Growing InSight

Developing a 3D growth model for the healthy radius and ulna using statistical shape modeling

Lizet de Vries November 2024

Growing InSight:

Developing a 3D growth model for the healthy radius and ulna using statistical shape modeling

Lizet de Vries

Student number: 4870352

November 21, 2024

Thesis in partial fulfillment of the requirements for the joint degree of Master of Science in

Technical Medicine

Leiden University – Delft University of Technology – Erasmus University Rotterdam

Master thesis project (TM30004 – 35 ECTS)

Dept. of Orthopedics & Sports Medicine Erasmus MC April 2024 – November 2024

Supervisors

Dr. J.W. (Joost) Colaris Medical supervision Prof. Dr. H.E.J. (DirkJan) Veeger Technical supervision D.F.R. (Derek) van Loon, MSc Daily supervision E.M. (Eline) van Es, MSc Additional supervision

Thesis committee members

Prof. Dr. H.E.J. (DirkJan) Veeger TU Delft Chair Dr. J.W. (Joost) Colaris Erasmus MC Dr. T.M. (Tom) Piscaer Erasmus MC E.M. (Eline) van Es, MSc Erasmus MC

An electronic version of this thesis is available at [http://repository.tudelft.nl/.](http://repository.tudelft.nl/)

Abstract

Objective

Insight into the three-dimensional (3D) healthy growth of the radius and ulna remains limited, as orthopedics rely primarily on simple two-dimensional radiographic measurements. In this study, we used statistical shape modeling (SSM) to explore 3D forearm shape variations during healthy growth and predict individual radial and ulnar anatomy over time.

Methods

We trained SSMs using CT segmentations of 117 healthy radii and 116 ulnae. Principal Component Analysis (PCA), with and without Procrustes scaling, extracted general shape variation, while we used Partial Least Squares Regression (PLSR) to create an average growth model and prediction models. We developed several methods using age, sex, bone length and/or initial anatomy to predict bone anatomy over time. The models were validated by comparing the predicted bone models to the original segmentations of follow-up samples of 22 radii and 19 ulnae. The primary outcome measure was Root Mean Squared Error (RMSE), complemented by bone length error and mean and maximum distance errors.

Results

Size differences dominated the unscaled PCA- and PLSR-based SSMs, while the scaled PCA captured more subtle shape variations. With PLSR, we extracted the longitudinal growth trajectory for boys and girls aged 4 to 18 in 3D from the cross-sectional data. Using initial anatomy and age as input, the validation samples were predicted with RMSEs of less than 1 mm. The lowest bone length errors were reached using age, sex, and bone length as predictors.

Conclusion

We successfully utilized SSM to investigate and predict healthy growth of the radius and ulna in 3D. The current models are clinically viable as reference anatomy in pre-operative planning, and the methodology shows promise to support diagnosing and treating forearm deformities.

Contents

1. Introduction

The radius and ulna are essential components of the upper extremity's complex anatomy, enabling movement and rotation of the arm, wrist, and hand – functions critical for daily activities (1). Skeletal growth of the forearm is a powerful process during childhood that not only prepares the bones and joints for a long lifespan but also has the ability to gradually correct deviations after fractures (2-4). Despite their importance, surprisingly little is known about the healthy growth patterns of the radius and ulna. In orthopedics, long bone growth is typically assessed using simple 2D radiographic measurements, while our interest lies in capturing these bones' three-dimensional (3D) shapes (5-11).

Statistical shape modeling (SSM) has proven useful for understanding 3D anatomical variations in bones during growth (12, 13). An SSM is a mathematical model that captures common shape variability within

a dataset (14). As Ambellan et al. stated, *"SSMs are the foundation for the analysis of anatomical cohort data, where characteristic shapes are correlated to demographic data."* (12). This makes SSMs particularly useful for exploring growth-related shape variation in a cohort of children's forearm bones across different ages. With a cross-sectional dataset consisting of samples from different individuals at various time points, it is possible to capture a longitudinal effect (15). The SSM then describes trends across the temporal range over multiple subjects, including their anatomical differences, representing the population's average *Figure 1 Visualization of a cross-sectional dataset with age* shape variation over time (Figure 1).

as temporal factor.

SSMs utilizing regression techniques can predict an individual's growth trajectory by applying the average pattern of shape change to their unique anatomical structure and/or demographic characteristics (13). This concept is the foundation of a personalized growth model in 3D, useful for understanding growth as well as for support in diagnosing and treating injuries.

Forearm fractures are among the most common injuries in children, accounting for nearly 40% of all pediatric fractures (16). Most of these fractures heal without complications. However, a potential complication is malunion of the radius, ulna, or both, which can result in curved forearms with pain, instability of the distal radial-ulnar joint, and impaired forearm rotation (17-19). In young patients, nature sometimes offers a helping hand by correcting deviations through growth (2-4). However, whether the deformity and its symptoms will resolve sufficiently through growth remodeling remains unpredictable. If SSM proves effective in mapping healthy growth of the radius and ulna, this approach could serve as a stepping stone for modeling growth and remodeling in pathological conditions. The ability to model and predict individual healthy anatomy in 3D could already be valuable in clinical practice. A patient's anatomy can be compared to the healthy model to identify abnormalities, and individual predictions can be useful in planning surgical interventions, where the contralateral healthy side of the patient usually serves as a reference. Access to an 'atlas' of healthy radii and ulnae based on individual characteristics like age and sex could eliminate the need for contralateral imaging, facilitating the planning of bilateral deformity corrections and reducing radiation exposure.

