Characterization of Normal Femoral Anatomy in Pediatric Patients using Statistical Shape Models

Hannicka Sahlstedt

2018

LUNDS UNIVERSITET

Master's Thesis in

Biomedical Engineering

Faculty of Engineering, LTH
Department of Biomedical Engineering

Supervisors: Assoc. Prof. Hanna Isaksson
Dr. Lorenzo Grassi

Abstract

It is still not well defined how the anatomy of the hip in healthy pediatric patients develop. Defining the span of normal anatomical development could help identify patients that deviate from healthy development and enable early diagnosis of certain pathological conditions of the hip.

To study the normality range of the 3D morphology of the hip for pediatric patients, a statistical shape model (SSM) can be used. Given a large enough set of hips the SSM can describe all the expected variations within the population.

In this MSc project SSMs were created for pediatric patients using CT scans of patients between the ages 7 and 20 years who had no reported hip pathologies. The method for creating the models was developed from an existing pipeline for construction of SSMs of the proximal femur of adult patients. The models created contained the variation of the proximal part of the femur for a healthy population of pediatric patients. The SSMs were used to identify anatomical differences during development by studying separate age and gender groups. When creating the SSMs all femurs were scaled so that the variation in size was not part of the model.

Significant differences in the morphological development were found between three age groups with 7 - 10, 12 - 15 and 16 - 17 year old patients. No significant differences were found between female and male patients.

The developed methods and results from this thesis may eventually be useful for identifying patients that are developing abnormally and thereby help the surgeon with clinical decision making.

Preface

This master's thesis was carried out during 2017 and completed in 2018. The thesis was done at the Department of Biomedical Engineering at the Faculty of Engineering, Lund University.

Acknowledgements

I could not have done this project alone, and there are some people I would like to thank for their help.

Foremost I want to thank my supervisors Hanna Isaksson and Lorenzo Grassi, for assisting me in every step of this project. I would like to thank Dr Sami Väänänen (Department of Applied Physics, University of Eastern Finland) for all the help he has provided during the project. I want to thank Dr Carl Johan Tiderius (Department of Orthopedics, Clinical Sciences, Lund, Lund University) and Dr Magnus Karlsson (Department of Orthopedics, Clinical Sciences, Malmö, Lund University) for their expertise and their time, required to collect the CT scans needed for this project. Thanks to Emilia Törnell and Amanda Wallin for their work, segmenting the CT scans. Their work allowed my project to proceed further than would have been possible without them.

Lastly I would like to thank the whole Biomechanics group for instantly making me feel part of the group. Thanks to them I have learned so much, not only related to my project but also about the whole world of academics.

Hannicka Sahlstedt
Lund, 2018

List of Abbreviations

BMD	Bone Mineral Density
CT	Computed Tomography
DXA	Dual Energy X-ray Absorptiometry
GPA	Generalized Procrustes Analysis
ICP	Iterative Closest Point
PCA	Principal Component Analysis
SD	Standard Deviation
SSAM	Statistical Shape and Appearance Model
SSM	Statistical Shape Model
TPS	Thin Plate Spline

Table of Contents

1 Introduction

1.1 Background

The normal morphological development of the hip for pediatric patients is not clinically well defined. Defining what normal anatomical development is could help identify patients that deviate from the normal and enable early diagnosis of certain pathological conditions of the hip (Siebelt et al., 2014; Nelson et al., 2016). Such conditions include epiphysiolysis, Perthes disease and hip dysplasia. These conditions all lead to high risk of hip osteoarthritis in early adult life.

There are different parameters that can be measured to quantify the morphology of the hip. Examples of these parameters are the Caput-Collum-Diaphyseal (CCD) angle, the lateral center edge (LCE) angle and Klein's line (Figure 1.1) (Beaulé et al., 2012).

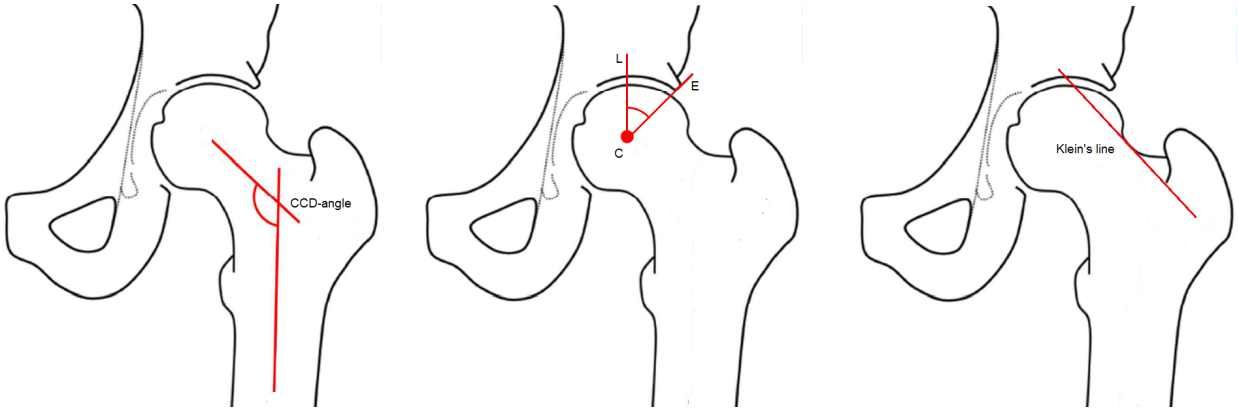

Figure 1.1: Examples of parameters used to quantify the hip morphology. Left: CCD-angle, the angle between the femoral neck and shaft. Middle: the LCE-angle, the angle that forms between the most superior point on the acetabular rim (E), the center of the femoral head (C) and a vertical line (L). Right: Klein's line, a line parallel to the upper edge of the femoral neck.

Many of the anatomical parameters used for morphological quantifications are defined in 3D. However, currently diagnostics of hip pathologies are often based on evaluations of femoral radiographs, thus on 2D projections of the bone (Nelson et al., 2016; Siebelt et al., 2014; Agricola et al., 2013; Beaulé et al., 2012). Measuring these parameters from 2D images may introduce uncertainties in the measurements as all parameters are not well represented from all angles.

Credible measurements of the hip rely on the 3D morphology (Clohisy et al., 2009). For a single radiograph the patient position and rotation when the scan is obtained has an impact on the accuracy with which quantifications of the hip can be made. One way to obtain a 3D representation of the patient femur is by computed tomography (CT) scans. However, the CT imaging imposes a higher radiation dosage compared to plain X-ray or dual energy X-ray (DXA) images. (Damilakis et al., 2010; Wren et al., 2005)

A low radiation dose is always desirable, and is of special importance when imaging pediatric patients. Children's developing organs and tissues are inherently more vulnerable to the damaging effects of radiation than adults. Even though individual risks are small, the radiation from CT scans increases the risk of diseases, such as leukemia and other forms of cancer, in a population (Zacharias et al., 2013; Almohiy, 2014; Pearce et al., 2012).

To study the normality range of the 3D morphology of the hip for pediatric patients, a statistical shape model (SSM) can be used. The SSM can be constructed from CT scans of a set of normal femurs. Since scans made for other purposes can be used no new radiation has to be imposed on patients to construct the models. Given a large enough set, the model will contain all the variation expected in the femoral shape of the population. From such a model all normal anatomies can be constructed (Cootes et al., 2000).

Within the Biomechanics group of the Department of Biomedical Engineering, Lund University, there has been an initiative to assess the normal development of the hip morphology in growing children. Assessments could further be used to establish if it is possible to identify individuals at risk of developing hip pathologies that need medical treatments and surgery in adult life. If so, can limits be found for when treatment is necessary? Assessments would be made using a clinically feasible tool that can describe and quantify hip morphology in 3 dimensions based on only one 2 dimensional radiographic image. The first step in this larger framework is to develop a SSM for the proximal part of the femur for children. This is also the aim of this MSc project.

An existing pipeline for creation of statistical shape and appearance models (SSAM) is available (Väänänen et al., 2015). The pipeline creates models, that contain the outer shape of the bone and the bone mineral density (BMD) value, of the femur for adult patients. In the current MSc thesis project, the existing pipeline is further developed and adapted to create SSMs for pediatric patients. These SSMs will contain the femoral shape, but no BMD values.

1.2 Aim

The aim of this project was to create SSMs of the proximal femur for pediatric patients of different ages and genders. Furthermore, the project aimed to identify anatomical differences during development by studying separate age and gender groups.

2 Theory

In this section, the theoretical background needed to understand the creation of the statistical models is provided. The theoretical approach for model creation is visualized in Figure 2.1. First, segmentations of the outer surface of a set of patient femurs were obtained. The segmented surfaces consisted of triangulated point clouds, describing initial meshes for each patient femur. On each patient femur landmarks were placed. The landmark placement was done by registering a template femur mesh to the outer shape of each patient femur. The template nodes were used as landmarks. The template mesh consisted of two parts, a surface mesh and a volume mesh. First the template surface mesh was registered to each initial femur mesh. This registration was made using an iterative closest point (ICP) algorithm. Next the template volume mesh was registered to the registered surface mesh. This registration was done using thin plate splines (TPS).

When all patient femurs were described by registered volume meshes the meshes were aligned with each other using generalized Procrustes analysis (GPA). The aligned volume meshes were analyzed using principal component analysis (PCA) to create a SSM (Figure 2.1).

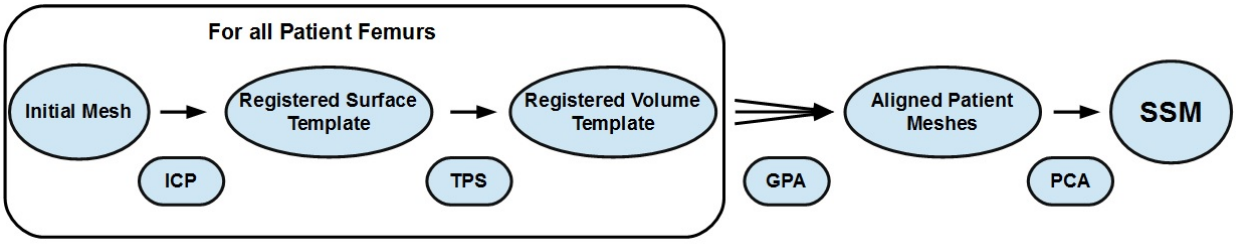

Figure 2.1: An overview of the project pipeline containing the steps explained in Theory.

2.1 Iterative Closest Point

Iterative closest point (ICP) is an algorithm that finds a transformation that aligns two point clouds. One point cloud is assigned as fixed and the other as moving. For each of the moving points, the closest point among the fixed points is found. A transformation that minimizes the distance from each moving point to a point in the fixed point cloud is calculated. Using the transformation all moving points are transformed. This procedure is iterated until convergence is reached. The points are assumed not to differ too much from alignment when the algorithm is started. If the shapes are oriented in the same direction just aligning the shape centers might be enough (Rusinkiewicz and Levoy, 2001). In this project an ICP algorithm was used to register a template surface mesh to the anatomy of a set of patient femurs. A nonrigid transformation is included in the registration.

2.2 Thin Plate Spline

Thin plate splines (TPS) is a method for finding a mapping from $\mathbb{R}^n \rightarrow \mathbb{R}^n$ given a set of landmarks and their locations after transformation. Based on the known locations of landmarks before and after transformation TPS finds the full mapping that best transforms

the whole $\mathbb{R}^n$ to align the landmarks with their transformed correspondences. The theory behind TPS can be found in (Bookstein, 1989). In this project TPS was used to find the mappings that best registered a template volume mesh to surface meshes describing the anatomy of a set of patient femurs.

2.3 Generalized Procrustes Analysis

Generalized Procrustes analysis (GPA) is a method used to find the similarity transformation that best aligns a set of shapes, with each other. To align the shapes corresponding landmarks must be placed on them. The transformation involves translation, rotation, reflection, and scaling. The method is further described in (Gower, 1975). In this project GPA was used to align a set of volume meshes, describing patient femur anatomies.

2.4 Principal Component Analysis

Principal component analysis (PCA) is a method that can be used to simplifying a set of variables. The simplification is done by generating a new set of variables, called principal components. All principal components are orthogonal to each other and each is a linear combination of the original variables.

Each principal component forms an axis in space. Due to the structure of the principal components if the original set of variables are projected onto the first principal component the maximum variance of the projected variable is obtained. When projecting the original value to each following principal component, the projected variable variance decreases. The percentage of variation explained by each principal component can be calculated.

By recreating the original variables using a linear combination of only a subset of principal components, an approximation of the original data is obtained (Jackson, 1991). In this project PCA was used to construct SSMs.

2.5 Statistical Shape Model

A statistical shape model (SSM) is a model constructed from a set of training objects with corresponding landmarks. PCA is performed on the landmarks of all training object. Each principal component obtained from the PCA describes a principal mode of variation for the set of training object. The first principal component, or principal mode of variation, contains the largest amount of variation. Each following principal mode contains a decreasing amount of variation. Each training object can be reconstructed as a linear combination of the model principal modes of variation. The scalar value that each principal mode is multiplied with in this linear combination is often referred to as the mode score (Cootes et al., 2000).

In this project SSMs were created for the proximal femur for pediatric patients.

3 Material and Methods

In this chapter the project methods used to create SSMs are presented. The methods pipeline is visualized in Figure 3.1. First the approach for creating meshes from CT images of patient femurs is described. It is then described how multiple meshes are used to create SSMs. The method used for the model creation is an adaptation of the method described by (Väänänen et al., 2015). Finally the methods for evaluating the models are explained.

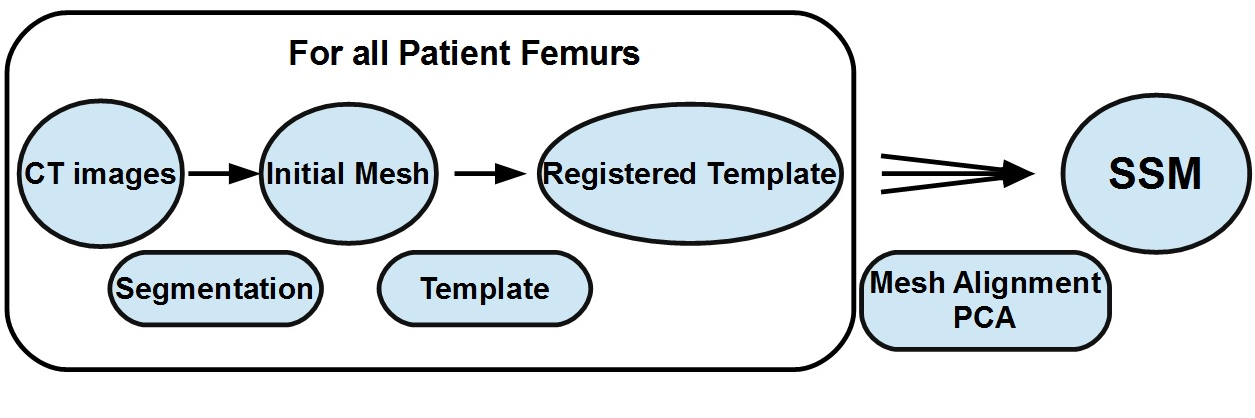

Figure 3.1: A visualization of the pipeline for the project methods.

