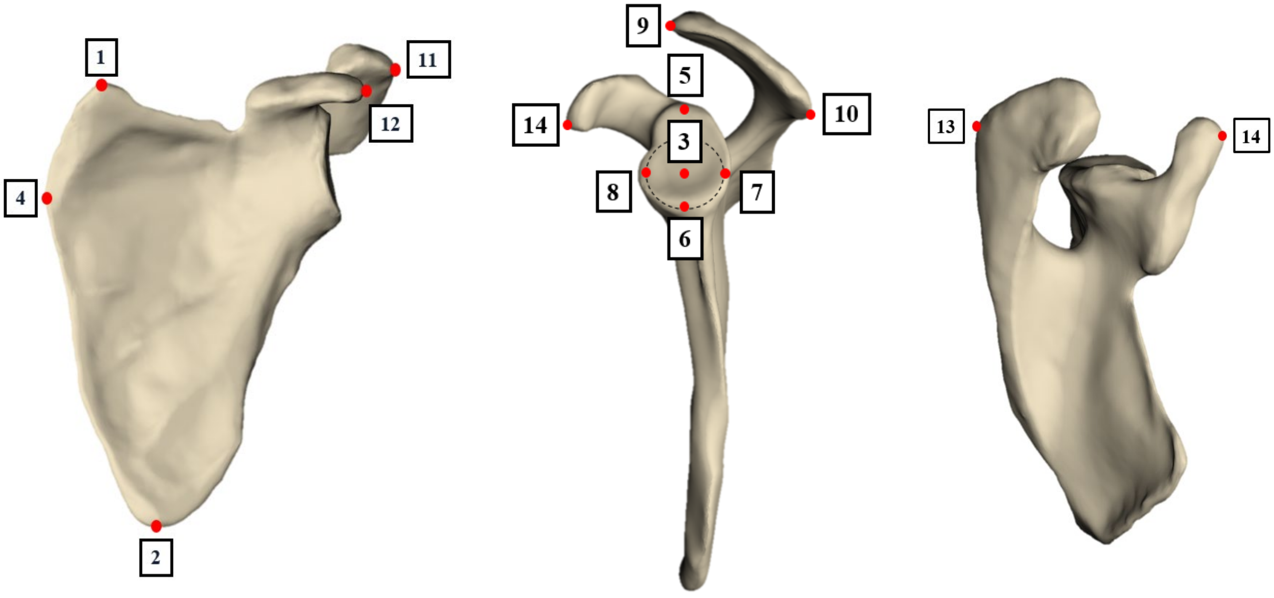


Figure 2. Schematic view of coronal and transverse planes along with three characteristic points and corresponding coordinate system for a left shoulder.

Table 2. Description of 14 anatomical points for measurement purposes.

Point Number	Anatomical Point Name	Point Description
P1	Superior border point	The most superior point on the scapula body while facing directly from posterior to anterior.
P2	Inferior border point	The most inferior point on the scapula body while facing directly from posterior to anterior.
P3	Glenoid center point	The center of the best-fitting inferior circle on the glenoid cavity while facing the glenoid from lateral to medial [119].
P4	Trigonum spinae point.	The point on the most medial part of scapular spine located at the intersection of the scapular spine and medial border of the scapular body while facing directly from anterior to posterior.
P5	Superior glenoid point	Position on superior glenoid rim where long head of biceps inserts while facing the glenoid from lateral to medial.
P6	Inferior glenoid point	Position most inferior on glenoid rim while facing the glenoid from lateral to medial.
P7	Posterior glenoid point	Position on posterior glenoid rim at most extreme point of lower width of glenoid while facing the glenoid from lateral to medial.
P8	Anterior glenoid point	Position on anterior glenoid rim at most extreme point of lower width of glenoid while facing the glenoid from lateral to medial.
P9	Acromion tip point_1	The most anterior point at the end of acromion process while facing the scapula from the lateral to medial.
P10	Acromion process point _1	The most inferio-posterior point of the acromion process while facing the scapula from the lateral to medial.
P11	Acromion tip point_2	The most lateral point at the end of acromion process while facing the scapula from posterior to anterior.
P12	Coracoid tip point_1	The most lateral point at the end of coracoid process while facing the scapula from anterior to posterior.
P13	Acromion process point_2	The most posterior-lateral point of the acromion process while facing the scapula from superior to inferior.
P14	Coracoid tip point_2	The most anterior point at the end of coracoid process while facing the scapula from both superior to inferior and lateral to medial.



(a) **(b)** **(c)**

Figure 3. Location of 14 points selected on a left scapula for the anatomical measurements is illustrated in three different views. (a): Anterior view, (b): Lateral view and (c): Superior view.

Table 3. List of length-related parameters for the anatomic measurements.

Parameter Number	Parameter Name	Parameter Description	Measurement Formula
1	Scapula Height	Distance between the most superior and inferior aspects of the scapula body.	$\lvert\bar{P}_1 - \bar{P}_2\rvert$
2	Scapula Width	Distance between the center of the glenoid and medial end of the spine where it intersects the medial scapular border.	$\lvert\bar{P}_3 - \bar{P}_4\rvert$
3	Glenoid Height	Height of the glenoid cavity along the superior-inferior (SI) direction or Y axis.	$\lvert Y_{P5} - Y_{P6}\rvert$
4	Glenoid Width	Distance between the most anterior and posterior points of the glenoid cavity and which is perpendicular to the SI axis.	$\lvert\bar{P}_7 - \bar{P}_8\rvert$
5	Acromion Length	Length of the acromion process from the anterior apex to the most posterior point.	$\lvert\bar{P}_9 - \bar{P}_{10}\rvert$
6	Lateral Acromion to Glenoid Center Distance	Distance between the most lateral point of the acromion and the center of the glenoid cavity along the medial-lateral direction or Z axis.	$Z_{P11} - Z_{P3}$
7	Coracoid Tip to Glenoid Center Distance	Distance between the most lateral point of the coracoid process to the center of the glenoid cavity along the medial-lateral direction or Z axis.	$Z_{P12} - Z_{P3}$
8	Posterior Inferior (PI) Acromion to Glenoid Center Distance	Distance between the most posterior point of the acromion process to the center of the glenoid cavity along the anterior-posterior (AP) direction or X axis.	$X_{P10} - X_{P3}$
9	Superior Anterior (SA) Acromion to Glenoid Center Distance	Distance between the most anterior point of the acromial tip to the center of the glenoid cavity along the anterior-posterior (AP) direction or X axis. Positive indicates acromial tip is anterior to glenoid center.	$X_{P9} - X_{P3}$

