DATA NOTE

Using statistical shape modelling of DXA images to quantify the shape of the proximal femur at ages 14 and 18 years in the Avon Longitudinal Study of Parents and Children [version 1; referees: 1 approved with reservations]

Monika Frysz [1,2], Jenny S. Gregory [3], Richard M. Aspden [3], Lavinia Paternoster [1], Jonathan H. Tobias [2]

[1]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2]Musculoskeletal Research Unit, Translational Health Sciences, University of Bristol, Bristol, UK
[3]Institute of Medical Science, University of Aberdeen, Aberdeen, UK

v1 **First published:** 06 Feb 2019, **4**:24 (https://doi.org/10.12688/wellcomeopenres.15092.1)
Latest published: 06 Feb 2019, **4**:24 (https://doi.org/10.12688/wellcomeopenres.15092.1)

Abstract

Hip shape is an important determinant of hip osteoarthritis and osteoporotic hip fracture; however, little is known about its development in childhood and adolescence. While previous studies largely focused on individual geometrical indices of hip geometry such as neck-shaft angle or femoral neck width, statistical shape modelling offers the means to quantify the entire contour of the proximal femur, including lesser trochanter and acetabular eyebrow. We describe the derivation of independent modes of variation (hip shape mode scores) to characterise variation in hip shape from dual-energy X-ray absorptiometry (DXA) images in the Avon Longitudinal Study of Parents and Children (ALSPAC) offspring, using statistical shape modelling. ALSPAC is a rich source of phenotypic and genotypic data which provides a unique opportunity to investigate the environmental and genetic influences on hip shape in adolescence, as well as comparison with adult hip shape.

Keywords

ALSPAC, hip shape, joint shape, statistical shape modelling

Open Peer Review

Referee Status: ?

	Invited Referees 1
version 1 published 06 Feb 2019	? report

1 **Carme Rissech**, University of the Balearic Islands, Spain

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Avon Longitudinal Study of Parents and Children (ALSPAC) gateway.

Corresponding author: Monika Frysz (monika.frysz@bristol.ac.uk)

Author roles: Frysz M: Data Curation, Formal Analysis, Funding Acquisition, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Gregory JS**: Conceptualization, Data Curation, Methodology, Software, Supervision, Writing – Review & Editing; **Aspden RM**: Conceptualization, Methodology, Software, Writing – Review & Editing; **Paternoster L**: Conceptualization, Methodology, Supervision, Writing – Review & Editing; **Tobias JH**: Conceptualization, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through a PhD Studentship to MF [105504] and the ALSPAC core programme grant [102215]. The UK Medical Research Council and Wellcome [102215] and the University of Bristol provide core support for ALSPAC. LP works in a unit that receives support from the UK Medical Research Council and the University of Bristol [MC_ UU_12013/4 & MC_UU_12013/5].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Frysz M, Gregory JS, Aspden RM *et al.* **Using statistical shape modelling of DXA images to quantify the shape of the proximal femur at ages 14 and 18 years in the Avon Longitudinal Study of Parents and Children [version 1; referees: 1 approved with reservations]** Wellcome Open Research 2019, **4**:24 (https://doi.org/10.12688/wellcomeopenres.15092.1)

First published: 06 Feb 2019, **4**:24 (https://doi.org/10.12688/wellcomeopenres.15092.1)

Introduction

Osteoarthritis (OA) and osteoporotic fractures are the most common age-related musculoskeletal diseases and are associated with significant healthcare burden. Previous studies suggest that hip shape is an important risk factor for both hip OA[1,2] and osteoporotic hip fracture[3]. Little is known, however, about its development in childhood and adolescence. Traditionally, hip shape architecture has been assessed by measuring lengths and angles. However, it has been recognized that single geometrical measurements are often correlated with measures of body size as well as other geometrical indices[4]. Statistical shape modelling is a method which uses a set of landmark points to describe an outline of an object. It provides the means for capturing global shape of an object as opposed to a single geometrical measurement and can represent a combination of several different aspects of proximal femur shape (e.g. variation in femoral neck (FN) along with variation in femoral head).

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a longitudinal birth cohort, which in the 1990s recruited pregnant women in South West England[5]. ALSPAC is a rich source of data, including phenotypic and genetic data collected for the mothers, fathers and children. It is uniquely suited for examining variation in hip shape in earlier life, based on hip dual-energy X-ray absorptiometry (DXA) scans obtained when the children were, on average, 14 and 18 years old. This data note describes the methodology and data used to quantify the shape of the proximal femur in ALSPAC offspring at these time points. In order to aid comparability with other studies and between the time points, an adult reference statistical shape model (SSM) template (based on 19,379 images) was applied to these data.

