

Magnetic Resonance Imaging and the distribution of bone marrow fat in hip osteoarthritis

Jennifer. S. Gregory, PhD,¹ Rebecca J. Barr, PhD,¹ Victor Varela, PhD,^{2,4} Trevor S. Ahearn, PhD,² Jennifer Lee Gardiner, PhD,^{3,5} Fiona J. Gilbert, MD,^{2,6} Thomas W. Redpath, PhD,² James D. Hutchison, PhD,¹ Richard M. Aspden, DSc.¹

¹Arthritis and Musculoskeletal Medicine, Institute of Medical Sciences,

²Aberdeen Biomedical Imaging Centre, Lillian Sutton Building,

School of Medicine, Medical Sciences and Nutrition, University of Aberdeen,

Foresterhill, Aberdeen, AB25 2ZD, UK. ³Wyeth Research, Collegeville, PA, USA

Current affiliations: ⁴9616 Castle Ridge Circle, Highlands Ranch, Colorado 80129,

USA, ⁵Bill & Melinda Gates Foundation, Seattle, WA, USA, ⁶Department of

Radiology, University of Cambridge School of Clinical Medicine, Box 218 Cambridge Biomedical Campus, Cambridge, UK.

Address for correspondence

Professor Richard M. Aspden

Arthritis and Musculoskeletal Medicine

School of Medicine, Medical Sciences and Nutrition

IMS Building

University of Aberdeen

Foresterhill

Aberdeen AB25 2ZD

UK

e-mail: r.aspden@abdn.ac.uk

Tel: +44 1224 437445

Grant support

This study was supported by an award (Ref: WHMSB-AU119) 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 Wyeth. The funder played no part in the design, execution, analysis or publication of this paper.

Running title: **Marrow fat distribution in hip OA by MRI**

Abstract

Purpose: Characterise the distribution of bone marrow fat in hip osteoarthritis (OA) using MRI and to assess its use as a potential biomarker.

Methods: 67 subjects (39 female, 28 male) with either total hip replacement (THA) or different severities of radiographic OA, assessed by Kellgren-Lawrence grading (KLG), underwent 3T MR imaging of the pelvis using the IDEAL sequence to separate fat and water signals. Six regions of interest (ROI) were identified within the proximal femur. Within each ROI the fractional-fat distribution, represented by pixel intensities, was described by its mean, standard deviation, skewness, kurtosis and entropy.

Results: Hips were graded: 12 as severe symptomatic (THA), 33 had KLG0 or 1, 9 were KLG2, 11 with KLG3 and 2 with KLG4 were analysed together. The fractional-fat content in the whole proximal femur did not vary with severity in males (mean (SD) 91.2 (6.0)%) but reduced with severity in females from 89.1 (6.7)% (KLG0,1), 91.5 (2.9)% (KLG2), 85.8 (16.7)% (KLG3,4) to 77.5 (11.9)% (THA) (ANOVA $P=0.029$). These differences were most pronounced in the femoral head where mean values fell with OA severity in both sexes from 97.9% (2.5%) (KLG0,1) to 73.0% (25.9%) (THA, $P<0.001$) with the largest difference at the final stage. The standard deviation and the entropy of the distribution both increased ($P<0.001$).

Conclusions: Descriptors of the fractional fat distribution varied little with the severity of OA until the most severe stage, when changes appeared mainly in the femoral head, and have, therefore, limited value as biomarkers.

Keywords: Osteoarthritis, hip, intramedullary fat, fractional fat content, imaging biomarker

INTRODUCTION

In 2013 the World Health Organisation classified osteoarthritis (OA) as the most common condition affecting the musculoskeletal system (1). Worldwide, it affects an estimated 9.6% of men and 18% of women over 60 (1) and in more economically developed countries OA has been reported in approximately 40% of those over 70 years of age (2). The incidence increases with age and, in an increasingly elderly population, the number of sufferers continues to increase. Little is known about the pathogenesis of primary OA and early detection is difficult. Treatment is limited to analgesia, exercise and weight loss, where appropriate, until joint destruction and pain are severe enough to warrant a surgical joint replacement.

Although traditionally thought of as a cartilage disease there is increasing recognition that OA is a disorder affecting the whole joint (3). There is a recognized link with obesity, especially with knee OA but also with hip and, curiously, hand OA (4-7). This latter finding indicates that increased weight-bearing is not the primary problem but that metabolic factors may be more important (8). Epidemiological studies have suggested a systemic aetiology independent of weight-bearing (9, 10). It has been proposed that generalised OA may be a systemic disorder affecting the whole musculoskeletal system driven by lipid metabolism (11). The degeneration and loss of cartilage have been the main focus of diagnostic and therapeutic studies, despite changes in the bone figuring highly in the radiographic signs; namely subchondral sclerosis, cyst formation and marginal osteophytosis. There is a proliferation of poorly mineralised subchondral bone (12) and increased bone containing an altered profile of growth factors in the iliac crest, remote from the weight-bearing regions of the joints (13).

Bone forming osteoblasts share a common mesenchymal stem cell precursor with adipocytes (14). Defective co-regulation and alterations in lipid metabolism are possible mechanisms for the bone pathologies observed in both OA and osteoporosis (11, 15). A greater concentration of fatty acids has been reported in cartilage of OA patients (16) and an increased fat content in bone marrow from osteoarthritic femoral heads (17). Not only was the mass of fat per unit volume of bone tissue doubled in tissue from OA patients (17), despite the reduced marrow space due to the bone proliferation, but fractional levels of (*n*-6) fatty acids, precursors to pro-inflammatory eicosanoids were also increased (17). A report of a pilot study using MRI, demonstrated a difference in the lipid fractions in femoral bone marrow and muscles around the hip joint between healthy volunteers and OA patients (18).

