Significant morphological change in osteoarthritic hips identified over 6-12 months using Statistical Shape Modelling

Barr, Rebecca J; Gregory, Jennifer S.; Yoshida, Kanako; Alesci, Salvatore; Aspden, Richard M.; Reid, David M

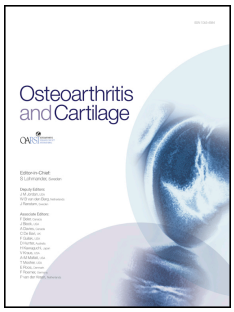
Published in:
Osteoarthritis and Cartilage

DOI:
10.1016/j.joca.2018.04.004

Publication date:
2018

Document Version
Peer reviewed version

Accepted Manuscript



Significant morphological change in osteoarthritic hips identified over 6-12 months using Statistical Shape Modelling

Rebecca J. Barr, Jennifer S. Gregory, Kanako Yoshida, Salvatore Alesci, Richard M. Aspden, David M. Reid

PII: S1063-4584(18)31146-4

DOI: 10.1016/j.joca.2018.04.004

Reference: YJOCA 4214

To appear in: *Osteoarthritis and Cartilage*

Received Date: 29 June 2017

Revised Date: 29 March 2018

Accepted Date: 7 April 2018

Please cite this article as: Barr RJ, Gregory JS, Yoshida K, Alesci S, Aspden RM, Reid DM, Significant morphological change in osteoarthritic hips identified over 6-12 months using Statistical Shape Modelling, *Osteoarthritis and Cartilage* (2018), doi: 10.1016/j.joca.2018.04.004.

Significant morphological change in osteoarthritic hips identified over 6-12 months using Statistical Shape Modelling.

Rebecca J Barr[1,2†], Jennifer S Gregory[1†], Kanako Yoshida[1], Salvatore Alesci[3], Richard M Aspden[1], David M Reid[1]

[†] both authors contributed equally to this paper

[1]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, UK,

Current affiliations: [2]Medicines Monitoring Unit (MEMO), Division of Molecular & Clinical Medicine, School of Medicine, University of Dundee; [3]Takeda Pharmaceuticals, Washington District of Columbia, USA.

Rebecca J Barr, r.y.barr@dundee.ac.uk

Jennifer S Gregory, j.gregory@abdn.ac.uk

Kanako Yoshida, kanako.yoshida@nhs.net

Salvatore Alesci, alescis@gmail.com

Richard M Aspden, r.aspden@abdn.ac.uk

David M Reid, d.m.reid@abdn.ac.uk

Corresponding Author
Professor R.M. Aspden
Institute of Medical Sciences
University of Aberdeen
Foresterhill
Aberdeen
AB25 2ZD
r.aspden@abdn.ac.uk
+44 1224 437445

Running title: **Shape changes in hip OA over 12 months**

Abstract

Objective Predicting who will develop osteoarthritis, assessing how rapidly their disease will progress and monitoring early responses to treatment are key to the development of therapeutic agents able to treat this crippling disease and to their future clinical use. Statistical Shape Modelling (SSM) enables quantification of variations in multiple geometric measures describing the whole hip joint to be considered in concert. This prospective study evaluates the responsiveness of SSM to changes in hip-shape within one year.

Methods Sixty-two people, mean age 67.1 yrs, were recruited. Dual-energy X-ray Absorptiometry images were taken at three timepoints (baseline, six and twelve months). Based on Kellgren-Lawrence grading (KLG)) of their baseline images, subjects were classified into control/doubtful OA: KLG<1 in both hips; moderate OA: KLG=2; and severe OA: KLG≥3 in their most severe hip. Morphology was quantified using SSM and changes in shape were assessed using generalised estimating equations. Standardized response means (SRM) were calculated for the first and second 6 month periods, then the full 12 months.

Results Disease severity ranged from KLG0-KLG4 in the 124 hips assessed at baseline. Three SSM modes (Modes 1, 3 and 4) were associated with OA severity. Across the whole cohort, SRM magnitudes ranged from 0.16 to 0.63. The greatest subgroup SRM (magnitude 0.91) was observed over 12 months in those subjects with moderate OA (KLG2).

Conclusions We have demonstrated that SSM can capture changes in hip shape over 6 and 12 months across the entire hip joint providing a sensitive measure of hip OA progression.

Keywords: osteoarthritis; statistical shape modelling; morphology; hip; radiographic; DXA

Introduction

Despite the high prevalence of osteoarthritis (OA) there is a dearth of treatments that can modify disease progression, leaving patients with symptomatic relief alone until a joint replacement becomes their only option. Among the reasons for the lack of Disease Modifying Osteoarthritis Drugs (DMOADs) are our current inabilities to predict who will develop OA, assess how rapidly their disease will progress and monitor treatment success.

