A framework for three-dimensional statistical shape modeling of the proximal femur in Legg-Calvé-Perthes disease

†Luke G. Johnson[1,2], †Joseph D. Mozingo[3], Penny R. Atkins[3,4,5], Seaton Schwab[6], Alan Morris[4], Shireen Y. Elhabian[4,5], David R. Wilson[7], Harry K.W. Kim[8], Andrew E. Anderson[3,4,6]

†Luke G. Johnson and Joseph D. Mozingo contributed equally to the development of this article.

[1]School of Biomedical Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[2]BC Children's Hospital Research Institute, Vancouver, British Columbia, Canada
[3]Department of Orthopaedics, University of Utah, Salt Lake City, Utah, USA
[4]Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah, USA
[5]School of Computing, University of Utah, Salt Lake City, Utah, USA
[6]Department of Biomedical Engineering, University of Utah, Salt Lake City, Utah, USA
[7]Department of Orthopaedics, University of British Columbia, Vancouver, British Columbia, Canada
[8]Texas Scottish Rite Hospital for Children, Dallas, Texas, USA

Correspondence: Andrew E. Anderson, 590 Wakara Way, Rm A100, Salt Lake City, UT 84108, USA. Email: Andrew.Anderson@hsc.utah.edu Phone: +1 801 587 5208

Running Title: Perthes femoral shape model

Acknowledgements: We acknowledge funding from the Canadian Institutes of Health Research (Funding reference 165956), Arthritis Society Canada (TGP22-0000000149), and National Institutes of Health (NIBIB-U24EB029011, NIBIB-R01EB016701, NIAMS-R01AR076120, NHLBI-R01HL135568, NIGMS-P41GM103545 and R24 GM136986).

1 Abstract

Purpose: The pathomorphology of Legg-Calvé-Perthes Disease (LCPD) is a key contributor to poor long-term outcomes such as hip pain, femoroacetabular impingement, and early-onset osteoarthritis. Plain radiographs, commonly used for research and in the clinic, cannot accurately represent the full extent of LCPD deformity. The purpose of this study was to develop and evaluate a methodological framework for three-dimensional (3D) statistical shape modeling (SSM) of the proximal femur in LCPD.

Methods: We developed a framework consisting of three core steps: segmentation, surface mesh preparation, and particle-based correspondence. The framework aims to address challenges in modeling this rare condition, characterized by highly heterogeneous deformities across a wide age range and small sample sizes. We evaluated this framework by producing a SSM from clinical magnetic resonance images of 13 proximal femurs with LCPD deformity from 11 patients between the ages of six and 12 years.

Results: After removing differences in scale and pose, the dominant shape modes described morphological features characteristic of LCPD, including a broad and flat femoral head, high-riding greater trochanter, and reduced neck-shaft angle. The first four shape modes were chosen for evaluation of the model's performance, together describing 87.5% of the overall cohort variance. The SSM was generalizable to unfamiliar examples with an average point-to-point reconstruction error below 1mm. We observed strong Spearman rank correlations (up to 0.79) between some shape modes, 3D measurements of femoral head asphericity, and clinical radiographic metrics.

Conclusion: In this study we present a framework, based on SSM, for the objective description of LCPD deformity in three dimensions. Our methods can accurately describe overall shape variation using a small number of parameters, and are a step towards a widely accepted, objective 3D quantification of LCPD deformity.

Keywords: Legg-Calvé-Perthes Disease, hip joint, pediatrics, statistical shape modeling, morphology

2 Introduction

Legg-Calvé-Perthes Disease (LCPD) is a pediatric hip disorder characterized by avascular necrosis of the femoral head (FH).[1] LCPD often results in a permanent residual deformity of the FH and secondary acetabular dysplasia, which can manifest in features such as an enlarged, widened and sometimes flat FH, short femoral neck, a high-riding greater trochanter, acetabular dysplasia, and hip joint incongruity. Patients with severe residual FH deformity often present with hip pain, limited range of motion, and femoroacetabular impingement during adolescence.[2, 3] Hip surgery is often indicated for these patients.[4, 5] An increased severity of residual FH deformity and acetabular dysplasia at skeletal maturity (as described by the radiographic Stulberg classification[6]) is also associated with greater risk of early-onset osteoarthritis (OA), with a reported rate of moderate-to-severe radiographic OA of more than 40% in LCPD patients as young as 29 years at follow-up,[7] compared to only 2% in the 50-54 year old general population.[8]

Quantifying the degree of LCPD deformity is an important part of patient assessment, treatment planning and follow-up, but there are limitations to current approaches. Radiographic metrics do not accurately depict the full extent of the pathomorphology of the deformity because they are based on planar projections of the three-dimensional (3D) anatomy. These measurements are also sensitive to positioning errors between patients and examinations, including difficulty in obtaining plain radiographs of true orthogonal views due to suboptimal leg positioning that can be caused by decreased hip motion.[9, 10] Computed tomography provides 3D images of the bones, but it requires considerable exposure to ionizing radiation and does not provide good visualization of soft tissues, like articular cartilage overlying the bone. Magnetic resonance imaging (MRI) requires no ionizing radiation and yields 3D images with good soft tissue visualization, making it suitable for ongoing follow-up imaging. However, a major limitation of all 3D imaging is that there is currently no widely accepted, objective way by which to quantify the 3D deformity in LCPD.

3D surface reconstructions from volumetric images have, however, been used to objectively describe anatomy and anatomical variation through statistical shape modeling (SSM).[11–13] A key advantage of SSM is that it provides a holistic representation of shape with independent continuous parameters (i.e., shape modes) without *a priori* assumptions regarding measurements believed to be most clinically relevant. Previous applications of SSM in the hip have included prediction of OA outcomes in older adults,[14, 15] identification of the most useful 2D clinical

measurements in cam-type femoroacetabular impingement syndrome,[16] and describing growth patterns in pediatric hip disorders.[17, 18] It follows from these examples that SSM should provide the foundation for an open and objective standard for describing 3D shape variation in LCPD. For such a foundation to be useful, it must perform well when the available datasets are small and highly heterogeneous, which is typical for studies of LCPD deformity.[19–21]

The accuracy and applicability of SSM relies fundamentally on the choice of landmark correspondences, particularly in highly heterogeneous morphology such as LCPD deformity. Placing landmarks by hand is impractical for dense point sets on 3D surfaces, and many automated methods rely on either brute force (e.g. matching closest nodes on meshes) or geometric comparisons, both of which introduce suboptimal correspondences that can confound statistical analysis.[22] Recently, entropy-particle-based methods for optimization of correspondence, as used in the open-source SSM software *ShapeWorks*, have demonstrated excellent performance.[12, 23, 24] To facilitate improved assessment of the long-term prognosis of this relatively rare condition, the present study extends entropy-particle-based SSM methods to also include an incremental optimization routine.[25] This incremental approach was designed to enable optimization of landmark correspondences when using a small, heterogenous dataset.