3.1 CT Images

The CT scans used in this project to create SSMs of the proximal femur, were in vivo scans acquired from patients between seven and twenty years of age. Images were obtained between the years 2000 and 2015. The patients included in the project were chosen as they were all without any reported hip-related pathology. The CT scans were extracted from existing databases and approved for use in this project with Ethical Permission (Number 256-2015, KVB 178-15). Since only existing CT scans that were made for other purposes were used, no new radiation doses were imposed on patients as part of this project.

In total 61 bilateral patient scans were available. They were divided into three age groups, 7 - 10, 12 - 15 and 16 - 17 years old. The age and gender distributions are presented in Table 3.1.

Ages	All Patients	Male patients	Female patients
All ages	61	38	23
7-10 y	11	8	3
12-15 y	30	20	10
16-17 y	20	10	10

Table 3.1: Number of patient CT scans available. The table shows age and gender distributions for the patients. All scans contain both left and right side femurs.

CT scans contained images taken in the transversal plane for all patients. The images had pixel sizes in the range $0.5-1\ mm$. The image slice spacing was in the range $0.5-1\ mm$ for all except three patients, from the age group 7 - 10 years, where the slice spacing was about $3\ mm$.

3.2 Image Segmentation

In order to create meshes of the patient femurs the bones had to be segmented from the CT scans. Manually segmented image masks of the femurs in the scans were available from a previous BSc project (Törnell and Wallin, 2017). The segmentations were done in the software Seg3D (CIBC, 2016).

3.3 Initial Meshes

The image masks obtained through manual segmentation were used to create an initial mesh that described the outer surface of each patient femur. The initial femur meshes were created in Matlab using the function *isosurface*. The function creates a surface mesh from a "pixel volume". 1000 uniformly distributed nodes were chosen from the surface of each initial femur mesh and saved for use in the first step of mesh registration.

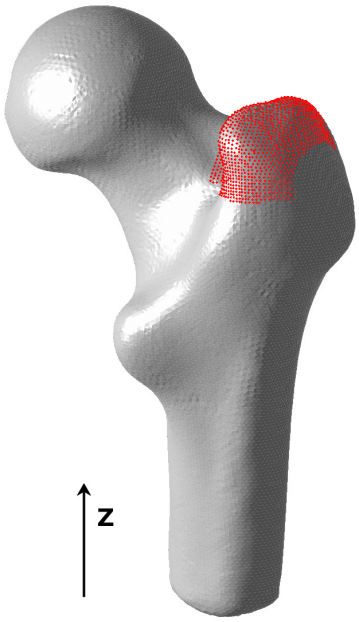

Figure 3.2: The template surface mesh. The nodes that define the greater trochanter region are marked in red.

3.4 Template Mesh Registration

In order to create a SSM for the femur all patient femur meshes had to have corresponding landmarks. The landmarks were defined by registering the same template mesh to each initial femur mesh. This resulted in the same node numbering, and nodes corresponding to approximately the same part of the femoral anatomy, for all patient femur meshes. The nodes were used as landmarks.

The registration of the template was done in two steps. First the outer shape, the template surface mesh was registered to each initial femur mesh. Then the template volume mesh was registered to the registered surface mesh. The details of the registrations are explained in this section.

3.4.1 Template Mesh

The template used for landmark placement consisted of a surface mesh and a volume mesh. The template surface mesh is shown in Figure 3.2. The template surface mesh contained 27543 nodes, and 55082 triangular faces. The template volume mesh contained 25958 nodes, and 138794 tetrahedral elements. The template mesh was created from the right side femur of an adult patient and was about 14 *cm* long, cut 6 *cm* below the tip od the lesser trochanter. The nodes that belonged to the greater trochanter region were defined for the template surface mesh (marked in red in Figure 3.2). The region information was needed for the registration.

3.4.2 Template Surface Mesh Registration

For a successful registration of the template surface mesh to the initial femur mesh, their geometries had to resemble each other. To satisfy this requirement the template surface mesh was cropped to only contain about as much of the proximal part of the femur as was present in the patient CT scans, and therefore initial femur meshes. Only nodes above a z-coordinate threshold, (see Figure 3.2) from the template, were used for registration. The threshold was chosen manually by the user after having been shown an image of the initial femur mesh next to the template surface mesh. An example of such an image is seen in Figure 3.3.

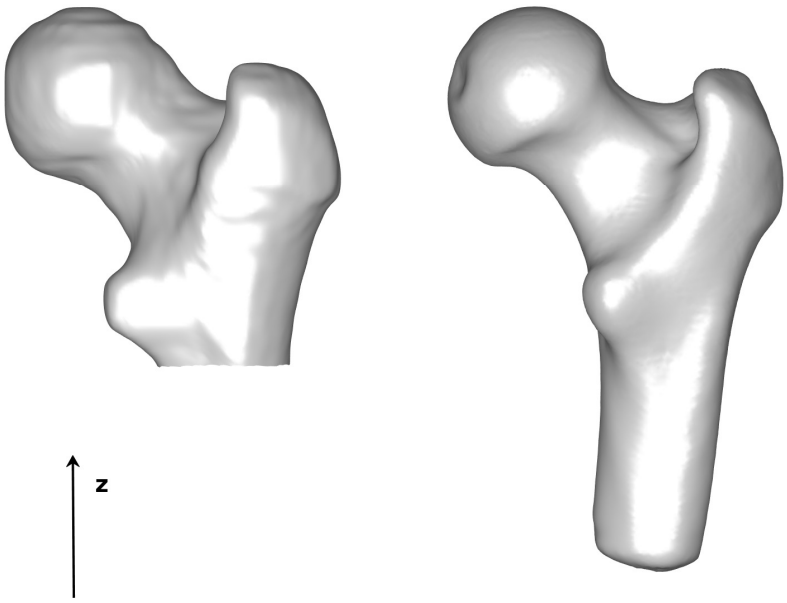

Figure 3.3: Initial mesh for a patient (left) next to the template surface mesh (right). The z-coordinate at which the template is cropped is manually chosen for each patient femur.

The 1000 uniformly distributed nodes chosen from the initial femur mesh were registered to the cropped template surface mesh. If the femur mesh and the template were from different sides, the template surface mesh nodes were first mirrored.

The template surface mesh was registered to the 1000 nodes from the initial femur mesh using an ICP algorithm. Figure 3.4 shows the registration steps. First the initial placement of both the template surface mesh and the 1000 nodes from the initial femur mesh is shown (Figure 3.4a). The first step of the registration was a translation of the nodes from the initial femur mesh to align their mean coordinate with the template surface mesh mean coordinate (Figure 3.4b). Next the nodes from the initial femur mesh were rigidly registered to the template surface mesh (Figure 3.4c), then the nodes were resized (Figure 3.4d), and an affine transformation was applied to register the nodes to the template (Figure 3.4e). These four transformation steps were combined into one transformation that was applied to the nodes from the initial femur mesh at their initial location (Figure 3.4f). Finally the inverse of the combined transformation was applied to the template surface mesh, registering it to the initial placement of initial femur mesh (Figure 3.4g).

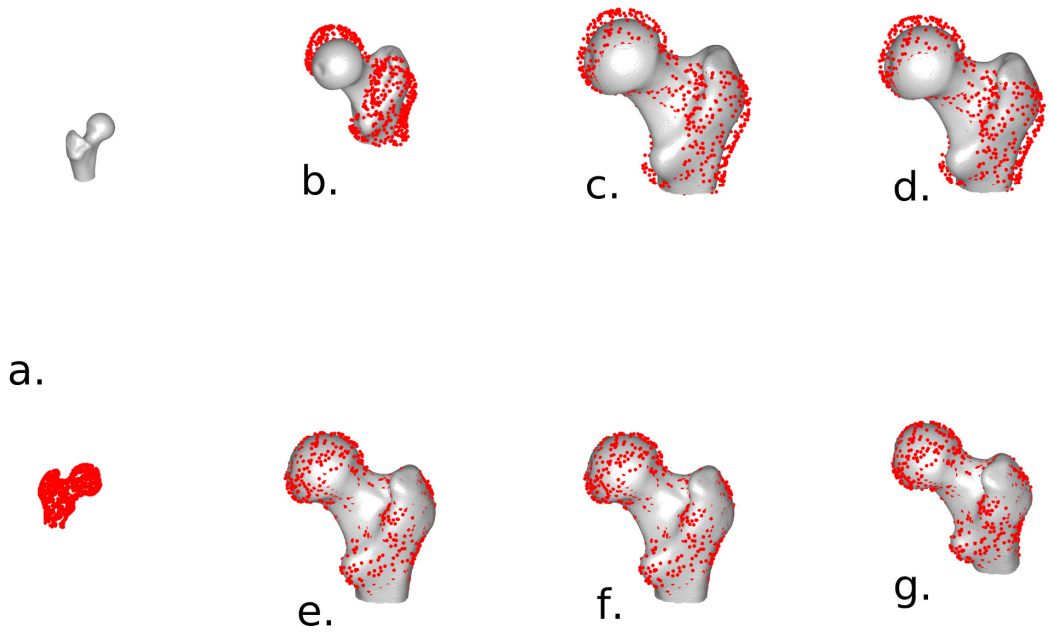

Figure 3.4: Registration of the template surface mesh to the 1000 nodes from the initial femur mesh (in red). a) The initial placement of the template and the nodes from the initial femur mesh. b) The nodes after translation and the template. c) The nodes after rigid registration and the template. d) The nodes after resizing and the template. e) The nodes after affine transformation and the template. f) The nodes after applying the combined transformation and the template. g) The initial placement of the nodes and the template after applying the inverse transformation.

At some locations the femur anatomy was more likely to differ from the registered template. The greater trochanter was such a region. This region was therefore treated in an additional registration step. The nodes, on the template surface mesh, that belonged to the greater trochanter region, were predefined (Figure 3.2). These nodes were used in a non-linear registration step. This created a closer registration of the greater trochanter region.

The trochanter region of the initial femur mesh was chosen as the nodes in a region close to the template trochanter region. The nodes from the initial femur mesh were chosen from the mesh registered to the template (the registration step in Figure 3.4f). The trochanter region of the template mesh was then registered to the trochanter nodes from the initial femur mesh. Figure 3.5 shows an example of an initial femur mesh and a registered surface mesh before and after the additional trochanter registration.

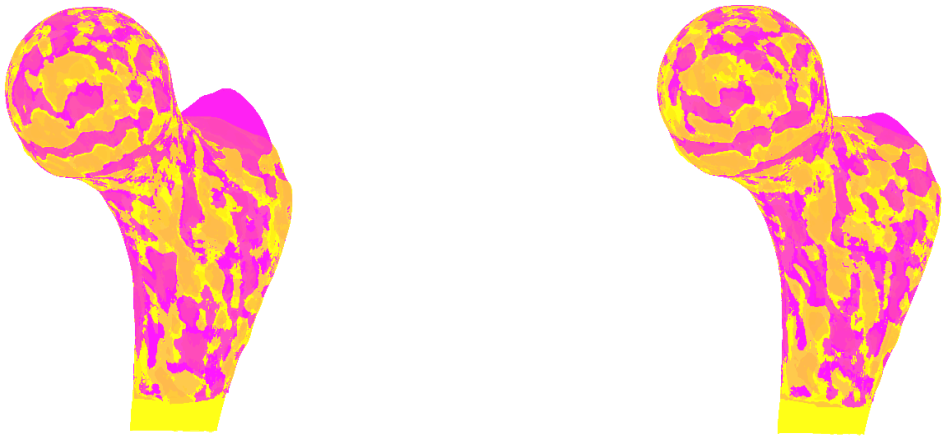

Figure 3.5: The template surface mesh (magenta) registered to the initial femur mesh (yellow) for one patient example. (Left) Registration before additional registration of the greater trochanter. (Right) Registration after additional registration of the greater trochanter.

The last step of the registration of the surface mesh was a non-rigid registration. This registration was performed using the Matlab function *nonrigidICP* (Manu, 2017). The function performed a finer registration of the template surface mesh to the initial femur mesh.

3.4.3 Template Volume Mesh Registration

To register the template volume mesh to the registered surface mesh, surface nodes were randomly chosen from the registered surface mesh. Initially 10000 nodes were chosen from the whole surface mesh. From these nodes that were not used in the surface registration step, since the mesh was cropped, were removed.

The nodes from the templates surface mesh, at its original location, and from the registered surface mesh respectively, were used to find the TPS transformation that best aligned the two node sets. The TPS transform was then used to register the whole template volume mesh to the registered surface mesh. Last, the registered volume mesh was distally cropped at the same threshold (z-coordinate) as the template surface mesh.

3.4.4 Evaluation of Template Mesh Registration

To quantify how well the volume meshes were registered to each initial femur mesh the distance from each surface node of the registered volume mesh to the surface of the initial femur mesh was calculated. The mean and standard deviation of the root mean square (RMS) distance and maximum distance for all mesh pairs were calculated. The calculations were also done for each age subgroup separately.

3.5 Statistical Shape Models

When volume meshes were created, for all patient femurs, a SSM could be created. The SSM could be constructed using any constellation of the registered volume meshes. A SSM was created from a set of femur volume meshes assigned as training objects. All training objects were aligned to each other using GPA. The GPA removed the scale of the objects. As a result the variance in size was not included in the SSM. Since the volume meshes were cropped not all template nodes were present in all training objects. Nodes that were missing in half or more of the training objects were excluded from the shape model.

The location of nodes that were missing in a training object, but used in the model, were estimated. The estimated node location was that of the corresponding node in the TPS registered volume mesh before cropping.

For the aligned training objects the mean and standard deviation of each node coordinate were calculated. The mean values and standard deviations were used to normalize all node coordinates. The SSM was created by performing PCA on the normalized coordinates.

SSMs were created for six patient groups, that contained different subgroups. One model contained all patients, one contained all male patients and one all female patients. The last three models contained all patients from each of the three age groups respectively. The models are presented in Table 3.2.

Model	Model Description	Model Name	#Patients (Femurs)
1	All patients	All	61 (122)
2	Patients age 7-10 years	7-10 y	11 (22)
3	Patients age 12-15 years	12-15 y	30 (60)
4	Patients age 16-17 years	16-17 y	20 (40)
5	All male patients	All Male	38 (76)
6	All female patients	All Female	23 (46)

Table 3.2: Descriptions of the SSMs created together with model names and number of patients in each model. Number of femurs in parentheses.

3.6 Model Analysis

To evaluate the models the amount of variance explained by each principal mode was calculated. Reduced models were formed by the first principal modes needed to explain at least 95% of the model variance.