We currently lack responsive measures, or biomarkers, which could be used to assess the risk of OA in individuals, detect early disease, monitor progression or evaluate therapies. Radiography and magnetic resonance imaging (MRI) offer several biomarkers reflecting structural changes in cartilage and bone. Radiographic joint space narrowing is currently the only biomarker accepted by the Federal Drugs Agency in the USA but its lack of sensitivity requires studies to have large sample sizes and prolonged duration. Measures of joint space width using MRI (19) and the application of newer MR methods, such as dGEMRIC to assess cartilage composition (20, 21), assessment of glycosaminoglycan content using saturation transfer (22) and T1rho (23), still focus on changes in cartilage, but the detection of bone marrow lesions (24, 25) is broadening the whole area of MR imaging biomarkers (26). In this study we used MRI to measure the fractional fat content within the bone marrow in the proximal femur in patients with different severities of

osteoarthritis and characterized its distribution using statistical measures of image texture with a view to assessing its potential as an imaging biomarker.

MATERIALS AND METHODS

Subject Recruitment

Patients with end-stage symptomatic OA were recruited from the pre-operative assessment clinic while attending hospital in preparation for a total hip arthroplasty (THA). At this clinic they were given an information pack to take home and invited to participate.

For comparison with this group we required individuals to represent a range of severities of OA. Because there is no recognised grading of severity for symptomatic OA we chose to use a radiographic definition based on Kellgren-Lawrence grading. This cross-sectional analysis formed part of a prospective study using subjects recruited from the local Radiology Information System (RIS). Computerised searches of the database identified subjects greater than 30 years-old 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 if any of the following were reported: surgical interventions (including joint prostheses and osteotomies), inflammatory arthropathies, congenital or developmental dysplasias, avascular necrosis, metabolic bone disease or absence of a formal report on the Radiology Information System.

Having identified possible subjects, letters were sent to the referring physician to seek their help in recruiting the subject into the study (no incentive was offered). The referring physician sent an information pack to the subject who was asked to

complete a form and return it to indicate interest in participating in the study. They were then invited to attend hospital for MR imaging. The radiographs of these individuals were graded using the Kellgren-Lawrence system (KLG) (27, 28) by a single independent reader with four years' experience who was blinded to clinical diagnosis. The most severely affected hip joint was used, except if both hips were graded the same when the right hip was chosen, in subsequent analyses.

MR Imaging

Imaging was done using a 3 Tesla Philips Achieva MRI scanner (Philips Medical Systems, Best, Netherlands). A protocol was developed based on the Dixon method (29) and tested on phantoms comprising tubes containing different ratios of soya oil and water. The implementation (Philips Healthcare Clinical Science) was based on the iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL) sequence (30) and used three gradient-echo images acquired sequentially, with the initial TE equal to 2.1 ms, followed by additional images having TE increased by an increment (Δ TE) of 0.76 ms and then by 2Δ TE. This corresponds to the initial TE value having water and fat in phase, with the early and late TE values having water and fat $\pm 120^\circ$ out of phase. The 3-echo Dixon method allows signals to be assigned to either water or lipid protons unambiguously following suitable data analysis of the three images corresponding to each TE value.

The sequence was used to acquire 5 slices in the coronal plane with a slice thickness of 5 mm, and slice separation of 0.5 mm. The acquired in-plane resolution was 2.5 mm with an acquisition time of 6 minutes 11 seconds. TR was 160 ms with an RF pulse angle of 20° . These values give a proton density weighting so that the observed relative fat/water signal ratios are not overly affected by differences in

relaxation times between the water and lipid protons. Relative fat/water signal values were calculated directly from the decomposed water-fat images using programs written in-house. The slice in which the femoral head had its largest diameter was chosen for analysis. Images were also examined and the presence of bone marrow lesions (BMLs) and cysts noted; both were taken to be regions of high signal in the 'water' images but BMLs were assumed to have diffuse edges whereas cysts presented with demarcated boundaries.

The method was calibrated against a phantom comprising a set of tubes with oil:water volume proportions from 70:30 to 100:0 in steps of 3. These were made up from soy oil and pure water as described by Bernard et al. (31). A Bland and Altman analysis of agreement (32) showed the MRI measures to underestimate the true value with a bias of -14% (95% CI -15.6%, -12.6%). In order to generate a correction factor the MRI value was regressed on the true value yielding

$$\text{MRIfat\%} = 0.797\text{oil\%} + 3.167 \quad (R^2 = 0.99)$$

where oil% is the percentage volume of the soy oil and MRIfat% is that calculated.

This equation was used to correct the values for the mean and standard deviation of the percentage fat calculated from the images in each region of interest as described below. The remaining calculated parameters of the distribution required no correction.

Statistical Analysis

Images of the hip displaying calculated fractional fat values were segmented into one acetabular and six femoral regions of interest (ROI) using statistical shape modelling (33-36) (Figure 1). Image texture in each ROI was obtained from the statistical

distribution of values of fractional fat content. The histogram of calculated pixel fat values, was characterised by calculating the mean, standard deviation, skewness, kurtosis and entropy of the distribution. Cysts were automatically detected (MRIfat% < 20%) and removed from each ROI before analysis. The entropy of the distribution in each area containing N pixels in which x_i is the number of pixels with intensity i was defined as

$$S = (1/N) \sum -x_i \ln(x_i)$$

Results were tested for normality (Shapiro-Wilk) and, if normally distributed, are presented as mean (standard deviation) otherwise as median [25%, 75%] values. Analysis of variance (ANOVA) was used to investigate the relationship between OA severity and the statistical descriptor of the fractional-fat distribution in each ROI separately and in the total femur, adding all the femoral areas together. If the normality test failed ($P < 0.05$) a Kruskal-Wallis ANOVA on ranks was performed. Tests were corrected for sex.

RESULTS

A total of 67 subjects (39 female and 28 male) with a mean age of 65.2 years (range 40.5 – 81.5) took part in this cross-sectional study; 12 were about to undergo THA and 55 were identified from the RIS and were classified from their grade of radiographic OA. From the radiographs, 6 hips were classified as KLG0 and 27 with KLG1, 9 with KLG2, 11 with KLG3 and 2 with KLG4. Precision of KL grading was determined in a previous study using quadratic-weighted kappa (QWK) with an intra-observer repeatability for this reader of 0.88 and inter-observer repeatability of 0.81

against a radiologist with more than ten years' experience (37). For analysis KLG0 and 1 were combined, as were KLG3 and 4. Table 1 shows the groups and results of tests for differences in the distribution of radiographic severity with sex, age and Body Mass Index (BMI). No difference was found between groups with age or BMI but Chi-squared tests showed they were not evenly distributed between sexes. Accordingly, subsequent tests for significance were corrected for sex.