The objectives of this study were to develop a framework for constructing expandable SSM of LCPD pathomorphology, evaluate its accuracy and sensitivity to training population size, and illustrate its use to quantify 3D anatomical variation in a small cohort of LCPD patients.

3 Materials and Methods

The shape modeling framework presented herein consists of three key steps: segmentation, surface mesh preparation, and particle-based correspondence. We present these steps below in the context of their application to our evaluation cohort of 11 patients with LCPD, and the analyses used to evaluate the resulting SSM.

3.1 Patients and Imaging

Eleven patients (eight male, three female) with LCPD were included in this Institutional Review Board approved study (Table 1). Two patients had bilateral LCPD. Patient ages ranged from six to 12 years old (mean 9 ± 2 years). The modified Waldenstrom stages of LCPD progression[26] ranged from IIa (early fragmentation stage) to IV (healed stage), and the modal stages were IIIa and IIIb (early and late reossification stages), containing four proximal femurs each.

Table 1: Patient demographics, natural history stage, and anatomical measurements.

Subject	Gender	Age (yr.)	Side	Modified Elizabethtown stage	Asphericity (mm)	NSA (degrees)	ATD (mm)
1	Male	10	L	IV	2.3	126	16.8
2	Male	10	L	IIIb	2.9	134	31
3	Male	6	L	IIIb	3.6	130	19.2
			R	IIIa	3.5	131	17.3
4	Female	9	R	IIIb	4.8	132	9.9
5	Male	9	L	IIIa	4.0	141	32.3
6	Male	10	L	IIIb	5.1	140	19.1
			R	IIIa	5.4	135	12
7	Female	12	R	IIb	5.6	162	43.8
8	Male	12	R	IIIa	5.8	114	0
9	Male	11	L	IIa	2.7	132	14.8
10	Male	9	L	IIb	4.2	152	25.8
11	Female	8	R	IIa	2.7	131	15.4
Total [mean (SD)]:		9 (2)			4.0 (1.2)	135.4 (11.8)	19.8 (11.2)

Abbreviations: NSA, neck-shaft angle; ATD, articular-trochanteric distance.

An MRI scan of the hip joints was obtained using a GE Hdxt 1.5 Tesla scanner (Waukesha, WI, USA).[19] A fat-suppressed 3D spoiled gradient-echo sequence was used to acquire images coronally (repetition time = 8.9 ms, echo time = 2.8 ms, flip angle= 10°, bandwidth = 20.8 kHz, slice thickness = 1.0 mm, matrix = 288 x 288). Each image volume was resampled based on its smallest voxel dimension (0.47 – 0.63 mm) to produce an isotropic image volume for segmentation.

3.2 Framework Step 1: Segmentation

From the resampled images of each patient, the affected proximal femur (or femurs for the two patients with bilateral LCPD) was segmented to include the femoral head and approximately two centimetres of the proximal femoral shaft below the tip of the lesser trochanter (Figure 1). Segmentation was performed manually by two segmentation experts (LJ and SS). An initial segmentation pass of each proximal femur, focusing on anatomical accuracy, (LJ, *3D Slicer* v4.13.0, Slicer community, www.slicer.org)[27] was followed by a refinement pass to correct any errors and remove voxel-scale surface roughness (SS, *Amira* v6.0.1, Thermo Fisher Scientific Inc., Waltham, MA, USA). Due to the patients' varied stages of ossification, the segmentation protocol included isolation of both bone and cartilage of the proximal femoral epiphysis and apophyses of the greater and lesser trochanters, to better reflect the full shape of the bone (Figure 1). To assess the repeatability of the segmentation protocol between raters and software packages, each rater separately completed both segmentation passes on one femur, and the Hausdorff distance was calculated between the two segmentation boundaries.