The diagnosis of radiographic OA is made using semi-quantitative scoring systems such as Kellgren-Lawrence grades (KLG), Croft or ACR [1-3]. These scoring systems assign a score to each joint based on a number of primary radiographic features of joint disease including: deformation of the femoral head, osteophytes, subchondral cysts, subchondral sclerosis and reduced joint space width. Although these scoring systems are adequate for radiographic disease diagnosis and broad stratification into control/doubtful, moderate and severe OA they are not sensitive enough to provide a useful tool for the identification of subjects likely to progress rapidly nor for treatment monitoring in clinical trials.

Joint space width (JSW), a surrogate marker of cartilage loss, is a quantitative measure of OA progression and is currently the only Federal Drugs Administration-approved measure of OA progression. JSW is used as an endpoint in clinical trials of all new DMOADs. However, joint space narrowing focuses on a single feature in a single tissue in a multi-tissue, multi-feature disease.

In recent years, magnetic resonance imaging has become the mainstay of much of the ongoing OA research on disease progression. There is no doubt that data gained from MR images are invaluable in providing detailed information on bone marrow lesions, cysts and cartilage erosion. However, MR imaging is not suitable for all hip OA patients; not only is it costly but many patients are simply too obese to fit in a standard MR scanner.

Dual energy X-ray Absorptiometry (DXA) is an imaging modality used in the diagnosis of osteoporosis. In addition to bone mineral density (BMD) data, modern high-resolution scanners, such as the GELunar iDXA, generate near radiographic quality images. These images are currently under-utilised, simply being used to check positioning when capturing BMD data. We have previously demonstrated that these images can be used to grade radiographic OA severity as repeatably as from radiographs [4].

Statistical shape modelling (SSM) uses Principal Component Analysis to describe variation in the complicated shapes of natural objects [5]. SSMs have previously been applied to images captured by different imaging modalities taken from a number of different sites within the human body including heart [5], brain [6], spine [7, 8], hip [9, 10] and leg [11]. In the context of this study the SSM is applied to hip images captured from an iDXA. This allows variations in multiple geometric measures to be considered in concert thus describing the hip joint as a whole. Each principal component, or mode of variation, describes a different aspect of hip shape. By applying the SSM to a series of hip images from the same individual taken at

intervals allows the investigation of how the shape of the hip changes as a function of time. In recent years, studies using SSM have shown that the resultant models can act as a biomarker for musculoskeletal disorders of the hip, predicting hip fractures in osteoporotic cohorts [10, 12-14] and incidence and progression of osteoarthritis, including total hip replacement [9, 15-21].

Standardized Response Mean (SRM) (the mean difference between two measurements divided by the standard deviation of the difference) is an effect size index that is frequently used in medicine to gauge clinical change over time. Biomarkers with higher responsiveness are desirable for clinical trials as this indicates trials could be shorter and more cost-effective.

In this study we examine the change in hip shape mode scores over a 12 month period using DXA images captured from a mixed-sex cohort with a range of OA severities at baseline.

Materials and methods

Subject recruitment

This is a prospective study using subjects recruited from the NHS Grampian Radiology Information System (RIS). Computerised searches of the database identified subjects aged >30 years who had undergone an anteroposterior pelvic radiograph or bilateral radiographs of the hips in the previous 12 months. Radiographic reports were examined by a clinician to

assess suitability for the study. Subjects were excluded based on the following criteria: surgical interventions (including total joint replacement and osteotomies), clear radiological evidence of inflammatory arthritis, congenital / developmental dysplasia, avascular necrosis, metabolic bone disease, or absence of a formal report on the RIS.

Following subject identification, letters were sent to the physician who initially referred the subject for the radiograph to seek their help in recruiting the subject into the study (no incentive was offered). The referring physician was asked to send an information pack to the subject. Subjects were asked to complete a contact form and return it, to indicate interest in participating in the study.

Written informed consent was obtained when subjects attended for bone density assessment, in accordance with the declaration of Helsinki. The Grampian Research Ethics Committees approved the study (ref. 06/S0801/116)

Radiographic grading

The radiographs of subjects who agreed to participate in the study were scored for OA severity in both hips, by a single reader blinded to the clinical diagnosis, using the Kellgren-Lawrence system (KLG) [1, 22]. Subjects were classified into control/doubtful OA according to the most severe hip or the right hip if both hips scored the same KLG: KLG no more than 1 in either hip; moderate OA: KLG no more than 2; and severe OA: KLG of at least 3 in one hip

based on the KLG of their most severe joint. Subjects recruited into the study underwent Dual Energy X-ray absorptiometry (DXA) scans (iDXA GE Lunar) of both hips on entry and after 6 and 12 months.

Statistical Shape Modelling

The detail of the statistical shape modelling technique (SSM) used has been described elsewhere [23]. In brief, Statistical Shape Modelling uses a set of "landmark points" to describe the outline of an object. Each landmark point refers to the same location in every image (for example the base of the lesser trochanter), allowing the variation in shape to be measured across different images, and all the images from both hips at all three visits (0, 6 and 12-month time-points) were included in the model.