Qualitatively, the images from severe OA patients appeared very different from the no-OA group with muscle delineation being obscured and a different distribution of marrow fat apparent in the femur. Representative images of each of the four grades of radiographic severity are shown in Figure 2. Characteristics of the fractional-fat distribution in the total femur are shown in Figure 3 and none was very sensitive to increasing severity of radiographic OA. Small differences were found only in the females for the mean (ANOVA $P = 0.029$) whilst differences in entropy bordered on traditional values for significance ($P = 0.055$). Closer examination using *post hoc* tests showed that the difference arose in females undergoing THA in which fractional fat content fell to 77.5% (11.9%) compared with 88.9% (8.0%) for those with no-OA to moderate OA. There was, therefore, a significant difference between the mean fat content of the total proximal femur in males and females in the THA group ($P = 0.022$).

Considering each region separately, differences in characteristics of the fractional fat distribution were found in the most proximal regions, especially in the femoral head (regions 1 and 2). The region most affected was ROI 1, the load-bearing area of the femoral head (Figure 4) in which the mean, standard deviation and entropy all changed with increasing severity of radiographic OA. No difference was found in the skewness or kurtosis of the fat distributions. There was no difference between males

and females and the mean value reduced from 97.9% (2.5%) in those with KLG0,1 to 73.0% (25.9%) in the THA group ($P < 0.001$), with the largest difference being found in this end-stage group. The standard deviation and the entropy of the distribution increased (both $P < 0.001$). In region 2 (Figure 5) the differences were similar but not as marked; the mean decreased from 95.7% (5.4%) to 84.6% (14.2%) ($P < 0.001$) and, again, the standard deviation ($P = 0.003$) and the entropy ($P = 0.007$) both increased. The differences became smaller in the more distal regions and only the average showed a significant difference with its median value in ROI 3 falling from 91.8% [84.7%, 95.5%] in the KLG0,1 group to 80.6% [70.8%, 85.0%] in the THA group ($P = 0.042$) and, in ROI 4, from 88.6% [80.4%, 91.9%] in the KLG0,1 group to 68.7% [66.6%, 82.2%] in the THA group ($P = 0.021$). In all of these, the largest differences appeared in the THA group. No significant differences in mean fractional fat content were found with severity in ROI 5 ($P = 0.52$), ROI 6 ($P = 0.19$) or the acetabulum ($P = 0.17$).

Cysts were identified in three, and bone marrow lesions (BMLs) in nine, of the images analysed. Measurements from the 'water' images showed that the fractional water content was highest in cysts (90.0% (2.0%)) and slightly lower in the BMLs (75% (18%)).

DISCUSSION

The appearance of images from THA patients was very different from that of the other groups. As well as less distinct muscle delineation there were apparent differences in marrow fat and hence the objective of this study was to quantify these and explore whether characteristics of the statistical distributions could be used as

imaging biomarkers. Calculations of the fractional fat content in the bone marrow from the MR signals support the results from a previous study (18) that there is a significantly smaller fractional-fat content in the proximal femur of patients with OA. This difference is only apparent in those with end-stage disease and occurs mainly in the femoral head, with the differences between the groups diminishing with increasing distance from the joint. The femoral medulla is the source of haematopoiesis, giving marrow in young individuals its red colour. With age there is a known reduction in red marrow with a corresponding increase in yellow, or fatty, marrow. The rate of change is dependent on sex but is usually complete by the age of 65 for both sexes (38). Red marrow is approximately 40% lipid and yellow marrow is ~80% lipid (39) and the values we have obtained are comparable with those previously reported for yellow marrow.

The method used three echo times and provided uniform fat saturation by using an iterative least squares method to estimate B_0 inhomogeneities. Asymmetrically placed echoes also improve the signal-to-noise ratio from that in the original Dixon method (40). While excellent agreement has been achieved for fat contents, generally less than 40%, comparing imaging with MR spectroscopy in liver steatosis, (41), the high fat content of bone marrow and the presence of trabecular bone complicates the measurement of fractional-fat by shortening the values of T_2^* for both water and fat leading to rapid gradient echo signal decay with TE (42). Recent studies have shown that water components with short T_2^* and either very short or long T_2 values have negligible contributions to the MR signal at the values for TE accessible in clinical gradient echo sequences (42). Consequently, MR proton-density fat-fraction images overestimate the percentage fat. Modelling T_2^*

corrections indicated that the greatest fat-fraction bias, just over 6%, occurred at fractional-fat values close to 50% and that this bias decreased with increasing fat content so that for fractional-fat values greater than about 85% the bias was less than 3% (43). These corrections were not available to us and this might explain why our values were unexpectedly high. So, while the values measured may not be accurate, this should not invalidate their comparison across the range of OA severities in which all data were acquired using the same methods.

A limitation of the method is that only fractional contents of water and fat can be measured, not absolute values. BMLs and cysts had high water contents and resulted in a reduction in the calculated mean fractional fat% and an increase in the standard deviation, especially in ROI 1 where they were most commonly found. The presence of BMLs and cysts, however, did not explain all the reduction in fractional fat content as we excluded them as far as possible and implementation of this measure as a biomarker would be easiest by taking values over the whole ROI. In doing this, however, a low mean fractional fat% might arise from BMLs as well as a generally low percentage fat content. A previous histological study indicated that the amount of bone marrow edema represented only a very small fraction of the head and neck region of examined specimens (44), but here, where we subdivided the head, they represented a considerable proportion of the ROI and had a marked effect on the values calculated for that region. This same study, using fractional MR signals calculated from fat and water suppressed proton density images, reported an increase in the fractional amount of water in OA from 42% water in the control group to 60% in the OA femora (44). While this indicates similar trends to our study, the values are hard to reconcile with our results and with the traditional observation of large amounts of fatty, yellow marrow in the proximal femur of elderly individuals.