In this study, we used statistical shape modeling to (1) explore the shape variation in 3D during healthy growth and (2) develop and compare models to predict individual anatomy of the healthy radius and ulna over time. Based on results from comparable studies involving different bones, our primary technical objective was to create a prediction model capable of approximating the original anatomy with a mean distance error of less than 2 mm.

2. Methods

2.1. Data

Upper extremity computed tomography (CT) scans of patients under 19 years old were obtained retrospectively from the research database of the Erasmus MC Orthopedics department. We used scans of healthy radii and ulnae, initially collected for comparison with the affected side in various forearm pathologies between 2017 and June 2024. Patients gave permission for the use of their clinical data for research purposes. Incomplete scans and bones with abnormalities were excluded. The training dataset used to develop the SSMs comprised 117 radii and 116 ulnae. The validation set included follow-up scans of 22 radii and 19 ulnae, acquired approximately one year after the initial CT scan. Sex and age at the moment of the initial scan and, if applicable, at the follow-up scan were collected from the electronic health records. The included unaffected forearm bones originated from 60 boys and 57 girls aged 3.4 to 18.8 (Figure 2).

Figure 2 Age and sex distribution in the study population.

2.2. Preprocessing

Development of an SSM involves several key steps (as outlined in Figure 3 and Appendix A), with a dataset of 3D models as the foundation. Two authors (EE, LV) segmented the radius and ulna from the CT scans in Mimics Research Suite (Materialise, Leuven, BE). We needed one-body 3D models to develop our SSMs. Therefore, samples without an ossified growth plate, thus virtually identified as consisting of more than one body, were adjusted in 3-Matic (Materialise, Leuven, BE) and MeshLab (version 2023.12) to obtain one part (Appendix B.1). We checked mesh quality using the Trimesh python library (nr. of bodies, volume, water tightness, winding consistency, nr. of vertices and faces, and maximum edge length) and remeshed all models with a maximum edge length of 1.0 mm as input parameter. The data consisted of left and right bones. All right-arm models were reflected to obtain only left variants.

Figure 3 Flowchart of the development and validation process of the statistical shape models in this study consisting of (1) preparing the datasets, (2) creating the statistical shape models and (3) validation of the prediction models. Abbreviations: PCA = Principal Component Analysis, PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares.

2.3. Point Distribution Model

Statistical shape modeling requires a Point Distribution Model (PDM) with dense point correspondence among all samples (Appendix A). We created PDMs for the radius and ulna using the open-source software ShapeWorks (University of Utah, Salt Lake City, USA). All models were rigidly aligned to a reference mesh using Iterative Closest Point (ICP) registration. Optimization parameters were tuned iteratively by trial and error. We started with a subset of the data and a small number of particles, increasing both while fine-tuning the parameters. The final number of particles was chosen based on a tradeoff between the level of anatomical detail and computational time. We performed optimization twice. Once without scaling, thereby retaining size differences, and the other approach included Procrustes registration with scaling. During this type of registration, the sum of the squared distances to the mean of the dataset is minimized by translation, rotation, and scaling to remove the overall scale from the model.

We used fixed-domain modeling to place the correspondences on the follow-up validation samples, using the preexisting shape models for the radius and ulna.

We assessed the geometric accuracy of the initial PDMs and those representing the validation samples. This involved calculating the Euclidian distance between each particle of the PDM and its nearest vertex on the corresponding original model and averaging this distance across all particles and samples.

2.4. Statistical shape modeling

We applied multiple SSM techniques to achieve the different objectives in this study (Figure 3, part 2). Principal Component Analysis (PCA), a commonly used method in similar research, was employed to describe and visualize overall shape variation in the dataset. We used Partial Least Squares Regression (PLSR) to extract longitudinal trends from the cross-sectional data for prediction purposes. Two regression techniques, standard PLSR and Canonical Powered PLS (CPPLS), along with various predictor variables were evaluated to identify the best approach for our goals.

2.4.1. Principal Component Analysis

Using ShapeWorks, we performed PCA on both the scaled and unscaled PDMs of the radius and ulna to identify the primary modes of shape variation. The variations represented in the first five modes of the PCA, called principal components (PC), were reported descriptively. We used linear regression models to assess which PCs were significantly associated with age differences. A p-value of 0.05 or lower was considered statistically significant.

The common measures for an SSM's general quality – compactness, generalization and specificity – were calculated for both the unscaled and the scaled SSMs. Compactness is the proportion of shape variance explained per principal component. Generalization uses a Leave-One-Out analysis to measure the SSM's ability to accurately represent samples not encountered during model development. Specificity evaluates the ability to generate valid novel samples, measured as the average distance of synthetic samples to their nearest member in the original SSM.

2.4.2. Partial Least Squares Regression

The unscaled PDMs were used to train PLSR-based SSMs in R (version 4.4.1) with the pls package (20). The primary five shape modes and the models' compactness were extracted.

Average growth model

CPPLS was used to incorporate both discrete and continuous variables (Algorithm 1 in Appendix B.2). The coordinates of the samples' particles from the PDM were set as the response variables (Y). Age, sex, and their interaction factor were used as predictors (X). We used this model to construct the average radial and ulnar growth trajectory for boys and girls aged 4 to 18. The bone lengths of the constructed 3D bone models were measured using the vertices of their oriented bounding boxes in Trimesh and reported.

Prediction models

The CPPLS model described was extended with the predictor variable bone length. The validation samples' geometries were predicted using the CPPLS prediction model's seventh component, the interaction factor between age, sex, and bone length (Algorithm 2 in Appendix B.2). We also evaluated the CPPLS-based model using only age and sex as predictors.

