Good sensitivity and specificity of ultrasound for detecting pseudotumors in 83 failed metal-on-metal hip replacements

Olli LAINIALA, Petra ELO, Aleksi REITO, Jorma PAJAMÄKI, Timo PUOLAKKA, and Antti ESKELINEN

Coxa Hospital for Joint Replacement, Tampere, Finland.
Correspondence: antti.eskelinen@coxa.fi
Submitted 2014-07-25. Accepted 2014-11-01.

Background and purpose — Ultrasound is used for imaging of pseudotumors associated with metal-on-metal (MoM) hips. Ultrasound has been compared with magnetic resonance imaging, but to date there have been no studies comparing ultrasound findings and revision findings.

Methods — We evaluated the sensitivity and specificity of preoperative ultrasound for detecting pseudotumors in 82 patients with MoM hip replacement (82 hips). Ultrasound examinations were performed by 1 of 3 musculoskeletal radiologists, and pseudotumors seen by ultrasound were retrospectively classified as fluid-filled, mixed-type, or solid. Findings at revision surgery were retrieved from surgical notes and graded according to the same system as used for ultrasound findings.

Results — Ultrasound had a sensitivity of 83% (95% CI: 63–93) and a specificity of 92% (CI: 82–96) for detecting trochanteric region pseudotumors, and a sensitivity of 79% (CI: 62–89) and a specificity of 94% (CI: 83–98) for detecting iliopsoas-region pseudotumors. Type misclassification of pseudotumors found at revision occurred in 8 of 23 hips in the trochanteric region and in 19 of 33 hips in the iliopsoas region.

Interpretation — Despite the discrepancy in type classification between ultrasound and revision findings, the presence of pseudotumors was predicted well with ultrasound in our cohort of failed MoM hip replacements.

Several authors have reported periarticular soft tissue lesions called pseudotumors in association with metal-on-metal (MoM) hip replacements (Pandit et al. 2008, Langton et al. 2011). Many patients with adverse reactions to metal debris (ARMD) present with elevated blood cobalt (Co) and chromium (Cr) levels and symptoms such as pain and discomfort in the hip and groin region (Toms et al. 2008, Hart et al. 2011). Some patients are asymptomatic, however, with normal metal ion levels—but they have still developed a pseudotumor (Toms et al. 2008, Wynn-Jones et al. 2011, Hart et al. 2012). Magnetic resonance imaging (MRI) and ultrasound (US) are the main radiological imaging modalities for pseudotumors in MoM hips (Ostlere 2011). MRI produces artifacts because of the metal implant. Metal artifact-reduction sequences have been used to reduce this, but residual artifact remains and obscures the tissues that are directly adjacent (Sofka et al. 2006). US images are not compromised by metal artifacts, and US provides images with good soft tissue resolution of both intracapsular lesions and extracapsular pseudotumors. However, US may have limited value in the evaluation of deep lesions (Ostlere 2011). Most of the recent publications on cross-sectional imaging of MoM hips have concentrated on evaluation of MRI as the main imaging tool in pseudotumor diagnostics. The sensitivity and specificity of US have been reported in relation to MRI (Nishii et al. 2014, Siddiqui et al. 2014, Garbuz et al. 2014), but to our knowledge, these have not been compared to revision findings.

We evaluated the sensitivity and specificity of US for detecting pseudotumors in a cohort of patients with failed MoM hip replacements. A secondary aim was to determine how well the preoperative US classification of pseudotumors would match the perioperative surgical findings. To achieve this, we compared preoperative US findings with perioperative surgical findings in patients with MoM hip replacements who underwent revision surgery at our institution.

Materials and methods

Study population
In 2010, the United Kingdom Medicines and Healthcare products Regulatory Agency announced a medical device alert regarding ASR hip replacement implants (Articular Surface
Replacement; DePuy Orthopaedics, Warsaw, IN) (MHRA 2010). After the announcement, we established a mass-screening program to identify possible articulation-related complications in patients who had undergone either ASR hip resurfacing (HR) or ASR total hip replacement (THR) at our institution. As part of the screening, all patients were referred for cross-sectional imaging. MRI was our primary imaging modality, but if MRI was contraindicated or could not be done due to patient-related factors, US was used. Systematic screening of MoM component models other than ASR began in January 2012. Screening included whole-blood metal ion measurements, Oxford hip score questionnaire, and clinical examination as with ASR patients, but systematic cross-sectional imaging was not performed. If a patient was symptomatic or if whole-blood Co or Cr levels were higher than 5 ppb (Hart et al. 2011).