The landmark points were placed using the active shape modelling toolkit (Visual Automation Limited, Manchester, UK), a software program that runs within MATLAB (The Math Works Inc, Natick, United states) software environment, and analysed using custom made software (SHAPE, Aberdeen University, Aberdeen, UK). The analysis performs a Procrustes transformation, to remove effects of overall size, before applying principal components analysis to generate a series of orthogonal modes that describe the variations in shape within the set of images. The modes are scaled to have a mean of zero and unit standard deviation for the whole image set. Each image is then given a score to describe how far the shape lies from the mean shape for each mode.

In this study we built a 55-point SSM. This model included the proximal femur, part of the acetabulum and osteophytes, allowing visualization of common radiographic features observed in OA (Figure 1).

Figure 1 here

Statistical Analysis

Comparison between the groups at baseline was performed using one-way ANOVAs for continuous variables (or Kruskal-Wallis ANOVA on Ranks if data were not normally distributed determined using the Shapiro-Wilk test) and a Chi-squared test for sex balance. To account for multiple measures from the same person (both hips, multiple visits) Generalized Estimating Equations (GEE) were used to investigate the relationship between shape, KLG and changes between visits after adjustment for age, sex and BMI (SPSS v22, IBM corp). Both hips were included at each time-point and the dependent variable was the mode of interest. Missing data were excluded. A separate GEE was run for each mode. Variables analysed were Baseline KLG, Visit (in months), age at baseline, sex, BMI and the interaction between visit and baseline KLG. An autoregressive correlation matrix (AR(1)) was used. The distribution was set to 'normal' and link to 'identity'. This generates a measure of the slope of the relationship, B, and its 95% confidence interval (CI) as well as a *P*-value, which we used merely as a guide. We took results to be statistically significant when the 95% CI did not contain zero. Where statistically significant ($P<0.05$) changes over time (with visit) were observed, Standardised Response Means (SRM) and 95% confidence intervals

were calculated using the hip with the highest KLG for each person (Medcalc version15, Ostend, Belgium). The SRM scales the difference between the means of two sequential measurements by the standard deviation of the differences and enables measures of different magnitudes to be compared. This allowed direct comparison of the responsiveness with other published imaging measures used in OA.

Results

In total, 62 subjects (37 female and 25 male) were recruited into this study. At baseline, 14 hips were graded as KLG 0, 49 as KLG 1, 34 as KLG 2, 16 as KLG 3 and 11 as KLG 4. KLG was determined from the baseline radiograph taken in the preceding 12 months. The average time between radiograph and baseline iDXA scan was 225 ± 104 days. The mean time between baseline and the 6-month DXA scans was 174 (17) days, between 6- and 12-month DXA scans was 182 (18) days and for baseline DXA to 12-month DXA scan was 357 (30) days. All three scans were obtained from 50 of the participants; eight had no 6-month DXA scan and a further four did not receive a 12-month DXA scan. Table 1 shows the results of tests for differences in the distribution of age, sex or BMI between the three severity groups (control/doubtful, moderate and severe OA) at baseline. BMI was not normally distributed (Shapiro-Wilk P=0.034), so Kruskal-Wallis ANOVA on Ranks was applied. No statistically significant differences were observed for age (P=0.37), BMI (P=0.084) or sex (P=0.067).

Table 1 near here

A scree plot [24], showing the amount of variance in the model described by each mode, was used to select the first 5 modes of variation. Mode 1 described 22.2% and mode 5 4.4% of the variance. Together these 5 modes explained over 55% of the total variance and each subsequent mode described less than 4% of the variance in the model.

These first 5 modes were analysed for association with KLG using GEE. After adjustment for age, sex and BMI, statistically significant associations with KLG were observed in three modes; mode 1 ($P < 0.0001$, B = –0.29 (95% CI = –0.381, –0.19)), mode 3 ($P = 0.001$, B = 0.29 (95% CI = 0.12, 0.46)) and mode 4 ($P = 0.014$, B = 0.18 (95% CI = 0.04, 0.32)). Of these, two also showed significant differences with visit, mode 1 ($P < 0.0001$, B = –0.036 (95% CI = –0.051, –0.021)) and mode 4 ($P < 0.0001$, B = –0.037 (95% CI = –0.054, –0.019)).

Examining these three modes further, by including interaction terms between KLG and visit, showed significant interactions for modes 3 ($P = 0.009$, B = –0.018 (95% CI = –0.032, –0.005)) and 4 ($P < 0.009$, B = –0.022 (95% CI = –0.038, –0.005)), but no interaction for mode 1 ($P = 0.26$). All three modes associated with KLG were significantly associated with sex and modes 1 and 4 were also linked to BMI ($P \leq 0.05$). Two modes (Modes 2 and 5) were associated with age ($P < 0.05$) but were not linked to any other input variables in the model.