Texture analysis is a common means of characterising features in images and there are numerous methods available that describe the distribution of pixel intensities in a digital image. Here we used simple measures of the statistical distribution and added 'entropy'; a statistical measure of randomness to describe the changes observed in each ROI. The mean and standard deviation of a Gaussian distribution are common descriptors and need no introduction, other than to note that they are sometimes called the first and second moments of the distribution. If a distribution is not quite Gaussian, skewness is a measure of whether it is lopsided, the third moment, and kurtosis measures whether it is more 'peaky' or slightly 'squashed', the fourth moment. Skewness is negative if there are more values than expected to the left of the distribution, is zero for a Gaussian and positive if there is a tail to right. Kurtosis takes the value of three for a normal distribution and a value of greater than three indicates a distribution with a sharper peak and fatter tails. The fractional-fat distributions measured here were skewed to the left with sharper peaks and broader tails than a strictly normal distribution but neither skewness nor kurtosis showed any significant differences with severity in any of the ROIs. The entropy and standard deviation of the distribution, however, did increase in the femoral head indicating that the fat distribution became broader and more random in the final stages of the disease.

These findings of a smaller mean fractional-fat content in severe OA, however, seemingly contradict previous laboratory findings (17) and observations from surgery in which fat is often expressed as the femoral head is excised. Measuring the total fat content and expressing it as a mass of fat per unit gross volume of bone tissue biopsy demonstrated a doubling of lipid content in bone cores from patients with OA (17); the water content was not measured. In this study, the MR signal only enables

the fractional lipid content to be measured as a percentage of total signal with the remainder assumed to be due to water. Comparing bone from patients with severe OA with that from less severe or no OA is further complicated by the increased amount of cancellous bone in OA (12), which will reduce the volume available for bone marrow and alter the MR signal. Previous studies, however, reported that bone in severe OA is hypomineralized and the mass fraction of water in the bone increases to about 24%, compared with 17% in normal bone matrix (12). So, along with more bone and smaller spaces between trabeculae there is an increase in water within the bone matrix. If, however, this water is closely associated with the trabecular bone it may be that it has a short T2* and will not contribute to the MR signal as described above.

A further limitation arises from the cross-sectional nature of this study. While it provides initial data on possible associations between marrow fractional-fat content and OA severity, a longitudinal study would be needed to show whether more subtle changes could be detected in individuals with the incidence and progression of OA. Radiographs were graded by only one reader, not directly involved in this study, which may have led to some ROA cases being mis-classified, although calibration in previous studies showed intra-and inter-observer repeatability for this reader (37) to fall into the 'almost perfect' ($QWK \geq 0.81$) as defined by Landis and Koch (45). Study numbers are also relatively small, especially in the KLG2 and 4 groups, and when further subdivided by sex led to some results being close to, but greater than, traditional measures denoting significance. Greater numbers may improve the statistical power but in the search for a biomarker, while there may be scientific value in finding a significant association, such results will be of limited use in individuals or small cohort studies. The calibration procedure we used has been used elsewhere

but it needs to be recognised that the types of lipids and their MR signal are unlikely to match those found in the bone and it simply serves to adjust the values closer to what might be expected *in vivo*. further studies are required to compare MR measurements of lipid and water with laboratory measurements in order to obtain absolute values of total and fractional lipid contents.

In conclusion, fat-enhanced MR images showed little variation with increasing severity of radiographic OA, differences were only found in patients with end-stage disease awaiting THA. The differences were then confined to the most proximal parts of the femur. Calculation of fractional-fat contents from these regions showed them to be significantly lower, by about 25%, in the THA group and they were accompanied by a change in the texture of the images within the bone, indicating a broader and more random fat distribution in the final stages of the disease. The differences we measured, however, occur late in the disease process and would be enhanced by contributions from cysts and BMLs if added to the lower fat content. They do not appear capable, however, of providing a novel imaging biomarker of disease incidence or sensitive enough to be able to monitor progression.

ACKNOWLEDGEMENTS

We are grateful to Mrs D. Younie for kindly arranging the imaging sessions and Mrs B. MacLennan (research radiographer) for acquiring the MR images. We also thank Dr S. Galea-Soler, Dr G. Waiter, Dr Zhiqing Wu, and Dr K. Yoshida for Kellgren-Lawrence grading, help and advice, and Mr G. Buchan for his expertise making the phantoms.

REFERENCES

1. Wittenauer R, Smith L, Aden K. Background Paper 6.12. Osteoarthritis. World Health Organisation, 2013.
2. Valdes AM, Spector TD. Genetic epidemiology of hip and knee osteoarthritis. *Nat Rev Rheumatol* 2011;7(1):23-32.
3. Radin EL, Burr DB, Caterson B, Fyhrie D, Brown TD, Boyd RD. Mechanical determinants of osteoarthrosis. *Seminars in Arthritis & Rheumatism* 1991;21(3 Suppl 2):12-21.
4. Cicuttini FM, Baker JR, Spector TD. The association of obesity with osteoarthritis of the hand and knee in women: a twin study. *J Rheumatol* 1996;23(7):1221-1226.
5. Oliveria SA, Felson DT, Cirillo PA, Reed JI, Walker AM. Body weight, body mass index, and incident symptomatic osteoarthritis of the hand, hip, and knee. *Epidemiology* 1999;10(2):161-166.
6. Yusuf E, Nelissen RG, Ioan-Facsinay A, Stojanovic-Susulic V, DeGroot J, van Osch G, et al. Association between weight or body mass index and hand osteoarthritis: a systematic review. *Ann Rheum Dis* 2010;69(4):761-765.
7. Carman WJ, Sowers M, Hawthorne VM, Weissfeld LA. Obesity as a risk factor for osteoarthritis of the hand and wrist: a prospective study. *Am J Epidemiol* 1994;139(2):119-129.
8. Aspden RM. Obesity punches above its weight in osteoarthritis. *Nat Rev Rheumatol* 2011;7(1):65-68.