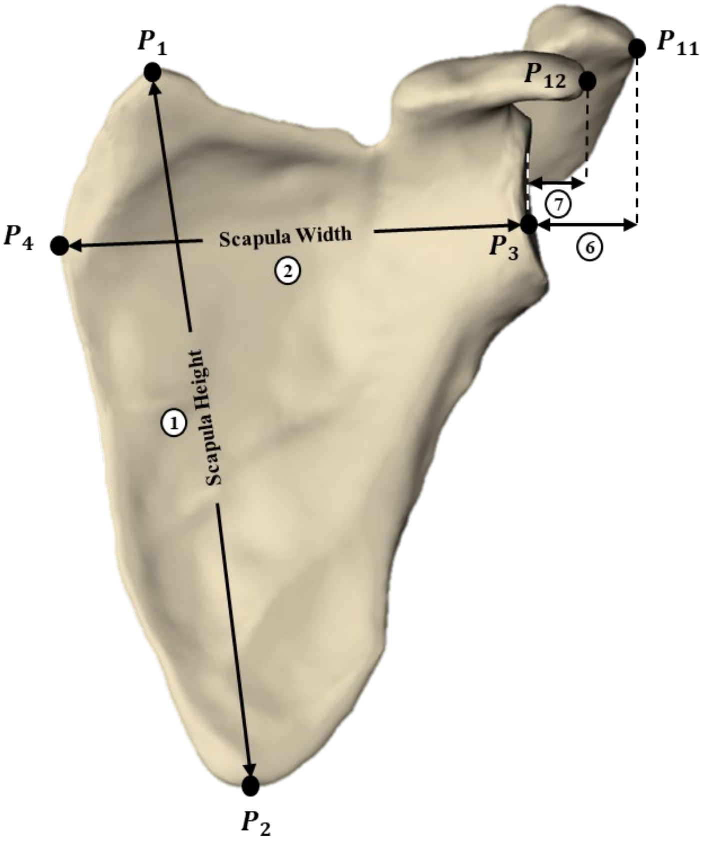


Figure 4. Illustration of anterior view of a left shoulder showing measurement parameter 1, 2, 6 and 7.

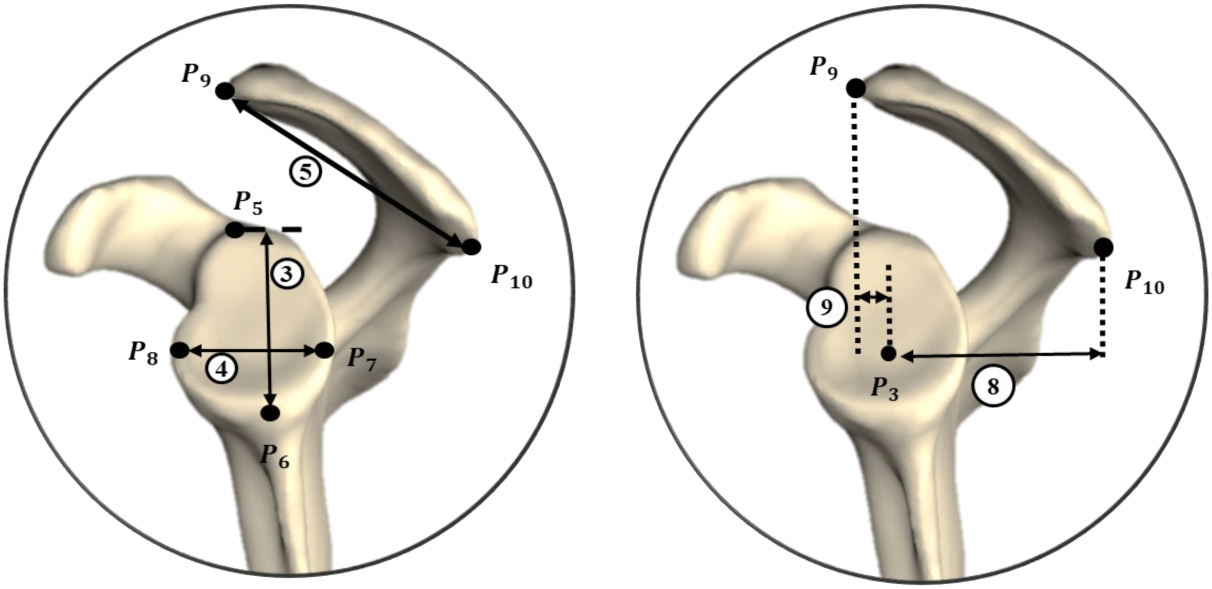


Figure 5. Illustration of lateral view of a left shoulder for measurement parameter 3, 4 and 5 in the left circle and measurement parameter 8 and 9 in the right circle.

Table 4. List of angle-related parameters for the anatomic measurements.