The training objects were reconstructed from the reduced models. The reconstructed objects were compared to the corresponding training objects. The comparison was quantified by calculating the distances from each surface node of the reconstructed object to the surface of the training object. The two objects' volumes were also compared. The volumetric difference was calculated as $|V_{training} - V_{reduced}|/V_{training}$, where $V_{training}$ was the volume of the training object and $V_{reduced}$ the volume of the object reconstructed using the reduced model.

The scores, for the four most important principal modes, for all training objects were compared. Comparisons were done for all subgroups within each SSM. The Kolmogorov-Smirnov test was done for the mode scores for all subgroups of patients in each model. This was done to test if the mode scores, for all groups and subgroups of training objects, were normally distributed. After normal distribution was confirmed, a t-test was done to see if the subgroups of scores were significantly different from each other. For both statistical tests, a significance level of 5% was used.

4 Results

4.1 Template Mesh Registration Accuracy

Segmentations of CT images were the basis for model creation. An example of a CT scan together with segmented femurs is seen in Figure 4.1. The segmentation is displayed in the software Seg3D.

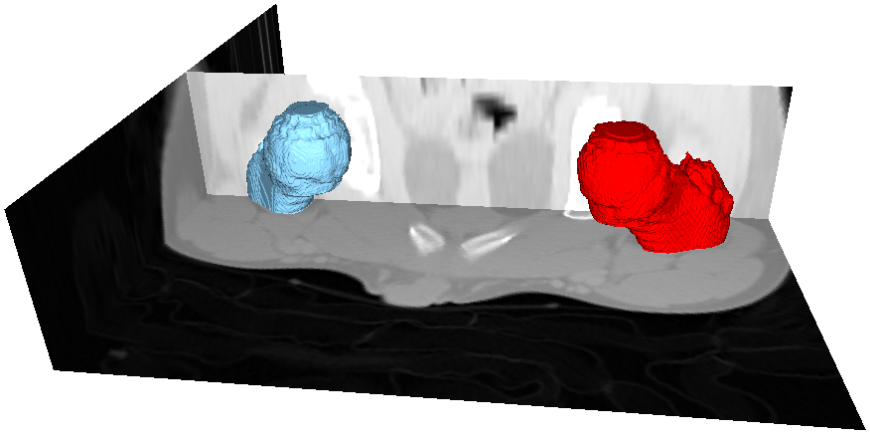

Figure 4.1: A segmented CT scan. The image shows a view from the software Seg3D.

The initial femur meshes were compared to the registered volume meshes. The mean RMS distances between the mesh pairs for all patients were smaller than 0.2 mm, with standard deviation on the same order of magnitude (Table 4.1). The mean was larger for the youngest patient group (7-10 years) than for the two older (12-15 and 16-17 years). The mean maximum distance between initial femur meshes and registered volume meshes were in the order of 2 mm (Table 4.1).

	All	7-10 y	12-15 y	16-17 y
RMS distance difference (mm)	0.17 ± 0.17	0.23 ± 0.28	0.15 ± 0.12	0.16 ± 0.15
Max distance difference (mm)	2.04 ± 2.21	2.04 ± 2.42	1.99 ± 2.02	2.11 ± 2.40

Table 4.1: Quantification of differences between initial femur mesh and registered template volume mesh. Data is presented as *mean* ± *standard deviation*.

The regions where the meshes differ can be visualized by superimposing the initial femur mesh and the template volume mesh or by displaying the calculated distance values as a color scale at each node in a plot of one of the meshes. An example of this for one representative patient is displayed in Figure 4.2.

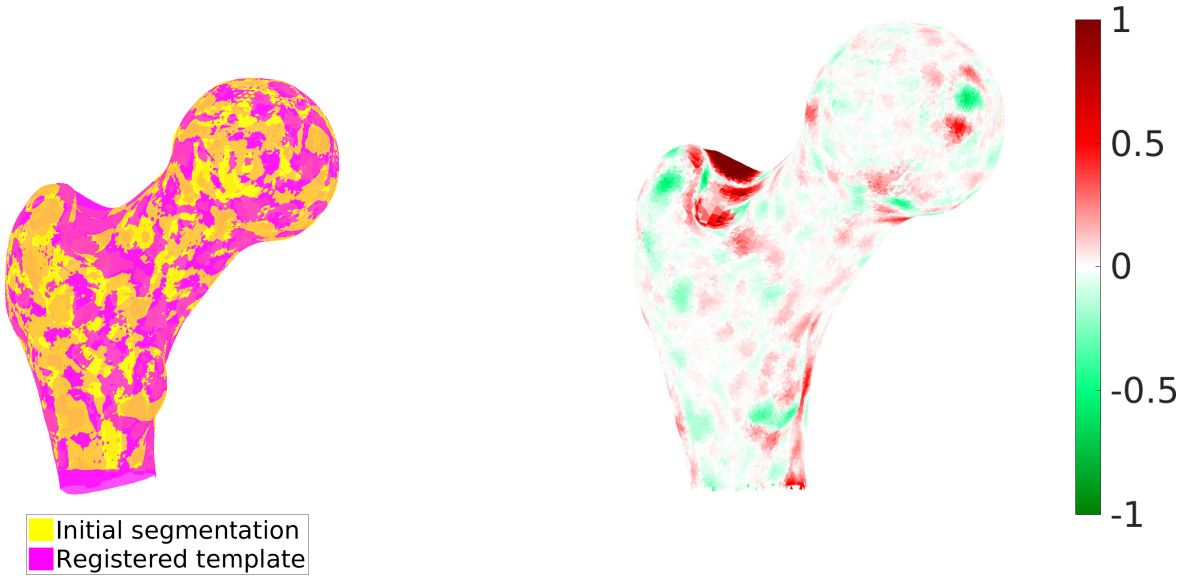

Figure 4.2: Meshes for the left femur from a male patient from age group 16-17 years. Left: The initial femur mesh (yellow) and registered template volume mesh (magenta) superimposed. Right: The registered template volume mesh with a color scale displaying distances (mm) to the initial femur mesh.

4.2 Statistical Shape Models

4.2.1 Model Variance

The percentage of shape variations explained by each principal mode of the created models are presented in Table 4.2.

Mode	All		7-10 y		12-15 y		16-17 y		All Male		All Female	
1	32.6	32.6	28.4	28.4	22.6	22.6	33.0	33.0	37.7	37.7	28.9	28.9
2	19.5	52.1	19.3	47.7	20.4	43.0	21.2	54.2	16.6	54.3	23.9	52.8
3	13.0	65.1	12.9	60.6	17.0	60.0	12.0	66.2	13.9	68.2	10.5	63.3
4	7.5	72.6	11.3	71.9	9.5	69.5	8.8	74.9	7.0	75.2	8.9	72.2
5	4.4	77.0	6.1	78.0	5.8	75.3	4.6	79.5	3.7	78.9	6.1	78.3
6	3.8	80.8	5.6	83.6	4.3	79.6	4.4	83.9	3.3	82.2	3.9	82.2
7	2.5	83.4	3.2	86.8	2.9	82.5	2.5	86.4	2.5	84.7	2.7	85.0
8	2.0	85.3	3.0	89.8	2.2	84.7	2.0	88.4	2.3	87.0	2.6	87.6
9	1.5	86.8	2.0	91.8	1.7	86.4	1.7	90.1	1.4	88.4	2.2	89.8
10	1.4	88.2	1.8	93.6	1.5	87.9	1.5	91.6	1.2	89.6	1.5	91.2
11	1.1	89.3	1.3	95.0	1.1	89.0	1.0	92.7	0.8	90.4	1.2	92.4
12	0.9	90.2	1.1	96.0	0.9	90.0	1.0	93.6	0.8	91.2	1.0	93.4
13	0.7	90.9	0.9	96.9	0.8	90.8	0.8	94.4	0.8	92.0	0.8	94.2
14	0.7	91.6	0.7	97.6	0.7	91.5	0.7	95.1	0.7	92.7	0.6	94.8
15	0.6	92.2	0.6	98.1	0.7	92.2	0.6	95.6	0.5	93.2	0.6	95.4
20	0.3	94.2	0.2	99.8	0.4	94.7	0.3	97.7	0.3	95.3	0.3	97.2
25	0.2	95.6	0.0	0.0	0.3	96.4	0.2	98.7	0.2	96.6	0.2	98.4
50	0.1	98.5	-	-	0.1	99.7	0.0	0.0	0.1	99.3	-	-
75	0.0	99.5	-	-	-	-	-	-	0.0	100.0	-	-

Table 4.2: Percentage of variation explained by each principal mode for different models together with the cumulative variation explained by the principal modes.

The number of principal modes needed to explain at least 95% of the variation of the SSM were chosen as a reduced shape model. The total number of principal modes, the number of principal modes in each reduced model and the percentage of principal modes used in reduced models are presented in Table 4.3.

Model	All	7-10 y	12-15 y	16-17 y	All Male	All Female
Reduced	23	11	21	14	19	15
Full	122	22	60	40	76	46
Reduced / Full (%)	18.9	50.0	35.0	35.0	25.0	32.6

Table 4.3: The number of principal modes in the reduced model explaining at least 95% of the model variation and the total number of principal modes for all SSMs. The percentage of the total number of principal modes needed to explain at least 95% of the model variance.

4.2.2 Reduced Model Reconstruction Accuracy

The mean RMS distance reconstruction errors for the reduced models were smaller than 0.5 *mm* (Table 4.4). For the model "7-10 y" the difference was even smaller.

The mean maximum distance difference reconstruction errors from the reduced models were smaller than 2.0 *mm*. For the model "7-10 y" the difference was smaller than 1.0 *mm* (Table 4.4).

The regions where the meshes differ can be visualized by superimposing the training object mesh and the corresponding mesh reconstructed using the reduced model or by displaying the calculated distance values as a color scale at each node in a plot of one of the meshes. An example of this for one representative patient is displayed in Figure 4.3.

	All	All Male	All Female
Volumetric difference (%)	0.54 ± 0.44	0.59 ± 0.43	0.74 ± 0.53
RMS distance difference (mm)	0.45 ± 0.07	0.46 ± 0.08	0.42 ± 0.09
Max distance difference (mm)	1.82 ± 0.62	1.91 ± 0.76	1.67 ± 0.59
	7-10 y	12-15 y	16-17 y
Volumetric difference (%)	0.83 ± 0.77	0.50 ± 0.41	0.52 ± 0.42
RMS distance difference (mm)	0.22 ± 0.07	0.40 ± 0.07	0.40 ± 0.08
Max distance difference (mm)	0.82 ± 0.32	1.64 ± 0.64	1.56 ± 0.49

Table 4.4: Quantifications of the differences between registered volume meshes and meshes reconstructed using reduced models.

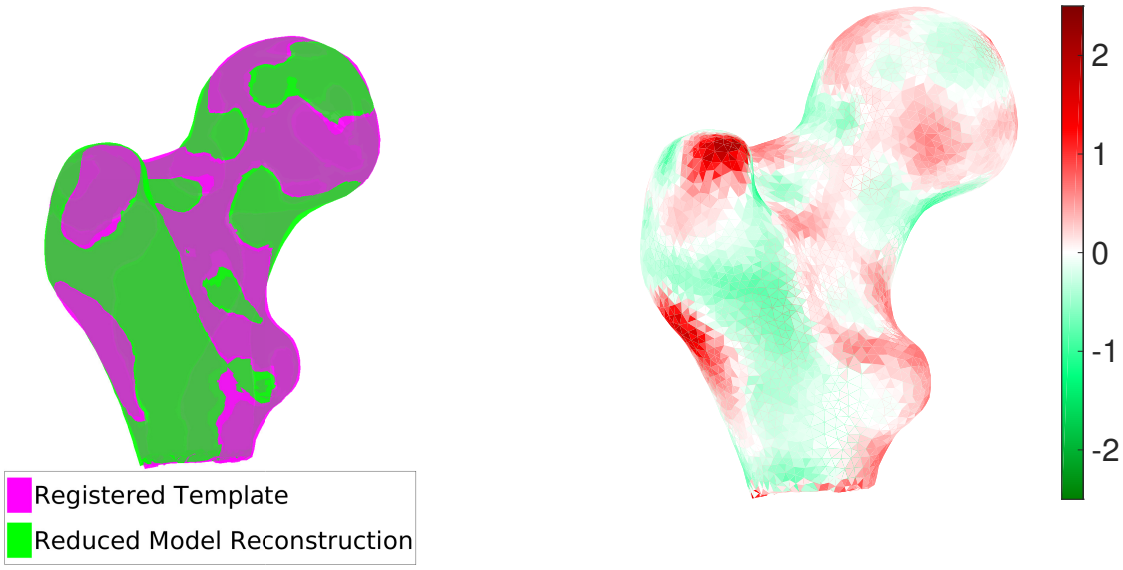

Figure 4.3: Meshes for the left femur, from a male patient, from age group 16-17 years, constructed using the model "All". Left: The registered template volume mesh (magenta) and the mesh reconstructed using the reduced SSM (green) superimposed. Right: The registered template volume mesh with a color scale displaying distance (mm) to the mesh reconstructed using the reduced SSM.

4.2.3 Principal Components

The Kolmogorov-Smirnov test showed that all subgroups from all models had normally distributed mode scores ($p < 0.05$), for principal modes 1-4.

T-tests showed no statistically significant differences between genders for any of the models ($p < 0.05$) (Table 4.5). For model "All" (containing all patients) the t-test showed a difference between the youngest age group and the two older ones for principal mode 1 and a difference between all age groups for principal mode 2 ($p < 0.05$). For model "All Male" (containing all male patients) the t-test showed a difference between the youngest and the two older age groups ($p < 0.05$, principal mode 1). For model "All Female" (containing all female patients) the t-test showed a difference between the youngest and the two older age groups ($p < 0.05$, principal mode 2) (Table 4.5).

Model	All				7-10 y				12-15 y				16-17 y				All Male				All Female			
Mode	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
7-10 vs 12-15	*	*	ns	ns		-				-				-			*	ns	ns	ns	ns	*	ns	ns
12-15 vs 16-17	ns	*	ns	ns		-				-				-			ns	ns	ns	ns	ns	ns	ns	ns
7-10 vs 16-17	*	*	ns	ns		-				-				-			*	ns	ns	ns	ns	*	ns	ns
M vs F	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns		-				-		

Table 4.5: Results from t-tests. Comparison of subgroups for each SSM. In the table (*) indicates that the subgroups are from different Gaussian distributions and (ns) that they are from the same ($p < 0.05$). The (-) indicates that no test was done.

The scores for the first two principal modes for all objects in each model are visualized in figures 4.4, 4.6, 4.8, 4.10, 4.12, 4.14 and 4.16. The shapes that correspond to variation of the two first principal modes for each model are visualized in Appendix A (Figures A.1 - A.6).

The mean and standard deviation for the first four principal modes in each model are visualized as error bars (Figures 4.5, 4.7, 4.9, 4.11, 4.13, 4.15 and 4.17). They are displayed both for all patients in a model and for each subgroup contained in a model.