9. Hart DJ, Doyle DV, Spector TD. Association between metabolic factors and knee osteoarthritis in women: the Chingford Study. *J Rheumatol* 1995;22(6):1118-1123.
10. Cooper C, Egger P, Coggon D, Hart DJ, Masud T, Cicuttini F, et al. Generalized osteoarthritis in women: pattern of joint involvement and approaches to definition for epidemiological studies. *J Rheumatol* 1996;23(11):1938-1942.
11. Aspden RM, Scheven BAA, Hutchison JD. Osteoarthritis is a systemic disorder involving stromal cell differentiation and lipid metabolism. *The Lancet* 2001;357:1118-1120.
12. Li B, Aspden RM. Composition and mechanical properties of cancellous bone from the femoral head of patients with osteoporosis or osteoarthritis. *J Bone Miner Res* 1997;12:641-651.
13. Dequeker J, Mokassa L, Aerssens J. Bone density and osteoarthritis. *J Rheumatol* 1995;22:98-100.
14. Grigoriadis AE, Heersche JN, Aubin JE. Differentiation of muscle, fat, cartilage, and bone from progenitor cells present in a bone-derived clonal cell population: effect of dexamethasone. *J Cell Biol* 1988;106(6):2139-2151.
15. Kruger MC, Horrobin DF. Calcium metabolism, osteoporosis and essential fatty acids: a review. *Prog Lipid Res* 1997;36(2-3):131-151.
16. Lippiello L, Walsh T, Fienhold M. The association of lipid abnormalities with tissue pathology in human osteoarthritic articular cartilage. *Metabolism* 1991;40(6):571-576.
17. Plumb MS, Aspden RM. High levels of fat and (n-6) fatty acids in cancellous bone in osteoarthritis. *Lipids Health Dis* 2004;3:12.

18. Gregory JS, Ahearn TS, Redpath TW, Gilbert FJ, Munro N, Hutchison JD, et al. MRI measurements of the fat content of bone and muscle in hip osteoarthritis. *Osteoarthritis Cartilage* 2007;15(Suppl. C):C192.
19. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, et al. Brief Report: Cartilage Thickness Change as an Imaging Biomarker of Knee Osteoarthritis Progression: Data From the Foundation for the National Institutes of Health Osteoarthritis Biomarkers Consortium. *Arthritis Rheumatol* 2015;67(12):3184-3189.
20. Williams A, Gillis A, McKenzie C, Po B, Sharma L, Micheli L, et al. Glycosaminoglycan distribution in cartilage as determined by delayed gadolinium-enhanced MRI of cartilage (dGEMRIC): potential clinical applications. *AJR Am J Roentgenol* 2004;182(1):167-172.
21. Crema MD, Hunter DJ, Burstein D, Roemer FW, Li L, Eckstein F, et al. Association of changes in delayed gadolinium-enhanced MRI of cartilage (dGEMRIC) with changes in cartilage thickness in the medial tibiofemoral compartment of the knee: a 2 year follow-up study using 3.0 T MRI. *Ann Rheum Dis* 2014;73(11):1935-1941.
22. Ling W, Regatte RR, Navon G, Jerschow A. Assessment of glycosaminoglycan concentration in vivo by chemical exchange-dependent saturation transfer (gagCEST). *Proc Natl Acad Sci U S A* 2008;105:2266-2270.
23. Keenan KE, Besier TF, Pauly JM, Smith RL, Delp SL, Beaupre GS, et al. T1rho Dispersion in Articular Cartilage: Relationship to Material Properties and Macromolecular Content. *Cartilage* 2015;6(2):113-122.
24. Roemer FW, Frobell R, Hunter DJ, Crema MD, Fischer W, Bohndorf K, et al. MRI-detected subchondral bone marrow signal alterations of the knee joint:

terminology, imaging appearance, relevance and radiological differential diagnosis.

Osteoarthritis Cartilage 2009;17(9):1115-1131.

25. Ahedi H, Aitken D, Blizzard L, Cicuttini F, Jones G. A population-based study of the association between hip bone marrow lesions, high cartilage signal, and hip and knee pain. Clin Rheumatol 2014;33(3):369-376.

26. Eckstein F, Peterfy C. A 20 years of progress and future of quantitative magnetic resonance imaging (qMRI) of cartilage and articular tissues-A personal perspective. Semin Arthritis Rheum 2015.

27. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494-502.

28. The Atlas of Standard Radiographs of Arthritis. Rheumatology 2005;44(suppl 4):iv43-iv72.

29. Dixon WT. Simple proton spectroscopic imaging. Radiology 1984;153(1):189-194.

30. Reeder SB, Pineda AR, Wen Z, Shimakawa A, Yu H, Brittain JH, et al. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. Magn Reson Med 2005;54(3):636-644.

31. Bernard CP, Liney GP, Manton DJ, Turnbull LW, Langton CM. Comparison of fat quantification methods: a phantom study at 3.0T. J Magn Reson Imaging 2008;27(1):192-197.

32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 1986;1(8476):307-310.

33. Cootes TF, Hill A, Taylor CJ, Haslam J. Use of Active Shape Models for Locating Structures in Medical Images. *Image Vision Computing* 1994;12(6):355-365.
34. Gregory JS, Waarsing JH, Day JS, 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(11):3634-3643.
35. Goodyear SR, Barr RJ, McCloskey E, Alesci S, Aspden RM, Reid DM, et al. Can we improve the prediction of hip fracture by assessing bone structure using shape and appearance modelling? *Bone* 2013;53(1):188-193.
36. Barr RJ, Gregory JS, Reid DM, Aspden RM, Yoshida K, Hosie G, et al. Predicting OA progression to total hip replacement: can we do better than risk factors alone using active shape modelling as an imaging biomarker? *Rheumatology (Oxford)* 2012;51(3):562-570.
37. 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(2):239-244.
38. Dunnill MS, Anderson JA, Whitehead R. Quantitative histological studies on age changes in bone. *J Pathol Bacteriol* 1967;94(2):275-291.
39. Vande Berg BC, Malghem J, Lecouvet FE, Maldague B. Magnetic resonance imaging of the normal bone marrow. *Skeletal Radiol* 1998;27(9):471-483.