To enable the use of an initial bone model as input for predicting geometry at a follow-up moment, we used PLSR with the particle coordinates as the predictor (X) and age as the response variable (Y) (Algorithm 3 in Appendix B.2). The scores were extracted and plotted against age for every shape mode. A second-order polynomial was fitted on the scores to extract the follow-up scores (Figure 4). The PLSR model was trained using ten, three and one component(s) to find the optimal trade-off between model simplicity and accuracy.

One-year follow-up samples of a subset of the training dataset were used to validate the prediction models and compare their performances (Figure 2 and Figure 3, part 3). We used a Leave-One-Out method: the prediction models were trained using the training dataset minus the sample of which the follow-up would be predicted. The predicted bones were compared to the original bone segmentations from the follow-up CTs. The primary outcome measure was the Root Mean Squared Error (RMSE), a standard measure for the difference between a predicted value or model and the actual one. We used mean and maximum distance error per vertex and bone length error as secondary outcomes. Also, the measured bone length errors of the radius and ulna were compared per validation sample to evaluate joint congruence. The per-vertex distance errors were visualized by color mapping them onto the mean radius and ulna.

Figure 4 Example of the score vs. age plots for the first two components of the Partial Least Squares Regression-based statistical shape model. The blue line is a second-order polynomial fit representing the average change in score over age. The green point is the predicted score based on geometry and age at the initial CT. The red point is calculated by applying the average trend to the predicted initial score until the age at the moment of the follow-up CT is reached. Abbreviations: FU = Follow Up.

3. Results

3.1. Point Distribution Model

The final PDMs of the radius and ulna consisted of 2048 particles and were in dense correspondence (Figure 5). The tuned parameters used for optimization can be found in Appendix C.1. We found geometric accuracies of 0.35 ± 0.01 mm for the initial and the fixed-domain modeled PDMs of the radius and ulna.

Figure 5 Point distribution models (PDMs) of the radius (left) and ulna (right) on different samples from the training set. Two particles are highlighted to visualize the dense correspondence of the PDMs.

3.2. Principal Component Analysis

3.2.1. Unscaled PCA

Applying PCA on the unscaled models of the radius and ulna resulted in very compact SSMs with 98% and 96% of the total variance captured in the first component alone. The models' generalization errors were 1 mm and 1.5 mm with one PC and decreased to <0.4 mm and <0.65 mm when using 10 PCs for the radius and ulna, respectively. Specificity reached a maximum of 1.1 mm and 1.4 mm when using up to 50 PCs (Figure 6).

Overall size differences were dominant in PC1 and significantly correlated with age (p < 0.001) (Figure 7 and Table 1). In every other PC of the radial SSM, less than 1% of the variance was captured, representing minor shape variations in the dataset. PC2 showed a slight misalignment along the long axis (Figure 7). The third PC corresponded mainly to variation in bone width, and the fourth and fifth to bowing and internal angulation in the anteroposterior and lateromedial direction, respectively (Appendix C.2). None of these components showed a significant association with age (Table 1). The second PC of the ulnar SSM was also influenced by the misalignment. The main anatomical variations represented by the third PC were ulnar bowing, bone width by the fourth, and depth of the trochlear notch by the fifth PC. There was a significant association between the fifth component and age ($p =$ 0.048).

Figure 6 Compactness, generalization and specificity plots for the first 50 principal components of the unscaled (left) and scaled (right) statistical shape models of the radius and ulna.

Figure 7 Shape variation (% of total variance) captured by the first and second PCs in the unscaled statistical shape model of the radius (left) and ulna (right), visualized as ±*2*s*, where* ^s *is the standard deviation. The asterisks indicate a significant association with age, with the arrows showing the direction of age change from - to +2*s*. Abbreviations: PC = principal component.*

p -values PC – age	Radius unscaled	Ulna unscaled	Radius scaled	Ulna scaled
PC1	$2.59e-33*$	$5.99e-36*$	$1.21e-5$ *	$2.16e-14*$
PC ₂	0.0915	0.261	0.201	0.0533
PC ₃	0.485	0.892	0.000909 *	$0.0239*$
PC4	0.902	0.292	$4.08e-14*$	$1.03e-5$ *
PC ₅	0.698	$0.0476*$	0.395	$0.00638*$

Table 1 P-values of the linear regression models between the first five principal components and age. The asterisks indicate a statistically significant association with age. Abbreviations: PC = principal component.

3.2.2. Scaled PCA

The scaled SSMs were less compact with 34% and 32% explained by the first PC for the radius and ulna, respectively. When more than 10 PCs were used, the proportion of the dataset's total variance represented in the scaled SSMs increased to over 90% (Figure 6). Errors in generalization and specificity were lower than in the unscaled model. As size differences were absent in these SSMs, more subtle shape variations covered a larger percentage of the total variance in the first couple of PCs (Figure 8). Bone width was the principal variation in the first PC of the scaled radial SSM. PC2 showed differences in internal angulation of the distal radius and PC3 in the coronal bowing. In the fourth PC, the maturity of the proximal and distal joints varied; in the fifth PC, mainly the distal radius joint surface area and prominence of the tuberosity. In the scaled ulnar SSM, the first PC also represented variation in bone width but, in contrast to the radius, getting thinner with age. Sagittal bowing was seen in the second component. In the third PC, the shaft width varied slightly. A variation in trochlear notch depth was found in the first four PCs. In the fifth, the coronoid process and radial notch varied in prominence, all representations of proximal ulnar maturity. Components one, three and four of the radial SSM and one, three, four and five of the ulnar SSM showed significant association with age (Table 1 and Appendix C.2).