MoM hip replacements were used in 2,904 operations at our institution between January 2001 and November 2011. Before May 2013, 433 MoM hips (in 397 patients) had undergone revision surgery. Pre-revision US imaging was available for 125 hips (117 patients). We excluded those hips that had been imaged with US more than 12 months before revision surgery; this decision was based on our previous study, in which the sensitivity of MRI to detect pseudotumors remained at the same level up to 1 year after imaging and decreased thereafter, most likely due to the evolving nature of pseudotumors (Lainiala et al. 2014). Thus, we assumed that US imaging performed more than 12 months before revision surgery would also be unreliable. 116 hips (in 109 patients) had undergone US less than a year before revision surgery. To reduce bias caused by previous imaging, we ruled out hips that had had MRI examination less than 1 year before US examination (30 hips, 27 patients). Of the remaining 86 hips (in 82 patients, 50 of whom were females), 22 had undergone MRI over 1 year before US. To avoid bias from clustered observations (i.e. 2 hips in the same patient analyzed as independent observations (Ranstam et al. 2011)), only the right hip of bilateral patients was analyzed. 78 revisions were performed due to ARMD, 2 for infection, 1 for aseptic loosening of the acetabular component, and 1 for malposition of the acetabular component, associated with pain and sensation of subluxation. 82 hips in 82 patients were included in the final analyses. Failure was considered to be secondary to ARMD if metallosis, macroscopic synovitis, and/or extracapsular pseudotumors were found during revision—and/or a moderate-to-large amount of perivascular lymphocytes along with tissue necrosis and/or fibrin deposition was seen in the histopathological sample (Reito et al. 2013). Component loosening and periprosthetic fracture had to be ruled out clinically and radiologically in order to set a diagnosis of ARMD. Infection was ruled out if all (at least 5) culture results from the samples obtained during revision surgery were negative.

The 64 THRs included were 35 ASR implants, 2 M2a-Magnum (Biomet, Warsaw, IN), 3 Mitch (Finsbury Orthopaedics, Leatherhead, UK), 9 Pinnacle (DePuy), 6 R3 (Smith and Nephew, Memphis, TN), 3 ReCap (Biomet), 4 Birmingham Hip Resurfacing (BHR; Smith and Nephew), 1 Durom (Zimmer, Warsaw, IN), and 1 Conserve Plus (Wright Medical Technology, Memphis, TN). There were also 18 hip resurfacings involving 9 ASR and 9 BHR implants. Mean age of the patient at the time of revision was 63 (20–81) years. Mean time between primary operation and revision was 5.4 (1.9–11) years and median time between US and revision was 4.1 (0.3–11) months (Table 1, see Supplementary data).

Revisions
All revisions were performed by or under the direct supervision of 4 experienced hip revision surgeons. Revision was considered if a pseudotumor with solid component was seen by US regardless of symptoms and/or blood Co and Cr levels, if a patient had a symptomatic hip and elevated blood metal ion levels, or if a patient had persistent and/or severe symptoms (Reito et al. 2013). All operations were performed according to a standard protocol involving the posterior approach, paying special attention to the excision of all abnormal metal-contaminated tissue.

Revision findings were retrieved from surgical notes and graded as fluid-filled, solid, or mixed-type based on the surgeon’s description of the consistency, content, and wall thickness of pseudotumors. We considered all extracapsular fluid-filled lesions or mass lesions with variable connection to the joint capsule to be pseudotumors. Cystic lesions with thin walls were graded as fluid-filled pseudotumors. Extracapsular lesions with no or only minor fluid-like component were graded as solid pseudotumors. Mixed-type was defined as being mainly fluid-filled, but also having thick walls and/or solid contents. We divided pseudotumors into iliopsoas (anterior) region pseudotumors and trochanteric (posterior) lateral region pseudotumors, which in our experience are the 2 most usual locations for extracapsular pseudotumors in MoM hips.