Figure 2 shows the variations in shape described by modes 1, 3 and 4. Decreasing Mode 1 values (from +2sd to –2sd) show all the classical features of radiographic osteoarthritis: increasing size of superior and inferior femoral head osteophytes, joint space narrowing and

deformation of the femoral head. In addition, there is also a reduction in neck shaft angle, some uncovering of the femoral head and an increase in neck width. Increasing Mode 3 values (from –2sd to +2sd) were associated with superior and inferior femoral head osteophytes and with uncovering of the femoral head. Increasing mode 4 values (from –2sd to +2sd) were associated with growth of a superior femoral head osteophyte but not inferior osteophytes and a decrease in neck shaft angle.

Figure 2 near here

The average mode 1 score decreased over the course of the 12-month follow-up and was also found to have a negative association with KLG. Whilst the average mode 4 score also decreased with visit this showed an inverse relationship with OA severity. Investigating the effect size of these changes SRM values were calculated from the hip with the highest KLG at baseline (used for severity grouping) over the first and second 6-month periods, then for the full 12 months for modes 1,3 and 4 (Table 2). Further analysis by severity group (Table 3) revealed the largest change in mode 1 was seen in those with moderate OA (12-month SRM=-0.91 [-1.36, -0.55]) followed by those with severe OA (12-month SRM=-0.61 [-1.26, -0.07]). In the control/doubtful OA group, mode 3 showed the largest changes (12-month SRM=0.66 [0.26, 1.06]). Whilst in the severe group mode 4 was the most sensitive (12-month SRM=-0.67 [-1.33, -0.10]).

Table 2 near here

Table 3 near here

Discussion

Data from this study demonstrate that SSM is a sensitive imaging biomarker that can capture morphological changes in the hip over a period as short as 6 months, but more reliably over 12 months. We observed a significant association between hip shape and OA disease severity in three modes (Modes 1, 3 and 4) in this model. There was a rise in SRM between 0-6 and 0-12 months when the cohort was analysed as a whole for each of the modes. The SRM values reported here (-0.63, 0.44 and -0.40 for modes 1, 3 and 4 respectively over 12 months) are greater than those reported elsewhere in the literature for a single measure over this timescale (reviewed in [25]), demonstrating improved responsiveness to disease progression. Further, analysis by severity grouping revealed the greatest change was observed over 12 months in the moderate OA group (Mode 1 SRM -0.91). In addition, even in those patients categorised as having severe OA using the KLG scoring system (KLG3 or 4) at baseline we were still able to demonstrate a progression over 12 months by shape changes captured by mode 1 (SRM -0.61) and mode 4 (SRM -0.67). Interestingly, mode 3 appeared to detect changes over 12 months in individuals starting with no or doubtful OA with a comparable SRM of 0.66. The variability between the SRMs for the first and the second 6-month periods indicates that although changes may be detected in six months this is not as reliable as over the full 12 months when the changes are more consistent. Whilst there are, as yet, no DMOADs, new drugs that are in the early stage of development are likely to work most effectively in mild to moderate OA subjects. When therapies become available to slow radiological progression in OA SSM would then

provide a sensitive means of measuring these changes, especially in early and moderate disease but also even in those with the most advanced OA.

Unsurprisingly, the shape model showed strong links with KLG, especially Mode 1 in which a lower score was associated with the growth of osteophytes, femoral head and neck deformation and joint space narrowing (Figure 2). One strength of the SSM approach is that all of these features are detected simultaneously and a score assigned on a continuous scale; they do not have to be assessed separately and graded on an ordinal scale. The patterns observed by shape modes presented here match well with geometrical features observed by other studies. For example, a decreased neck shaft angle, as seen in Mode 3 and Mode 4 was identified as a risk factor for OA by Doherty et al 2008 [26]. Deformation of the femoral head, such as the non-spherical shape observed with low mode 1 scores has also been identified previously in OA [26] as has the further deformation that occurs with the progression of OA, with the femoral head becoming increasingly mushroom-shaped [15, 19, 20]. This sometimes-dramatic change in shape is included in the semi-quantitative Kellgren-Lawrence grading system for moderate and severe cases. The uncovering of the femoral head, as seen in mode 1 and 3 is a feature also captured in the Chingford cohort using the Hip Morph software [27]. Statistical Shape Modelling allows quantification of these effects and identifies feature that co-occur. This enables changes over periods as short as 6 to12 months to be measured (as shown here) or, as found previously, over a period up to 5 years [19].