Methods

ALSPAC Data

ALSPAC is a longitudinal birth cohort which recruited a total of 14,541 pregnant women with expected delivery date between 1st April 1991 and 31st December 1992. Of these pregnancies, 69 have no known birth outcome, and of the remaining 14,472 pregnancies, 195 were twin, 3 were triplet and 1 was quadruplet accounting for 14,676 known foetuses. These pregnancies resulted in 14,062 live births, of which 13,988 children were alive at 1 year of age.

In addition to the initial enrolment that took place between 1991 and 1992, further recruitment took place when the children were, on average, 7 years old, and another from age 8 onwards to which eligible children and those not initially enrolled were also invited. This resulted in a total of 15,247 pregnancies enrolled. Since recruitment these children have been followed up at regular intervals; questionnaire and clinical assessment data have been collected. Moreover, additional data on siblings, mothers and their partners, have also been collected.

Hip DXA scans

Hip DXA scans collected during two assessment clinics, Teen Focus (TF) 2 and TF 4, were used to quantify the shape of proximal femur. TF 2 was performed between January 2005 and September 2006. The target age for attendance was 13.5 years (mean age at attendance was 13.8 years, range 12.5–15.1 years). TF 4 clinic started in December 2008 and was completed by early to mid-2011. The target age for attendance was 17.5 years (mean age at attendance was 17.8 years, range 16.2–19.8 years).

Of 11,351 individuals invited to the TF 2 clinic, 6,147 attended and a total of 6,162 images were available to align in Shape software (please note that for quality purposes a number of individuals were re-invited and duplicate scans were performed), of which 4,468 were available for SSM. Of 10,101 individuals invited to the TF 4 clinic, 5,217 attended and 4,746 images were available to align in Shape, of which 4,413 were available for final modelling. For details regarding image exclusion please refer to Table 1.

Statistical shape model (SSM)

Raw hip DXA images were securely transferred to collaborators in Aberdeen for image processing and uploaded into Shape software (University of Aberdeen). Each image was marked up with a set of landmark points (please refer to Figure 1, which shows the placement of landmark points, and Table 2, which describes the anatomical positions of each of the key landmark points (shown in red in Figure 1)).

Following point placement, Procrustes analysis was used to estimate the mean shape. The aim of this step is, first, to remove any translational, rotational and scaling information and then align each image as closely as possible. After completing the alignment, principal component analysis (PCA) was performed using the coordinates of each point to build the SSM, producing a set of orthogonal modes of variation known as principal components (referred to as hip shape modes (HSMs)). These modes together explain 100% of variance in the data set, with the first HSM accounting for the largest amount of variance and subsequent HSMs accounting for less variance. Each HSM has a mean of zero and unit standard deviation (SD), and each image

Table 1. Avon Longitudinal Study of Parents and Children offspring hip shape data.

	Age 14	Age 18
Description	**N**	
Total number of images uploaded in shape	6,162	4,746
Excluded twins, sibs and re-invites	171	115
Excluded images without genetic or TF4 data	1,255*	NA
Excluded images due to poor image quality	268	218
Total hips aligned	**4,468**	**4,413**
Of those, with genetic data	3,929	3,198
Of those, with data at both adolescent time points	3,188	

*Due to delay in image acquisition and given the time constrains, halfway through image alignment it was decided to restrict alignment of the remaining images to those who had both, genetic data and DXA image acquired at TF 4 clinic.

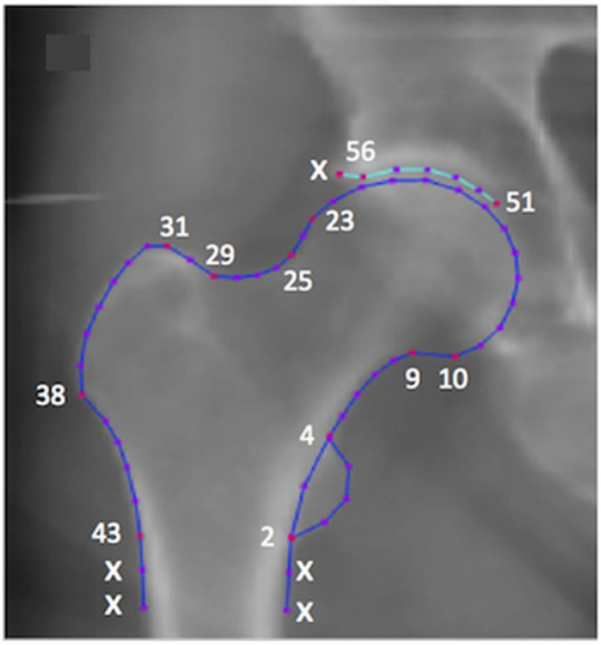

Figure 1. Outline of proximal femur shape and key landmark point positions used to derive 53-point SSM. Please note points 0, 1, 44, 45, and 57 [marked with x] were not included in the final model.

and, consequently, each individual is assigned a set of values for each HSM which describes the number of SDs away from the mean shape.