Ultrasound examination
Each examination was performed with a Logiq E9 ultrasound machine (GE Healthcare) by 1 of 3 musculoskeletal radiologists. In most cases, an ML 6-15-D linear transducer was used (4.5–15.0 MHz, 13 × 58 mm footprint, field of view (FOV) 50 mm, and depth of field (DOF) 8 cm). However, in cases with poor visibility due to obesity, a 9L-D linear transducer (2.4–10.0 MHz, 14 × 53 mm footprint, FOV 45 mm, and DOF 12 cm) or a C1-5 convex transducer (1.6–6.0 MHz, 17 × 75 mm footprint, FOV 65 degrees, and DOF 35 cm) was used. An anterior approach was used to evaluate the hip joint, the iliopsoas muscle, and the tendon region. A lateral approach was used to evaluate the greater trochanteric and deep fascia region. Posteriorly, the hip joint and the trochanteric and gluteal regions were examined. Pseudotumors were classified retrospectively according to the same classification as for the
revision findings based on the consistency, content, and wall thickness of pseudotumors described by the radiologist (Figures 1–3). Intracapsular findings were not classified systematically and were therefore excluded from analysis. Doppler was not used in examinations. Synovial fluid aspirates were not acquired routinely during the US examination.

**Statistics**

Sensitivity, specificity, positive predictive value, and negative predictive value with 95% confidence intervals (CIs) were calculated for detection of trochanteric and iliopsoas-region pseudotumors with US (Herbert 2013). Kappa coefficient was calculated for statistical comparison of differences in classification between US and revision surgery findings. To assess the effect of time between US examination and revision surgery on the calculated sensitivity and specificity, we divided the study population into 2 equal-sized groups based on the median time between US and revision (3.5 months). Sensitivity, specificity, and kappa coefficient with 95% CIs were calculated for each group separately. We used IBM SPSS Statistics version 20.

**Ethics**

The institutional review board approved the study (April 27, 2011; R11006) and the procedures followed were in accordance with the Helsinki Declaration.

**Results**

Pseudotumors were found in 46 of 82 hips during the revision surgery. A trochanteric region pseudotumor was found in 13 hips and an iliopsoas region pseudotumor was found in 23 hips. In 10 hips, a pseudotumor was found in both the iliopsoas region and the trochanteric region.

| Table 2. Summary of test characteristics |
|----------------------------------------|
| Trochanteric region | Iliopsoas region |
|---------------------|------------------|
| Sensitivity (95% CI) | 83% (63–93)      | 79% (62–89)      |
| Specificity (95% CI) | 92% (82–96)      | 94% (83–98)      |
| Positive predictive value (95% CI) | 79% (59–91)      | 90% (74–96)      |
| Negative predictive value (95% CI) | 93% (84–97)      | 87% (76–94)      |

**Trochanteric region pseudotumors**

Preoperative US showed 19 of the 23 posterolaterally located pseudotumors found during revision surgery. The 4 pseudotumors that were not detected with US were fluid-filled and thin-walled (here, mean time between US and revision was 4.5 (2.7–7.3) months). 4 fluid-filled pseudotumors and 1 mixed-type pseudotumor were also seen in preoperative US but they were not found during revision surgery (here, mean time between US and revision was 4.5 (2.1–7.3) months for the fluid-filled pseudotumors and the actual time was 4.4 months for the mixed-type pseudotumor). Thus, US had a sensitivity of 83% (95% CI: 63–93) and a specificity of 92% (95% CI: 82–96) for detecting pseudotumors in the trochanteric region (Table 2). Cross-tabulation of US and revision findings in trochanteric region is given in Table 3.

14 of the 18 fluid-filled pseudotumors found in revision surgery were correctly classified in US also. None of the 4 mixed-type pseudotumors were correctly graded in preoperative US examination. The kappa coefficient calculated from Table 3 was 0.64 (good agreement; 95% CI: 0.47–0.80).