Joint space width (JSW) is the current gold standard for clinical trials assessing structural progression of OA. Whilst much of the focus has been on the knee a number of studies have investigated hip OA progression and these provide a useful effect size comparator. Few are as good over a period as short as 6-12 months as those presented in this study. A systematic review published in 2011 pooled 11 studies measuring minimum JSW in a meta-analysis [25]. The overall SRM for the pooled data was 0.66. This is similar to the SRM for our mode 1 data, although the average time to follow-up was not reported for the pooled data and the studies included ranged from 1-8 years, whereas our SRM for mode 1 was achieved after just 1 year of follow-up. Data from individual studies produce SRMs that range from 1.75 to 0.27 [28-32]. Findings from a 1-year study in which subjects started with KLG 2 or 3, reported a reduction in minimum JSW of 0.52 (0.49) mm (SRM 1.05) [33], which is comparable with 0.91 observed in the current study when analysing those patients with a KLG 2 at baseline. Traditionally, the terms 'small,' 'medium,' and 'large' were used by Cohen to provide a qualitative assessment of the effect size. He pointed out, however, that these should be treated as relative not only to each other but also to the research method being employed in any given investigation [34]. Accordingly, we have used SRMs in order to be able to compare our results with those from studies of JSW, which is measured on a different scale.

We are not the only group to use DXA scans to predict OA progression, although others have used DXA images to measure specific elements of hip geometry and morphometry [17, 35]. The use of DXA images has a number of key advantages over radiographs. The radiation dose of a bilateral hip iDXA is ~20 µSv, compared with ~700 µSv for a bilateral hip

radiograph. In addition, patient positioning is also routinely more reproducible than in most hip or pelvic radiography as position is important for accurate measurement of hip bone mineral density, the most common use of DXA scanners. This minimizes possible effects due to internal or external rotation. While the current modality of choice for OA imaging is MRI, as it provides a 3D image, MRI scans are more expensive and require longer appointments compared with DXA. In addition, modern DXA scanners, such as the iDXA, can take patients weighing up to 450 lb, which is a distinct advantage in a disease where obesity is common; the narrow aperture of an MRI scanner can make it difficult to scan patients whose BMI exceeds ~35 kg/m^2.

OA is gradually undergoing a paradigm shift away from being a disease of articular cartilage to one of the whole joint or even the whole body. In this respect the use of SSM enables changes across the whole joint, including contact alignment or coverage, anatomy, cartilage thickness and osteophytes to be captured. This may contribute to its greater sensitivity to subtle morphometric changes that appear to provide better biomarkers for incidence and progression. Capturing these changes across the entire joint also paves the way to explore further genetic factors affecting joint shape [36-38].

As with any study ours has limitations. We classified our patients into 3 groups based on radiographic OA, which is not necessarily coincident with symptomatic clinical OA. Patients recruited for this study were those who had undergone a pelvic or bilateral hip radiograph in the previous 12 months and it is possible, therefore, that those in the control/doubtful

group, whilst not exhibiting radiographic OA, may have had clinically symptomatic OA. KLG was determined from the baseline radiograph and these routine healthcare radiographs were taken at some time in the preceding 12 months, as described above, which may have affected group assignment. However, we have previously published data comparing KLG on recruitment radiographs and baseline DXA images in this cohort that demonstrates that this time gap had little effect on KLG repeatability [4], highlighting the lack of sensitivity of KLG as a method for monitoring OA disease progression. It might also have been useful to have taken matching radiographs at the 6-month time point as this would have allowed us to do a head to head comparison of change over time DXA vs radiographs using SSM, this was not done, however, in an effort to minimise radiation exposure in this cohort. Sex and BMI both approached conventional standards for statistical significance ($0.10>P>0.05$) and it is possible that a larger cohort might have led to dividing by sex or having to include BMI as a covariate. In a recent study of a birth cohort of 1633 individuals imaged using DXA between the ages of 60-64 years we found that all except one of the first ten modes generated in that study differed between men and women [39] and that higher BMI throughout adulthood and greater gains in BMI were associated with a shorter femoral neck and a wider and flatter femoral head [40]. There was also an interaction between sex and BMI in the current study with men having a greater BMI than women. This is something that can be explored in future, larger, studies having shown in the current study that, at least in principle, it is possible to detect changes over 12 months irrespective of sex or BMI.

In conclusion, we have demonstrated that SSM can capture changes in hip shape across the entire joint over a period as short as 6 months, but more reliably, 12 months. This provides a

sensitive measure of hip OA progression. While similar results may be expected from carefully positioned radiographic imaging, the use of DXA images demonstrates the utility of this low-dose imaging modality for the assessment of OA.

Acknowledgements

We are grateful to all the study participants. We thank Lana Gibson and Jennifer Scott for their expertise with the iDXA scanner as well as iDXA precision data.

Funding source

This study was supported by an award (Ref: WHMSB_AU068/071) from the Translational Medicine Research Collaboration – a consortium made up of the Universities of Aberdeen, Dundee, Edinburgh and Glasgow, the four associated NHS Health Boards (Grampian, Tayside, Lothian and Greater Glasgow & Clyde), Scottish Enterprise and initially Wyeth, now Pfizer.