The mean and standard deviation for all principal modes needed to explain at least 95% of the model variation are presented in Appendix A (Tables A.1 - A.6). The mean and standard deviations were calculated both for all patients in the model and separately for each subgroup contained in the model.

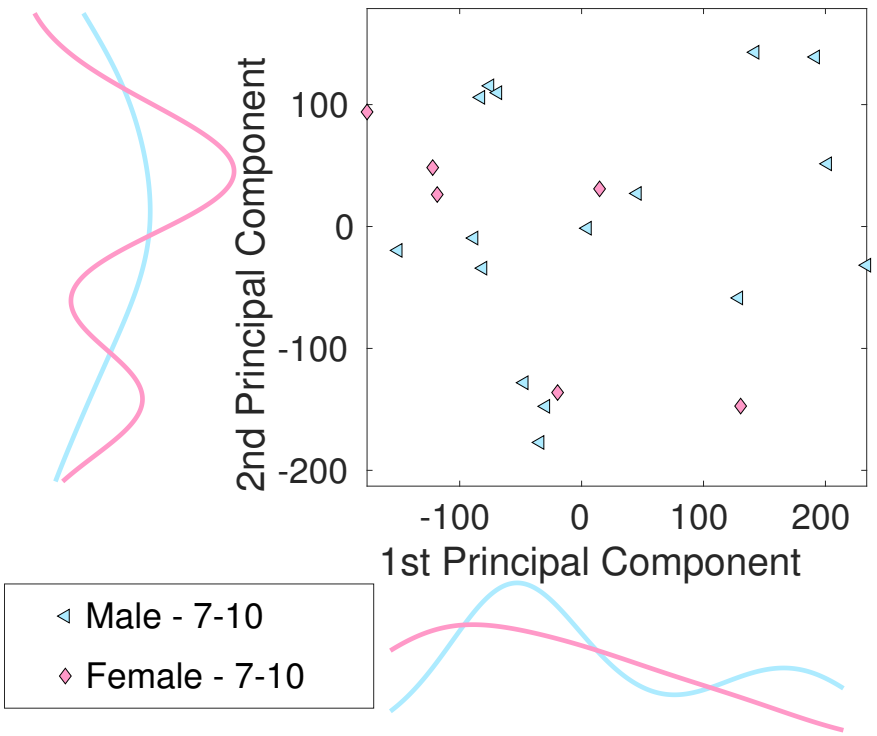

Figure 4.4: Model "7-10 y". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on gender.

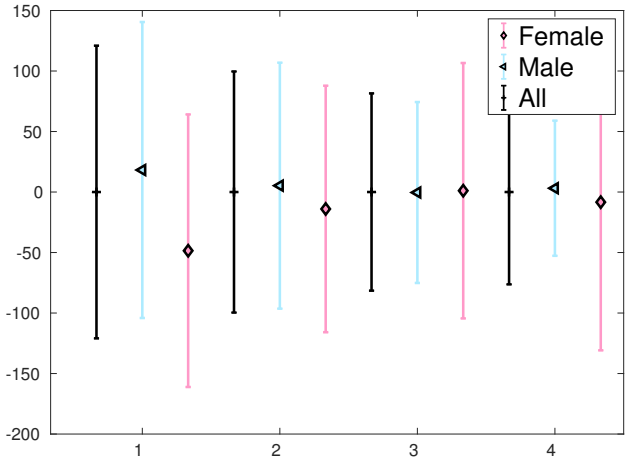

Figure 4.5: Model "7-10 y". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on gender.

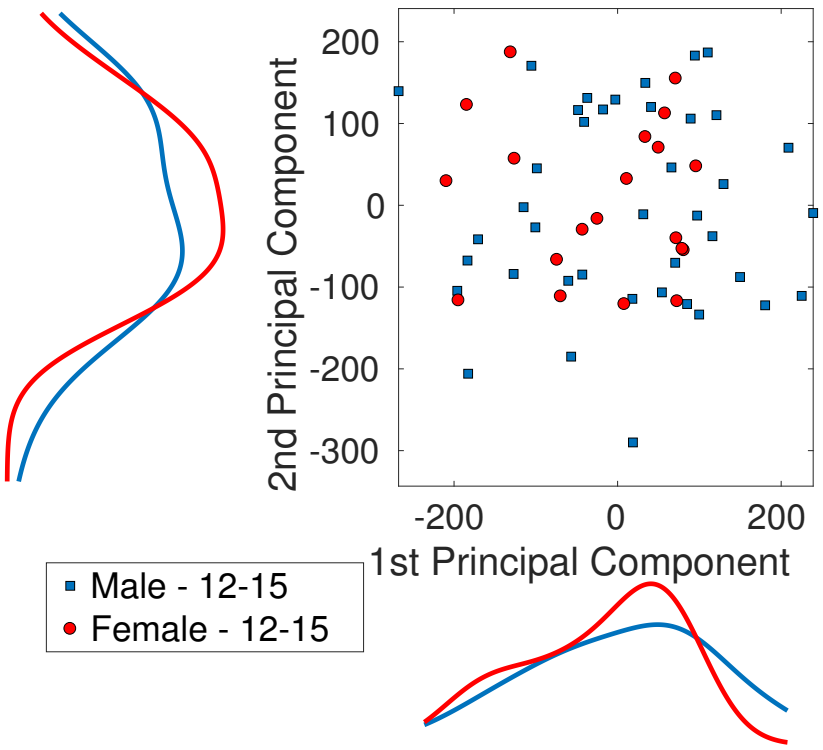

Figure 4.6: Model "12-15 y". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on gender.

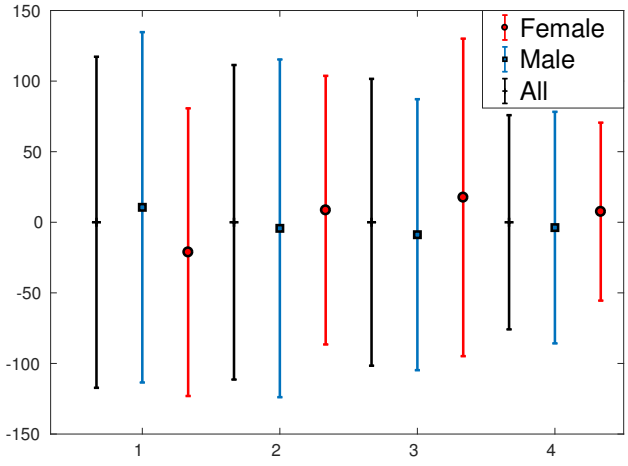

Figure 4.7: Model "12-15 y". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on gender.

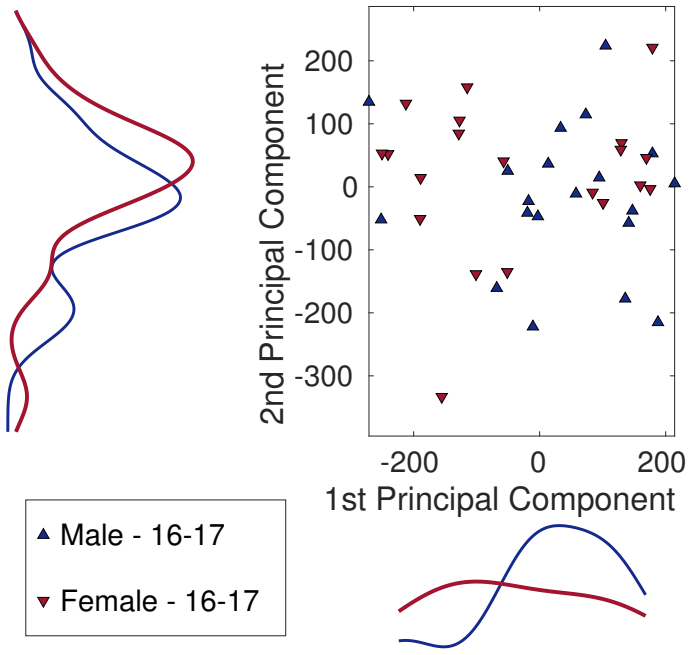

Figure 4.8: Model "16-17 y". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on gender.

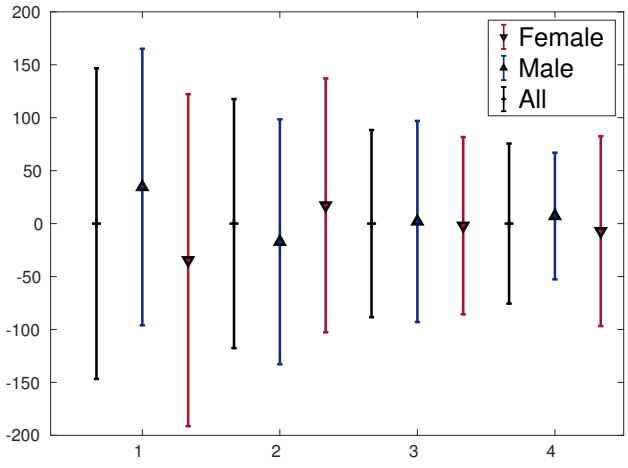

Figure 4.9: Model "16-17 y". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on gender.

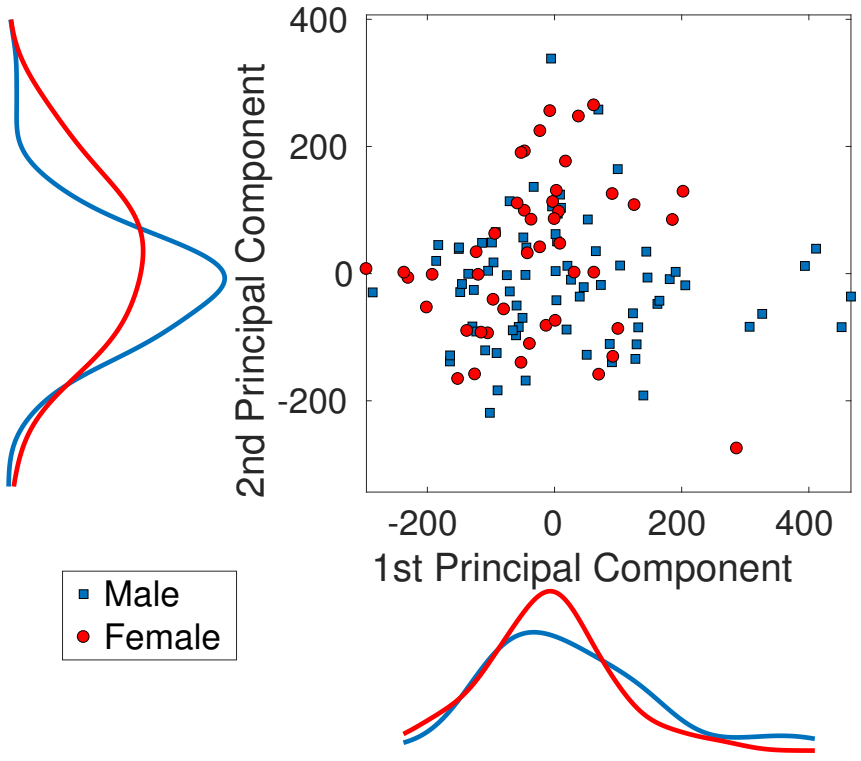

Figure 4.10: Model "All". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on gender.

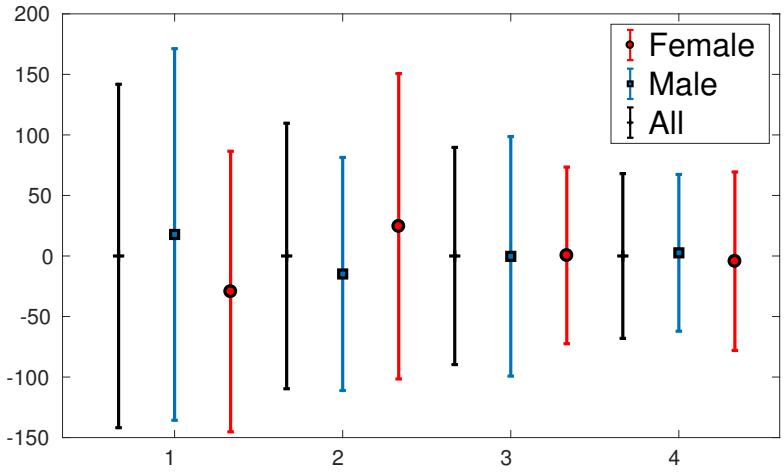

Figure 4.11: Model "All". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on gender.

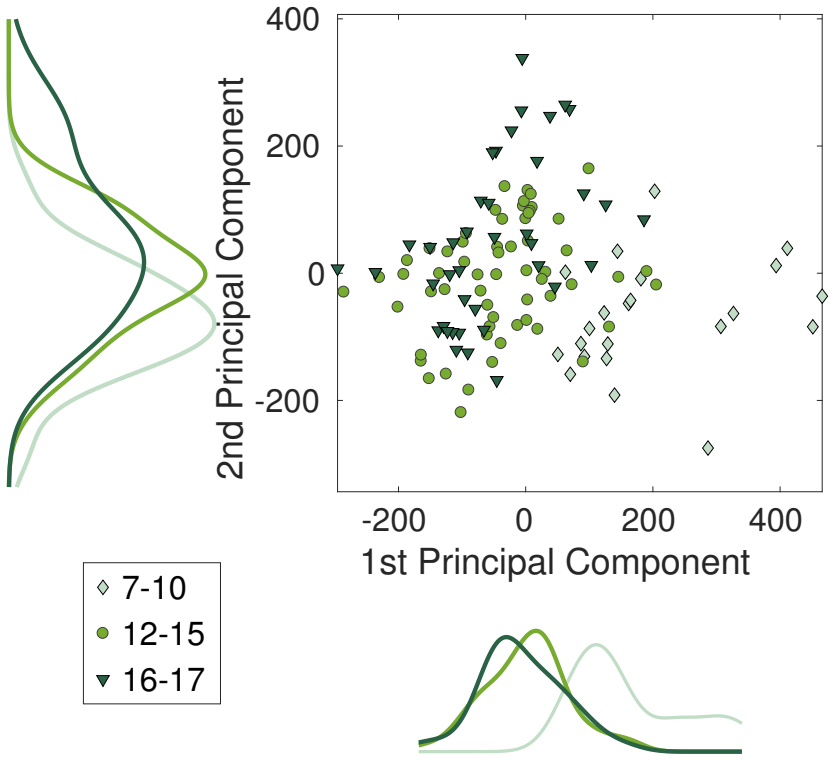

Figure 4.12: Model "All". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on age.