Parameter Number	Parameter Name	Parameter Description	Measurement Formula
10	Fulcrum Axis	Angle between the line connecting the most anterior point of the coracoid to the most posterior-lateral point of acromion and the medial lateral direction or Z axis.	$\cos^{-1}\frac{(\bar{P}_{13}-\bar{P}_{14}).\bar{Z}}{\lvert\bar{P}_{13}-\bar{P}_{14}\rvert}$
11	Inclination Angle	Angle between the superior-inferior (SI) glenoid axis and the medial-lateral direction or Z axis.	$\cos^{-1}\frac{(\bar{P}_{5}-\bar{P}_{6}).\bar{Z}}{\lvert\bar{P}_{5}-\bar{P}_{6}\rvert}$
12	Version Angle	Angle between the anterior-posterior (AP) glenoid axis and the medial-lateral direction or Z axis minus 90 degrees. 0 degrees indicates neutral version.	$\cos^{-1}\frac{(\bar{P}_{7}-\bar{P}_{8}).\bar{Z}}{\lvert\bar{P}_{7}-\bar{P}_{8}\rvert}$
13	Acromial Tilt Angle	Projected angle on the sagittal plane (X-Y plane) between the acromion process and the line connecting the most posterior point of the acromion to the most anterior point of the coracoid process.	$\cos^{-1}\frac{(\bar{P}_{9}-\bar{P}_{10}).(\bar{P}_{14}-\bar{P}_{10})}{\lvert\bar{P}_{9}-\bar{P}_{10}\rvert\lvert\bar{P}_{14}-\bar{P}_{10}\rvert}$ $Z_{\bar{P}_9,\bar{P}_{10},\bar{P}_{14}}=0$
14	Critical Shoulder Angle (CSA)	Angle between the SI glenoid axis and the connecting line the most lateral point of the acromion to the most inferior point of the glenoid cavity.	$\cos^{-1}\frac{(\bar{P}_{5}-\bar{P}_{6}).(\bar{P}_{11}-\bar{P}_{6})}{\lvert\bar{P}_{5}-\bar{P}_{6}\rvert\lvert\bar{P}_{11}-\bar{P}_{6}\rvert}$
15	SI Glenoid-Acromion Angle	Projected angle on the sagittal plane (X-Y plane) between the acromion process and superior-inferior (SI) glenoid axis.	$\cos^{-1}\frac{(\bar{P}_{5}-\bar{P}_{6}).(\bar{P}_{9}-\bar{P}_{10})}{\lvert\bar{P}_{5}-\bar{P}_{6}\rvert\lvert\bar{P}_{9}-\bar{P}_{10}\rvert}$ $Z_{\bar{P}_5,\bar{P}_6,\bar{P}_9,\bar{P}_{10}}=0$
16	SI Glenoid-Acromion-Coracoid Angle	Projected angle on the sagittal plane (X-Y plane) between the superior-inferior (SI) glenoid axis and the line connecting the most posterior point of acromion to the most anterior point of the coracoid.	$\cos^{-1}\frac{(\bar{P}_{6}-\bar{P}_{5}).(\bar{P}_{14}-\bar{P}_{10})}{\lvert\bar{P}_{6}-\bar{P}_{5}\rvert\lvert\bar{P}_{14}-\bar{P}_{10}\rvert}$ $Z_{\bar{P}_5,\bar{P}_6,\bar{P}_{10},\bar{P}_{14}}=0$

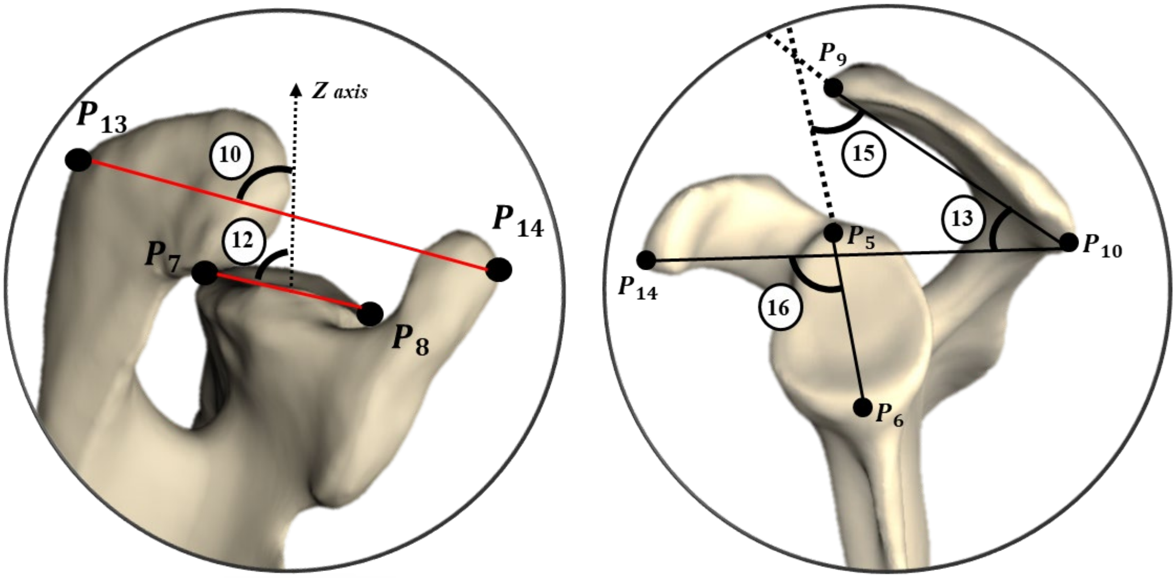


Figure 6. Illustration of superior view of a left shoulder for measurement parameter 10 and 12 in the left circle and lateral view for measurement parameter 13, 15 and 16 in the right circle.