**Iliopsoas-region pseudotumors**

Preoperative US revealed 26 of the 33 anteriorly located pseudotumors found in revision surgery. All 7 pseudotumors that were not detected with US were fluid-filled (here, mean time between US and revision was 3.8 (1.3–6.1) months). 1 solid
and 2 fluid-filled pseudotumors seen in preoperative US were not found during revision surgery (here, time between US and revision was 3.2 months for the solid pseudotumor and 2.1 and 8.0 months for the fluid-filled pseudotumors). US had a sensitivity of 79% (95% CI: 62–89) and a specificity of 94% (95% CI: 83–98) for detecting pseudotumors located in the iliopsoas region (Table 2). Cross-tabulation of US and revision findings in the iliopsoas region is given in Table 4.

8 of the 21 fluid-filled pseudotumors and 6 of the 8 mixed-type pseudotumors found in revision surgery were accordingly classified in preoperative US also. Of the 4 solid pseudotumors found in revision, 3 were classified as mixed-type and 1 was classified as fluid-filled by US. The kappa coefficient calculated from Table 4 was 0.52 (moderate agreement; 95% CI: 0.38–0.66).

### The effect of time between US and revision surgery on sensitivity, specificity, and kappa coefficient

There was no statistically significant difference when we compared the patients with 3.5 months or less between the US and revision to the patients for whom the time interval in question exceeded 3.5 months.

### Discussion

To our knowledge, this is the first study to compare pseudotumors detected by US to those encountered in revision surgery. To date, the sensitivity and specificity of US have been estimated in relation to MRI (Nishii et al. 2014, Siddiqui et al. 2014, Garbuz et al. 2014). Siddiqui et al. (2014) compared US findings in 19 MoM hips to findings by MRI, which they considered to be the gold standard. They reported poor sensitivity in detecting pseudotumors with US relative to MRI (69%) and recommended that MRI should be used as the first-line examination because of the higher accuracy. They also stressed the anatomical information provided by MRI, which can be used for preoperative planning. Nishii et al. (2014) also evaluated the sensitivity and specificity of US by comparing it to MRI. They found a sensitivity of 74% and a specificity of 92%, with US failing to detect 7 of 27 abnormal lesions detected with MRI and MRI failing to detect 3 of 23 lesions seen with US. The authors stated that they considered MRI to be a more reliable screening method but that they considered US to be a primary screening tool due to its better availability, lower cost, and possibly more reliable detection of small lesions. Garbuz et al. (2014) evaluated the sensitivity and specificity of US and MRI in 40 MoM hips by determining the agreement between them using MRI as gold standard. They found a sensitivity of 100% and a specificity of 96% for US and a sensitivity of 92% and a specificity of 100% for MRI, and they argued that US should be used as the primary imaging modality due to the significantly lower costs.

In the present study, 11 fluid-filled pseudotumors were found at revision but not at the preoperative US, and 6 fluid-filled pseudotumors were seen at the US but not at revision. 1 mixed-type pseudotumor seen by US in the trochanteric region and 1 solid pseudotumor seen in the iliopsoas region were also not found during revision surgery. All cases with false-negative US findings had thin-walled and fluid-filled pseudotumors. In a recent study, Almousa et al. (2013) tried to analyze the natural history of inflammatory pseudotumors and they found that asymptomatic pseudotumors frequently increased and decreased in size, with occasional remission of small masses. In our previous study, we found low sensitivity for MRI images that were over 1 year old, most likely due to the developing nature of lesions (Lainiala et al. 2014). The actual change in lesions might explain the false negatives and false positives, and also some of the misclassified cases in the present study. In this study, there was no statistically significant difference in sensitivity and specificity between the patients with ≤ 3.5 months and > 3.5 months between the US and the revision, but there was a small number of patients in the subgroup analysis. Furthermore, the lesions lay deep in patients with excessive subcutaneous tissue, which may also explain some false-negative findings.

Even though it has been suggested that asymptomatic fluid-filled pseudotumors may be of less importance clinically (Hart et al. 2012), the natural history of these lesions is still unclear. Almousa et al. (2013) found that some of the asymptomatic pseudotumors increased in size, transforming from cystic to solid, and they found abductor and iliopsoas muscle damage...
in a few cases with increasing pseudotumor size. Furthermore, Grammatopolous et al. (2009) reported poor results for MoM hip revisions performed due to pseudotumors. They also speculated that one of the reasons for the poor outcome of such revisions might be the excessive tissue resection needed in cases with vast soft tissue abnormalities. It therefore seems reasonable to follow up patients with cystic lesions by repeated cross-sectional imaging. In our opinion, this is important to detect and revise aggressively expanding lesions early enough to minimize soft tissue destruction.