Figure 8 Shape variation (% of the total variance) captured by the first five PCsin the scaled statistical shape model of the radius (top) and ulna (bottom), visualized as [±]*2*s*, where* ^s *is the standard deviation. The asterisks indicate a significant association with age, with the arrows showing the direction of age change from - to +2*s*. Abbreviations: PC = principal component.*

3.3. Partial Least Squares Regression

After performing PCA, we discovered a misalignment along the long axis in the PDMs (Figure 7, PC2). Before performing Partial Least Squares Regression (PLSR), we improved the alignment of all data (training and validation, particles and meshes) using landmark alignment followed by ICP. For every sample, the particles furthest away from each other were selected as landmarks.

The first shape mode of the PLSR models of the radius and ulna explained 95.1% and 93.1% of the total variance. With >2 shape modes for the radius and >5 for the ulna, the SSM covered 99% of all age-

related shape variation (Figure 9). The first shape mode mainly represented overall size differences (Figure 10). The second shape mode captured variation in radial bowing, internal angulation of the distal end, and an increasing circumference of the proximal end. In mode three, the bone width of the radius varied, especially at the distal part, as the joint surface enlarged. The fourth mode showed variance in bone width and internal angulation. The latter was also seen in mode five, but the other way around. In the ulna, variations in maturity of the proximal joint were seen in all shape modes. The third mode showed distal coronal bowing differences with slight shaft length differences. In shape mode five, we also found variation in bowing but in sagittal direction.

Figure 9 Compactness plot for the first ten shape modes of the Partial Least Squared Regression statistical shape model of the radius and ulna.

Figure 10 Shape variation (% of the total variance) captured by the first five shape modes in the PLSR-based statistical shape model of the radius (top) and ulna (bottom), visualized as $±2\sigma$ *, where* σ *is the standard deviation.*

3.3.1. Average growth model

With the average growth model, based on the CPPLS-based SSM, we could visualize the average growth trajectory of boys and girls from 4 to 18 years old. Bone length in this model increased almost linearly over the years. The radius and ulna grew with an average of 9.0 and 9.8 mm per year for boys and 7.1 and 7.7 mm per year for girls. On average, the male forearm bones were longer than the female bones (Appendix C.3).

3.3.2. Prediction models

In general, the PLSR-based prediction model with initial shape input, including ten shape modes, and the CPPLS-based model using age, sex, and bone length as predictors performed best. We found the lowest RMSE using the PLSR-based prediction model (Table 2). The error was less than 1 mm for both the radius and ulna. The CPPLS-based prediction resulted in an RMSE of 0.86 mm for the radius and 1.16 mm for the ulna. This model predicted bone lengths with an accuracy of less than 1 mm. Meanwhile, the predicted validation samples from the PLSR showed a mean absolute bone length error of 2.82 mm and 3.90 mm for the radius and ulna. The worst predictions regarding RMSE are visualized in Figure 11 and compared to their original bone model. Deviations in the PLSR-based predictions involve mainly bone length differences. In the CPPLS-based predictions, minor angulation differences are the primary cause of the higher RMSEs.

Figure 11 Predictions of the radius and ulna for the samples with the highest Root Mean Squared Errors using the PLSR-based prediction model (left, pink) and the CPPLS-based prediction model (right, blue), compared to the original bone models (white). Abbreviations: PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares.

When we used only the first three shape modes of the PLSR-based model, the RMSE and mean bone length error for the radius increased slightly to 0.90 mm and 2.83 mm, respectively (Appendix C.4). For the ulna, a larger increase to an RMSE of 1.20 mm and a bone length error of 4.21 mm was seen. When using only the primary shape mode, RMSE was above 1 and 1.3 mm, and bone length error was above 3 and 4 mm for the radius and ulna, respectively. The CPPLS model using only age and sex as predictors performed the worst with RMSEs of more than 1.5 mm and mean bone length errors of over 12 mm.

Table 2 Root Mean Squared Errors and bone length errors for the best performing PLSR- and CPPLS-based SSM.

Abbreviations: PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares, RMSE = Root Mean Squared Error, SD = standard deviation.

Comparing the bone length error of the radius and ulna per validation sample revealed a mean absolute difference of 1.49 mm for the PLSR model, ranging from -2.68 to 3.77 mm. The complete CPPLS model had an absolute mean difference of 0.76 mm, ranging from -2.42 mm to 1.14 mm (Appendix C.4).

We visualized the average distance error among the validation set for every vertex on the mean mesh as a color map for the two best-performing prediction models (Figure 12). The mean per-vertex distance for the radius ranged from 0.25 to 1.54 mm for the PLSR-based model and from 0.27 to 1.75 mm for the CPPLS-based model. In the ulnar dataset, the mean distances across the vertices were between 0.2 and 2.3 mm and between 0.38 and 2.0 mm for the PLSR and CPPLS models, respectively. The maximum per-vertex distance error between the predicted and original radius among the validation samples was 5.0 mm when using PLSR and 5.5 mm with the CPPLS-based model (Figure 13). We found slightly higher maximum errors of respectively 7.0 mm and 8.0 mm for the ulna. The meshes predicted using PLSR with age and the initial geometry as input had the largest deviation from the original bone models at the proximal and distal joint surfaces due to bone length differences. CPPLS-based prediction showed the largest distance errors in the epiphyseal areas of the radius and ulna. Around the head of the proximal radius, the sigmoid notch, the coronoid process and the anterior part of the olecranon, the predicted meshes differed more than average from the original ones.

Figure 12 Mean distance errors per vertex visualized on the average radius (top) and ulna (bottom) for the PLSR-based SSM (left) and CPPLS-based SSM (right). For each model, a histogram with color legend is plotted on the left with the per-vertex distances in mm. Abbreviations: PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares.

Figure 13 Maximum distance errors per vertex visualized on the average radius (top) and ulna (bottom) for the PLSR-based SSM (left) and CPPLS-based SSM (right). For each model, a histogram with color legend is plotted on the left with the per-vertex distances in mm. PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares.