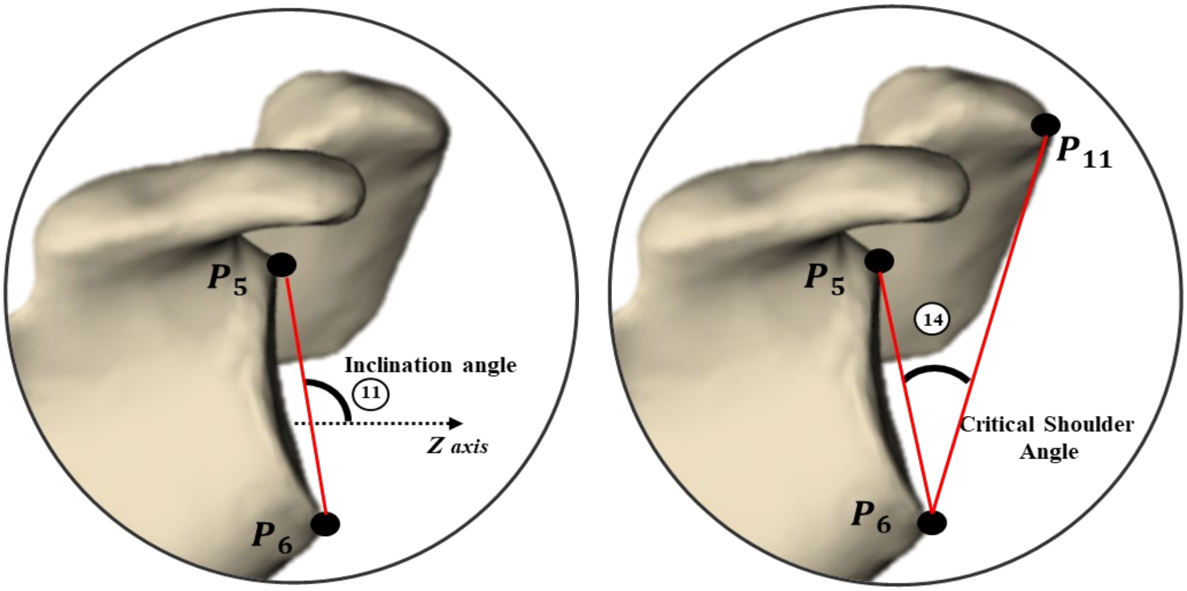


Figure 7. Illustration of anterior view of a left shoulder for Inclination angle (measurement parameter 11) in the left circle and Critical Shoulder Angle (measurement parameter 14) in the right circle.

Furthermore, to evaluate the characteristic effect of each mode of variation of the SSM and SSIM on shape, and to help assess the degree to which each mode of variation effects one or more bone characteristics simultaneously, the weight of each of the modes of variation used to generate each of the 10,000 instances generated by each model were correlated to each of the 16 measurements taken for that instance (i.e. for the SSM, each of the 7 mode weights were correlated to each of the 16 measurements for all 10,000 generated instances).

Supplemental Results

Characteristic Variation Captured by Principal Components of Statistical Models

The characteristics variations captured by Modes 1-to-7 of the SSM and SIM are shown in Figure 8-12, respectively, and by Modes 1-3 for the SSIM in Figure 10.

Regarding the SSM, Mode 1 represents homogenous size changes. Regarding the SSM, Mode 2 represents changes in scapular breadth as indicated by the largest deformations occurring at the medial border. The largest deformations seen in Mode 3 are located in the inferior glenoid region corresponding to changes in glenoid erosion, which are expected due to the patients in this cohort all having Walch-type B2 and B3 glenoid erosions. Modes 4 & 5 show changes in coracoid and acromion morphology with Mode 4 shifting the coracoid in the superior-inferior direction and the anterior acromial tip shifting antero-posterior; while with Mode 5, the entire acromion shifts posterior due to changes in the angle of projections of the scapular spine as indicated by the high deformation shown along its entire length. Mode 6 exhibits changes primarily in the extent to which the acromion projects laterally; while Mode 7 captures changes in the supero-inferior inclination of the glenoid, which is a secondary characteristic of the type of erosion in this patient cohort.