The funder had no involvement in the collection, analysis and interpretation of data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

Dr J.S. Gregory was the holder of an MRC New Investigator award (Ref: G0901242)

Contributions

Rebecca J Barr: Conception and design, Analysis and interpretation of the data, Drafting of the article, Final approval of the article, Obtaining of funding

Jennifer S Gregory: Conception and design, Analysis and interpretation of the data, Drafting of the article, Final approval of the article, Obtaining of funding

Kanako Yoshida, Critical revision of the article for important intellectual content, Final approval of the article, Collection and assembly of data

Salvatore Alesci: Conception and design, Analysis and interpretation of the data, Critical revision of the article for important intellectual content, Final approval of the article, Obtaining of funding

Richard M Aspden: Conception and design, Critical revision of the article for important intellectual content, Final approval of the article, Obtaining of funding, Integrity of the work as a whole, from inception to finished article.

David M Reid: Conception and design, Critical revision of the article for important intellectual content, Final approval of the article, Obtaining of funding

Competing interests

Dr Salvatore Alesci was an employee of Wyeth at the time of the study. All other authors have no competing interests to declare.

References

1. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957; 16: 494-502.
2. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. American Journal of Epidemiology 1990; 132: 514-22.
3. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis and Rheumatism 1991; 34: 505-14.
4. Yoshida K, Barr RJ, Galea-Soler S, Aspden RM, Reid DM, Gregory JS. Reproducibility and Diagnostic Accuracy of Kellgren-Lawrence Grading for Osteoarthritis Using Radiographs and Dual-Energy X-ray Absorptiometry Images. J Clin Densitom 2015; 18: 239-44.
5. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models- Their Training and Application. Computer Vision and Image Understanding 1995; 61(1): 38-59.
6. Cootes TF, Taylor CJ. Anatomical statistical models and their role in feature extraction. Br J Radiol 2004; 77 Spec No 2: S133-9.
7. Meakin JR, Gregory JS, Aspden RM, Smith FW, Gilbert FJ. The intrinsic shape of the human lumbar spine in the supine, standing and sitting postures: characterization using an active shape model. J Anat 2009; 215: 206-11.
8. Pavlova AV, Meakin JR, Cooper K, Barr RJ, Aspden RM. The lumbar spine has an intrinsic shape specific to each individual that remains a characteristic throughout flexion and extension. Eur Spine J 2014; 23 Suppl 1: S26-32.
9. Barr RJ, Gregory JS, Reid DM, Aspden RM, Yoshida K, Hosie G, et al. Predicting osteoarthritis progression to total hip replacement: can we do better than risk

374 factors alone using active shape modelling as an imaging biomarker? Rheumatology
375 (Oxford) 2012; 51: 562-70.
376 10. Goodyear SR, Barr RJ, McCloskey E, Alesci S, Aspden RM, Reid DM, et al. Can we
377 improve the prediction of hip fracture by assessing bone structure using shape and
378 appearance modelling? Bone 2013; 53: 188-93.
379 11. Varzi D, Coupaud SA, Purcell M, Allan DB, Gregory JS, Barr RJ. Bone morphology of
380 the femur and tibia captured by statistical shape modelling predicts rapid bone loss
381 in acute spinal cord injury patients. Bone 2015; 81: 495-501.
382 12. Gregory JS, Testi D, Stewart A, Undrill PE, Reid DM, Aspden RM. A method for
383 assessment of the shape of the proximal femur and its relationship to osteoporotic
384 hip fracture. Osteoporosis International 2004; 15: 5-11.
385 13. Baker-Lepain JC, Luker KR, Lynch JA, Parimi N, Nevitt MC, Lane NE. Active shape
386 modeling of the hip in the prediction of incident hip fracture. Journal Of Bone And
387 Mineral Research 2011; 26: 468-74.
388 14. Whitmarsh T, Fritscher KD, Humbert L, del Rio Barquero LM, Roth T, Kammerlander
389 C, et al. Hip fracture discrimination from dual-energy X-ray absorptiometry by
390 statistical model registration. Bone 2012; 51: 896-901.
391 15. Nelson AE, Liu F, Lynch JA, Renner JB, Schwartz TA, Lane NE, et al. Association of
392 incident symptomatic hip osteoarthritis with differences in hip shape by active shape
393 modeling: the Johnston County Osteoarthritis Project. Arthritis Care Res (Hoboken)
394 2014; 66: 74-81.
395 16. Betancourt MC, Linden JC, Rivadeneira F, Rozendaal RM, Zeinstra SM, Weinans H, et
396 al. Dual energy x-ray absorptiometry analysis contributes to the prediction of hip
397 osteoarthritis progression. Arthritis Res Ther. 2009; 11: R162.