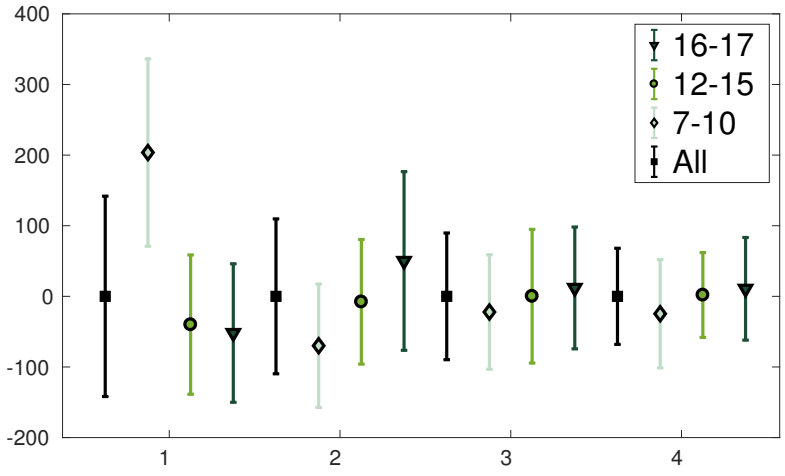

Figure 4.13: Model "All". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on age.

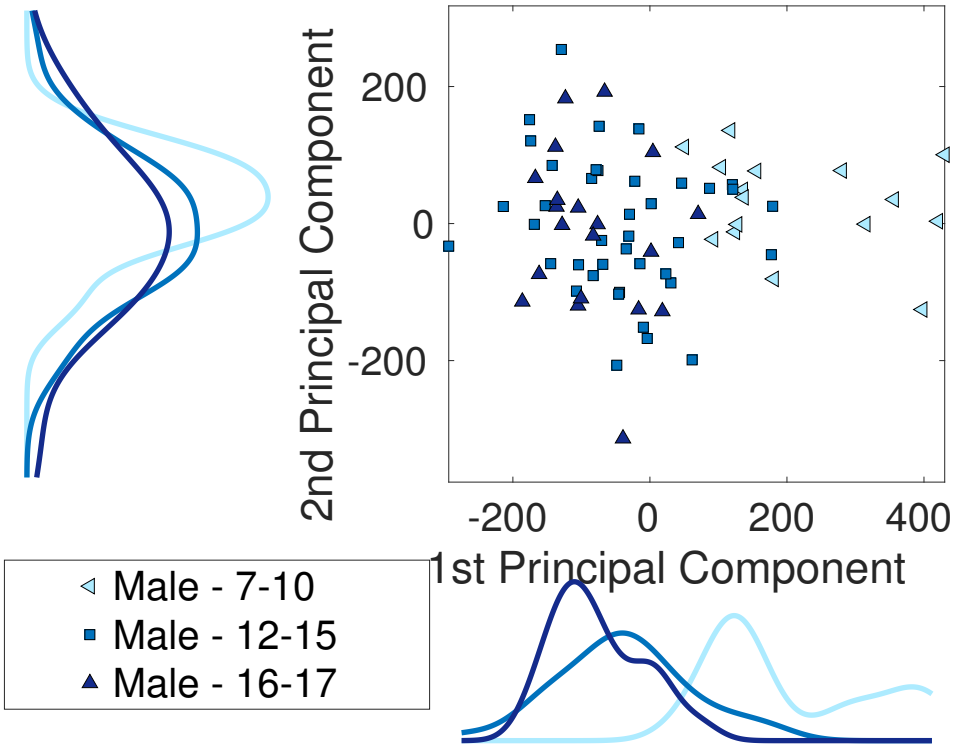

Figure 4.14: Model "All Male". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on age.

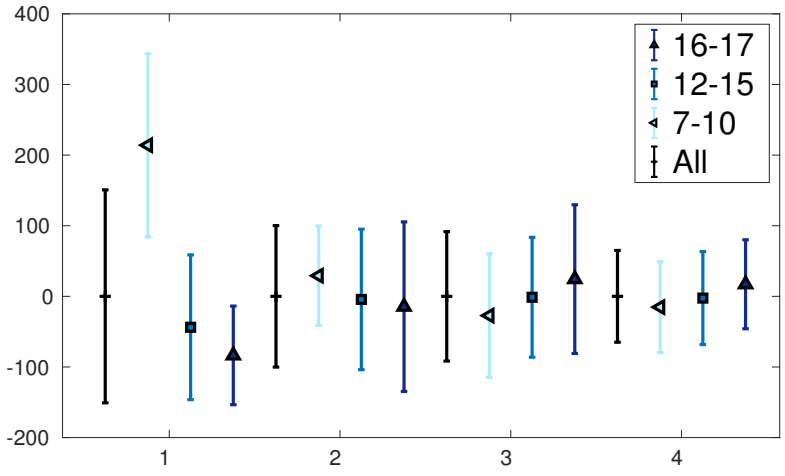

Figure 4.15: Model "All Male". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on age.

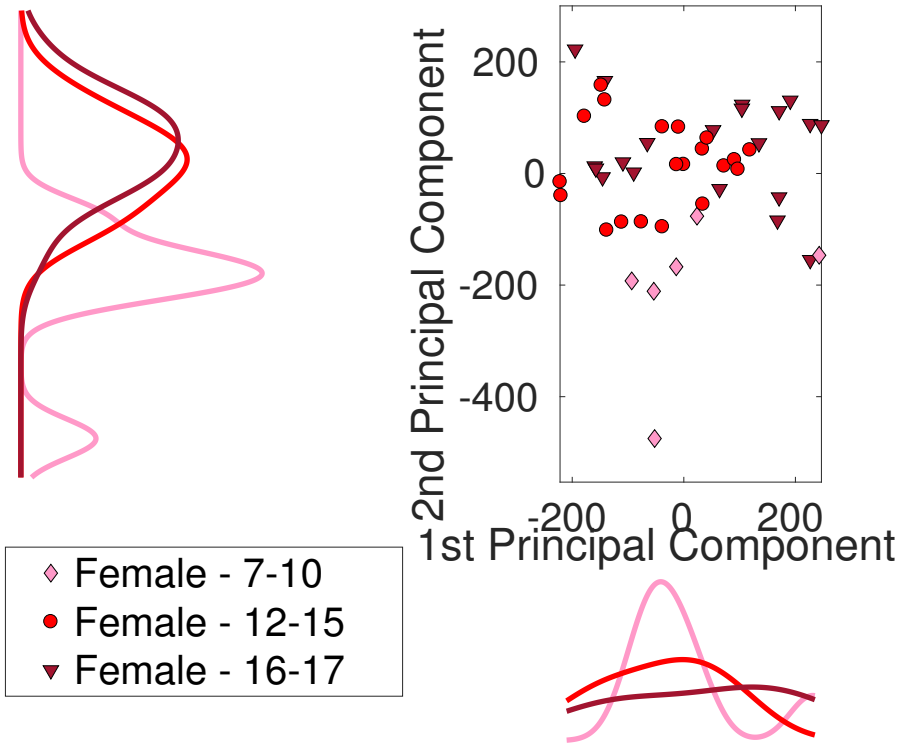

Figure 4.16: Model "All Female". The distributions of scores for all objects in the model for the two first principal components. The scores are split in subgroups based on age.

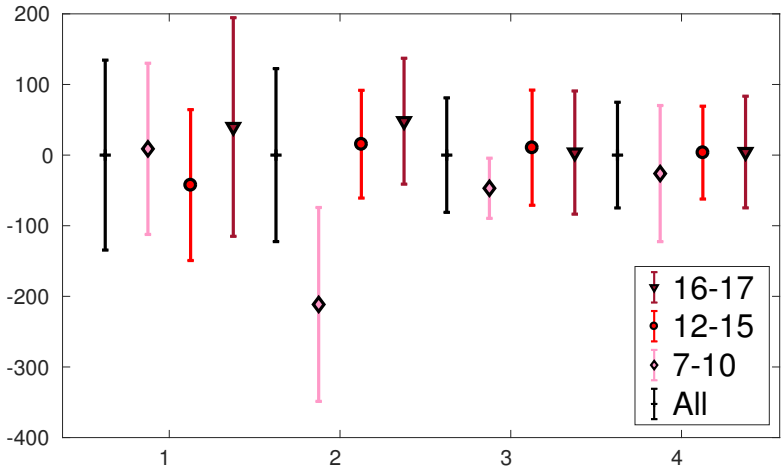

Figure 4.17: Model "All Female". The mean and standard deviation for the first 4 principal modes. Both for all scores in the model and split in subgroups based on age.

5 Discussion

This project aimed to construct SSMs for pediatric patients. Such models were created. The models helped identify when hip development occurs in pediatric patients. Significant anatomical differences were found between the three age groups defined. No significant gender based differences were found.

5.1 Template Mesh Registration Accuracy

To determine how well the template mesh registered to the initial meshes the distances between the mesh pairs were calculated (Table 4.1). The mean maximum distance between the initial mesh and the registered template mesh was on the order of 2 mm. That value can be compared to the size of the pixels in the CT images, which ranged from 0.5 to 3 mm. The largest differences between the meshes were equivalent to a maximum of 4 pixels. The largest maximum distance were found among the oldest patients. Older patients generally have larger femurs than younger patients, thus making the relative error smaller. However, in most cases mesh differences (RMS distances) were smaller than one pixel. The largest mean RMS distance was found for the youngest patient group. This can be explained by the fact that the CT scans' resolution was sometimes lower for this patient group. The registered mesh has a smoother surface than the low resolution segmentations making differences between them greater. Another aspect that contributed to this was that the registered template was the femur from an adult patient. Therefore it was likely to differ more from younger patients (see Figure 5.1 for a comparison), than from older patients. The registration error can be considered small, and the registered template meshes representative for the femurs' anatomies.

In addition to the distance differences being small, they were evenly distributed over the mesh surface (e.g. Figure 4.2). Had the differences been focused to only one part of the anatomy, it might have been problematic, but this was not the case.

5.2 Statistical Shape Models

The number of principal modes that explained 95% of the model variance were chosen as reduced models. Comparison of a registered volume mesh and the same mesh reconstructed using a reduced model gave mean RMS distances between the mesh surfaces slightly smaller than 0.5 mm (Table 4.4). That was smaller than one pixel from the CT-images. The reconstruction errors can be considered small enough to study the anatomical development of the hip of pediatric patients. The lengths of the patient femurs used in this project were ranging from 7 cm, for the youngest patients, to 12 cm, for older patients. Even for the shortest femurs, reconstruction errors of this magnitude are insignificant when parameters such as the CCD-angle, LCE-angle and Klein's line (see Figure 1.1) are measured. Hence the number of principal modes chosen were enough for this purpose.

The model that contained the smallest patient group was the model "7-10 y". The model contained 11 patients, 22 femurs. Such a small number of patients did definitely not contain the whole normality range for the age group. This model needed half of the principal modes to explain 95% of the model variation (Table 4.3). Since left and right side femurs from the

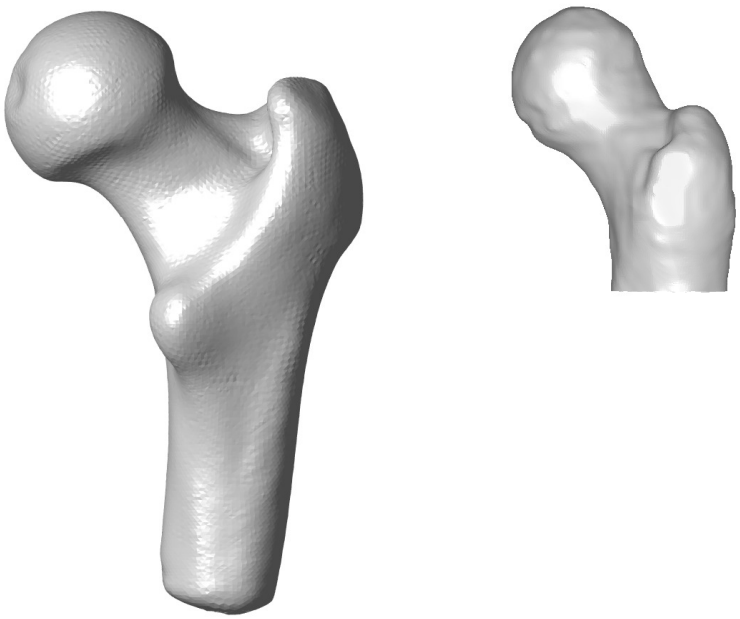

Figure 5.1: A qualitative comparison of the anatomical difference between the template mesh (left) and the anatomy of a patient from the youngest age group (right). The meshes are to scale.

same patient are likely anatomically similar, the large number of modes indicated that the model did not include all possible variation expected for the age group. A larger number of patients would be preferable when creating the SSM.

The largest model, "All" (that contained all patients), was created from 61 patients, 122 femurs. At least this amount of patients would have been preferable to have in all subgroups, when creating the SSMs.

The model that needed the largest number of principal modes to explain at least 95% of the model variation used 23 principal modes (Table 4.3). Depending on the purpose of the model, the number of principal modes might be too large to be computationally feasible. For this project the number of principal modes did not pose a problem. Therefore no further tests were done to see if reconstructions would have been sufficiently accurate for a smaller amount of principal modes.

The femur anatomies were described by volume meshes. The only information needed for the creation of the SSMs was the shape of the femur surface. It would have been enough to construct the SSMs using surface meshes. The volume nodes could introduce uncertainty in the model. E.g. if the volume nodes would be distorted in a way making their distribution denser in one part of the femur than another, this would be captured by the SSM principal modes. It would have no real impact on the femur anatomy.

5.3 Statistical Analysis

The Kolmogorov-Smirnov test showed that all patient subgroups were normally distributed. The smallest subgroup (females, 7-10 years, Table 3.1) only had 6 samples. With such a small sample set it is generally difficult to confirm, or rule out, normal distribution.

For model "All", the t-test comparing the age groups showed a significant difference between the youngest patients (7-10 years) and the two older age groups (12-15 and 16-17 years) in principal mode 1 and a significant difference between all of the age groups in principal mode 2 (Table 3.1). This indicates that the femoral anatomy develops as children grow older.

T-tests found no significant difference between male and female patients, in any of the models. When comparing the models "All Male" and "All Female" a significant difference was found between the youngest patients (7-10 years) and the two older age groups (12-15 years and 16-17 years). This indicates that the development of the femoral anatomy of children primarily relates to age and not gender.

For the models "7-10 y" and "All Female" tests were done on the smallest patient group (females, 7-10 years). With such a small sample size (n=6) significant differences are challenging to find, and it may well be that a larger patient group would be needed to confirm the outcomes.

The variation in size of the patients' femurs was removed when the SSMs were created. Therefore the differences that were found between age groups were not due to the size of femurs. The tests showed that there was a significant difference between the anatomies of the femur of pediatric patients of different ages, independent on the femoral size. There was no significant difference between the femurs of female and male patients.