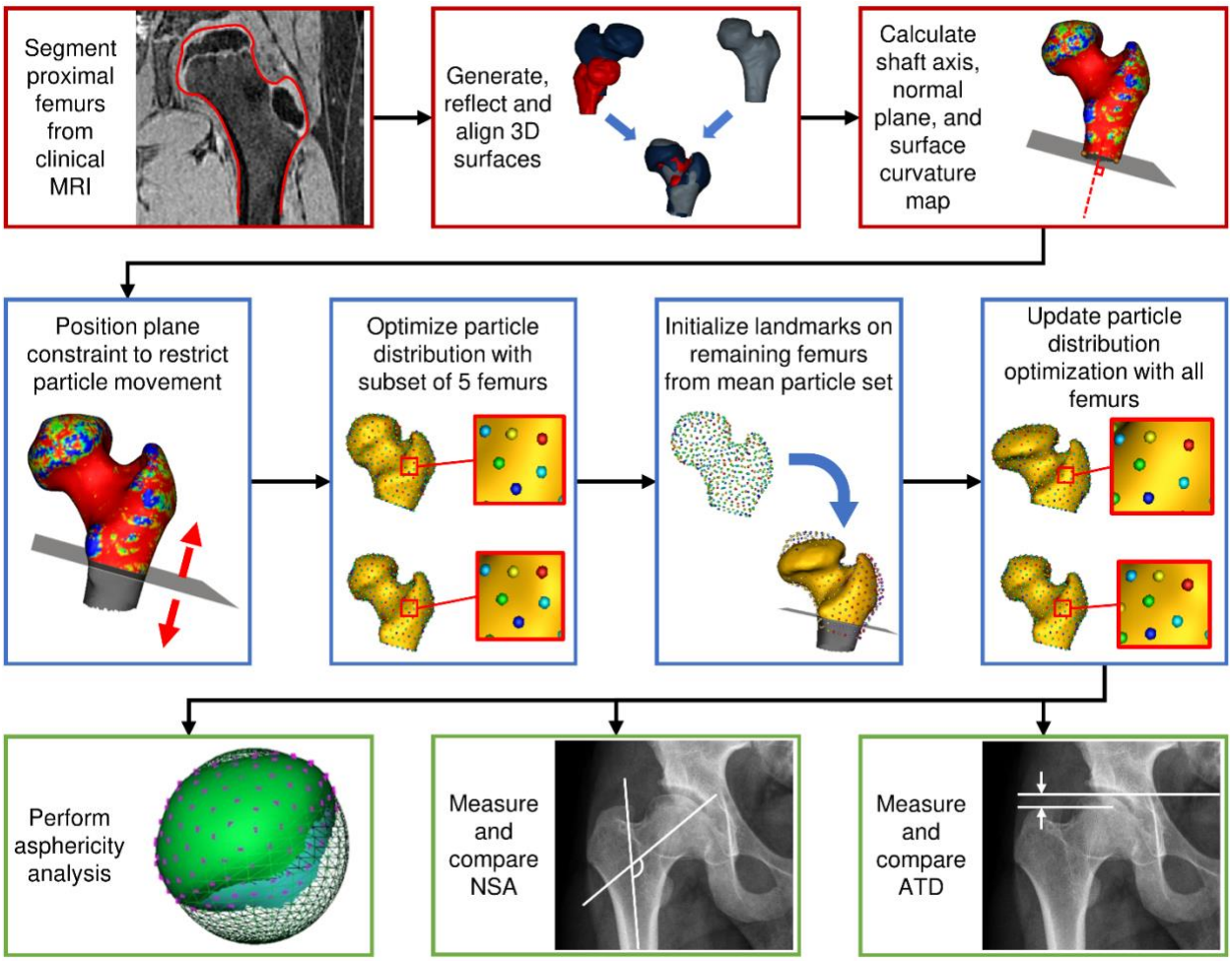


Fig. 1 Key stages of the modeling and analysis pipeline for this statistical shape model (SSM) framework, including segmentation and surface mesh preparation steps (top row, red), the particle-based correspondence step using an incremental optimization routine in *ShapeWorks* (middle row, blue), and analysis of anatomical variation (bottom row, green).

3.3 Framework Step 2: Surface Mesh Preparation

Three-dimensional surfaces of the proximal femurs were generated in *Amira*, using a previously published iterative smoothing and decimation protocol.[28] Due to the small size of the pediatric anatomy, meshes were scaled to three times their native resolution to reduce the effective strength of particle repulsion and improve correspondence optimization in the subsequent step. Meshes were reflected, if left-sided, and aligned via the iterative closest point algorithm using *CloudCompare* v2.11 (www.cloudcompare.org)[29] (Figure 1).

Cutting planes were defined to constrain particle placement to consistent regions of the proximal femur. To do so, the best-fit cylinder to the femoral shaft was determined for each mesh using custom *MATLAB* code (*MATLAB* R2022a, The MathWorks Inc, Natick, Massachusetts, USA). The orientation of the long axis of the cylinder defined the normal vector of the plane (Figure 1). Next, Gaussian curvature maps were computed for each mesh using *MeshLab* (v2022.02, Visual Computing Lab, ISTI-CNR, Pisa, Italy), to guide the proximal/distal placement of the plane in the particle-based correspondence step.[30]

3.4 Framework Step 3: Particle-based Correspondence

We used *ShapeWorks* v6.3.0 (shapeworks.sci.utah.edu)[12] on an ASUS GL552V laptop (Intel Core i7-6700-HQ, Windows 10) to generate a particle-based correspondence model of the proximal femur. After importing the femur

surface meshes and cutting plane orientations generated in the previous steps, the proximal/distal position of the plane on each femur was standardized by using the distal part of the lesser trochanter physis when visible on MRI in conjunction with the division between the lesser trochanter and femoral shaft created by the curvature map (Figure 1).

We utilized a two-stage incremental optimization approach to establish particle correspondence across the femoral meshes. Incremental optimization routines begin with initially fitting a model to a subset of the most similar shapes, before incrementally adding outlier shapes. In small datasets with large shape variance, incremental optimization can achieve better particle correspondence and model compactness compared to a model optimized on the whole dataset at once.[25] In the current study the first stage involved placing correspondence particles (n=512) on a subset of five meshes using a fully automated hierarchical splitting strategy and entropy-based optimization (Figure 1).[24] Generalized Procrustes analysis removed the effect of pose and scale during optimization of particle position. The mean particle coordinates from the first stage were used to initialize particle locations on the remaining eight meshes, after which the optimization routine was rerun in order to establish correspondence on all 13 surfaces (Figure 1).

Principal component analysis (PCA) was utilized to consolidate the dimensionality of the model (equal to the number of particles multiplied by the three spatial dimensions) to a set of linearly uncorrelated modes, which describe the dominant shape variations among the cohort. Patient-specific PCA component scores were used to describe the shape of each femur relative to the variation captured by each mode. Here, PCA scores can be understood as the weights for each variable when calculating the principal component.

3.5 Framework Evaluation

We evaluated our SSM framework performance using the three standard metrics: compactness, the proportion of total variability explained by a chosen number of modes; generalization, the ability of the model to represent shapes that were not part of the training set; and specificity, the ability of the model to generate only valid shapes and to differentiate between shapes in different categories.[31] To test our model's sensitivity to changes in training set size, we extended the generalization calculation from a leave-one-out to a leave-N-out cross validation using custom Python code. Values of generalization were calculated for N = 1 to N = 9 (an effective training set size of 12 and 4 respectively); the case for N = 1 is equivalent to the standard generalization metric.