4. Discussion

4.1. Interpretation of results

We performed statistical shape modeling using PCA and PLSR-based techniques to develop 3D growth and prediction models for the radius and ulna. We successfully met the technical objective of predicting individual forearm bone geometry with a mean error of less than 2 mm.

Our dataset was parametrized into a PDM with a geometric accuracy of 0.35 mm. We think this was small enough not to significantly influence the SSMs' quality or performance. Unscaled PCA resulted in a very compact model where size differences dominated the variance in shape. As anticipated, the scaled SSMs' compactness was lower, and more nuanced shape variations were revealed (13). Variations in bone length, width, and maturity of the epiphyseal regions of both forearm bones and radio-ulnar bowing of the radius were significantly associated with age differences. No studies have assessed the 3D shape variation of the pediatric radius or ulna to date. Therefore, we compared our outcomes to the literature on SSMs of other bones during growth. Despite using a smaller dataset with slightly lower geometric parametrization accuracy, our PCA-based SSMs produced high-quality models with generalization and specificity errors no greater than 1.5 mm, comparable to the literature (21-23).

Using the canonical powered version of PLSR (CPPLS), we created an average growth model, offering a valuable 3D perspective on the growth of the radius and ulna in boys and girls. This offers an advantage over traditional 2D measurements, which tend to oversimplify complex anatomical shape variations. Some reconstruction errors occurred, particularly in the models of younger children, due to registration issues during optimization and the small population of the younger age groups.

We predicted the geometry of individuals' forearm bones using various prediction models and compared their accuracy. The PLSR-based method, which used initial shape input, performed best regarding RMSE. As expected, bone length was most accurately predicted by the CPPLS-based method, using age, sex, and bone length as predictors. However, this approach relies on prior knowledge of the length of the bone to be predicted, which was derived from the available 3D models in this study. While Carman et al. used pelvic, tibial, and fibular bone lengths measured in the clinic, its feasibility and accuracy for the radius and ulna would need to be researched (24).

All prediction models in this study, even those based on only one shape mode, showed lower RMSE than reported in comparable studies(24, 25). The mean per-vertex distances were lower than those reported by another study on a femur SSM (0.2 – 2.3 mm in this study compared to 0.5 – 3.0 mm in McKinsey et al.) (25). The largest per-vertex distances across all samples in the validation set were almost similar, with a maximum of 5.5 mm for the radius and 8.0 mm for the ulna, compared to 6.9 mm for McKinsey's model.

An advantage of the CPPLS method is its ability to easily incorporate additional predictor variables, such as ethnic origin, menarche, body length, weight, or arm span, which may enhance predictive accuracy (26). However, our CPPLS method could not incorporate initial personal anatomy as input and relied on linear relationships. Growth is a nonlinear process; some children grow faster earlier in childhood, while others grow until a higher age than average. Consequently, the CPPLS model based on only age and sex performed the worst. The PLSR-based prediction model could predict anatomical shape details very well, as demonstrated in the per-vertex distance colormaps (Figures 12 and 13). The largest deviations were observed at the joint surfaces due to a misprediction of bone length. A limitation of the reversed regression method in the PLSR model is that it allows for only one demographic variable – in this case,

age – to be included in the regression. On the other hand, an advantage is that a nonlinear polynomial can be fitted to account for differences in shape variation depending on age, such as the decreasing growth rate in older children. Nevertheless, the PLSR method proved ineffective for accurately predicting bone length at a later moment, as the initial shape does not provide information about the individual's growth stage. A more effective approach would be to combine the strengths of both methods, integrating the nonlinearity and initial shape input from the PLSR model with the ability of CPPLS to include multiple predictor variables. Applying a spline to the predictors and/or PDM could introduce more flexibility within the SSM, allowing the model to better capture shape changes that evolve nonlinearly with age. Furthermore, utilizing skeletal age instead of chronological age may improve prediction accuracy by accounting for the individual's growth stage and residual growth potential over time. Unfortunately, insufficient imaging was available for the samples in our dataset to determine skeletal age reliably.

When considering different models, it is essential to consider the purpose of the prediction. Also the accuracy measures should represent the feasibility of a prediction to serve its purpose. If the focus is on understanding the growth trajectory in 3D, the detailed anatomy around the epiphyses becomes more significant. However, for surgical correction of the radius or ulna, factors like bone length, angulation and ulnar variance are more important, as the anatomical shape of the joints can barely be changed. For the latter application, a both-bone SSM would be helpful to better understand and predict the positional relationship between the radius and ulna.

4.2. Strengths and limitations

This was the first study to explore growth in healthy forearm bones in 3D using statistical shape modeling. Predicting the individual anatomy of developing bones has been done before, but not for the radius and ulna. This study's strength also lies in its use of various SSM techniques and the comparison of different approaches to prediction. This allowed for a comprehensive evaluation of which techniques offer the best accuracy for specific clinical applications.

A limitation of our study was the lack of detail incorporated in the SSMs. We constructed the PDMs using 2048 particles. This number was chosen based on a tradeoff between the level of detail and computation time. With this number of particles, we could perform large-scale shape variation observations and predictions. However, more data points are needed to gain insight into more refined processes during growth, such as the closure of the growth plates.

Furthermore, the initial registration of the segmented bones in ShapeWorks could have been improved. After performing PCA, we discovered, by examining the second component, that the ICP algorithm struggled to align the bone models properly along the long axis. Therefore, we aligned the data more accurately before PLSR using automatic landmark registration followed by ICP. In future research, ensuring proper alignment of all 3D models during parametrization is recommended. Guiding the registration with (automatically identified) anatomical landmarks is a possible solution (27), with the benefit of the PDM also being anatomically accurate. How well the parametrization or PDM represents the bone shapes can be measured by geometric and anatomical accuracy. We did not annotate anatomical landmarks and were therefore limited to using geometric accuracy. Although the particles did not always represent the exact same anatomical location across samples, we ensured the PDMs closely represented the bone surfaceswith a geometric accuracy of 0.35 mm. This is sufficient to explore and predict shape variation.