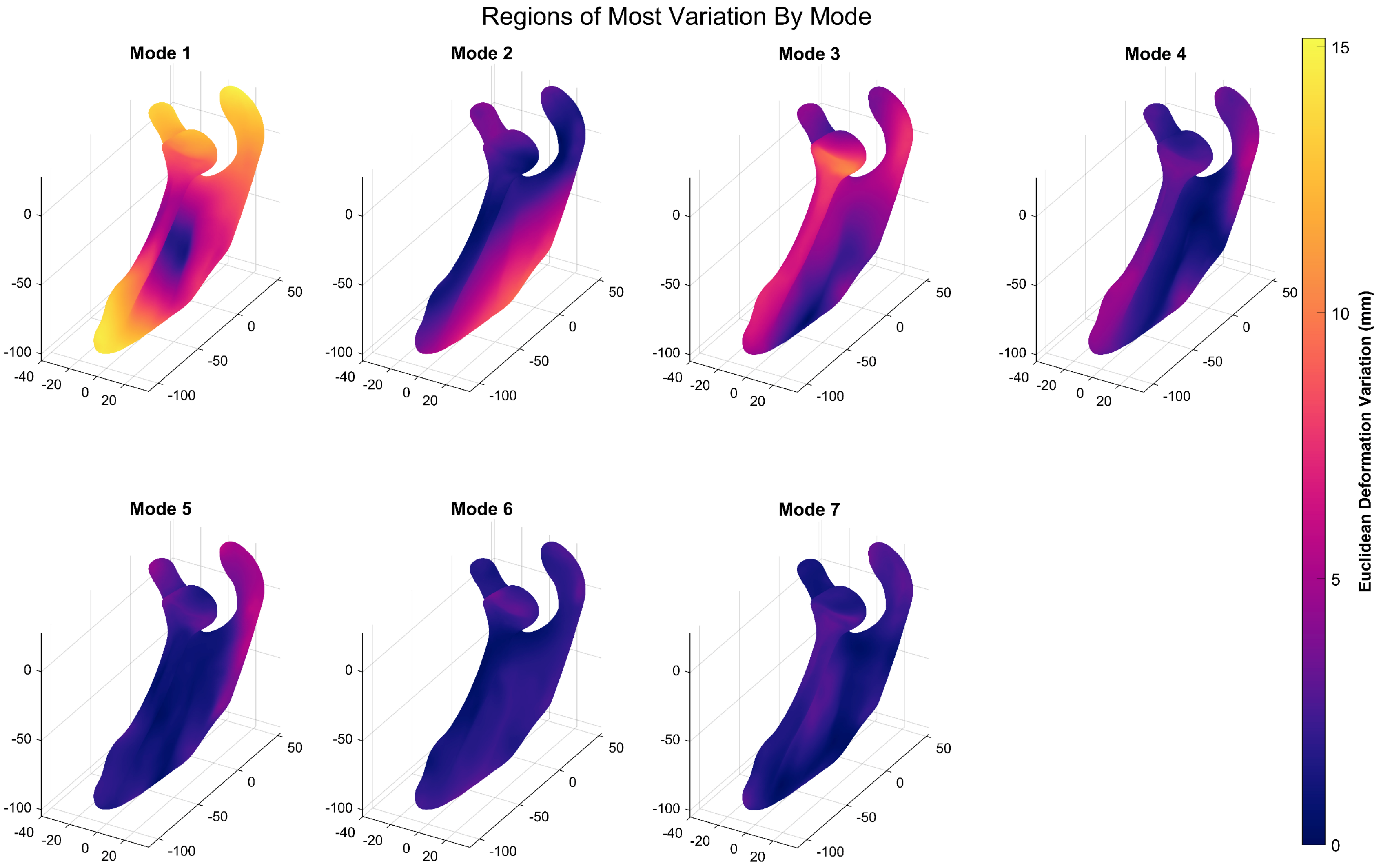


Figure 8 - SSM Characteristic Shape Changes. The scapula in each pane represents the average shape with changes in shape shown by the colormap variation. The colormaps represent the magnitude of deformation when the weighting for the given mode of variation is set to 1.96 SDs.

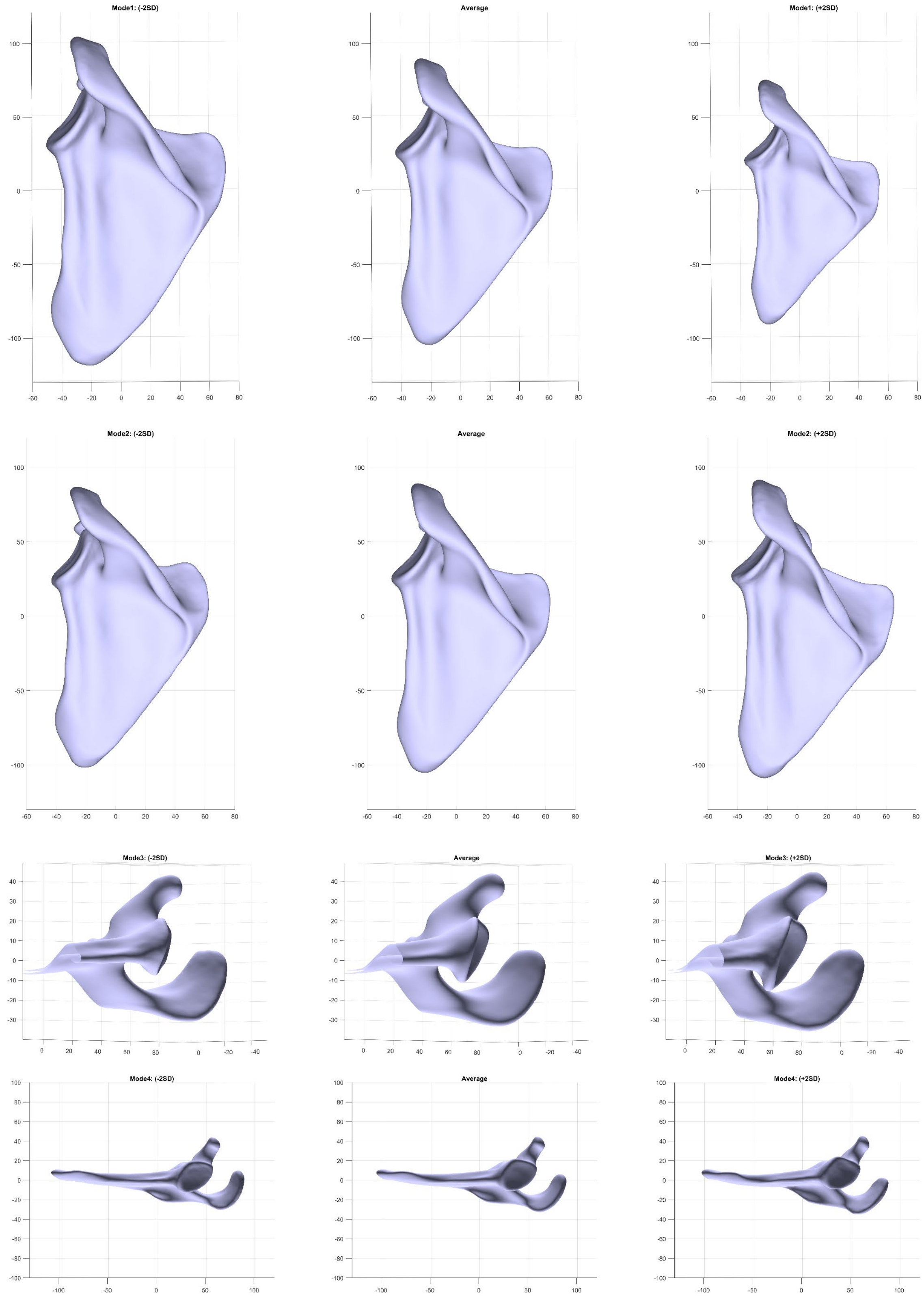


Figure 9 - SSM Mode 1 to 4 Variations. Instances of the SSM model are shown with -2SD, 0SD (i.e. mean), +2SD weights for modes 1 to 4.