40. Chen CA, Lu W, John CT, Hargreaves BA, Reeder SB, Delp SL, et al. Multiecho IDEAL gradient-echo water-fat separation for rapid assessment of cartilage volume at 1.5 T: initial experience. *Radiology* 2009;252(2):561-567.
41. Hines CD, Frydrychowicz A, Hamilton G, Tudorascu DL, Vigen KK, Yu H, et al. T(1) independent, T(2) (*) corrected chemical shift based fat-water separation with multi-peak fat spectral modeling is an accurate and precise measure of hepatic steatosis. *J Magn Reson Imaging* 2011;33(4):873-881.
42. Karampinos DC, Melkus G, Baum T, Bauer JS, Rummeny EJ, Krug R. Bone marrow fat quantification in the presence of trabecular bone: initial comparison between water-fat imaging and single-voxel MRS. *Magn Reson Med* 2014;71(3):1158-1165.
43. Karampinos DC, Ruschke S, Dieckmeyer M, Eggers H, Kooijman H, Rummeny EJ, et al. Modeling of T2* decay in vertebral bone marrow fat quantification. *NMR Biomed* 2015;28(11):1535-1542.
44. Taljanovic MS, Graham AR, Benjamin JB, Gmitro AF, Krupinski EA, Schwartz SA, et al. Bone marrow edema pattern in advanced hip osteoarthritis: quantitative assessment with magnetic resonance imaging and correlation with clinical examination, radiographic findings, and histopathology. *Skeletal Radiol* 2008;37(5):423-431.
45. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174.

Table 1. Numbers, age, sex and BMI of participants showing no differences in age or BMI with severity of OA (ANOVA) but a significant difference in the sex distributions between each group (Chi-squared) with too many females and too few males in the KLG0,1 group which was reversed in the mild and moderate groups.

	N	Sex	Age	BMI
		N female (%)	Mean (SD)	(kg / m²)
KLG0,1	33	25 (76%)	62.6 (11.3)	27.8 (4.2)
KLG2	9	3 (33%)	65.6 (7.1)	27.4 (4.0)
KLG3,4	13	4 (31%)	67.7 (6.6)	27.2 (3.0)
THA	12	7 (58%)	67.8 (10.5)	28.8 (4.7)
Total	67	39 (58%)	65.2 (10.1)	27.8 (4.0)
<i>P</i> -value		0.015	0.19	0.79

Figure legends

Figure 1 Segmentation into regions of interest was done using statistical shape modelling to define one acetabular and 6 femoral regions

Figure 2. Representative MR images for each of the four groups classified either by radiographic KLG or by THA.

Figure 3. Graphs of (a) mean, (b) standard deviation, (c) skewness, (d) kurtosis and (e) entropy of the fractional fat distribution in the total proximal femur for THA compared with degrees of radiographic OA.

Figure 4. Graphs of (a) mean, (b) standard deviation and (c) entropy of the fractional fat distribution in the most superior region (ROI 1) of the femoral head for the different groups.

Figure 5. Graphs of (a) mean, (b) standard deviation and (c) entropy of the fractional fat distribution in the femoral head (ROI 2) for the different groups.

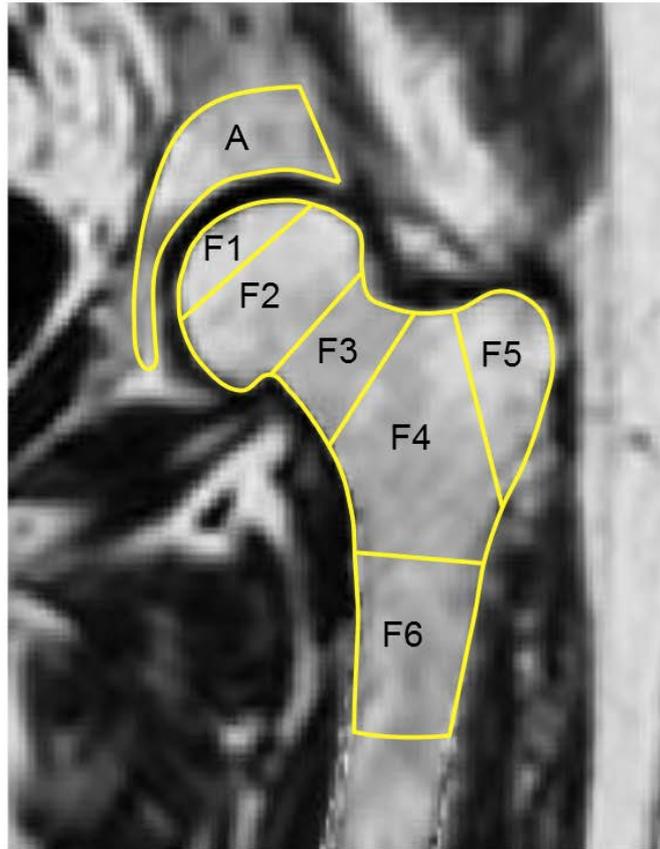


Figure 1. Segmentation into regions of interest was done using statistical shape modelling to define one acetabular and 6 femoral regions

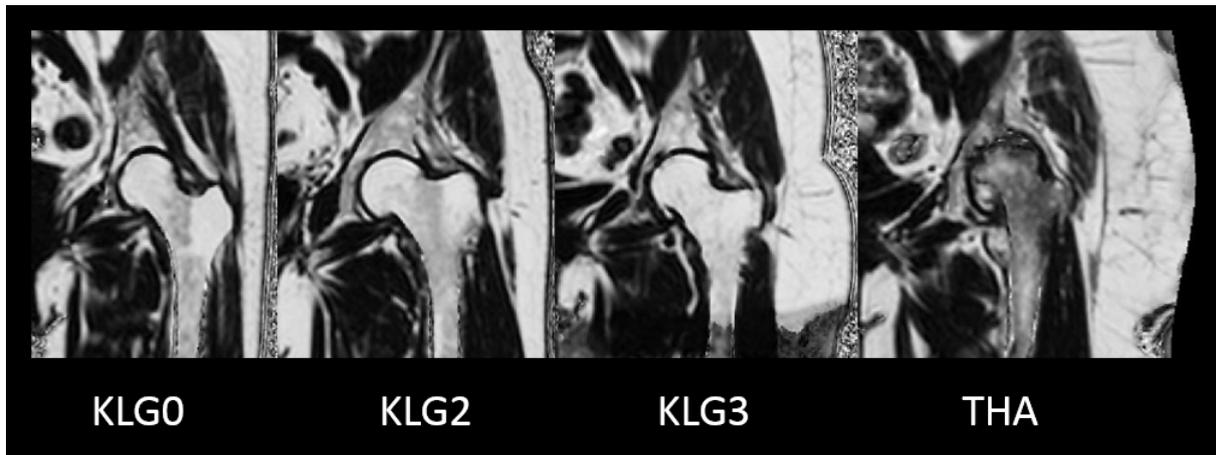


Figure 2. Representative MR images for each of the four groups classified either by radiographic KLG or by THA.