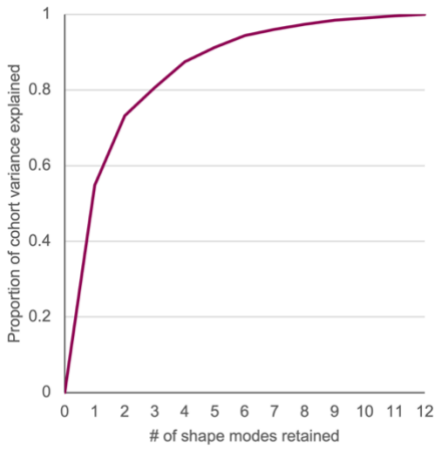
3.6 Analysis of Anatomical Variation

To explore how the current SSM framework may be applied alongside alternative 2D or 3D measures of deformity, we compared the PCA component scores for each femur in our evaluation cohort with a 3D measure of femoral head asphericity and 2D measurements from each patient's corresponding clinical radiographs. Gaussian curvature was computed for the reconstructed mean mesh in *MeshLab*, and used to extract the region corresponding to the femoral head in *CloudCompare* (Figure 1). The indices of the correspondence particles that resided within the mean femoral head region were then identified and used to calculate the best-fit sphere for each femoral head in the cohort using custom *MATLAB* code. Asphericity was expressed as the root-mean square error of the sphere fit (Table 1). Neck-shaft angle (NSA) and articulotrochanteric distance (ATD) were measured on anterior-posterior radiographs for each patient femur (Figure 1, Table 1). Last, Spearman's rank correlation coefficient was quantified to examine the relationship between asphericity, NSA and ATD with PCA component scores.

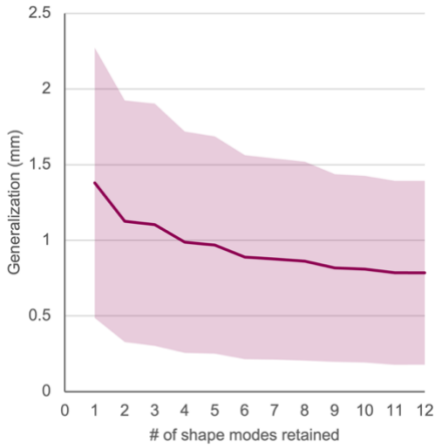
4 Results

Figure 2 shows good performance of this SSM framework when applied to the evaluation cohort. The model is compact, representing almost 90% of the total variation with just four modes (Figure 2a). The model generalizes to less than 1mm with four modes (Figure 2b), with the generalization curve flattening after the ninth mode. The model's specificity is effectively constant (Figure 2c), with a range of 1.12-1.14mm.

a.



b.



c.

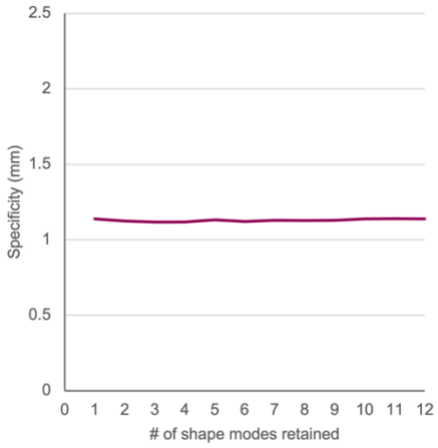


Fig. 2 Results of SSM evaluation when up to 12 shape modes are retained, presented using the three standard evaluation metrics: **a) compactness**, the cumulative proportion of cohort variance explained; **b) generalization**, the average point-to-point error when reconstructing unfamiliar shapes (shaded area indicates ±1 standard deviation); **c)**

specificity, the average point-to-point difference between shapes randomly generated from the SSM and their closest training set shapes.

Leave-N-out cross validation demonstrated the sensitivity of this framework to changes in the training set size when four shape modes are included (Figure 3). Increasing N from 1 to 9 (reducing the effective training set size from 12 to 4 femurs) increased the mean point-to-point reconstruction error (generalization) from 0.99mm to 1.40mm. The mean Hausdorff distance between segmentations completed separately by each rater (LJ and SS) on one femur was 0.38 mm (maximum 3.35 mm, 95th percentile 1.13 mm).

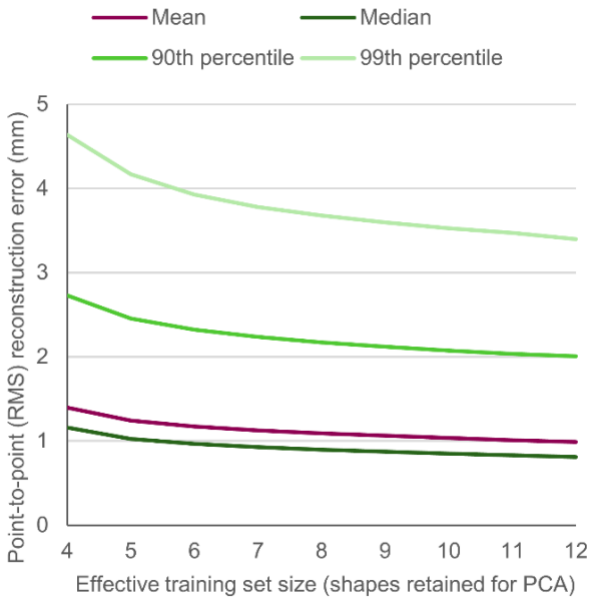


Fig. 3 Results of leave-N-out cross-validation evaluation of framework performance, showing the impact of changes in the effective training set sample size on the mean (purple), median (dark green), 90th and 99th percentile (green and light green, respectively) point-to-point reconstruction error using four shape modes.

The first four modes of variation were selected for further analysis, as these modes accounted for at least 5% of the total variation observed in the SSM. Together, these modes accounted for 87.5% of the total cohort variability, with generalization and specificity of 0.99mm and 1.12mm respectively. Mode I described the oblateness (the degree of compression of a sphere, here along the axis of the femoral neck, to form an ellipsoid) and width of the femoral head in the anterior-posterior and medial-lateral directions, vertical angle of the lateral greater trochanter, and ATD (Figure 4). Mode II described the slope of the superior femoral head, size of the femoral head and greater trochanter, ATD, NSA, prominence of the vastus ridge, and depth of the fovea (Figure 4). Mode III described the height of the femoral head and vastus ridge, shape of the intertrochanteric crest, and width of the femoral neck (Figure 4). Lastly, mode IV described anterior protrusion and asphericity of the femoral head, prominence of the posterior greater trochanter, and NSA (Figure 4).