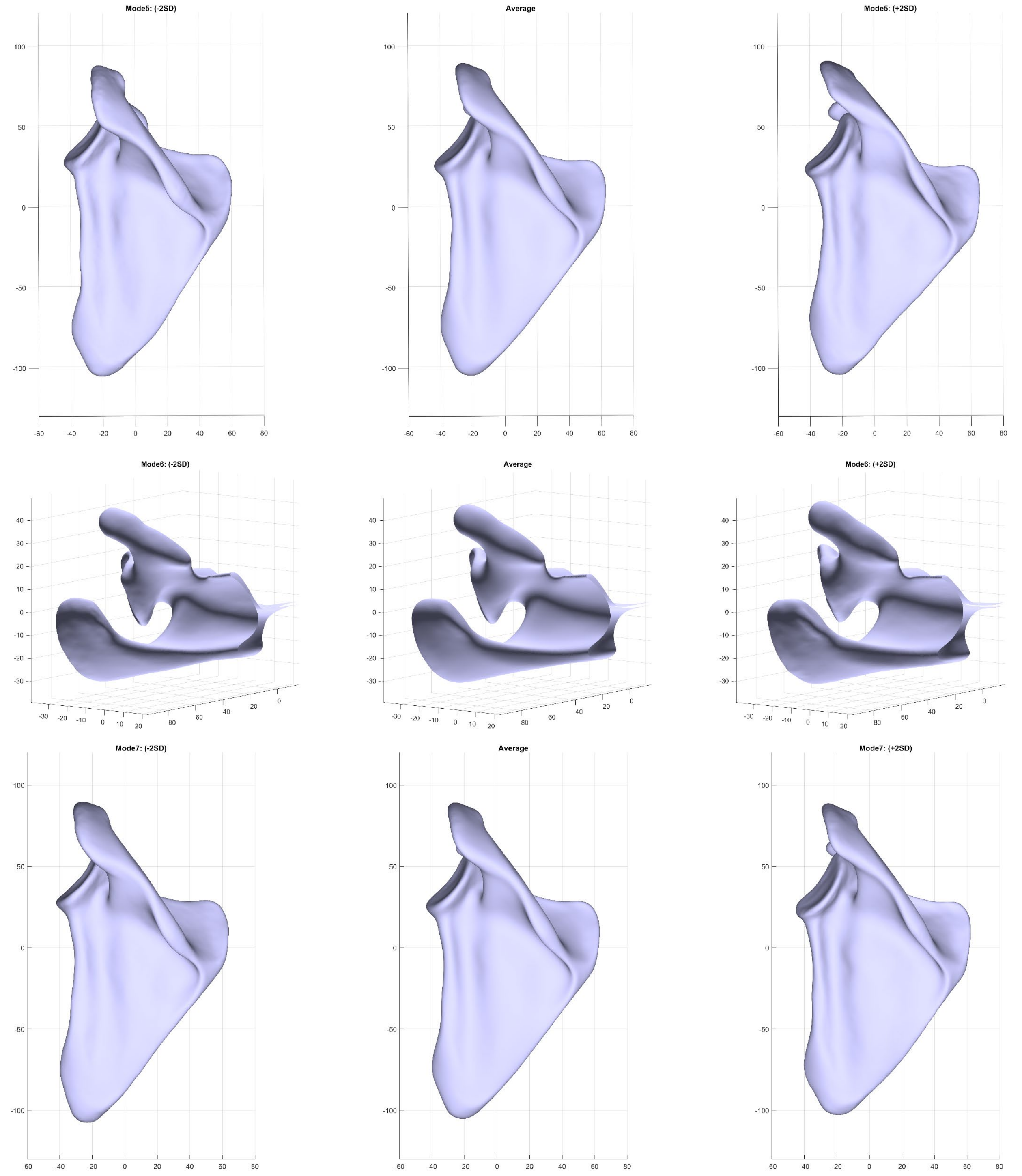


Figure 10 - SSM Mode 5 to 7 Variations. Instances of the SSM model are shown with -2SD, 0SD (i.e. mean), +2SD weights for modes 5 to 7.

Regarding the SIM, Mode 1 captures homogeneous intensity changes across the scapula as shown in Figure 9. Mode 1 is displayed separate from Modes 2 – 7 as the highly regional changes in those modes are better displayed using a different display method. Regarding the SIM, Mode 2 captures broad changes in the cortical shell and glenoid vault bone quality that have opposing changes in the interior cancellous bone of the lateral and medial border, coracoid, and scapular spine (N.B. 'opposing changes' indicates, for example, a decrease in intensity in one region while another region increases and vice versa). Mode 3 represents changes throughout the glenoid vault and interior of the acromion with opposing changes of lesser magnitude in the cortical shell of the scapular blade. Mode 4 captures large, highly localized changes in the infero-posterior glenoid surface and vault with lesser changes to the acromion and coracoid as well as opposing changes of smaller magnitude in the inferior scapular pillar (i.e. medial border). Mode 5 captures moderate changes in the posterior glenoid vault cortical shell and root of the scapular spine where it meets the glenoid vault as well as on the superior aspect of the spine, which are opposed by changed in the acromion bone quality. Mode 6 captures significant changes in the anterior glenoid vault along its entire supero-inferior length and the interior of the scapular pillar with opposing changes to the root of the spine at the posterior glenoid vault. Mode 7 captures localized changes in the acromion and coracoid. It can also be seen that across the 6 modes, changes generally become more diffuse and of smaller magnitude.