Total femur

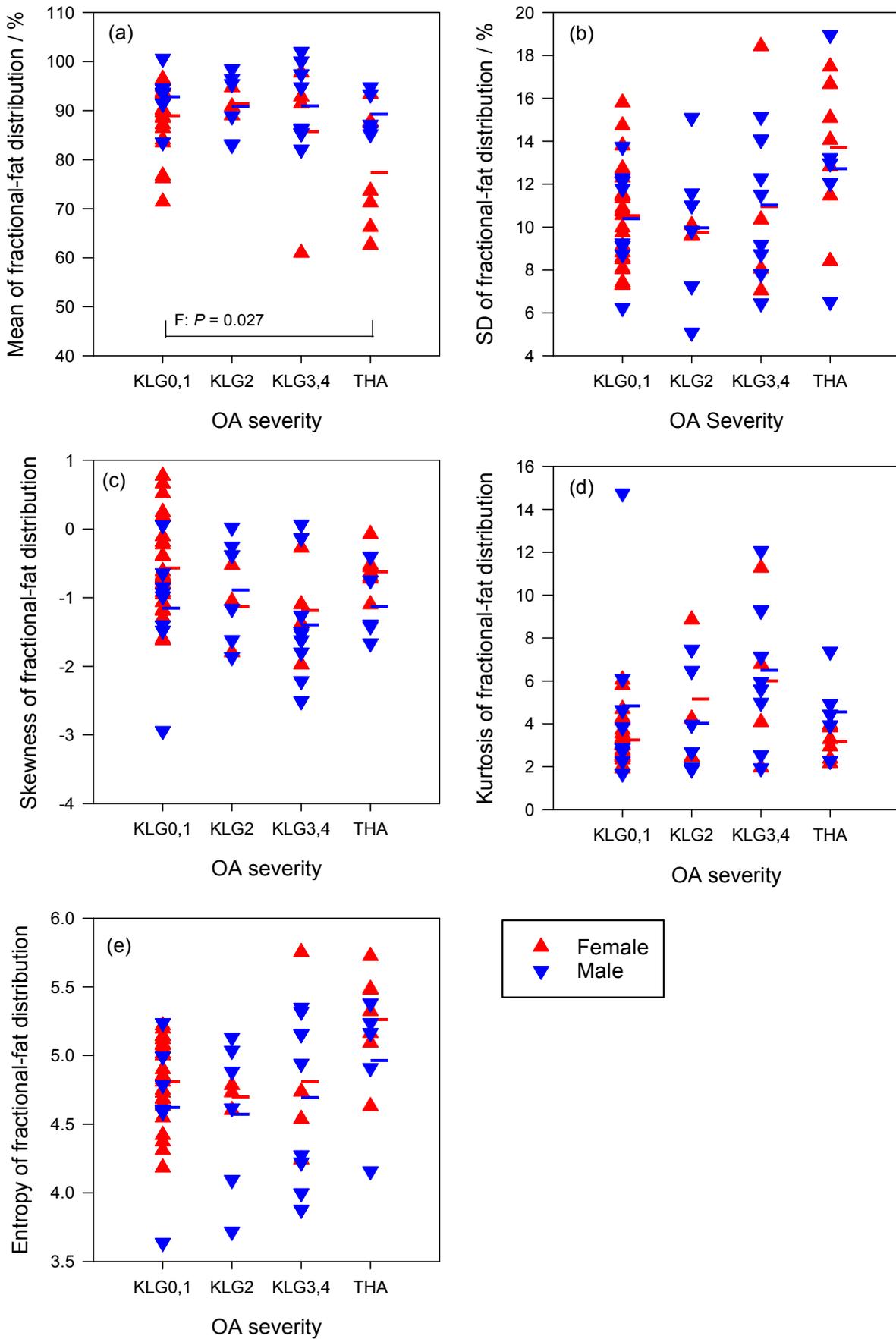


Figure 3. Graphs of (a) mean, (b) standard deviation, (c) skewness, (d) kurtosis and (e) entropy of the fractional fat distribution in the total proximal femur for THA compared with degrees of radiographic OA.