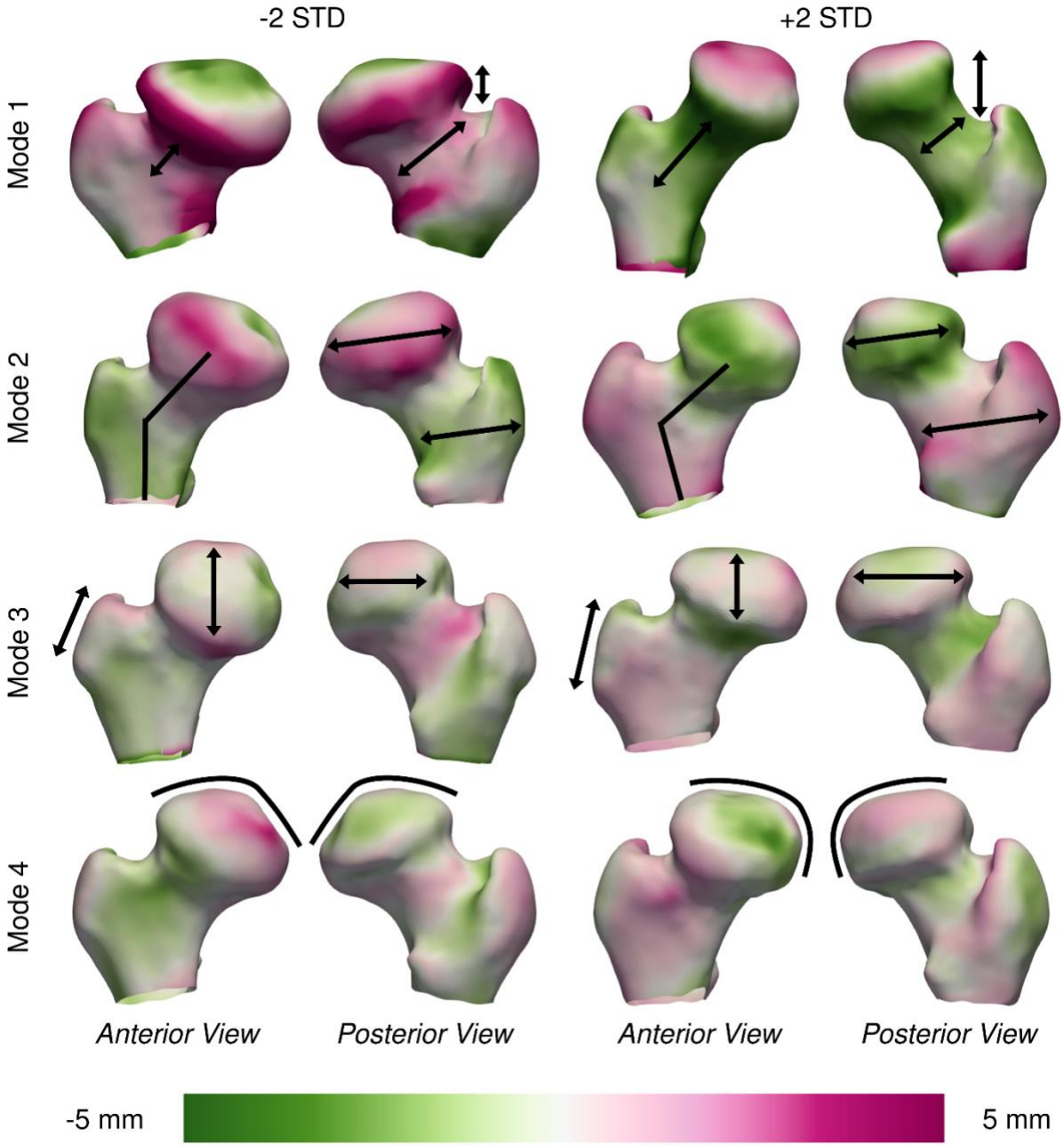


Fig. 4 The first four principal component analysis (PCA) modes of the evaluation statistical shape model (SSM). Surface distance plots show shapes with each mode set to ±2 standard deviations (SD) from the cohort mean shape, and the deviation of each surface from the mean shape is represented by a color gradient from green (negative displacement) to magenta (positive displacement). Arrows qualitatively represent notable areas of variation captured by each mode.

Shape mode scores were calculated for each femur in the evaluation cohort and compared with the corresponding femoral head asphericity, NSA, and ATD measurements for that femur (Figure 5). Correlations between shape mode scores and these parameters ranged from very weak to strong. Asphericity and NSA were most strongly correlated with mode IV (Spearman's rank correlation coefficients of 0.79 and 0.63 respectively), whereas ATD was most strongly correlated with mode II (Spearman's rank correlation coefficient of -0.63). Other notable correlations where

Spearman's rank correlation coefficient was stronger than ±0.4 included NSA with Mode II (-0.43) and ATD with mode I (0.58).