Applying external adult reference SSM template to adolescent data

One of the limitations of statistical shape modelling is the lack of comparability of HSMs with other datasets and studies, since each SSM is unique to that particular set of images. One way of overcoming this limitation is to apply a set of pre-defined HSMs, previously obtained from a reference population. An SSM template based on a reference set generated from a GWAS meta-analysis of hip shape from five cohorts (based on 19,379 images)[6], was applied to both adolescent datasets in order to directly compare hip shape between adolescent time points as well as with adult hip shape. See Table 3 for details regarding cohorts contributing to the adult reference SSM. Briefly, the reference model was built as described above and the eigenvectors were saved and used to calculate the mode scores for subsequent models (without adding the new image to the reference model or changing it in any way).

Reproducibility of point placement

A set of 100 images, collected during TF 4 clinic, were randomly selected and marked 2 months after completing the initial point placement in ALSPAC adolescents. The same set of images was also marked by a second marker. Intra- (within-) and inter-observer (between-observer) repeatability of manual point placement was measured as the difference in pixels between coordinates of 58 points. The intra- and inter-observer reliability assessed by mean point-to-point repeatability was 1.22 and 1.78 pixels, respectively. A cut off median point-to point difference of less than or equal to 3 was previously considered as accurate[7]. Whilst the initial model was based on a 58-point model, this was subsequently modified to a 53-point model due to high variability in points placed at the acetabular overhang and medial and lateral femoral shaft, in both adolescent and adult SSM templates.

Dataset

The first ten HSM scores generated for adolescent data collected at ages 14 and 18 years, using external adult reference SSM, are available in the ALSPAC resource. Similarly to previously

Table 2. Description of the key landmark points shown in red in Figure 1.

Point number	Anatomical feature
2	Medial femoral shaft meets inferior lesser trochanter (often maps to point 46, depending on position)
4	Medial femoral shaft meets superior lesser trochanter
9	Change in curvature: lateral inferior curvature of femoral head at point where it meets femoral neck
10	Change in curvature: medial inferior curvature of the femoral head
23	Change in curvature: superior lateral femoral head curvature
25	Change in curvature: inferior lateral femoral head where meets the superior femoral neck
29	Inferior greater trochanter slope where it meets superior femoral neck
31	Medial superior greater trochanter
38	Inferior lateral greater trochanter
43	Lateral femoral shaft
46	Inferior lesser trochanter (often maps to point 2)
51	Acetabular eyebrow medial end (end of brightest line)
56	Acetabular eyebrow lateral end

Table 3. Cohorts contributing to the adult reference statistical shape model.

Cohort	N	Gender	Mean age (SD) of participants
ALSPAC mothers	4,603	Females	47.9 (4.3)
Framingham	3,088	Males and females	63.3 (11.0)
MrOS	5,924	Males	74.0 (6.0)
SOF	1,715	Females	72.8 (4.6)
Twins UK	4,049	Males and females	52.5 (13.5)
Total	19,379		

ALSPAC, Avon Longitudinal Study of Parents and Children; MrOS, Osteoporotic fractures in men study; SOF, Study of Osteoporotic Fractures.