Lastly, there was considerable variation in the anatomy of the bone models, particularly in the proximal ulna. We chose to synthetically attach the detached epiphyseal parts to the ulna, which may have resulted in a loss of insight into the ossification process. The shape variance in the proximal ulna also reduced the feasibility of bone length error as a validation measure. A slight bone length difference between the radius and ulna in predictions will clinically be irrelevant if both bones are over- or underestimated by the same amount, as the proximal and distal radio-ulnar joints would remain congruent. However, even the best-performing model showed discrepancies of up to 2.5 mm between the radius and ulna length predictions. This error might also have been caused by the variance in the proximal ulna, which would mean the misalignment of the joints may have been smaller. Using a bone length measurement from proximal to distal radio-ulnar joint or incorporating both bones in one SSM could improve assessing prediction accuracy.

4.3. Future prospects

We believe several clinical applications of the 3D growth and prediction models are already within reach, requiring only minor improvements to the dataset and the applied SSM techniques. While developing the SSM incurs computational costs during optimization, the remaining modeling steps are rapid. Our model can predict a patient's radial or ulnar anatomy without the need for imaging in just seconds. This prediction can serve as a reference for preoperative planning of corrective osteotomies, which is currently unfeasible in cases of bilateral deformities. Additionally, the average growth model can provide a 'normal' growth pattern and assist in identifying deviations from the expected trend. This is useful, for example, in cases of epiphyseal damage, helping to determine residual growth and thus the optimal timing for intervention.

Further refinement of the models is necessary to enhance their clinical viability. This includes developing methods and metrics for comparing a patient's bone geometry against the reference growth model, allowing clinicians to detect deviations from the typical growth pattern. The same applies to the prediction models. While technical metrics, such as RMSE, are essential for evaluating and comparing the quality of SSMs, there is a need for measures that address specific clinical applicability. Exploring outcome measures tailored to clinical practice and easily interpretable by clinicians – like radiologic standards – would be a valuable direction for future research. Although our models demonstrated a low mean error, further refinements in predicting specific anatomical regions could help achieve clinically accepted tolerances, such as the 2 mm threshold for intra-articular fracture treatment.

A potential future application of SSM is reconstructing 3D bone models from 2D imaging (28). Specific landmarks or measurements from the 2D image could be used as predictors in the CPPLS regression model. This application can potentially reduce radiation exposure while retaining the advantages of 3D modalities. This method would also enable the collection of large, longitudinal datasets, as it would rely on widely available 2D imaging, virtually converted into 3D models.

This study focused on healthy growth, but SSM techniques could also be applied to datasets of deformed bones, incorporating bone remodeling alongside average growth. Growing bones have the capacity to remodel and gradually correct deformities (2-4, 29). However, whether remodeling will be sufficient to restore anatomical geometry is unpredictable. While we want to avoid invasive surgery as much as possible, postponing an intervention only to find it necessary later is also undesirable. The ability to predict whether growth will eliminate the need for corrective surgery would be valuable. Achieving this will likely require more samples and longitudinal data, for which the 2D to 3D conversion could be beneficial.

4.4. Conclusion

In this study, we successfully developed a 3D growth model of the healthy radius and ulna for children aged 4 to 18 years using statistical shape modeling. Based on a cross-sectional dataset, PCA and PLSR revealed longitudinal shape variations during growth. The regression-based shape models allowed for predicting individual radial and ulnar geometry over time. While further validation and the development of clinical outcome measures are necessary for implementing SSMs in pediatric orthopedic practice, these techniques show promise in supporting the diagnosis and treatment of forearm deformities.

5. References

1. Bair MM, Zafar Gondal A. Anatomy, Shoulder and Upper Limb, Forearm Radius. StatPearls. Treasure Island (FL): StatPearls Publishing

Copyright © 2024, StatPearls Publishing LLC.; 2024.

2. Fuller D, McCullough C. Malunited fractures of the forearm in children. The Journal of Bone & Joint Surgery British Volume. 1982;64-B(3):364-7.

3. Johari AN, Sinha M. Remodeling of forearm fractures in children. J Pediatr Orthop B. 1999;8(2):84-7.

4. Price CT, Scott DS, Kurzner ME, Flynn JC. Malunited forearm fractures in children. Journal of pediatric orthopedics. 1990;10(6):705-12.

5. de Pablos J, Franzreb M, Barrios C. Longitudinal growth pattern of the radius after forearm fractures conservatively treated in children. J Pediatr Orthop. 1994;14(4):492-5.

6. Knapik DM, Kolaczko JG, Drummond I, Liu RW. Longitudinal analysis of pediatric distal radius alignment parameters in a cohort of serial radiographs. Clin Anat. 2021;34(3):365-70.

7. Pritchett JW. Growth and Development of the Distal Radius and Ulna. Journal of Pediatric Orthopaedics. 1996;16(5):575-7.

8. Wu C, Wang D, Mo Y, Zhang Z, Ning B. Characteristics of the length of the radius and ulna in children. Frontiers in Pediatrics. 2022;10.

9. Gindhart PS. Growth standards for the tibia and radius in children aged one month through eighteen years. American Journal of Physical Anthropology. 1973;39(1):41-8.