17. Waarsing JH, Rozendaal RM, Verhaar JAN, Bierma-Zeinstra SMA, Weinans H. A statistical model of shape and density of the proximal femur in relation to radiological and clinical OA of the hip. Osteoarthritis and Cartilage 2010; 18: 787-94.

18. Agricola R, Reijman M, Bierma-Zeinstra SMA, Verhaar JAN, Weinans H, Waarsing JH. Total hip replacement but not clinical osteoarthritis can be predicted by the shape of the hip: A prospective cohort study (CHECK). Osteoarthritis and Cartilage 2013; 21: 559-64.

19. Gregory JS, Waarsing JH, Day J, Pols HA, Reijman M, Weinans H, 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: 3634-43.

20. Lynch JA, Parimi N, Chaganti RK, Nevitt MC, Lane NE. The association of proximal femoral shape and incident radiographic hip OA in elderly women. Osteoarthritis Cartilage 2009; 17: 1313-8.

21. Agricola R, Leyland KM, Bierma-Zeinstra SM, Thomas GE, Emans PJ, Spector TD, et al. Validation of statistical shape modelling to predict hip osteoarthritis in females: data from two prospective cohort studies (Cohort Hip and Cohort Knee and Chingford). Rheumatology (Oxford) 2015; 54: 2033-41.

22. The Atlas of Standard Radiographs of Arthritis. Rheumatology 2005; 44: iv43-iv72.

23. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models - Their Training and Application. Computer Vision and Image Understanding 1995; 61: 38-59.

24. Cattell RB. The scree test for a number of factors. Multivariate Behavioural Research 1966; 1: 245-76.

25. Chu Miow Lin D, Reichmann WM, Gossec L, Losina E, Conaghan PG, Maillefert JF. Validity and responsiveness of radiographic joint space width metric measurement in hip osteoarthritis: a systematic review. Osteoarthritis Cartilage 2011; 19: 543-9.

26. Doherty M, Courtney P, Doherty S, Jenkins W, Maciewicz RA, Muir K, et al. Nonspherical femoral head shape (pistol grip deformity), neck shaft angle, and risk of hip osteoarthritis: A case-control study. Arthritis and Rheumatism 2008; 58: 3172-82.

27. Nicholls AS, Kiran A, Pollard TCB, Hart DJ, Arden CPA, Spector T, et al. The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: A nested case–control study. Arthritis and Rheumatism 2011; 63: 3392-400.

28. Lane NE, Nevitt MC, Hochberg MC, Hung YY, Palermo L. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. Arthritis Rheum 2004; 50: 1477-86.

29. Conrozier T, Jousseaume CA, Mathieu P, Tron AM, Caton J, Bejui J, et al. Quantitative measurement of joint space narrowing progression in hip osteoarthritis: a longitudinal retrospective study of patients treated by total hip arthroplasty. Br J Rheumatol 1998; 37: 961-8.

30. Papaloucas CD, Ward RJ, Tonkin CJ, Buckland-Wright JC. Cancellous bone changes in hip osteoarthritis: A short-term longitudinal study using fractal signature analysis. Osteoarthritis and Cartilage 2005; 13: 998-1003.

31. Pavelka K, Gatterova J, Gollerova V, Urbanova Z, Sedlackova M, Altman RD. A 5-year randomized controlled, double-blind study of glycosaminoglycan polysulphuric acid complex (Rumalon) as a structure modifying therapy in osteoarthritis of the hip and knee. Osteoarthritis Cartilage 2000; 8: 335-42.

32. Lequesne M, Winkler P, Rodriguez P, Rahlfs VW. Joint Space Narrowing in Primary Osteoarthritis of the Hip - Results of a 3 Year Controlled Trial. Arthritis and Rheumatism 1992; 35: S135.

33. Chevalier X, Conrozier T, Gehrmann M, Claudepierre P, Mathieu P, Unger S, et al. Tissue inhibitor of metalloprotease-1 (TIMP-1) serum level may predict progression of hip osteoarthritis. Osteoarthritis Cartilage 2001; 9: 300-7.

34. Cohen J. Statistical power analysis for the behavioral sciences, Lawrence Erlbaum Associates 1988.

35. Castaño Betancourt MC, Van der Linden JC, Rivadeneira F, Rozendaal RM, Bierma Zeinstra SM, Weinans H, et al. Dual energy x-ray absorptiometry analysis contributes to the prediction of hip osteoarthritis progression. Arthritis Research and Therapy 2009; 11.

36. Lindner C, Thiagarajah S, Wilkinson JM, Panoutsopoulou K, Day-Williams AG, arc OC, et al. Investigation of association between hip osteoarthritis susceptibility loci and radiographic proximal femur shape. Arthritis Rheumatol 2015; 67: 2076-84.