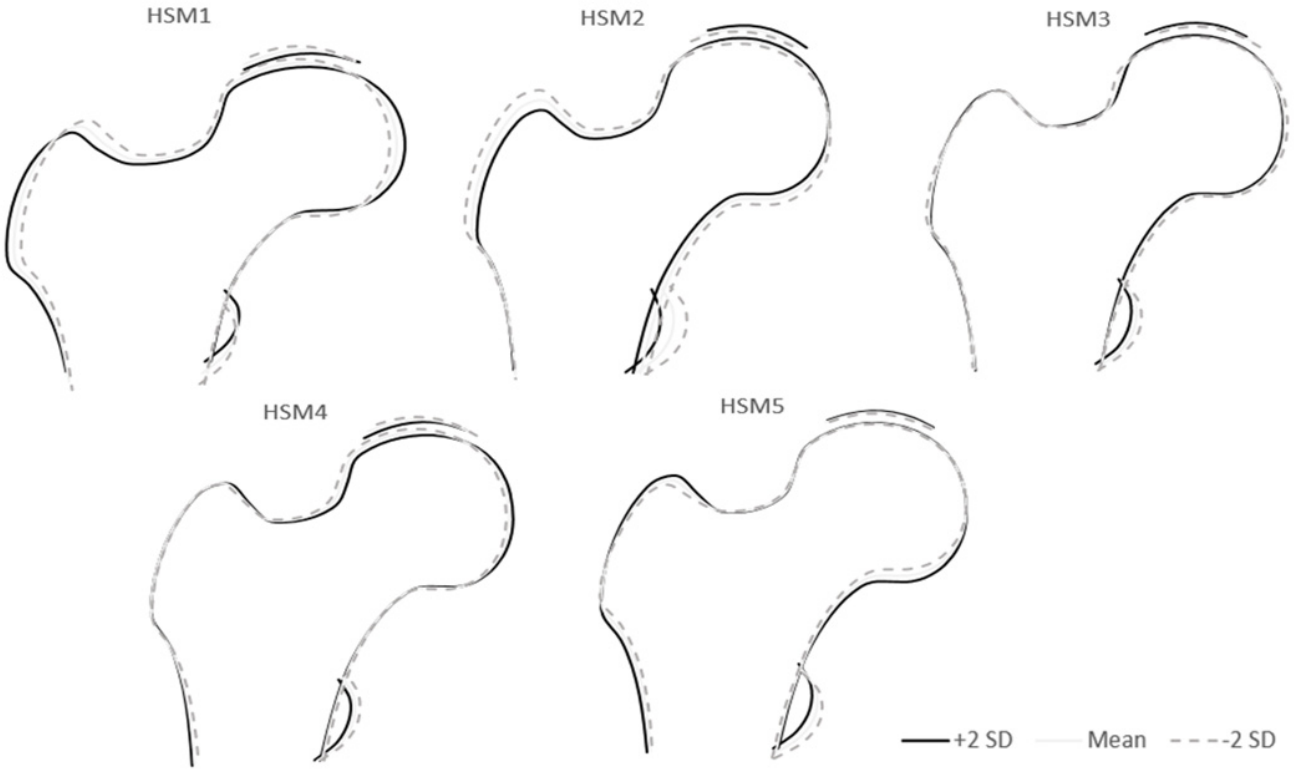

Figure 2. Variation in hip shape described by modes 1-5, based on adult reference SSM.