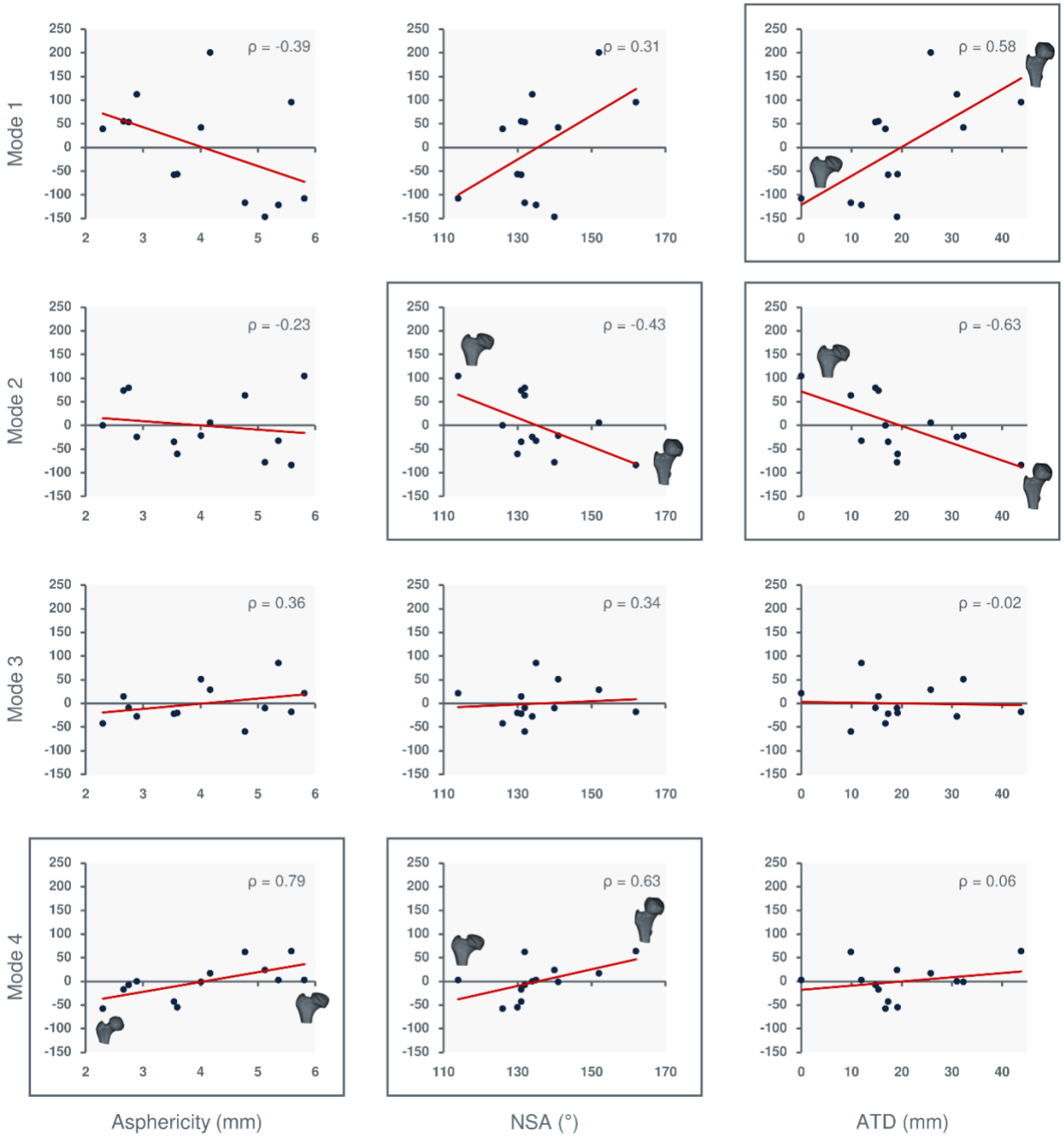


Fig. 5 Analysis of anatomical variation. Shape mode scores for each proximal femur in the evaluation cohort are plotted on the vertical axes (rows from top to bottom representing modes 1-4), against femoral head asphericity (left column), neck-shaft angle (NSA, middle column), and articular-trochanteric distance (ATD, right column). Spearman's rank correlation coefficient (ρ) is shown for each relationship, and relationships stronger than ρ = ±0.4 are indicated with a plot boundary and depictions of the femurs with the highest and lowest horizontal axis values.

5 Discussion

In this study, we developed and implemented a new methodological framework for constructing SSM of highly heterogeneous LCPD anatomy, based on progressive optimization of particle correspondences. We evaluated this framework by producing a 3D SSM of proximal femurs from clinical MR images of patients in stage IIa-IV LCPD, which provided a compact, accurate, and objective description of 3D shape and shape variation in LCPD. The first four PCA components of this model met our threshold for inclusion (>5% of overall cohort variance); these modes described morphological changes characteristic of LCPD consistent with clinical observation such as an enlarged and widened FH and a high-riding greater trochanter. Quantitative associations between PCA component scores and radiographic measurements demonstrated how to interpret SSM-based findings relative to clinical radiographic measurements of LCPD that are more familiar to clinicians than SSM-based measurements.

The contribution of this work is a framework, which includes protocols for image volume segmentation, surface mesh preparation, and generation of a particle-based correspondence model. The main steps of the current framework are derived from existing protocols,[13, 28] with the key innovation of incremental optimization of particle correspondences.[25] As well as improving performance in small and heterogeneous datasets, incremental optimization also enables incremental expansion of the model using data from multiple sources. Open science practices are a powerful tool to facilitate research into rare conditions such as LCPD, and every step in the framework can be carried out using free and open source software (including *Octave*, an open-source implementation of *MATLAB*). Collectively, the described framework and SSM results of this study provide a basis for developing more standardized tools to evaluate LCPD deformity in 3D. Due to its compact but comprehensive description of morphology, this modeling framework provides opportunities for future SSM research to relate patient-specific morphology with biomechanics in affected and healthy participants, or to longitudinally quantify how LCPD hips remodel in modified Waldenstrom stage IV until skeletal maturity.

In our small cohort of LCPD hips, the SSM framework produced a model with good evaluation results. A compactness of 87.5% with four modes demonstrates efficient representation of the cohort's variability. A generalization below 1mm with four or more modes suggests that the model was robust when applied to unfamiliar data, and was not over-fit to the training set. The specificity curve of the model did not meaningfully change with the number of modes considered, which may be an artifact of the small size and heterogeneity of the training set. To calculate specificity, *ShapeWorks* generates shapes by assigning random normally distributed shape mode weights. However, no individual shape in our evaluation cohort is close to the average shape in the first mode (Figure 5). As a result, the point-to-point distance between the large number of shapes generated near the mean and the relatively distant "nearest" training set examples dominates the specificity calculation for all modes. The resultant specificity of 1.1mm suggests that the model could nonetheless effectively discriminate between shapes and shape categories.

The accuracy of this evaluation model was better than other hip SSMs produced using larger training sets. For example, Ziaeipoor *et al.* reported a median generalization error of 1.1-4.6mm using a training set of 18 normal femurs, and Whitmarsh *et al.* reported a mean reconstruction error of 1.1mm using 12 modes with a training set of 85 normal, osteopenic and osteoporotic femurs.[32, 33] In addition, our leave-N-out cross-validation analysis suggests that the framework's performance is robust to changes in training set sample size. Collectively, these measures of SSM performance provide confidence in using the described framework to further study LCPD pathomorphology.

Our results were broadly similar to studies by Chan *et al.*, who produced an SSM that included some LCPD proximal femurs alongside those with slipped capital femoral epiphysis (SCFE) and no pathology.[17, 18] This previous model exhibited similar features in shape modes associated with LCPD when compared to the present study, including short and wide femoral necks and flattened heads as noted by Chan and colleagues. Similarity between the previous and present models was particularly notable in mode I, whereas subsequent modes differ with more characteristics of SCFE present in the previous model. We felt it was important to produce a new model using only LCPD proximal femurs to evaluate our SSM framework as it is unclear how much of the variation described by each mode can be attributed to LCPD as opposed to SCFE, particularly as LCPD proximal femurs made up only a small proportion (one ninth) of the whole cohort in the study by Chan and colleagues ($N_{LCPD} = 5$, $N_{SCFE} = 19$, $N_{healthy} = 21$).