10. Sallam AA, Briffa N, Mahmoud SS, Imam MA. Normal Wrist Development in Children and Adolescents: A Geometrical Observational Analysis Based on Plain Radiographs. Journal of Pediatric Orthopaedics. 2020;40(9):e860-e72.

11. Stahl EJ, Karpman R. Normal growth and growth predictions in the upper extremity. J Hand Surg Am. 1986;11(4):593-6.

12. Ambellan F, Lamecker H, von Tycowicz C, Zachow S. Statistical Shape Models: Understanding and Mastering Variation in Anatomy. Adv Exp Med Biol. 2019;1156:67-84.

13. de Vries LE, van Loon DFR, van Es EM, Veeger HEJ, Colaris JW. Exploring shape changes in healthy bone growth through 3D spatiotemporal statistical shape models: a scoping review [Manuscript submitted for publication]. 2024.

14. Cootes TF, Taylor CJ, Cooper DH, Graham J, editors. Training Models of Shape from Sets of Examples1992; London: Springer London.

15. Middleton KM, Duren DL, McNulty KP, Oh H, Valiathan M, Sherwood RJ. Cross-sectional data accurately model longitudinal growth in the craniofacial skeleton. Scientific Reports. 2023;13(1):19294.

16. Mäyränpää MK, Mäkitie O, Kallio PE. Decreasing incidence and changing pattern of childhood fractures: A population-based study. Journal of Bone and Mineral Research. 2010;25(12):2752-9.

17. Högström H, Nllsson BE, Wlllner S. Correction with Growth Following Diaphyseal Forearm Fracture. Acta Orthopaedica Scandinavica. 1976;47(3):299-303.

18. Kruppa C, Bunge P, Schildhauer TA, Dudda M. Low complication rate of elastic stable intramedullary nailing (ESIN) of pediatric forearm fractures: A retrospective study of 202 cases. Medicine. 2017;96(16):e6669.

19. Schmittenbecher PP. State-of-the-art treatment of forearm shaft fractures. Injury. 2005;36(1):S25-S34.

20. Liland KHM, Bjørn-Helge; Wehrens, Ron. pls: Partial Least Squares and Principal Component Regression. R package version 2.8-3; 2023.

21. Klop C, Becking AG, Klop C, Koolstra JH, Lobé NHJ, Maal TJJ, et al. A three-dimensional statistical shape model of the growing mandible. Scientific Reports. 2021;11(1):18843.

22. O' Sullivan E, van de Lande LS, El Ghoul K, Koudstaal MJ, Schievano S, Khonsari RH, et al. Growth patterns and shape development of the paediatric mandible - A 3D statistical model. Bone Rep. 2022;16:101528.

23. O' Sullivan E, van de Lande LS, Oosting AC, Papaioannou A, Jeelani NO, Koudstaal MJ, et al. The 3D skull 0-4 years: A validated, generative, statistical shape model. Bone Rep. 2021;15:101154.

24. Carman L, Besier TF, Choisne J. Morphological variation in paediatric lower limb bones. Scientific Reports. 2022;12(1):3251.

25. McKinsey K, Thompson A, Dsouza R, Bertocci G. Development of a model to predict 3D femur morphology in infants and young children. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. 2023;11(3):922-32.

26. Luk KD, Saw LB, Grozman S, Cheung KM, Samartzis D. Assessment of skeletal maturity in scoliosis patients to determine clinical management: a new classification scheme using distal radius and ulna radiographs. Spine J. 2014;14(2):315-25.

27. van Loon DFR, van Es EM, Eygendaal D, Veeger DHEJ, Colaris JW. Automatic identification of radius and ulna bone landmarks on 3D virtual models. Computers in Biology and Medicine. 2024;179:108891.

28. Patil A, Kulkarni K, Xie S, Bull AMJ, Jones GG. The accuracy of statistical shape models in predicting bone shape: A systematic review. The International Journal of Medical Robotics and Computer Assisted Surgery. 2023;19(3):e2503.

29. Gascó J, de Pablos J. Bone remodeling in malunited fractures in children. Is it reliable? J Pediatr Orthop B. 1997;6(2):126-32.

30. Cates JE, Elhabian SY, Whitaker RT, editors. ShapeWorks: Particle-Based Shape Correspondence and Visualization Software2017.

31. Jolliffe IT. Principal Component Analysis: Springer; 2002.

32. Höskuldsson A. PLS regression methods. Journal of Chemometrics. 1988;2.

33. Tommaso Mansi. Human-Computer Interaction [cs.HC]. École Nationale Supérieure des Mines de Paris ENEt-. Image-based physiological and statistical models of the heart: application to tetral- ogy of Fallot.: École Nationale Supérieure des Mines de Paris; 2010.

Appendices

A. Background of statistical shape modeling

A statistical shape model (SSM) is a geometric model that represents a dataset of 3D objects by capturing the average shape and its variations. The development process of an SSM consists roughly of parametrization, defining all objects by the same system to make them comparable, and comparing all shapes to identify the geometric variability in the dataset.

Parametrization

Before parametrization, all input meshes should be aligned. The goal of parametrization is to define the surfaces of all meshes with particles in a point cloud, called a Point Distribution Model (PDM). Every particle in the PDM should be placed at a similar anatomical location on every mesh but not necessarily at the same location in the coordinate system. The process of finding the coordinates for the PDM's particles on every mesh is called optimization.

Shapeworks (University of Utah, Salt Lake City, USA) is an open-source software package that includes a correspondence optimization algorithm to achieve optimal correspondence between the surface points of every sample in a dataset. The positions of the points are updated iteratively by gradient descent on an energy function balancing two terms, representing correspondence and uniform distribution, respectively (30). The algorithm aims to find the optimal trade-off between model compactness and accurate shape representation. It can be tuned using input parameters to balance the desired level of detail and correspondence.