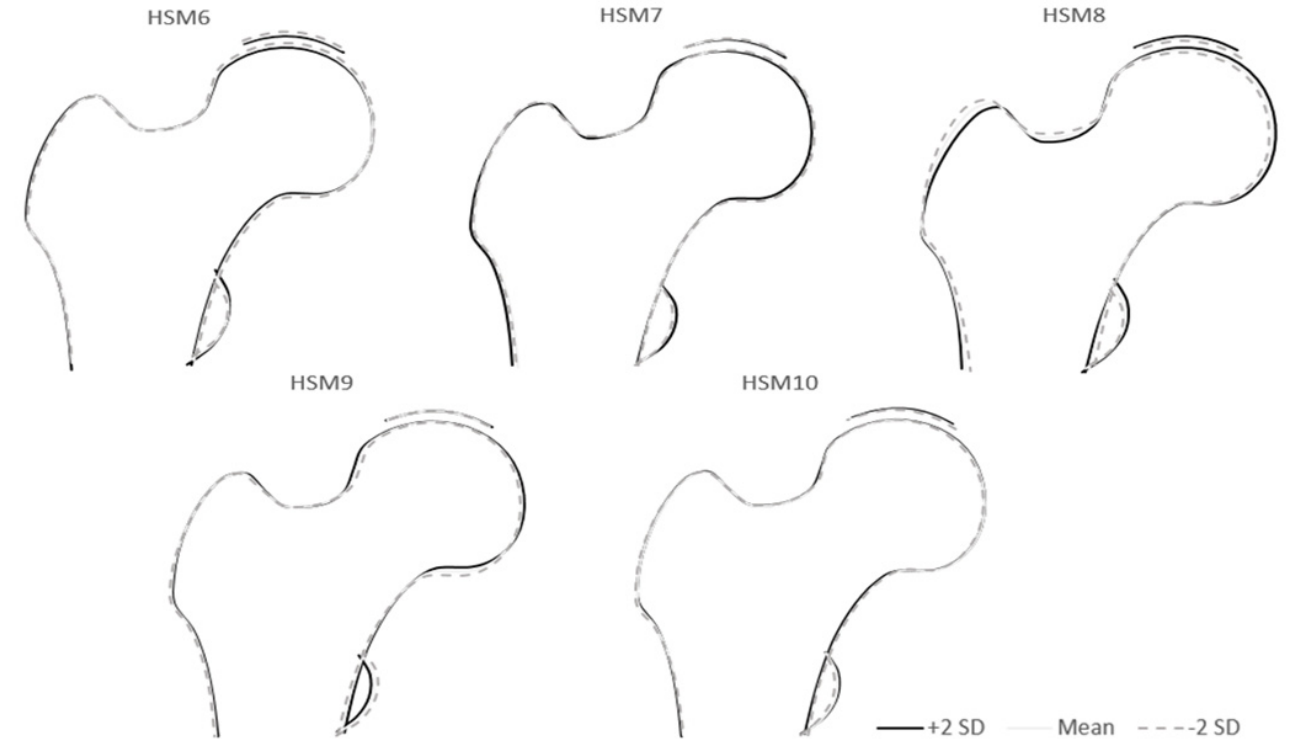

Figure 3. Variation in hip shape described by modes 6-10, based on adult reference SSM.

Table 4. Variation described by the top ten modes based on adult reference SSM. Please refer to Figure 2 and Figure 3 for graphical representation of each mode.

HSM (% of variation)	Key features described by each mode: +2 SDs (solid line) -2 SDs (dashed line)
1 (42%)	Positive scores (solid line) - Loss of femoral head curvature - Narrower FN Negative scores (dashed line) - Wider FN Smaller NSA
2 (13%)	Positive scores (solid line) - Narrower FN and femoral shaft - Smaller greater trochanter - Smaller femoral head (inferior aspect proximal to lesser trochanter) Negative scores (dashed line) - Wider FN - Larger greater and lesser trochanters
3 (8.5%)	Positive scores (solid line) - Smaller lesser trochanter - Narrower FN Negative scores (dashed line) - Wider FN - Larger lesser trochanter
4 (6.1%)	Positive scores (solid line) - Larger femoral head (medial aspect) - Narrower FN - Smaller lesser trochanter Negative scores (dashed line) - Cam-type deformity - Wider FN
5 (4.1%)	Positive scores (solid line) - Larger femoral head (inferior aspect proximal to lesser trochanter) - Larger greater trochanter - Wider FN Negative scores (dashed line) - Smaller femoral head (inferior aspect proximal to lesser trochanter) - Narrower FN - Larger lesser trochanter
6 (3.4%)	Positive scores (solid line) - Narrower FN Negative scores (dashed line) - Wider FN
7 (2.6%)	Positive scores (solid line) - Wider femoral shaft Negative scores (dashed line) - Narrower femoral shaft - Smaller lesser trochanter
8 (2.5%)	Positive scores (solid line) - Larger femoral head - Narrower FN - Smaller greater trochanter Negative scores (dashed line) - Smaller femoral head - Wider FN - Larger greater trochanter
9 (1.8%)	Positive scores (solid line) - Smaller femoral head (inferior aspect proximal to lesser trochanter) - Smaller lesser trochanter Negative scores (dashed line) - Larger femoral head (inferior aspect proximal to lesser trochanter) - Larger lesser trochanter
10 (1.5%)	Positive scores (solid line) - Larger lesser trochanter Negative scores (dashed line) - Smaller lesser trochanter

FN, femoral neck; NSA, neck-shaft angle.

published literature[7–9] the first 10 modes, which together explain 86% of variance in the adult SSM, were selected. Figure 2 and Figure 3 provide graphical representation and Table 4 provides summary of the features described by each HSM. Compared to mean = 0 and SD = 1 when using the data as its own reference, when using the adult reference SSM (based on adult data with age ranging from 48 to 74 years), means for the first ten HSMs ranged from -1.14 to 2.26 at age 14 and from -1.5 to 2.42 at age 18, whereas SDs ranged from 0.42 to 0.97 at age 14 and from 0.41 to 0.91 at age 18 (Table 5).

When the adult reference SSM was applied to ALSPAC mothers' images, means for HSMs 2–9 were close to 0 (ranging from -0.35 to 0.34) and SDs were close to 1 (ranging from 0.8 to 1), whereas mean and SD HSM1 score were 1.45 and 0.5, respectively.

The differences in means and SDs could be due to sex and/or age differences (i.e. mothers were on average 48 years old, therefore more closely resembling the ages of cohorts included in the reference model as opposed to ALSPAC offspring). The deviation away from the mean was particularly noted for HSM1, which is likely to reflect scanner differences between ALSPAC and other cohorts in the adult reference set. Different pixel spacing in the Lunar Prodigy scanner (used to acquire DXA scans in ALSPAC) relative to other scanners alters the aspect ratio (ratio between image height and width), and therefore HSM1 reflects these differences. An attempt was made to correct for these differences; however, some residual differences still remain.