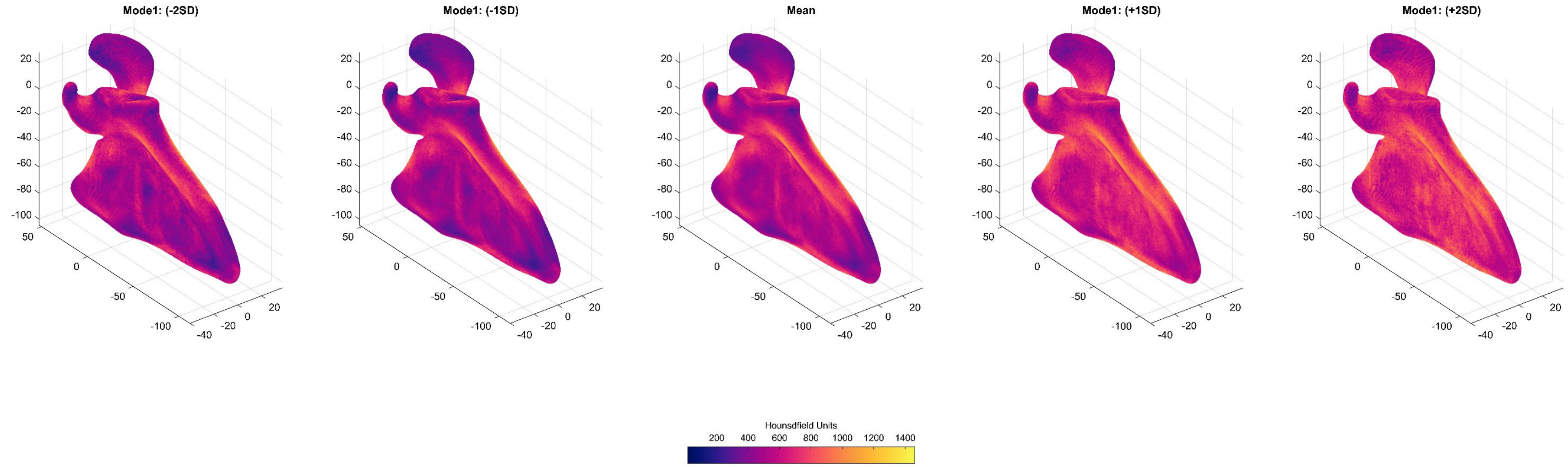


Figure 11 - SIM Mode 1 Characteristic Bone Property Changes. The colormaps represent the HU intensity for various weighting levels for Mode 1. It can be seen that this results in largely homogeneous changes in intensity as the weighting changes.

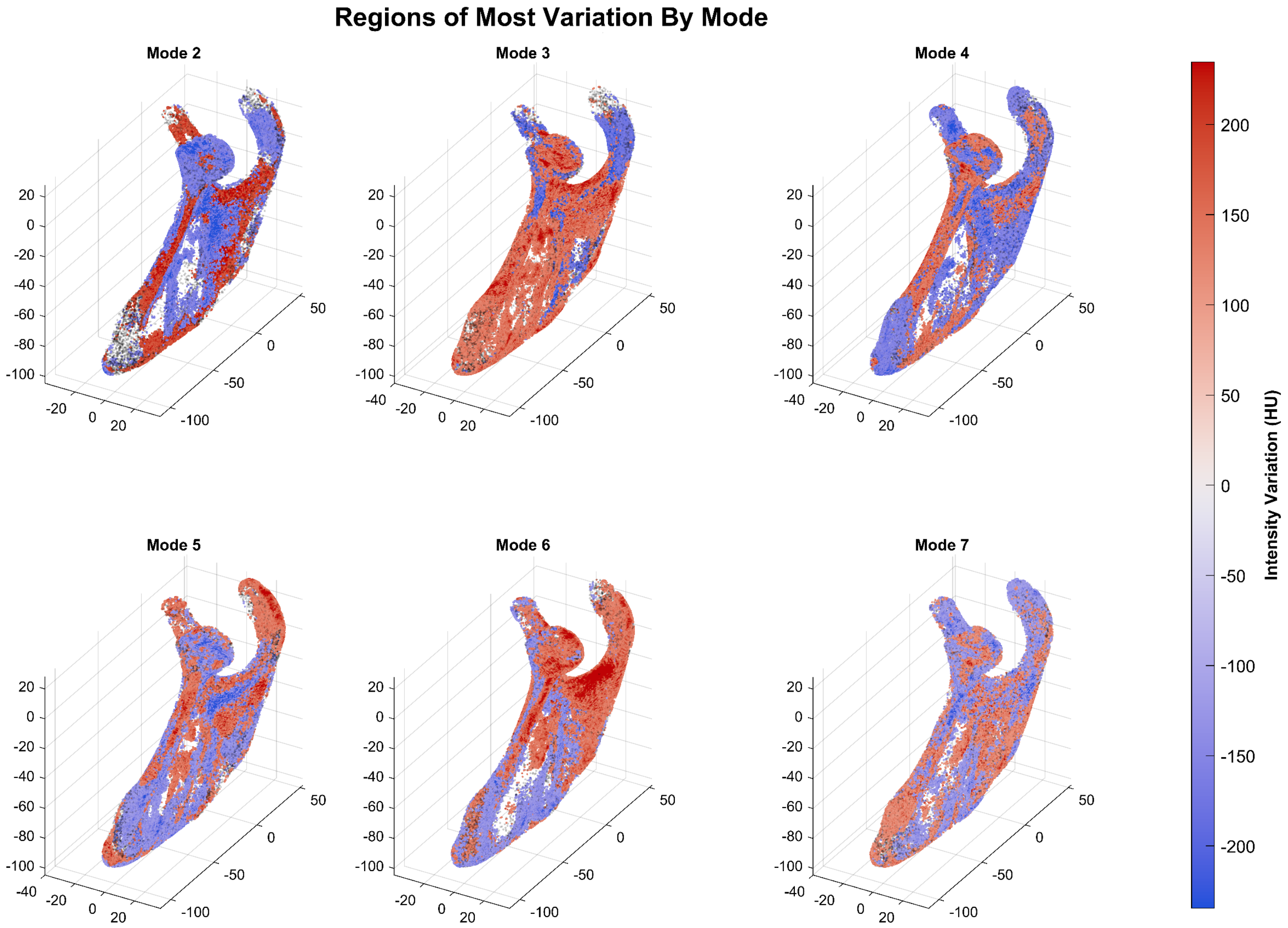


Figure 12 - SIM Characteristic Bone Property Changes. The colormaps represent the change in intensity when the weighting for the given mode of variation is set to 1.96 SDs. Given that change happens throughout the bone volume, to enable improved visualization, only points with increases or decreases in intensity in the top 20% of absolute change for each mode (i.e. each figure pane) are displayed. Furthermore, to improve readability of the colorbar, any point whose increase or decrease is in the top 10% across all mode (i.e. top 10% of all changes shown in entire figure) is shown as having maximal red or blue saturation, respectively (i.e. if the change is above a given HU threshold the color saturation remains constant). Note that the sign of the weightings of each mode of variation is arbitrary and thus a + or - SD does not inherently infer increases and decreases in intensity, respectively.