5.4 Future Work

Interpretation of SSMs can be challenging and the anatomies related to variations of the SSMs can be difficult to define. The most common shape variations are not directly linked to the parameters commonly used to quantify hip morphology, such as the CCD-angle, LCE-angle and Klein's line (Figure 1.1). A way to help interpret the models would be to measure anatomical parameters on the SSMs. All parameter values expected in a population could be measured on femurs constructed by the models. Parameters could be defined for each SSM. Assuming nodes retain their anatomical placement during variation of principal modes the parameters could easily be measured for all model variations.This way of parameter measurement promotes the use of volume meshes, since the volume nodes provide more possibilities for defining parameters.

Many of the anatomical parameters used to quantify the anatomy of the hip require information about both the femur and the acetabulum, the concave surface of the pelvis that meets the femur. One such angle is the LCE-angle (Figure 1.1). Construction of a SSM of the acetabulum in relation to the femur would make it possible to measure more parameters.

To create a SSM capable of representing the whole normality range of the proximal femur for pediatric patients the training objects need to contain all normal variations of the anatomy. A larger number of patients would be preferable when constructing SSMs. Espe-

cially for the model "7-10 y" containing only 11 patients, 22 femurs (Table 4.3). Since CT imaging of pediatric patients is avoided when possible, especially for the youngest patients, finding suitable scans to use could pose some problem. Here scans from 61 patients have been used. However, permission is granted for use of 200 CT scans, and appropriate scans are currently collected. With the implemented pipeline it should be easy to create more femur volume meshes and create new SSMs including more patients.

The SSMs need to be validated in order to establish how many femurs are actually needed to explain all the expected variations within a population. To validate how well the models generalize to new femur anatomies a leave-one-out approach could be implemented. This means constructing models using all but one training object and then measure how well the model can generate the training object that was left out. The leave-one-out procedure would be repeated for all training object. This would give a measure of how well the model generates new versions of femurs from pediatric patients with normal anatomy. Evaluation of the models are necessary to conclude how well they represent the whole normality range for the anatomies. Unfortunately there was not enough time to validate the models as part of this project.

A way of further developing the SSMs would be to add the thickness of the cortical (outer, denser) bone of the femur. Since the CT scans used for model constructions were not calibrated, it was not known how the gray scale of the images correspond to bone mineral density. Had the CT scans been calibrated BMD values could have been added to the volume nodes of the SSMs.

However the cortical bone is significantly denser than the trabecular (inner, less dense) bone and could be segmented, and the thickness could be described by the SSMs. Including the cortical thickness in the SSM could be interesting, to study how it develops during growth.

SSMs of the femur can create accurate 3D reconstructions, from a single or a pair of radiographs, accounting for the patient position and rotation in the radiograph (Väänänen et al., 2015). Reconstructions using the models could produce the 3D morphology of the hip, avoiding the radiation dosage imposed by CT imaging. Accurate 3D models allow a more reliable quantification of the hip geometry than only the 2D projections of the hip. This application of the models is a coming step in the larger study that this MSc project is part of but was not within the scope if this project.

6 Summary and Conclusion

SSMs of the proximal part of the femur for pediatric patients were created. The SSMs were created from volume meshes of the femurs of pediatric patients. The meshes were created from segmented CT scans. The differences between the meshes and the segmentations were on the same order of magnitude as the CT scan resolution. Therefore they were deemed representative for patient anatomies. The SSMs, containing the anatomical variations for the pediatric patients, were examined to identify when femoral development occurs. Patients were compared based on age and gender. The variation in femoral size was removed from the models.

Significant anatomical differences between patients belonging to three different age groups (7 - 10, 12 - 15 and 16 - 17 years) were found. The differences were independent of femur size. No significant differences were found between female and male patients.

The creation of the SSMs was the first step of a study that could help identifying normal anatomical development of the hip and patients deviating from the normal. The pipeline for SSM creation allows for more patients to be added to models. Creating models able to explain even more of the normal variations in femurs during growth may eventually help the surgeon with clinical decision making.

References

R. Agricola, M. P. Heijboer, S. M. A. Bierma-Zeinstra, J. A. N. Verhaar, H. Weinans, and J. H. Waarsing. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study. *Annals of the Rheumatic Diseases*, 72(6):918–923, 2013.

H. Almohiy. Paediatric computed tomography radiation dose: A review of the global dilemma. *World journal of radiology*, 6(1):1–6, 2014.

P. E. Beaulé, K. Hynes, G. Parker, and K. A. Kemp. Can the alpha angle assessment of cam impingement predict acetabular cartilage delamination? *Clinical Orthopaedics and Related Research*, 470(12):3361–3367, 2012.

F. L. Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

CIBC, 2016. Seg3D: Volumetric Image Segmentation and Visualization. Scientific Computing and Imaging Institute (SCI), Download from: http://www.seg3d.org.

J. C. Clohisy, J. C. Carlisle, R. Trousdale, Y. J. Kim, P. E. Beaule, P. Morgan, K. Steger-May, P. L. Schoenecker, and M. Millis. Radiographic evaluation of the hip has limited reliability. *Clinical Orthopaedics and Related Research*, 467(3), 2009.

T. Cootes, E. Baldock, and J. Graham. *Image Processing and Analysis*. Ed. R. Baldock and J. Graham, Oxford University Press, 2000.

J. Damilakis, J. E. Adams, G. Guglielmi, and T. M. Link. Radiation exposure in X-ray-based imaging techniques used in osteoporosis. *European Radiology*, 20(11):2707–2714, 2010.

J. C. Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, 1975.

J. E. Jackson. *User's Guide to Principal Component Analysis*. John Wiley & Sons, Inc., 1991.

Manu, 2017. nonrigidICP downloaded from MATLAB Central File Exchange: https://se.mathworks.com/matlabcentral/fileexchange/41396-nonrigidicp.

A. E. Nelson, J. L. Stiller, X. A. Shi, K. M. Leyland, J. B. Renner, T. A. Schwartz, N. K. Arden, and J. M. Jordan. Measures of Hip Morphology are Related to Development of Worsening Radiographic Hip Osteoarthritis Over 6 to 13 Year Follow-Up: The Johnston County Osteoarthritis Project. *Osteoarthritis and Cartilage*, 24(3):443–450, 2016.

M. S. Pearce, J. A. Salotti, M. P. Little, K. McHugh, C. Lee, K. P. Kim, N. L. Howe, C. M. Ronckers, P. Rajaraman, A. W. Craft, L. Parker, and A. B. De González. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: A retrospective cohort study. *The Lancet*, 380(9840):499–505, 2012.

S. Rusinkiewicz and M. Levoy. Efficient Variants of the ICP Algorithm. *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pages 145–152, 2001.

M. Siebelt, R. Agricola, H. Weinans, and Y. J. Kim. The role of imaging in early hip OA. *Osteoarthritis and Cartilage*, 22(10):1470–1480, 2014.

E. Törnell and A. Wallin. 3D-mätningar på friska barnhöfter påvisar anatomiska skillnader beroende av kön och ålder. In H. Isaksson and T. Jansson, editors, *Proceedings of Clinical Innovation*, volume 4. Department of Biomedical Engineering Faculty of Engineering, Lund University, 2017.

S. P. Väänänen, L. Grassi, G. Flivik, J. S. Jurvelin, and H. Isaksson. Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. *Medical Image Analysis*, 24(1):125–134, 2015.

T. A. L. Wren, X. Liu, P. Pitukcheewanont, and V. Gilsanz. Bone densitometry in pediatric populations: Discrepancies in the diagnosis of osteoporosis by DXA and CT. *Journal of Pediatrics*, 146(6):776–779, 2005.

C. Zacharias, A. M. Alessio, R. K. Otto, R. S. Iyer, G. S. Philips, J. O. Swanson, and M. M. Thapa. Pediatric CT: Strategies to Lower Radiation Dose. *American Journal of Roentgenology*, 200(5):148–155, 2013.

A Results from Statistical Shape Models

This appendix contains visualizations of the variations associated with the first two principal modes for each of the SSMs created in this project. The visualizations are done for modes cores $pc_1 = \{-2SD(pc_1), 0, 2SD(pc_1)\}$ and $pc_2 = \{-2SD(pc_2), 0, 2SD(pc_2)\}$ for the two principal modes respectively (Figures A.1-A.6). These images can be seen as a compliment to figures 4.4, 4.6, 4.8, 4.10, 4.12, 4.14 and 4.16, containing the score values for the two first principal modes for all objects in each of the models.

The Appendix also contains tables with the mean and standard deviations for all principal modes needed to describe at least 95% of the model variation for each model. (Tables A.1 - A.6.)

A.1 Model "All"

Mode	All	7-10 y	12-15 y	16-17 y	All Male	All Female
1	0.0 ± 141.8	203.6 ± 132.6	-40.0 ± 98.6	-52.0 ± 98.1	17.8 ± 153.5	-29.4 ± 115.9
2	0.0 ± 109.7	-69.9 ± 87.4	-7.8 ± 88.2	50.1 ± 126.5	-14.9 ± 96.2	24.5 ± 126.1
3	0.0 ± 89.7	-22.2 ± 81.1	0.2 ± 94.7	11.9 ± 86.3	-0.3 ± 98.9	0.5 ± 72.9
4	0.0 ± 68.0	-24.7 ± 76.6	1.9 ± 60.1	10.7 ± 72.7	2.6 ± 64.7	-4.3 ± 73.7
5	0.0 ± 52.2	-15.4 ± 63.7	-1.6 ± 41.4	10.8 ± 58.4	14.1 ± 42.6	-23.4 ± 58.2
6	0.0 ± 48.2	10.3 ± 56.6	3.3 ± 50.9	-10.6 ± 37.1	1.2 ± 49.5	-2.0 ± 46.5
7	0.0 ± 39.6	33.7 ± 42.6	-7.8 ± 33.9	-6.8 ± 36.8	2.5 ± 38.5	-4.1 ± 41.3
8	0.0 ± 35.1	2.9 ± 52.0	2.0 ± 30.4	-4.6 ± 30.5	3.6 ± 36.3	-5.9 ± 32.5
9	0.0 ± 30.1	6.0 ± 36.4	-10.6 ± 27.0	12.7 ± 25.2	-1.0 ± 30.7	1.6 ± 29.3
10	0.0 ± 29.5	-0.8 ± 41.6	3.0 ± 23.9	-4.1 ± 29.4	10.7 ± 20.3	-17.6 ± 33.8
11	0.0 ± 26.3	6.2 ± 20.5	-6.6 ± 29.3	6.5 ± 22.2	-2.7 ± 26.3	4.4 ± 25.9
12	0.0 ± 23.5	2.0 ± 32.1	1.0 ± 19.8	-2.7 ± 23.6	1.0 ± 22.7	-1.7 ± 24.9
13	0.0 ± 20.8	5.2 ± 25.6	-1.1 ± 20.2	-1.3 ± 18.9	1.5 ± 21.1	-2.5 ± 20.4
14	0.0 ± 20.5	-0.3 ± 31.2	2.0 ± 12.2	-2.9 ± 23.1	-0.6 ± 16.9	1.0 ± 25.4
15	0.0 ± 18.5	1.1 ± 25.2	-1.5 ± 15.7	1.6 ± 18.6	-2.8 ± 17.8	4.7 ± 18.9
16	0.0 ± 17.9	-2.6 ± 21.3	-1.2 ± 17.9	3.2 ± 15.9	-1.8 ± 18.9	3.0 ± 15.9
17	0.0 ± 16.2	3.5 ± 23.2	-1.4 ± 14.2	0.1 ± 14.5	-0.6 ± 17.1	1.0 ± 14.8
18	0.0 ± 16.2	4.8 ± 18.5	-4.7 ± 11.4	4.4 ± 19.2	-0.8 ± 17.1	1.3 ± 14.5
19	0.0 ± 15.0	2.4 ± 17.1	-2.4 ± 13.6	2.3 ± 15.6	-1.3 ± 14.8	2.1 ± 15.3
20	0.0 ± 14.3	-0.1 ± 17.0	-0.0 ± 14.2	0.1 ± 13.3	-1.9 ± 14.3	3.2 ± 14.0
21	0.0 ± 13.9	-1.7 ± 13.4	-0.3 ± 14.2	1.3 ± 13.9	0.4 ± 13.6	-0.6 ± 14.5
22	0.0 ± 13.1	-0.5 ± 18.4	-0.2 ± 13.0	0.6 ± 9.6	0.6 ± 13.3	-1.0 ± 12.8
23	0.0 ± 12.6	-0.6 ± 14.2	-1.2 ± 13.0	2.2 ± 11.2	-0.3 ± 12.8	0.5 ± 12.4

Table A.1: Model "All". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

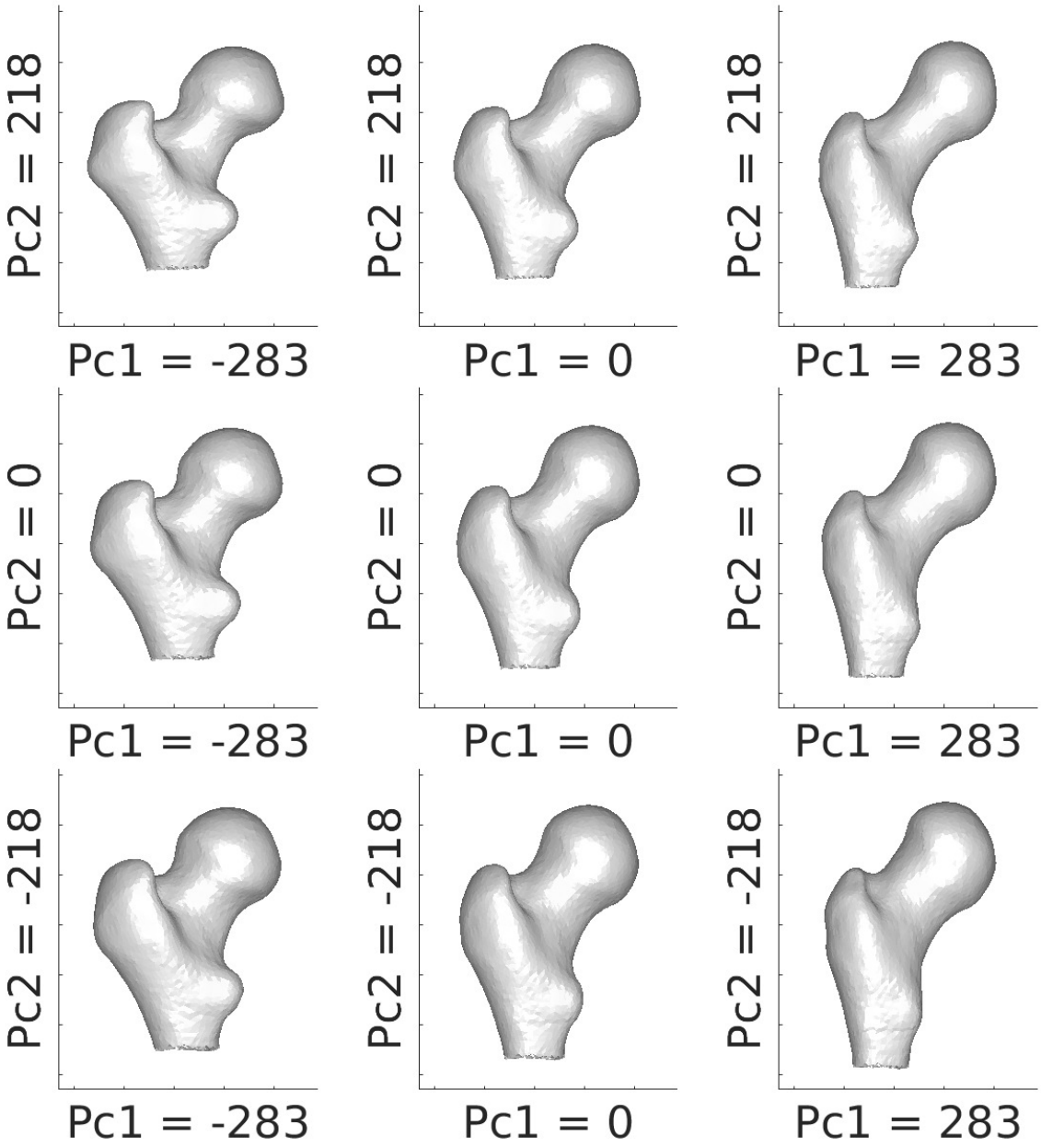

Figure A.1: Model "All". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.