Whilst direct comparison of the modes across the time points is an added advantage of applying an external reference SSM, one of the potential issues that may arise is that previously independent HSMs might no longer be independent of each other. In order to quantify the extent of the potential loss of independence, after applying SSM based on the combined adult reference model to adolescent data Matrix Spectral Decomposition was performed using the matSpD tool to compute the number of independent modes. The top ten HSMs based on adult reference SSM at both time points were first correlated (Table 6 and Table 7) and tested for independent number of variables (HSMs) using matSpD. The results showed close to 10 (9.6) independent variables for both time points suggesting that the loss of independence is unlikely to materially affect the results.

Table 5. Mean HSM scores for the top ten HSMs based on ALSPAC adolescent and mothers' images, after applying adult reference SSM (compared with mean=0 and SD=1 when data from each time point included as its own reference).

	Age 14	Age 18	Mothers
HSM	Mean (SD)	Mean (SD)	Mean (SD)
1	2.26 (0.42)	2.42 (0.41)	1.45 (0.53)
2	0.57 (0.76)	0.23 (0.85)	-0.01 (0.90)
3	-0.19 (0.68)	0.10 (0.66)	-0.31 (0.92)
4	0.87 (0.68)	0.36 (0.73)	0.32 (0.77)
5	-1.14 (0.79)	-1.50 (0.84)	-0.35 (0.94)
6	0.27 (0.68)	0.27 (0.86)	-0.01 (1.00)
7	-0.25 (0.63)	0.02 (0.70)	-0.14 (0.87)
8	0.39 (0.97)	0.02 (0.91)	0.06 (0.95)
9	0.22 (0.76)	-0.21 (0.91)	0.34 (0.95)
10	-1.09 (0.59)	-1.04 (0.77)	0.11 (0.92)

SSM methodology offers a powerful approach to study subtle changes in hip morphology and it has been successfully applied to study variation in hip shape associated with the incidence[10,11] and progression of OA[12], as well as associations with hip fracture[13] in adult cohorts. Whilst a major drawback of the methodology is that as each model is data-driven, the HSMs generated are unique to the sample used, thus the results cannot be directly cross-compared with other studies. One of the key strengths of hip shape data in ALSPAC offspring is the application of an adult reference SSM to hip DXA images at ages 14 and 18 years, which allows direct comparisons of associations with HSMs between these time points.

Ethical approval and consent

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees, full details of the approvals obtained are available from the study website (http://www.bristol.ac.uk/alspac/researchers/research-ethics/).

Written informed consent was obtained from parents, and children were invited to give consent where appropriate. Study members have the right to withdraw their consent for elements of the study or from the study entirely at any time.

Data availability

ALSPAC data access is through a system of managed open access. The steps below highlight how to apply for access to the data included in this data note and all other ALSPAC data. The dataset generated in this data note has been deposited within the ALSPAC data resource and is linked to ALSPAC project number B1274. Please quote this number to request required variables which have been described in this dataset (HSMs generated at ages 14 and 18 years).

1. Please read the ALSPAC access policy (PDF, 627kB) which describes the process of accessing the data and samples in detail, and outlines the costs associated with doing so.

2. You may also find it useful to browse our fully searchable research proposals database, which lists all research projects that have been approved since April 2011.

3. Please submit your research proposal for consideration by the ALSPAC Executive Committee using the online process. You will receive a response within 10 working days to advise you whether your proposal has been approved.

Table 6. Correlation matrix for the top ten HSM scores at age 14 to assess the number of independent variables using matrix Spectral Decomposition (matSpD) which showed strong evidence for nearly all variables (9.6) to be independent.

	HSM1	HSM2	HSM3	HSM4	HSM5	HSM6	HSM7	HSM8	HSM9	HSM10
HSM1	1	0.1853	0.0371	0.0375	0.4698	-0.198	0.1578	-0.272	-0.2019	-0.1227
HSM2	0.1853	1	0.4216	0.131	0.3872	0.0883	0.054	-0.118	0.3098	-0.1471
HSM3	0.0371	0.4216	1	0.2081	0.1451	-0.0381	0.1772	0.144	0.2597	-0.1564
HSM4	0.0375	0.131	0.2081	1	0.0924	-0.1778	0.248	0.1602	0.2208	-0.2277
HSM5	0.4698	0.3872	0.1451	0.0924	1	-0.2271	0.0095	-0.0648	0.3164	-0.0647
HSM6	-0.198	0.0883	-0.0381	-0.1778	-0.2271	1	-0.0972	0.1324	-0.2759	-0.0347
HSM7	0.1578	0.054	0.1772	0.248	0.0095	-0.0972	1	-0.3302	0.2572	0.0019
HSM8	-0.272	-0.118	0.144	0.1602	-0.0648	0.1324	-0.3302	1	-0.191	0.0862
HSM9	-0.2019	0.3098	0.2597	0.2208	0.3164	-0.2759	0.2572	-0.191	1	-0.1126
HSM10	-0.1227	-0.1471	-0.1564	-0.2277	-0.0647	-0.0347	0.0019	0.0862	-0.1126	1