Regarding the SSIM, Mode 1 captures overall size change as well as medial-lateral shifting of the superior glenoid. Mode 1 also captures intensity changes that are largely homogeneous across the entire scapula. Mode 2 captures changes in medial-lateral scapular width and minor medial-lateral shifting of the superior glenoid. Mode 2 also captures changes in intensity across the scapular body and acromion with opposing changes in the lateral root of the spine. Mode 3 captures posterior glenoid erosion, and coracoid and acromion superior-inferior shifting. Mode 3 also captures intensity changes in the inferior scapular pillar and the scapular spine with opposing changes in the anterior side of the glenoid vault.

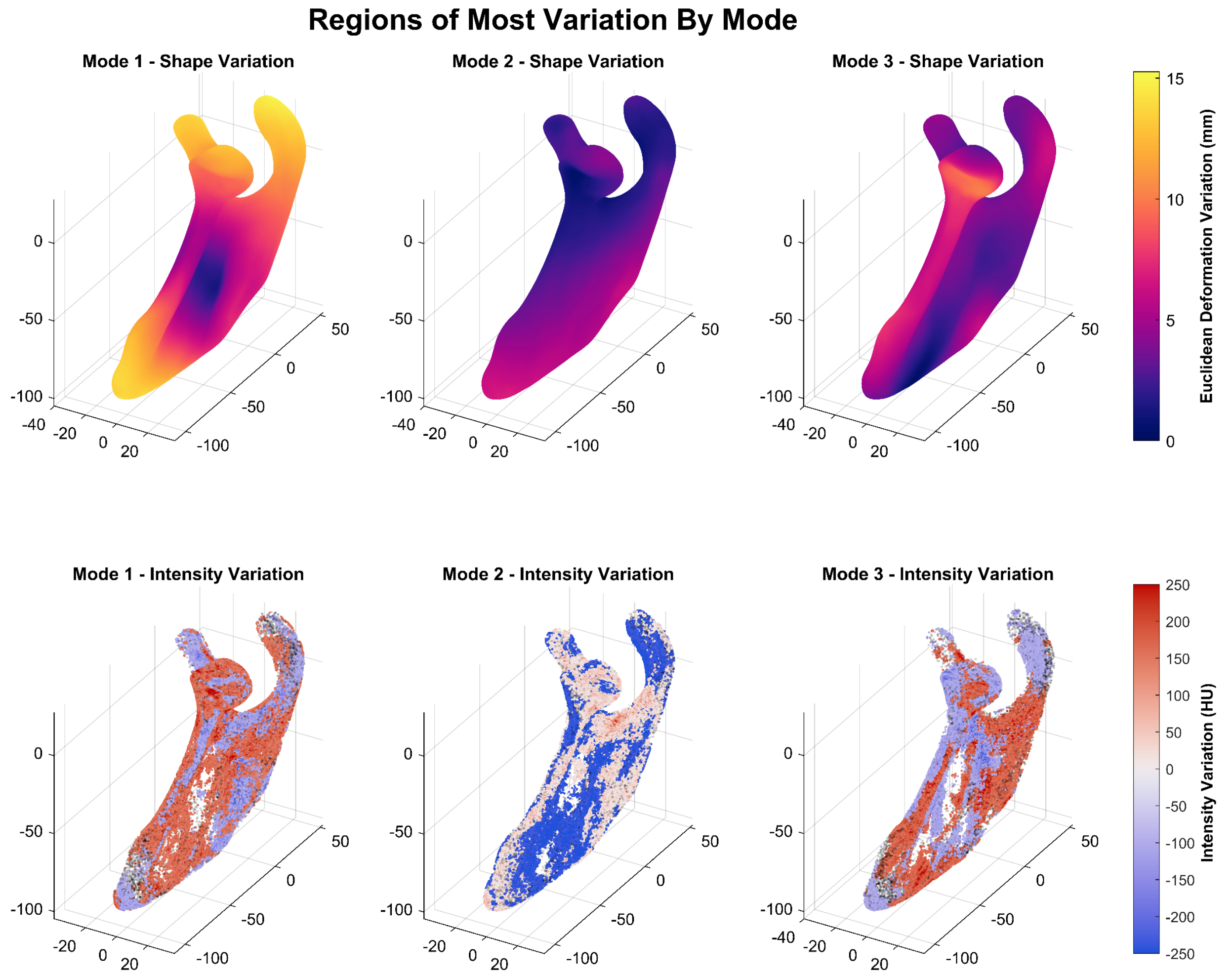


Figure 13 – SSIM characteristic shape and intensity variations. These shape and intensity subplots were created using the same process as for the SSM and SIM in the previous two figures.

Table 5 provides a list of the Shape and Intensity mode of variation combinations that demonstrate the strongest correlations and provides the amount of variance explained by each mode.

Table 5. 10 Strongest Correlations between Shape and Intensity Modes of Variation

Rank	**Shape Mode (%Variance Explained)**	**Intensity Mode (%Variance Explained)**	**Correlation (r)**
1	30 (0.09)	15 (2.1)	0.563
2	22 (0.20)	1 (14.9)	0.490
3	9 (1.14)	22 (1.73)	-0.488
4	6 (2.36)	25 (1.61)	-0.456
5	20 (0.23)	14 (2.2)	0.432
6	14 (0.51)	20 (1.88)	-0.427
7	14 (0.51)	14 (2.2)	0.420
8	7 (1.61)	26 (1.57)	0.410
9	4 (4.56)	12 (2.44)	0.398
10	12 (0.66)	25 (1.61)	0.397