Magnetic Resonance Imaging and the distribution of bone marrow fat in hip osteoarthritis
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 different severities of OA, assessed by Kellgren-Lawrence grading (KLG) or total hip replacement (THA), 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: 33 as no-OA (KLG0 or 1), 9 as Mild OA (KLG2), 11 as moderate OA (KLG3) and 14 as severe (KLG4 or THA). 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)% (no-OA), 91.5 (2.9)% (mild), 85.8 (16.7)% (moderate) to 77.5 (11.9)% (severe) (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%) (no-OA) to 73.0% (25.9%) (severe, $P<0.001$) 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 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. 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. There is a recognized link with obesity, especially with knee OA but also with hip and, curiously, hand OA (3-6). This latter finding indicates that increased weight-bearing is not the primary problem but that metabolic factors may be more important (7). Epidemiological studies have suggested a systemic aetiology independent of weight-bearing (8, 9). It has been proposed that generalised OA may be a systemic disorder affecting the whole musculoskeletal system driven by lipid metabolism (10). 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 (11) and increased bone containing an altered profile of growth factors in the iliac crest, remote from the weight-bearing regions of the joints (12).
Bone forming osteoblasts share a common mesenchymal stem cell precursor with adipocytes. Defective co-regulation and alterations in lipid metabolism are possible mechanisms for the bone pathologies observed in both OA and osteoporosis (10, 13). A greater concentration of fatty acids has been reported in cartilage of OA patients (14) and an increased fat content in bone marrow from osteoarthritic femoral heads (15). Not only was the mass of fat per unit volume of bone tissue doubled in tissue from OA patients (15), 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 (15). Reports of pilot studies using MRI, have demonstrated a difference in the lipid fractions in femoral bone marrow and muscles around the hip joint between healthy volunteers and OA patients (16).

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 (17) and newer MR methods such as dGEMRIC (18, 19), assessment of glycosaminoglycan content using saturation transfer (20) and T1rho (21) still focus on changes in cartilage, but the detection of bone marrow lesions (22, 23) is broadening the whole area of MR imaging biomarkers (24). In this study we use MRI to measure the fractional fat content and its distribution within the bone marrow in the proximal femur in patients with different severities of osteoarthritis with a view to assessing its potential as an imaging biomarker.
Materials and methods

Subject recruitment

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 imaging. In addition, patients with end-stage OA were recruited from the pre-operative assessment clinic at which they were given an information pack to take home as above and invited to participate.

The study was approved by local Ethics Committee (reference 08/S0802/170) and each subject gave their informed consent in writing according to the declaration of Helsinki.

Radiographic grading
The radiographs of patients who consented to participate in the study were assessed by a single reader blinded to clinical diagnosis. Each hip was assigned a grade for OA severity using the Kellgren-Lawrence system (KLG) (25-27).

We classified subjects as no-OA (KLG0 or 1), mild OA (KLG2), moderate OA (KLG3) and severe OA (KLG4), based on the KLG of their most severe hip joint. If both hips were graded the same the right hip was used in subsequent analyses. Those undergoing total hip arthroplasty (THA) were not given a radiographic grade but were included with the severe OA group.

**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 (28) and tested on phantoms comprising tubes containing different ratios of soya oil and water. The implementation was a 3-echo technique based on the IDEAL sequence (29) and used three gradient-echo images acquired sequentially, with the initial TE equal to 2.1 ms, with additional images having TE increased by an increment (ΔTE) of 0.76 ms and then by 2ΔTE. This corresponds to the initial TE value having water and fat in phase, with the early and late TE values having water and fat ±120° 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.

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. (30). A Bland and Altman analysis of agreement (31) 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 were segmented into one acetabular and six femoral regions of interest (ROI) using statistical shape modelling (32-35) (Figure 1). The statistical distribution of values of fractional fat content, given by 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

\[
S = \frac{1}{N} \sum -x_i \ln(x_i)
\]

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 Kruskall-Wallis ANOVA on ranks was performed.

Tests were corrected for sex.