Table 7. Correlation matrix for the top ten HSM scores at age 18 to assess the number of independent variables using matrix Spectral Decomposition (matSpD) which showed strong evidence for nearly all variables (9.6) to be independent.

	HSM1	HSM2	HSM3	HSM4	HSM5	HSM6	HSM7	HSM8	HSM9	HSM10
HSM1	1	0.141	0.2264	-0.0047	0.4621	-0.2515	0.0537	-0.1779	-0.1618	-0.0226
HSM2	0.141	1	0.3793	0.1983	0.4458	-0.1167	0.1083	-0.1985	0.3159	-0.0712
HSM3	0.2264	0.3793	1	0.4535	0.1827	-0.1872	0.3169	-0.0169	0.0756	-0.1303
HSM4	-0.0047	0.1983	0.4535	1	0.0864	-0.1524	0.1849	0.204	0.1695	-0.2213
HSM5	0.4621	0.4458	0.1827	0.0864	1	-0.3191	0.0347	-0.1862	0.4001	-0.0575
HSM6	-0.2515	-0.1167	-0.1872	-0.1524	-0.3191	1	-0.1257	0.1897	-0.3383	-0.0189
HSM7	0.0537	0.1083	0.3169	0.1849	0.0347	-0.1257	1	-0.1477	0.2756	0.1138
HSM8	-0.1779	-0.1985	-0.0169	0.204	-0.1862	0.1897	-0.1477	1	-0.1628	0.1194
HSM9	-0.1618	0.3159	0.0756	0.1695	0.4001	-0.3383	0.2756	-0.1628	1	-0.0967
HSM10	-0.0226	-0.0712	-0.1303	-0.2213	-0.0575	-0.0189	0.1138	0.1194	-0.0967	1

If you have any questions about accessing data, please email alspac-data@bristol.ac.uk.

The ALSPAC data management plan describes in detail the policy regarding data sharing, which is through a system of managed open access.

Grant information

This work was supported by the Wellcome Trust through a PhD Studentship to MF [105504] and the ALSPAC core programme grant [102215].

The UK Medical Research Council and Wellcome [102215] and the University of Bristol provide core support for ALSPAC. LP works in a unit that receives support from the UK Medical Research Council and the University of Bristol [MC_ UU_12013/4 & MC_UU_12013/5].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We are also grateful to Denis Baird for marking up images for repeatability testing.

References

1. Gregory JS, Waarsing JH, Day J, *et al.*: **Early identification of radiographic osteoarthritis of the hip using an active shape model to quantify changes in bone morphometric features: can hip shape tell us anything about the progression of osteoarthritis?** *Arthritis Rheum.* 2007; **56**(11): 3634–43.
PubMed Abstract | Publisher Full Text

2. Baker-LePain JC, Lane NE: **Relationship between joint shape and the development of osteoarthritis.** *Curr Opin Rheumatol.* 2010; **22**(5): 538–43.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Gregory JS, Testi D, Stewart A, *et al.*: **A method for assessment of the shape of the proximal femur and its relationship to osteoporotic hip fracture.** *Osteoporos Int.* 2004; **15**(1): 5–11.
PubMed Abstract | Publisher Full Text

4. Gregory JS, Aspden RM: **Femoral geometry as a risk factor for osteoporotic hip fracture in men and women.** *Med Eng Phys.* 2008; **30**(10): 1275–1286.
PubMed Abstract | Publisher Full Text

5. Boyd A, Golding J, Macleod J, *et al.*: **Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children.** *Int J Epidemiol.* 2013; **42**(1): 111–27.
PubMed Abstract | Publisher Full Text | Free Full Text

6. Baird DA, Evans DS, Kamanu FK, *et al.*: **Identification of Novel Loci Associated With Hip Shape: A Meta-Analysis of Genomewide Association Studies.** *J Bone Miner Res.* 2018.
PubMed Abstract | Publisher Full Text