Analysis of shape variation

Principal Component Analysis (PCA) is often used to construct SSMs. It is an unsupervised machine learning algorithm to mathematically describe variation in a dataset (31). The resulting model describes the probability of specific shape occurrences in terms of an average shape and a set of dominant modes, called Principal Components, that account for the main patterns of shape variation. Due to the unsupervised nature of PCA, shape variations captured by the model can result from various factors. It is a valuable method for gaining insight into all types of variation within a dataset. One can create an SSM with and without scaling. Unscaled models include size, which is very relevant during growth. Scaled models, not dominated by the large proportion of variation due to size differences, expose more subtle shape variations in the dataset (13).

Regression techniques can be helpful if one is specifically interested in shape variation related to a particular variable-dependent process, such as growth. Partial Least Squares Regression (PLSR) is a supervised method that allows the extraction of modes of shape variation related only to changes in a specific variable. It integrates PCA and linear regression to predict response variables (Y) from a set of predictors (X). The model's basis vectors, scores and loadings, maximize the variance within both variable sets and their covariance (32, 33).

Prediction

A trained SSM represents the population's average shape and variations but can also predict 'new' realistic anatomical geometries. The response can be predicted by giving demographic predictor values for a specific individual as input to the PLSR model.

B. Methods – additional material

1. Growth plate attachment

We synthetically attached the proximal ulnar growth plate to the olecranon using the following process: (1) cutting the bone at metaphyseal height, (2) performing dilation followed by erosion on the epiphyseal part using uniform mesh resampling with a precision of 1 and an offset of ± 7 mm, (3) merging the processed part and the original epiphyseal part, (4) merging the new one-bodied part with the original diaphyseal bone part.

Figure A.1 Visualization of the preprocessing steps to attach the growth plate to the ulna

2. Algorithms

Algorithm 1: Average growth model

```
Input training: Characteristics (age, sex) & Particle coordinates of training set 
Input prediction: Characteristics validation samples
Output: Predicted particle coordinates validation samples
n = number of training samples
p = number of particles
```

```
particles = concatenated particle coordinates per sample (n * 3p)age = age per sample (n)
sex = sex per sample as dummy variable (n)
model cppls \leftarrow cppls(particles \sim age*sex)
Predicted particles <- predict(model cppls, new data)
```
Algorithm 2: CPPLS prediction model

Input CPPLS training: Characteristics (age, sex, bone length) & Particle coordinates of training set Input CPPLS prediction: Characteristics (age, sex, bone length) validation samples Output: Predicted follow-up particle coordinates validation samples

```
data_LOO = training set with initial data of validation sample left out
dataval init = validation sample initial data
data val FU = validation sample follow-up data
```

```
## Train a CPPLS model and predict the FU geometry based on characteristics
model cppls <- cppls(particles ~ age*sex*bone length, data=data LOO)
Predicted particles cppls <- predict(model cppls, data val FU)[1, 1:6144, 7]
```
Algorithm 3: PLSR prediction model

Input PLSR training: Characteristics (age) & Particle coordinates of training set Input PLSR prediction: Initial particle coordinates & Characteristics (age) validation samples Output: Predicted follow-up particle coordinates validation samples

```
data LOO = training set with initial data of validation sample left out
dataval init = validation sample initial data
data val FU = validation sample follow-up data
## Train a PLSR model and find the scores for the validation sample's initial 
particles
ncomp = 10 #or 3 or 1
model plsr \langle - plsr(age \sim particles, ncomp = ncomp, data = data LOO)
scores <- model plsr.scores[]
Predicted scores <- predict(model plsr, newdata=data val init, type="scores")
## Fit polynomial regression on score x age data per component. Calculate the 
initial offset of score from the polynomial for the validation sample, 
predict the FU on the polynomial and add the offset to find the FU_score.
Fit \leq lm(scores \sim poly(age, 2), data = df(data LOO, scores))
Offset \leq- Predicted_scores - predict(fit, newdata = data val init$age)
FU scores <- predict(fit, newdata = data val FU$age) + Offset
```
Predicted particles plsr <- FU scores %*% model plsr.loadings + mean_coordinates

C. Results – additional material

1. Parameter tuning

Table B.1 Final parameters used for optimization in ShapeWorks

An explanation of the parameters can be found on the [ShapeWorks](https://sciinstitute.github.io/ShapeWorks/latest/workflow/parameters.html) website.

Figure C.1 Shape variation (% of the total variance) captured by the first five principal components in the unscaled statistical shape model of the radius (top) and ulna (bottom), visualized as $\pm 2\sigma$ *, where* σ *is the standard deviation. The asterisks indicate a statistically significant association with age, with the arrows showing the direction of age change from - to +2*s*. Abbreviations: PC = principal component.*

2. Principal Component Analysis

 -36

 $-1500 -$

Figure C.2 Score-age plots for the principal components significantly associated with age differences of the unscaled and scaled statistical shape models of the radius and ulna. The red line represents the fitted linear model, and its p-value is given in the top right corner for every principal component. Abbreviations: PC = principal component.

3. Average growth model

Table C.2 Bone lengths in the average growth model of the radius and ulna for males and females separately.

Figure C.3 Bone lengths over age for the average growth models of the male's and female's radius (top) and ulna (bottom).

4. Prediction model

Table C.3 Root Mean Squared Errors and bone length errors for all PLSR- and CPPLS-based SSMs. Abbreviations: PLSR = Partial Least Squares Regression, CPPLS = Canonical Powered Partial Least Squares, RMSE = Root Mean Squared Error, SD = standard deviation, Abs. = absolute value.

Bone length error difference radius vs ulna