Correlations between PCA component scores and standard radiographic measures were evaluated to interpret each significant mode of variation relative to existing clinical measures of deformity. In our evaluation cohort, NSA appeared to prominently change in modes 1 and 2 of the model (Figure 4). However, the correlation between mode I

principal component score and clinically measured NSA was weak compared to ATD, a closely related measurement (Figure 5). This may be caused by difficulty in measuring NSA in patients with short, wide femoral necks, asymmetric femoral anteversion,[34] and an unclear femoral head center. This example demonstrates a potential application of the framework: to evaluate radiographic measurements against a 3D baseline shape description, which has been explored and validated in previous studies of the proximal femur.[13, 16] Harris *et al.* (2013) applied a similar analysis to SSM of adult hips with cam FAI deformity,[13] and reported "moderate to weak" associations between key radiographic metrics and shape modes (up to ρ=0.403 for Mode 1 and alpha angle). The stronger associations (up to ρ=0.79) found in the present study demonstrate the large size of LCPD deformity compared to FAI.

Representation of 3D pathomorphology in terms of continuous shape modes could improve the assessment of long-term prognosis in the clinic. For example, in 2D SSM of adult hips, shape modes representing femoral neck length, femoral neck width, and head-neck offset are predictive of total hip replacement.[15] Similarly, individual modes or combinations of modes from SSM of LCPD deformity could predict the timing and severity of long-term outcomes and inform preventative management strategies.

This study does have some considerations and limitations. First, we studied a small group of hips to evaluate the framework in a cohort representative of LCPD morphology, and to illustrate how the resulting models may be related to existing clinical measures. Readers should not draw clinical conclusions from these findings, and we have not performed hypothesis testing for that purpose. A full assessment of LCPD pathomorphology should include age groups at various stages of development, each consisting of patients with a representative range of disease severities. The influence of factors such as the modality of the source imaging, resolution, and contrast on the performance of the current framework should also be considered as part of future research. Second, the field of view of the retrospective MR images did not include the distal femur or full hemipelvis, without which the model cannot describe key features such as femoral version[34] or the pose-dependent relationship between the femoral head and acetabulum. Future work making use of this SSM framework should ensure that all relevant anatomy is included in the field of view, including both femurs to the knee joint and the full pelvis. Finally, the framework presented herein does not itself separate changes in shape due to pathology from changes due to normal growth, the latter being substantial in the population of interest. We believe that accurate representation of shape for the whole range of deformity severity and at all developmental stages of growth will allow for more detailed studies of the relationship between deformity and growth in future. Our framework report in this study is well placed to provide that shape representation.

In summary, the methodological framework for SSM proposed in this study provides an accessible, compact and accurate representation of the 3D pathomorphology of LCPD. This expandable framework represents a step towards an objective standard for descripting 3D LCPD morphology, which could help researchers and clinicians better understand how pathomorphology affects long-term outcomes in patients. This understanding will facilitate the development of patient-specific treatment guidelines for residual proximal femoral deformity, particularly in patients who are at high risk of developing early-onset OA.

6 References

1. Kim HKW (2010) Legg-Calvé-Perthes Disease. J Am Acad Orthop Surg 18:676–686. https://doi.org/10.5435/00124635-201011000-00005

2. Joseph B, Mulpuri K, Varghese G (2001) Perthes' disease in the adolescent. J Bone Jt Surg 83:715–720. https://doi.org/10.1302/0301-620X.83B5.10663

3. Paley D (2011) The Treatment of Femoral Head Deformity and Coxa Magna by the Ganz Femoral Head Reduction Osteotomy. Orthop Clin North Am 42:389–399. https://doi.org/10.1016/j.ocl.2011.04.006

4. Tannast M, MacIntyre N, Steppacher SD, Hosalkar HS, Ganz R, Siebenrock KA (2013) A systematic approach to analyse the sequelae of LCPD. HIP Int 23:61–70. https://doi.org/10.5301/hipint.5000071

5. Clohisy JC, Nepple JJ, Ross JR, Pashos G, Schoenecker PL (2015) Does Surgical Hip Dislocation and Periacetabular Osteotomy Improve Pain in Patients With Perthes-like Deformities and Acetabular Dysplasia? Clin Orthop Relat Res 473:1370–1377. https://doi.org/10.1007/s11999-014-4115-7

6. Stulberg SD, Cooperman DR, Wallensten R (1981) The natural history of Legg-Calve-Perthes disease. J

Bone Jt Surg 63A:1095–1108

7. Larson AN, Sucato DJ, Herring JA, Adolfsen SE, Kelly DM, Martus JE, Lovejoy JF, Browne R, DeLaRocha A (2012) A Prospective Multicenter Study of Legg-Calvé-Perthes Disease. J Bone Jt Surg 94:584–592. https://doi.org/10.2106/JBJS.J.01073

8. Dagenais S, Garbedian S, Wai EK (2009) Systematic review of the prevalence of radiographic primary hip Osteoarthritis. Clin Orthop Relat Res 467:623–637. https://doi.org/10.1007/s11999-008-0625-5

9. O'Connor JD, Hill JC, Beverland DE, Dunne NJ, Lennon AB (2021) Influence of preoperative femoral orientation on radiographic measures of femoral head height in total hip replacement. Clin Biomech (Bristol, Avon) 81:105247. https://doi.org/https://dx.doi.org/10.1016/j.clinbiomech.2020.105247

10. Lechler P, Frink M, Gulati A, Murray D, Renkawitz T, Bücking B, Ruchholtz S, Boese CK (2014) The influence of hip rotation on femoral offset in plain radiographs. Acta Orthop 85:389–395. https://doi.org/10.3109/17453674.2014.931196

11. Johnson LG, Bortolussi-Courval S, Chehil A, Schaeffer EK, Pawliuk C, Wilson DR, Mulpuri K (2023) Application of statistical shape modeling to the human hip joint: a scoping review. JBI Evid Synth 21:. https://doi.org/10.11124/JBIES-22-00175