37. Baker-LePain JC, Lynch JA, Parimi N, McCulloch CE, Nevitt MC, Corr M, et al. Variant alleles of the Wnt antagonist FRZB are determinants of hip shape and modify the relationship between hip shape and osteoarthritis. Arthritis and Rheumatism 2012; 64: 1457-65.

38. Waarsing JH, Kloppenburg M, Slagboom PE, Kroon HM, Houwing-Duistermaat JJ, Weinans H, et al. Osteoarthritis Susceptibility Genes Influence the Association Between Hip Morphology and Osteoarthritis. Arthritis and Rheumatism 2011; 63: 1349-54.

468 39. Pavlova AV, Saunders FR, Muthuri SG, Gregory JS, Barr RJ, Martin KR, et al. Statistical
469 shape modelling of hip and lumbar spine morphology and their relationship in the
470 MRC National Survey of Health and Development. Journal of Anatomy 2017; 231:
471 248-59.
472 40. Muthuri SG, Saunders FR, Hardy RJ, Pavlova AV, Martin KR, Gregory JS, et al.
473 Associations between body mass index across adult life and hip shapes at age 60 to
474 64: Evidence from the 1946 British birth cohort. Bone 2017; 105: 115-21.

475

476 **Table 1.** Distribution of age, sex and BMI for each severity group comprising N individuals.

477 Significance differences between groups were tested using Chi-Squared (sex), Kruskal-Wallis

478 (BMI) and one-way ANOVA (age).

Severity Group	**N**	**Sex** **N female (%)**	**BMI** **Median (25% - 75%)**	**Age** **Mean (sd)**
Control/doubtful	20	15 (75%)	32.7 (29.2, 35.5)	64.9 (8.0)
Moderate	20	13 (65%)	29.1 (25.9, 32.7)	67.5 (7.2)
Severe	22	9 (40.9%)	28.5 (25.9, 33.9)	68.7 (10.5)
Total	62	37 (59.7%)	29.7 (27.3, 33.9)	67.1 (8.8)
***P*-value**		0.067	0.084	0.37

479

480

481 **Table 2**. Standardized Response Means (SRMs) and 95% confidence intervals for modes 1, 3
482 and 4 for time periods 0-6 months, 6-12 months and 0-12 months.

483

	0-6 months	6-12 months	0-12 months
Mode 1	-0.41 [-0.68, -0.05]	-0.23 [-0.50, 0.05]	-0.63 [-0.88], -0.34]
Mode 3	0.23 [-0.04, 0.47]	0.35 [0.05, 0.64]	0.44 [0.19, 0.68]
Mode 4	-0.18 [-0.47, 0.08]	-0.44 [-0.73, -0.15]]	-0.40 [-0.71, -0.11]

484

485

Table 3. Standardized Response Means (SRMs) and 95% confidence intervals for modes 1, 3 and 4 for time periods 0-6 months, 6-12 months and 0-12 months when data are split by severity (control/doubtful, moderate and severe OA). The number of individuals remaining in each group at each time point are given by N for the first mode.

		Control/doubtful	Moderate	Severe
Mode 1	0-6 months	-0.11 [-0.61, 0.42]	-0.64 [-1.28, 0.06]	-0.43 [-1.00, 0.04]
	(N)	19	19	16
	6-12 months	-0.38 [-0.99, 0.13]	0.08 [-0.47, 0.59]	-0.33 [-0.84, 0.20]
	(N)	19	16	15
	0-12 months	-0.40 [-0.84, 0.07]	-0.91 [-1.36, -0.55]	-0.61 [-1.26, -0.07]
	(N)	19	16	15
Mode 3	0-6 months	0.07 [-0.41, 0.52]	0.35 [-0.10, 0.78]	0.23 [-0.34, 0.71]
	6-12 months	0.68 [0.16, 1.19]	0.04 [-0.56, 0.48]	0.40 [-0.17, 0.99]
	0-12 months	0.66 [0.26, 1.06]	0.33 [-0.26, 0.74]	0.39 [-0.25, 0.76]
Mode 4	0-6 months	0.12 [-0.38, 0.60]	-0.27 [-0.77, 0.20]	-0.43 [-0.92, 0.17]
	6-12 months	-0.52 [-0.93, -0.03]	-0.33 [-0.79, 0.14]	-0.44 [--1.15, 0.16]
	0-12 months	-0.18 [-0.65, 0.34]	-0.39 [-0.99, 0.09]	-0.67 [-1.33, -0.10]

493 **Figure captions**

494 **Figure 1.** Design for the hip SSM, comprising the proximal femur, part of the acetabulum
495 and osteophytes. In the absence of the feature identified in colour, (e.g an osteophyte) the
496 points collapse back on to the primary outline.

497

498 **Figure 2**. Variations in hip shape identified by SSM. Mode 1 (a), Mode 3 (b) and Mode 4 (c)
499 showed significant associations with KLG. Solid red line indicates +2sd, dotted blue line
500 denotes -2sd.

501