**Results**

A total of 67 subjects (39 female and 28 male) took part in this cross-sectional study, 55 of whom were identified from the RIS and were classified from their grade of radiographic OA. From the radiographs, 33 hips were classified as no-OA (KLG0 or 1), 9 as mild OA (KLG2), 11 as moderate (KLG3) and 2 as severe (KLG4). To the severe group were added 12 patients about to undergo total hip arthroplasty for primary OA. Table 1 shows the results of tests for differences in the distribution of 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 with sex. Accordingly, subsequent tests for significance were corrected for sex.
Representative images of each of the four grades of severity are shown in Figure 2. Characteristics of the fractional-fat distribution in the total femur are shown in Figure 3 and were not very sensitive to increasing severity. 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 with the most severe OA 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 with severe OA (\( P = 0.022 \)).

Considering each region separately, significant differences in 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 OA. There was no difference between males and females and the mean value reduced from 97.9% (2.5%) in those with no-OA to 73.0% (25.9%) in the severe OA group (\( P < 0.001 \)), with the largest change occurring at the final severity stage. 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 less significant 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 no-OA group to 80.6% [70.8%, 85.0%] in those with severe OA (\( P = 0.042 \)) and, in ROI 4, from 88.6% [80.4%, 91.9%] in no-
OA to 68.7% [66.6%, 82.2%] in severe OA ($P = 0.021$). In all of these, the largest differences appeared with the final stage from moderate to severe OA. No significant differences in mean 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; BMLs had diffuse edges whereas cysts were taken to be those regions of high signal in the ‘water’ images with demarcated boundaries. 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**

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. Calculations of the fractional fat content in the bone marrow from the MR signals support the results from a previous study (16) that there is a significantly smaller fractional-fat content in the proximal femur of patients with OA. This difference is most apparent in those with severe 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. Red marrow is approximately 40% lipid and yellow marrow is
80% lipid (36) and the values we have obtained are comparable with those previously reported for yellow marrow.

A limitation of the Dixon method is that only fractional contents of water and fat can be measured, not absolute values. The ratio of fat to water is, therefore, smaller in the OA femoral heads than the controls. BMLs and cysts had high water contents and resulted in a reduction in the calculated mean 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 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 (37), 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 (37). 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.

Further limitations arise 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.
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 (38). 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 (15) 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 (15); 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 (11), 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 (11). So, along with more bone and smaller spaces between trabeculae there is an increase in water within the bone matrix that will contribute to the MR signal. 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 show little variation with increasing severity of OA until the final stages of the disease; 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 severe OA 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.
References


10. Aspden RM, Scheven BAA, Hutchison JD. Osteoarthritis is a systemic disorder involving stromal cell differentiation and lipid metabolism. The Lancet 2001;357:1118-1120.


15. Plumb MS, Aspden RM. High levels of fat and (n-6) fatty acids in cancellous bone in osteoarthritis. Lipids Health Dis 2004;3:12.


33. 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...


**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 more females in the no-OA group and too few in the mild and moderate groups.

<table>
<thead>
<tr>
<th>Severity (KLG)</th>
<th>N</th>
<th>Sex</th>
<th>Age</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N female (%)</td>
<td>Mean (SD)</td>
<td>(kg / m²)</td>
</tr>
<tr>
<td>No-OA (KLG0,1)</td>
<td>33</td>
<td>25 (76%)</td>
<td>62.6 (11.3)</td>
<td>27.8 (4.2)</td>
</tr>
<tr>
<td>Mild (KLG2)</td>
<td>9</td>
<td>3 (33%)</td>
<td>65.6 (7.1)</td>
<td>27.4 (4.0)</td>
</tr>
<tr>
<td>Moderate (KLG3)</td>
<td>11</td>
<td>4 (36%)</td>
<td>67.7 (6.6)</td>
<td>27.3 (3.3)</td>
</tr>
<tr>
<td>Severe (THA or KLG4)</td>
<td>14</td>
<td>7 (50%)</td>
<td>69.0 (10.1)</td>
<td>28.5 (4.4)</td>
</tr>
<tr>
<td>Total</td>
<td>67</td>
<td>39 (58%)</td>
<td>65.2 (10.1)</td>
<td></td>
</tr>
</tbody>
</table>

*P*-value: 0.029, 0.19, 0.89
**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 images for each of the four degrees of severity classified either by radiographic KLG or by THA. No-OA (KLG0 or 1), Mild OA (KLG2), Moderate OA (KLG3), Severe OA (KLG4 or 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 as function of OA severity.

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 as function of OA severity.

Figure 5. Graphs of (a) mean, (b) standard deviation and (c) entropy of the fractional fat distribution in the femoral head (ROI 2) as function of OA severity.