12. Cates J, Elhabian S, Whitaker R (2017) ShapeWorks: Particle-Based Shape Correspondence and Visualization Software. In: Statistical Shape and Deformation Analysis. Elsevier, pp 257–298

13. Harris MD, Datar M, Whitaker RT, Jurrus ER, Peters CL, Anderson AE (2013) Statistical shape modeling of cam femoroacetabular impingement. J Orthop Res 31:1620–1626. https://doi.org/10.1002/jor.22389

14. Turmezei TD, Treece GM, Gee AH, Sigurdsson S, Jonsson H, Aspelund T, Gudnason V, Poole KES (2020) Quantitative 3D imaging parameters improve prediction of hip osteoarthritis outcome. Sci Rep 10:4127. https://doi.org/https://dx.doi.org/10.1038/s41598-020-59977-2

15. Agricola R, Reijman M, Bierma-Zeinstra SMA, Verhaar JAN, Weinans H, Waarsing JH (2013) Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: A prospective cohort study (CHECK). Osteoarthr Cartil 21:559–564. https://doi.org/10.1016/j.joca.2013.01.005

16. Atkins PR, Shin Y, Agrawal P, Elhabian SY, Whitaker RT, Weiss JA, Aoki SK, Peters CL, Anderson AE (2019) Which Two-dimensional Radiographic Measurements of Cam Femoroacetabular Impingement Best Describe the Three-dimensional Shape of the Proximal Femur? Clin Orthop Relat Res 477:242–253. https://doi.org/10.1097/CORR.0000000000000462

17. Chan EF, Farnsworth CL, Koziol JA, Hosalkar HS, Sah RL (2013) Statistical shape modeling of proximal femoral shape deformities in Legg–Calvé–Perthes disease and slipped capital femoral epiphysis. Osteoarthr Cartil 21:443–449. https://doi.org/10.1016/j.joca.2012.12.007

18. Chan EF, Farnsworth CL, Klisch SM, Hosalkar HS, Sah RL (2018) 3-dimensional metrics of proximal femoral shape deformities in Legg-Calvé-Perthes disease and slipped capital femoral epiphysis. J Orthop Res 36:1526–1535. https://doi.org/10.1002/jor.23791

19. Standefer KD, Dempsey M, Jo CH, Kim HKW (2017) 3D MRI Quantification of Femoral Head Deformity in Legg–Calvé–Perthes Disease. J Orthop Res 35:2051–2058. https://doi.org/10.1002/jor.23484

20. Pienkowski D, Resig J, Talwalkar V, Tylkowski C (2009) Novel three-dimensional MRI technique for study of cartilaginous hip surfaces in Legg-Calvé-Perthes disease. J Orthop Res 27:981–988. https://doi.org/10.1002/jor.20909

21. Memiş A, Varlı S, Bilgili F (2021) A novel approach for computerized quantitative image analysis of proximal femur bone shape deformities based on the hip joint symmetry. Artif Intell Med 115:102057. https://doi.org/10.1016/j.artmed.2021.102057

22. Styner MA, Rajamani KT, Nolte L-P, Zsemlye G, Székely G, Taylor CJ, Davies RH (2003) Evaluation of 3D Correspondence Methods for Model Building. In: Taylor C, Noble JA (eds) Information Processing in Medical Imaging. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 63–75

23. Goparaju A, Iyer K, Bône A, Hu N, Henninger HB, Anderson AE, Durrleman S, Jacxsens M, Morris A,

Csecs I, Marrouche N, Elhabian SY (2022) Benchmarking off-the-shelf statistical shape modeling tools in clinical applications. Med Image Anal 76:102271. https://doi.org/10.1016/j.media.2021.102271

24. Cates J, Fletcher PT, Styner M, Shenton M, Whitaker R (2007) Shape modeling and analysis with entropy-based particle systems. Inf Process Med Imaging 20:333–45. https://doi.org/10.1007/978-3-540-73273-0_28

25. Scientific Computing and Imaging Institute (2022) Incremental Supershapes: Building a Shape Model Incrementally. http://sciinstitute.github.io/ShapeWorks/latest/use-cases/multistep/incremental_supershapes.html. Accessed 29 Aug 2023

26. Joseph B, Varghese G, Mulpuri K, Rao K.L. N, Sreekumaran Nair N (2003) Natural evolution of Perthes disease: A study of 610 children under 12 years of age at disease onset. J Pediatr Orthop 23:590–600. https://doi.org/10.1097/01241398-200309000-00005

27. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller J V., Pieper S, Kikinis R (2012) 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 30:1323–1341. https://doi.org/10.1016/j.mri.2012.05.001

28. Anderson AE, Peters CL, Tuttle BD, Weiss JA (2005) Subject-Specific Finite Element Model of the Pelvis: Development, Validation and Sensitivity Studies. J Biomech Eng 127:364–373. https://doi.org/10.1115/1.1894148

29. CloudCompare community (2023) CloudCompare (verson 2.11) [GPL software]. www.cloudcompare.org

30. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G (2008) Meshlab: an open-source mesh processing tool. In: Eurographics Italian chapter conference. Salerno, Italy, pp 129–136

31. Marzola A, Robilotta C, Volpe Y, Governi L, Furferi R (2021) Statistical Shape Model: comparison between ICP and CPD algorithms on medical applications. Int J Interact Des Manuf 15:85–89. https://doi.org/10.1007/s12008-020-00725-1

32. Whitmarsh T, Humbert L, De Craene M, Del Rio Barquero LM, Frangi AF (2011) Reconstructing the 3D Shape and Bone Mineral Density Distribution of the Proximal Femur From Dual-Energy X-Ray Absorptiometry. IEEE Trans Med Imaging 30:2101–2114. https://doi.org/10.1109/TMI.2011.2163074

33. Ziaeipoor H, Taylor M, Martelli S (2020) Population-Based Bone Strain During Physical Activity: A Novel Method Demonstrated for the Human Femur. Ann Biomed Eng 48:1694–1701. https://doi.org/10.1007/s10439-020-02483-3

34. Novais EN, Nunally KD, Ferrer MG, Miller PE, Wylie JD, Dodgen WT (2021) Asymmetrically increased femoral version with high prevalence of moderate and severe femoral anteversion in unilateral Legg-Calvé-Perthes disease. J Child Orthop 15:503–509. https://doi.org/10.1302/1863-2548.15.200247