A.2 Model "7-10 y"

Mode	All	Male	Female
1	0.0 ± 120.9	18.2 ± 122.3	-48.5 ± 112.6
2	0.0 ± 99.6	5.3 ± 101.6	-14.0 ± 101.9
3	0.0 ± 81.5	-0.4 ± 74.7	1.1 ± 105.5
4	0.0 ± 76.3	3.1 ± 55.8	-8.4 ± 122.4
5	0.0 ± 55.8	-1.6 ± 59.6	4.3 ± 48.9
6	0.0 ± 53.9	11.7 ± 50.0	-31.1 ± 55.7
7	0.0 ± 40.4	-0.7 ± 40.9	1.9 ± 42.6
8	0.0 ± 39.4	11.3 ± 39.9	-30.2 ± 15.7
9	0.0 ± 32.3	8.6 ± 31.9	-22.9 ± 22.0
10	0.0 ± 30.6	-7.2 ± 31.7	19.1 ± 18.2
11	0.0 ± 26.1	3.3 ± 24.0	-8.7 ± 31.8
12	0.0 ± 23.4	-2.3 ± 24.5	6.1 ± 20.9

Table A.2: Model "7-10 y". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

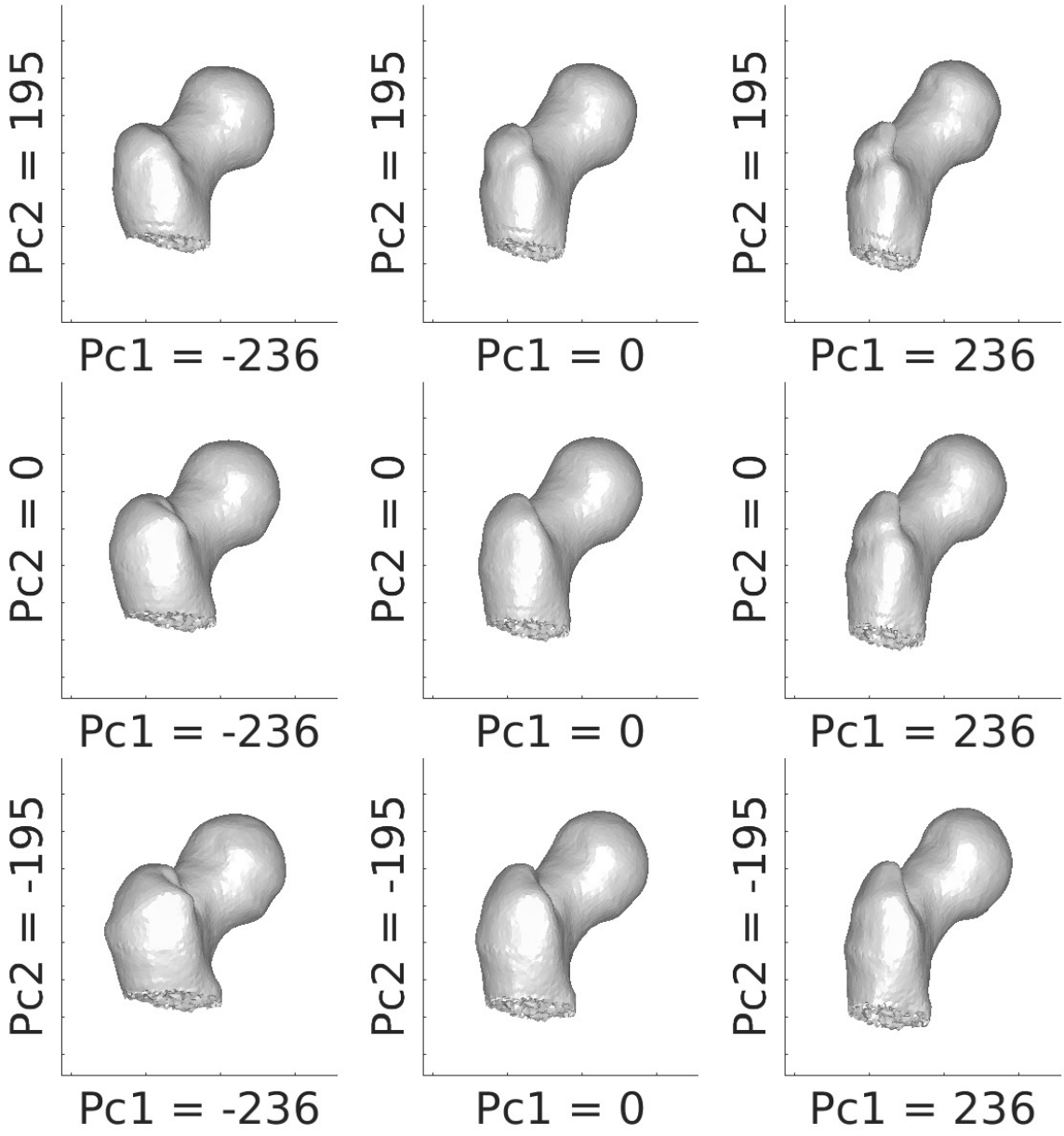

Figure A.2: Model "7-10 y". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.

A.3 Model "12-15 y"

Mode	All	Male	Female
1	0.0 ± 117.2	10.6 ± 124.1	-21.2 ± 101.9
2	0.0 ± 111.4	-4.3 ± 119.6	8.6 ± 95.1
3	0.0 ± 101.6	-8.8 ± 96.0	17.6 ± 112.5
4	0.0 ± 75.9	-3.8 ± 82.0	7.5 ± 63.0
5	0.0 ± 59.4	-10.0 ± 58.2	19.9 ± 57.9
6	0.0 ± 51.1	6.0 ± 52.4	-12.0 ± 47.2
7	0.0 ± 42.3	-5.9 ± 42.9	11.8 ± 39.5
8	0.0 ± 36.2	-11.1 ± 27.5	22.2 ± 41.8
9	0.0 ± 31.9	-0.3 ± 32.7	0.6 ± 31.0
10	0.0 ± 30.7	7.5 ± 32.0	-15.0 ± 21.6
11	0.0 ± 25.9	-0.6 ± 25.9	1.2 ± 26.4
12	0.0 ± 23.9	4.9 ± 22.7	-9.8 ± 24.0
13	0.0 ± 22.6	4.1 ± 21.8	-8.2 ± 22.4
14	0.0 ± 21.2	0.4 ± 19.8	-0.9 ± 24.2
15	0.0 ± 20.2	-1.5 ± 21.6	3.1 ± 17.2
16	0.0 ± 19.7	1.5 ± 20.5	-2.9 ± 18.3
17	0.0 ± 17.9	0.2 ± 18.3	-0.5 ± 17.6
18	0.0 ± 17.2	0.8 ± 19.0	-1.6 ± 13.3
19	0.0 ± 16.6	1.1 ± 16.2	-2.1 ± 17.5
20	0.0 ± 16.2	-0.5 ± 16.7	1.0 ± 15.6
21	0.0 ± 15.2	-0.6 ± 15.0	1.3 ± 15.8

Table A.3: Model "12-15 y". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

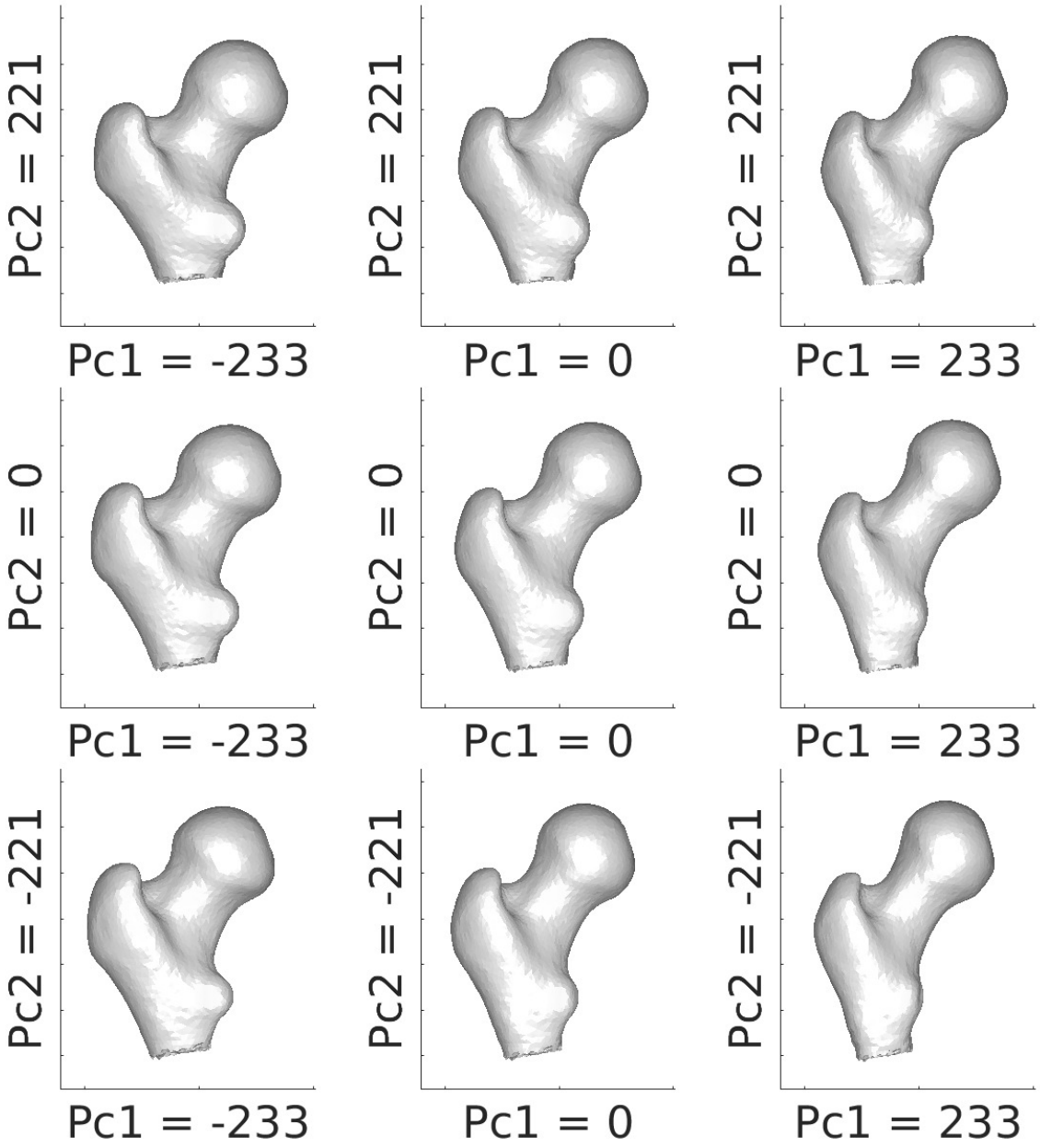

Figure A.3: Model "12-15 y". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.

A.4 Model "16-17 y"

Mode	All	Male	Female
1	0.0 ± 146.7	34.5 ± 130.6	-34.5 ± 156.8
2	0.0 ± 117.6	-17.2 ± 115.7	17.2 ± 119.9
3	0.0 ± 88.4	2.0 ± 95.0	-2.0 ± 83.7
4	0.0 ± 75.6	7.1 ± 59.8	-7.1 ± 89.7
5	0.0 ± 54.7	11.7 ± 24.7	-11.7 ± 72.4
6	0.0 ± 53.4	21.2 ± 46.2	-21.2 ± 52.7
7	0.0 ± 40.4	25.6 ± 33.6	-25.6 ± 29.0
8	0.0 ± 35.7	-8.2 ± 28.0	8.2 ± 41.2
9	0.0 ± 33.6	-1.1 ± 31.3	1.1 ± 36.5
10	0.0 ± 31.8	-0.3 ± 34.6	0.3 ± 29.6
11	0.0 ± 25.8	2.5 ± 25.1	-2.5 ± 26.9
12	0.0 ± 25.1	1.4 ± 27.2	-1.4 ± 23.5
13	0.0 ± 22.2	0.7 ± 25.2	-0.7 ± 19.4
14	0.0 ± 21.3	-1.4 ± 24.2	1.4 ± 18.5