7. Faber BG, Baird D, Gregson CL, *et al.*: **DXA-derived hip shape is related to osteoarthritis: findings from in the MrOS cohort.** *Osteoarthritis Cartilage.* 2017; **25**(12): 2031–2038.
PubMed Abstract | Publisher Full Text | Free Full Text

8. Pavlova AV, Saunders FR, Muthuri SG, *et al.*: **Statistical shape modelling of hip and lumbar spine morphology and their relationship in the MRC National Survey of Health and Development.** *J Anat.* 2017; **231**(2): 248–259.
PubMed Abstract | Publisher Full Text | Free Full Text

9. Baird DA, Paternoster L, Gregory JS, *et al.*: **Investigation of the Relationship Between Susceptibility Loci for Hip Osteoarthritis and Dual X-Ray Absorptiometry-Derived Hip Shape in a Population-Based Cohort of Perimenopausal Women.** *Arthritis Rheumatol.* 2018; **70**(12): 1984–1993.
PubMed Abstract | Publisher Full Text

10. Castaño-Betancourt MC, Rivadeneira F, Bierma-Zeinstra S, *et al.*: **Bone parameters across different types of hip osteoarthritis and their relationship to osteoporotic fracture risk.** *Arthritis Rheum.* 2013; **65**(3): 693–700.
PubMed Abstract | Publisher Full Text

11. An H, Marron JS, Schwartz TA, *et al.*: **Novel statistical methodology reveals that hip shape is associated with incident radiographic hip osteoarthritis among African American women.** *Osteoarthritis Cartilage.* 2016; **24**(4): 640–646.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Ahedi HG, Aspden RM, Blizzard LC, *et al.*: **Hip Shape as a Predictor of Osteoarthritis Progression in a Prospective Population Cohort.** *Arthritis Care Res (Hoboken).* 2017; **69**(10): 1566–1573.
PubMed Abstract | Publisher Full Text

13. Baker-LePain JC, Luker KR, Lynch JA, *et al.*: **Active shape modeling of the hip in the prediction of incident hip fracture.** *J Bone Miner Res.* 2011; **26**(3): 468–474.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Referee Status: ?

Version 1

Referee Report 08 March 2019

https://doi.org/10.21956/wellcomeopenres.16465.r34986

? **Carme Rissech**
Unit of Anatomy, Department of Medical Sciences, Faculty of Medicine, University of the Balearic Islands, Palma de Mallorca, Spain

From the title, it seems that the objective of the paper is to quantify the shape of the human proximal femur at ages 14 and 18 years by using the DXA scans from the Avon Longitudinal Study of Parents and Children. However, at the end of the introduction section, in the objectives, the authors say: “This data note describes the methodology and data used to quantify the shape of the proximal femur in ALSPAC offspring at these time points” (14 and 18 years of age). This lack of alignment between the title and objectives confuses the reader. In addition, the Introduction section starts talking about osteoarthritis and osteoporothic fractures, which confuses the reader even more. If the goal of this paper is to describe the methodology and data that in other studies the authors will use, the authors should focus on that. Therefore, the authors should clearly define their objectives and construct the title of the paper and the Introduction section according these objectives. Both of them (Title and Introduction) should lead the reader to the final objectives of the paper. I think that in addition to describing the methodology used to analyse the proximal shape of the femur by statistical shape modelling and the landmarks used, the authors should also explain how they will use all of this in their analyses. I wonder why they do not construct different models for boys and girls. It is well known that female and male femur each follow divergent growth trajectories which are clearly marked from 12 years of age onward (Pujol et al., 2016[1]). How are they going to use these ten 10 PCs on future papers? How will the application of the external adult reference statistical shape model template to adolescents and mothers aid comparability with other studies and ages? What results do they think can obtain from the application of these data in their future analyses? I think that all of this should be better explained and discussed in the paper.

Other comments:

- What is the difference between Geometric Morphometrics and Statistical shape modelling? This and the advantages to use Statistical shape modelling in front of Geometric Morphometrics should be explained in the Introduction section.
- What do key landmarks mean in Statistical shape modelling?
- In Material and Methods section, when the authors describe the final chosen sample of the Avon Longitudinal Study of Parents and Children used for their study, they should indicate the final number of boys and girls to be analysed.

References

1. Pujol A, Rissech C, Ventura J, Turbón D: Ontogeny of the male femur: Geometric morphometric

analysis applied to a contemporary Spanish population.*Am J Phys Anthropol*. 2016; **159** (1): 146-63
PubMed Abstract | Publisher Full Text

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Anatomy and Forensic Anthropology

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.