Region f1

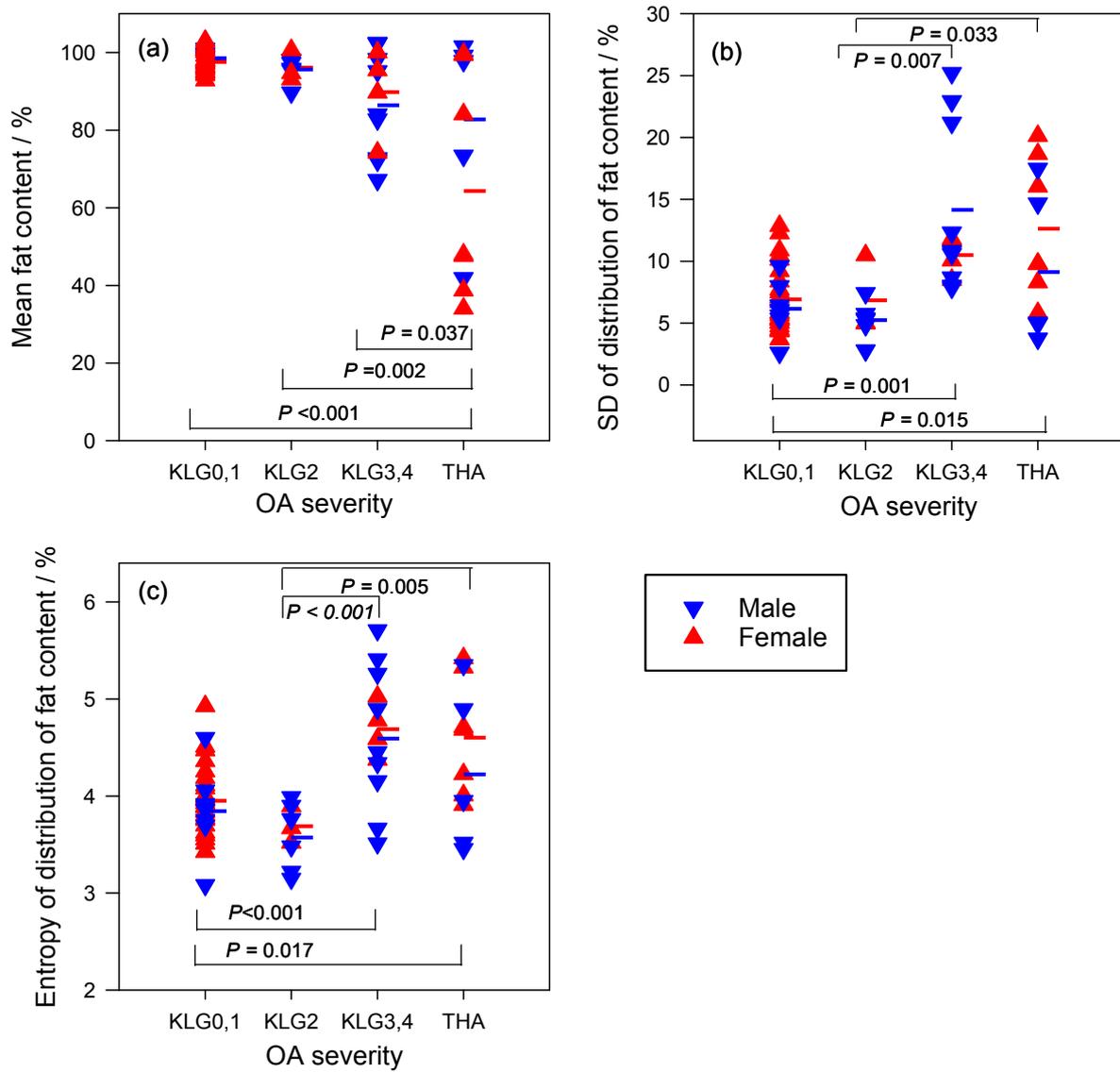


Figure 4. Graphs of (a) mean, (b) standard deviation and (c) entropy of the fractional fat distribution in the most superior region (ROI 1) of the femoral head for the different groups.

Region f2

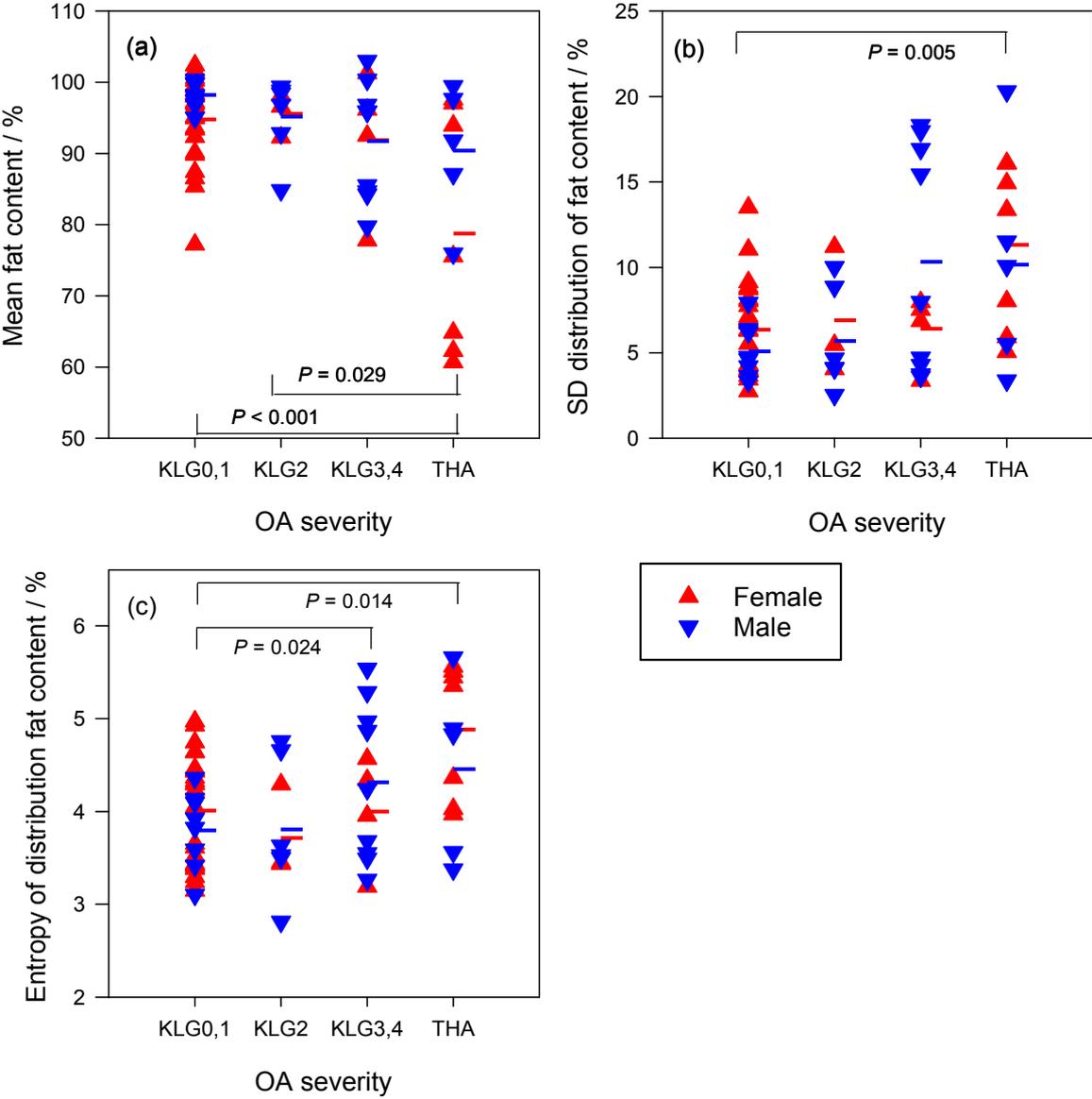


Figure 5. Graphs of (a) mean, (b) standard deviation and (c) entropy of the fractional fat distribution in the femoral head (ROI 2) for the different groups.