Table A.4: Model "16-17 y". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

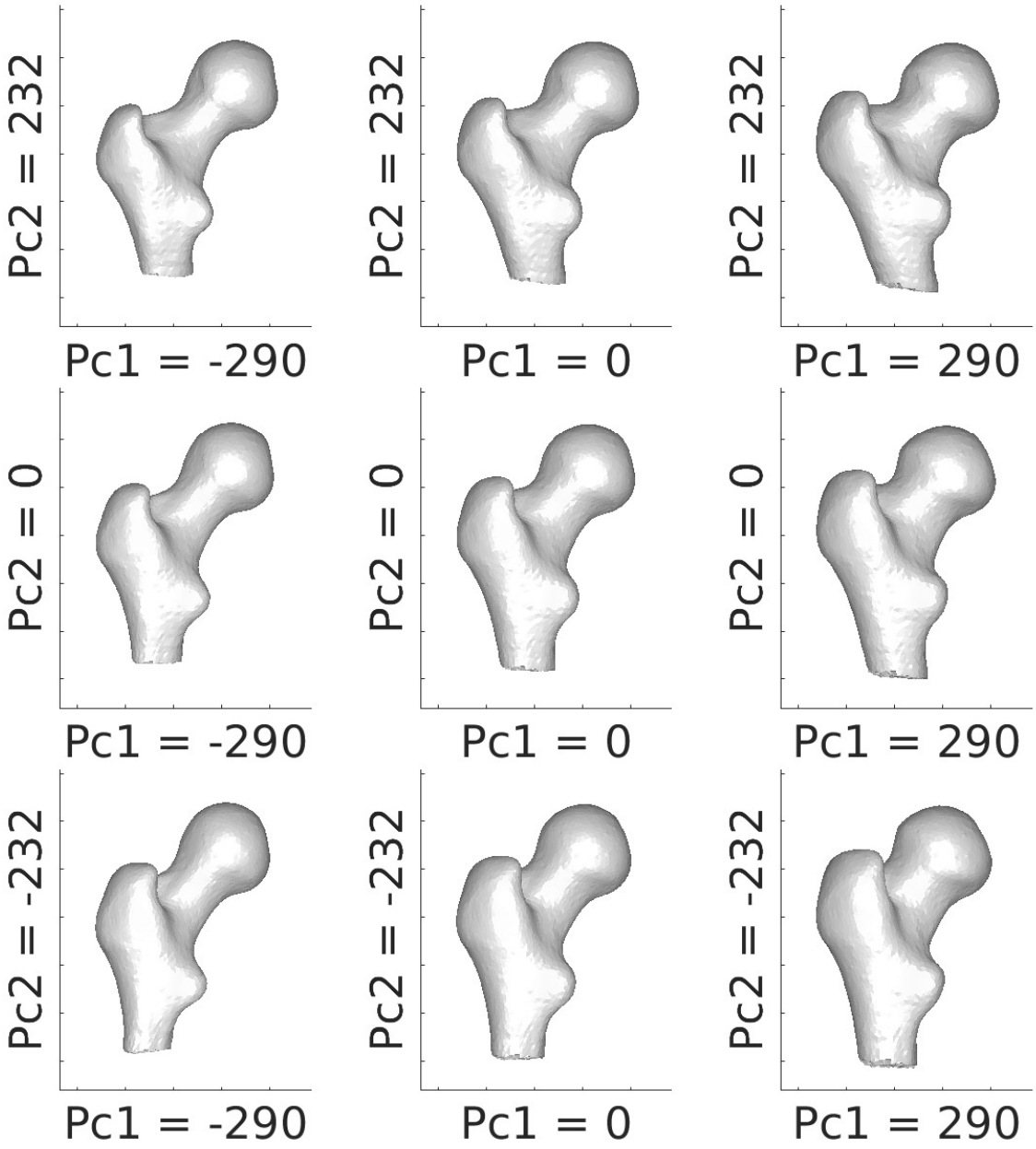

Figure A.4: Model "16-17 y". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.

A.5 Model "All Male"

Mode	All Male	7-10 y	12-15 y	16-17 y
1	0.0 ± 150.8	214.0 ± 129.9	-43.8 ± 102.5	-83.6 ± 69.8
2	0.0 ± 100.1	29.2 ± 70.3	-4.4 ± 99.4	-14.6 ± 120.0
3	0.0 ± 91.7	-27.2 ± 87.5	-1.3 ± 84.9	24.4 ± 105.3
4	0.0 ± 65.0	-15.3 ± 64.5	-2.4 ± 65.9	17.1 ± 63.0
5	0.0 ± 47.4	-9.5 ± 64.6	3.7 ± 45.8	0.1 ± 33.8
6	0.0 ± 44.6	-17.9 ± 51.9	-9.1 ± 35.2	32.6 ± 40.4
7	0.0 ± 39.0	26.4 ± 44.0	-13.4 ± 37.6	5.6 ± 24.4
8	0.0 ± 36.9	1.3 ± 47.9	-8.3 ± 34.6	15.5 ± 26.9
9	0.0 ± 29.4	-9.5 ± 41.5	8.8 ± 26.5	-10.0 ± 16.4
10	0.0 ± 26.8	2.3 ± 28.0	-0.8 ± 29.6	-0.2 ± 20.2
11	0.0 ± 22.6	-7.9 ± 25.7	4.3 ± 23.4	-2.3 ± 16.7
12	0.0 ± 21.6	6.1 ± 31.0	-1.6 ± 14.9	-1.7 ± 24.1
13	0.0 ± 21.3	-0.6 ± 29.6	-3.5 ± 17.5	7.5 ± 19.5
14	0.0 ± 20.7	-0.7 ± 31.0	1.4 ± 14.6	-2.2 ± 22.0
15	0.0 ± 18.0	-1.3 ± 25.8	0.1 ± 17.1	0.8 ± 12.3
16	0.0 ± 17.6	-1.3 ± 21.9	1.8 ± 16.5	-2.6 ± 16.3
17	0.0 ± 16.8	-1.8 ± 17.8	0.0 ± 17.9	1.4 ± 14.1
18	0.0 ± 16.0	1.3 ± 15.9	-0.7 ± 17.2	0.4 ± 14.1
19	0.0 ± 15.1	1.6 ± 16.4	-3.0 ± 15.3	4.8 ± 12.9
20	0.0 ± 13.9	1.1 ± 16.6	-0.4 ± 13.5	-0.1 ± 13.0

Table A.5: Model "All Male". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

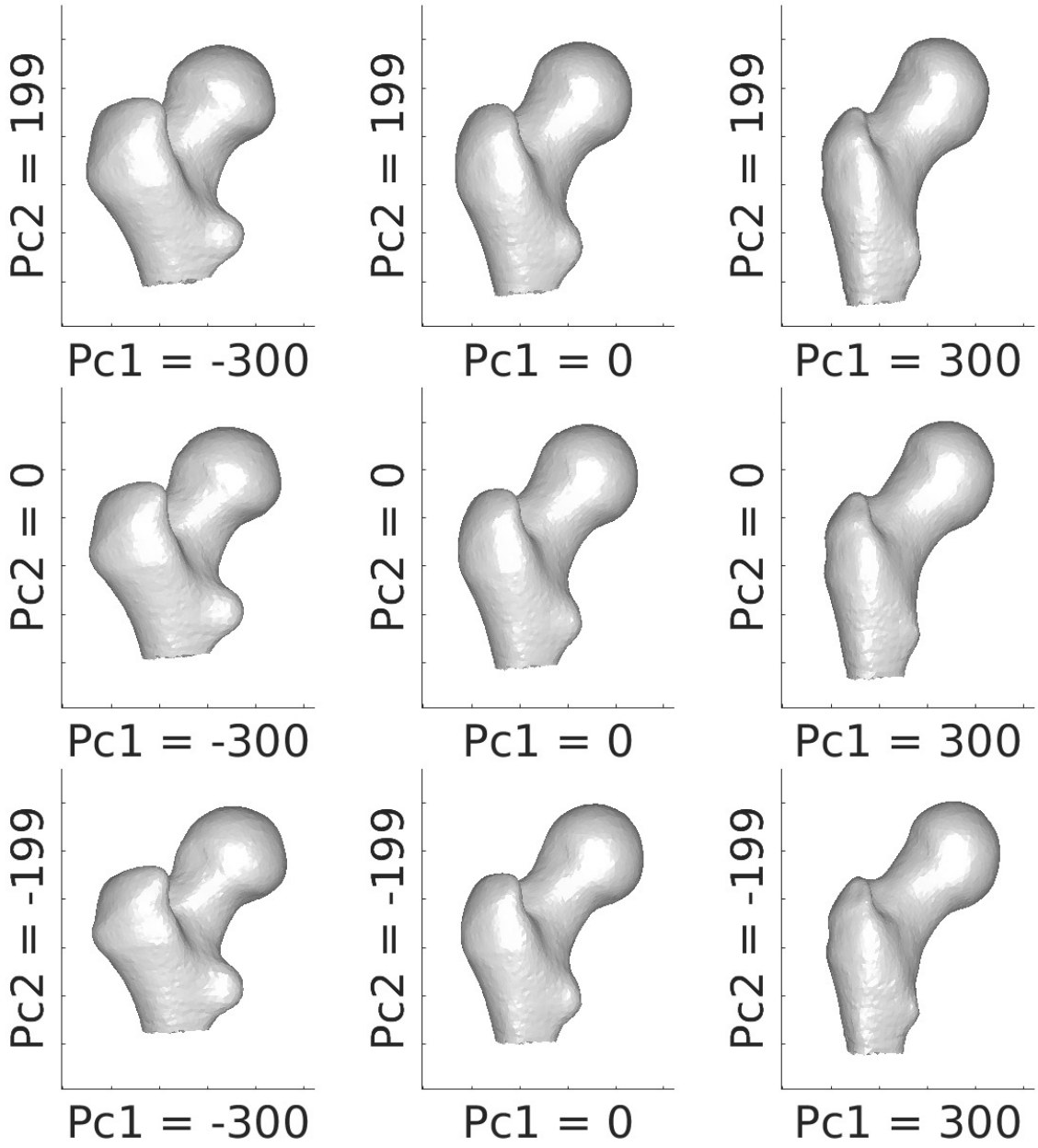

Figure A.5: Model "All Male". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.

A.6 Model "All Female"

Mode	All Female	7-10 y	12-15 y	16-17 y
1	0.0 ± 134.5	8.9 ± 121.2	-42.4 ± 106.8	39.8 ± 154.8
2	0.0 ± 122.5	-211.5 ± 137.3	15.4 ± 76.3	48.0 ± 89.1
3	0.0 ± 81.1	-47.0 ± 42.6	10.5 ± 81.5	3.6 ± 87.2
4	0.0 ± 74.8	-26.2 ± 96.4	3.5 ± 65.7	4.3 ± 79.1
5	0.0 ± 61.9	-10.4 ± 122.9	4.6 ± 36.1	-1.5 ± 61.1
6	0.0 ± 49.6	59.1 ± 31.7	-8.5 ± 49.1	-9.2 ± 43.3
7	0.0 ± 41.4	-37.6 ± 39.0	11.2 ± 34.0	0.0 ± 43.9
8	0.0 ± 40.2	7.9 ± 40.8	3.9 ± 43.4	-6.3 ± 37.8
9	0.0 ± 37.1	-5.7 ± 42.6	-2.6 ± 16.7	4.3 ± 49.7
10	0.0 ± 30.1	11.1 ± 47.9	-4.6 ± 31.7	1.3 ± 21.9
11	0.0 ± 26.9	-7.4 ± 33.6	-4.5 ± 24.7	6.8 ± 26.9
12	0.0 ± 25.3	-6.4 ± 40.0	0.9 ± 26.1	1.0 ± 19.8
13	0.0 ± 21.9	-3.6 ± 18.7	0.5 ± 25.5	0.6 ± 19.8
14	0.0 ± 19.7	-1.9 ± 9.9	-4.8 ± 16.8	5.4 ± 23.6
15	0.0 ± 19.2	1.2 ± 23.6	2.8 ± 22.5	-3.1 ± 14.2

Table A.6: Model "All Female". The $mean \pm SD$ for each principal mode needed to describe at least 95% of the model. The values are presented both for the full model and for all subgroups contained in the model.

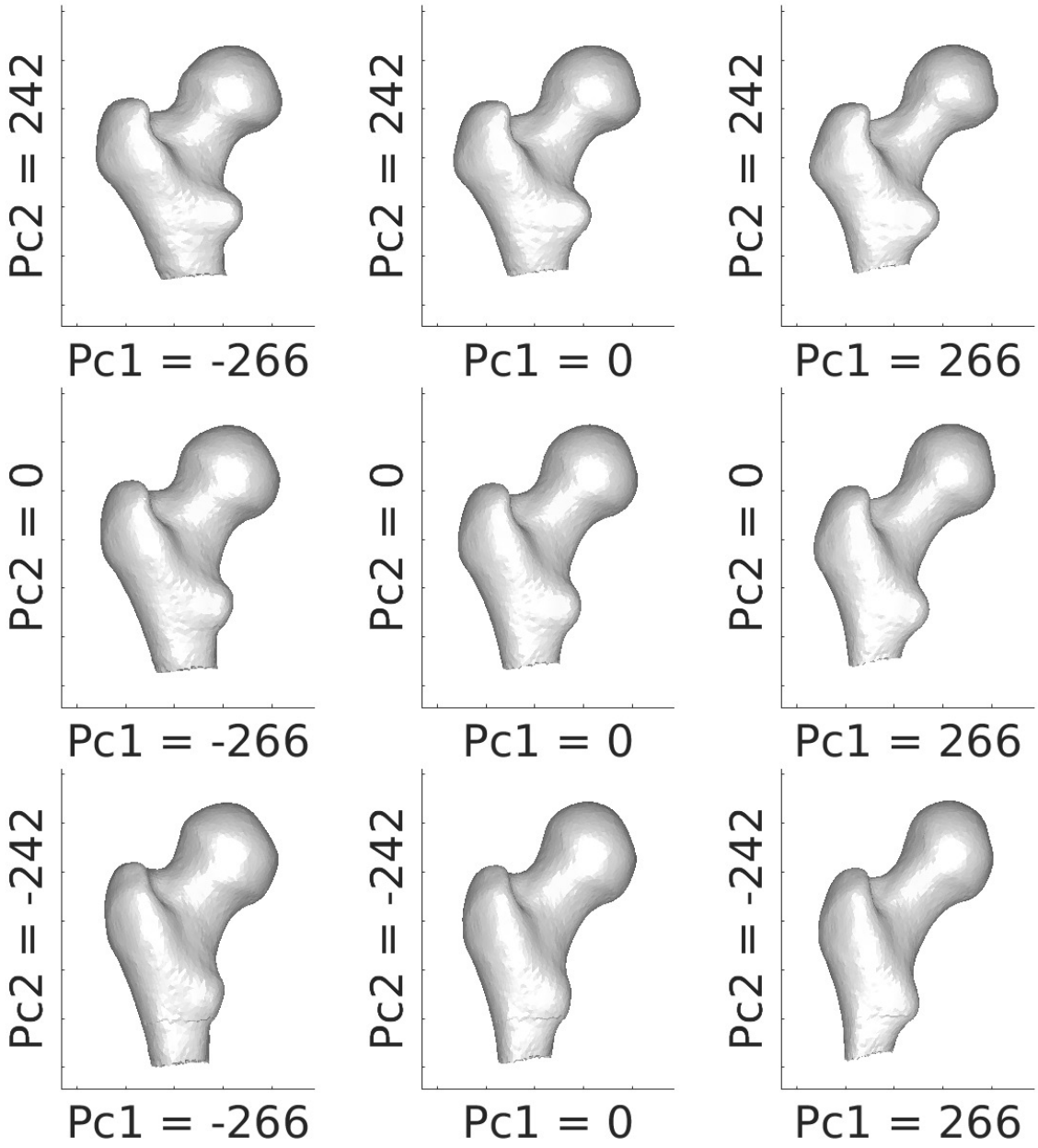

Figure A.6: Model "All Female". Each row of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 1. The score values are $\{-2SD(pc_1), 0, 2SD(pc_1)\}$. Each column of subfigures show reconstructions from the model for variation of the score corresponding to principal mode 2. The score values are $\{-2SD(pc_2), 0, 2SD(pc_2)\}$.