The study had some limitations. Each patient was imaged only once. Thus, we were unable to assess inter- and intra-observer reliability. At the time that the US examinations and revisions were performed, there was no published pseudotumor classification for US or perioperative findings. On the basis of typical findings encountered during revision surgeries of failed MoM hips at our institution and previous MRI classifications (Anderson et al. 2011, Hart et al. 2012, Hauptfleisch et al. 2012), we decided to classify pseudotumors as being fluid-filled, mixed-type, or solid. A similar description was used for US findings. Due to the retrospective analysis of the US findings, the re-grading of pseudotumors was not done in blinded fashion and may have been biased. We tried to reduce the bias caused by previous imaging results seen by radiologists who performed the US examinations by excluding the patients with MRI performed shortly before US. The most important reason for misclassification of pseudotumors in US is probably the lack of a prospective grading scheme for the perioperative findings — i.e. different surgeons may have described the lesions differently. Even though the type of pseudotumor was often misclassified, we consider that reporting of the presence or absence of pseudotumors was reliable since our institution recognized the problem with MoM hip-related pseudotumors early, and the surgeons were aware that soft tissue pathologies might be encountered. Our study cohort included revised patients only. Thus, the prevalence of pseudotumors was certainly higher in this study cohort than in the whole MoM population. Moreover, we do not know the number of possible false-negative findings in asymptomatic patients who also had normal blood metal ion levels and were therefore not considered for revision surgery. Due to this fact, the sensitivity of US may appear higher than it really is. The surgeons performing the revision operations were aware of the US findings, which may have affected the way in which they described the lesions that they found, but this is an in-built problem with this type of study.

In summary, the presence of pseudotumors was predicted well with US in our cohort of failed MoM hips. However, there was discrepancy in the classification of lesions in US examination and in revision surgery.

Supplementary data
Table 1 is available at the Acta Orthopaedica website (www.actaorthop.org), identification number 7791.
Ostlere S. How to image metal-on-metal prostheses and their complications. AJR Am J Roentgenol 2011; 197 (3): 558-67.
Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons C L, Ostlere S, Athanasou N, Gill H S, Murray D W. Pseudotumours associated with metal-on-metal hip resurfacing. J Bone Joint Surg Br 2008; 90 (7): 847-51.
Ranstam J, Karrholm J, Pulkkinen P, Makela K, Espehaug B, Pedersen A B, Mehnert F, Furnes O. Statistical analysis of arthroplasty data. II. guidelines. Acta Orthop 2011; 82 (3): 258-67.
Reito A, Puolakka T, Elo P, Pajamaki J, Eskelinen A. High prevalence of adverse reactions to metal debris in small-headed ASRTM hips. Clin Orthop Relat Res 2013; 471 (9): 2954-61.
Siddiqui I A, Sabah S A, Satchithananda K, Lim A K, Cro S, Henckel J, Skinner J A, Hart A J. A comparison of the diagnostic accuracy of MARS MRI and ultrasound of the painful metal-on-metal hip arthroplasty. Acta Orthop 2014; 85(4): 375-82
Sofka C M, Potter H G, Adler R S, Pavlov H. Musculoskeletal imaging update: Current applications of advanced imaging techniques to evaluate the early and long-term complications of patients with orthopedic implants. HSS J 2006; 2 (1): 73-7.
Toms A P, Marshall T J, Cahir J, Darrah C, Nolan J, Donell S T, Barker T, Tucker J K. MRI of early symptomatic metal-on-metal total hip arthroplasty: A retrospective review of radiological findings in 20 hips. Clin Radiol 2008; 63 (1): 49-58.
Wynn-Jones H, Macnair R, Wimhurst J, Chirodian N, Derbyshire B, Toms A, Cahir J. Silent soft tissue pathology is common with a modern metal-on-metal hip arthroplasty. Acta Orthop 2011; 82 (3): 301-7.