Are ultrasound and magnetic resonance imaging of value in assessment of Achilles tendon disorders? A two year prospective study

K M Khan, B B Forster, J Robinson, Y Cheong, L Louis, L Maclean, J E Taunton


Objectives: To (a) compare ultrasound (US; including grey scale and colour and power Doppler) and magnetic resonance imaging (MRI; with high resolution and fat saturation sequences) with a clinical yardstick in the evaluation of chronic Achilles tendinopathy, and (b) examine whether either imaging method predicted 12 and 24 month clinical outcome.

Methods: Forty five patients with symptoms in 57 Achilles tendons were diagnosed with tendinopathy by an experienced sports medicine doctor. All patients underwent US examination (12 MHz probe) with colour and power Doppler, and 25 consecutive patients also underwent MRI with high resolution T1 weighted and STIR sequences.

Results: US identified abnormal morphology in 37 of the 57 symptomatic tendons (65%) and normal morphology in 19 of 28 asymptomatic tendons (68%). Baseline US findings did not predict 12 month clinical outcome. The addition of colour and power Doppler did not improve the diagnostic performance of US. MRI identified abnormal morphology in 19 of 34 symptomatic tendons (56%) and normal morphology in 15 of 16 asymptomatic tendons (94%). Lesser grades of MR signal abnormality at baseline were associated with better clinical status at 12 month follow up.

Conclusions: US and MRI show only moderate correlation with clinical assessment of chronic Achilles tendinopathy. Graded MRI appearance was associated with clinical outcome but US was not.

Both ultrasound (US) and magnetic resonance imaging (MRI) can provide excellent morphological information in patients with overuse injury of the Achilles tendon. Imaging can undoubtedly distinguish the various structural abnormalities that can be found around the Achilles insertion.

Despite these apparent benefits, the extreme sensitivity of MRI, in particular, means that structural abnormalities detected by imaging may not correlate precisely with symptoms. The few studies that have examined both asymptomatic and symptomatic Achilles tendons have found varying degrees of overlap of imaging appearances between these clinically “normal” and “abnormal” tendons. Thus, it is acknowledged that the pathophysiology of chronic Achilles tendon disorders is complex and the nomenclature confusing.

The traditional role of imaging in clinical practice has included diagnosis and monitoring of clinical progress. There are now both academic and economic pressures to justify use of radiological services. To our knowledge, only one prospective study provides data to suggest that imaging of patients with overuse Achilles tendon disorders may improve patient outcome. Furthermore, no such study has confirmed that radiological monitoring of patient progress has a clinical or a cost benefit. Previous studies of patients with Achilles tendon disorders have not all used high frequency transducers in US assessment, and no dedicated Achilles studies have investigated the use of power and colour Doppler. Early MRI studies did not have access to the optimal surface coils and scan parameters now available for tendon assessment.

Therefore we undertook a prospective clinicoradiological study using optimised US and MRI bilaterally in patients suffering from Achilles tendon disorders of varying severity. The purpose was to compare the two modalities against a clinical yardstick in the assessment of the severity of the condition and determine whether either method was predictive of outcome.

PATIENTS AND METHODS

Study design

Patients underwent clinical assessment and bilateral imaging with US and MRI at baseline. A 12 month follow up consisted of clinical assessment and bilateral US, but no MRI. A two year follow up consisted of clinical assessment by telephone. The study was approved by the local institutional ethics review board, and all study participants provided written informed consent.

Clinical examination

We recruited 45 consecutive patients (27 men, 18 women; mean age 42 years, range 20–66) who attended our secondary referral based university sports medicine centre because of symptomatic Achilles tendinopathy (mean duration 21 months, range 0.5–120). At entry into the study (baseline), 12 patients had bilateral symptoms, and 33 patients had unilateral symptoms for a total of 57 symptomatic and 33 asymptomatic tendons. Of the asymptomatic tendons, 28 had never been symptomatic, and five had been painful previously but not for at least six months. These five tendons are not included in this report. Thus the report includes 85 tendons from 45 patients, with five tendons excluded. The clinical diagnosis was made by the referring clinician and confirmed by a board certified sports medicine doctor. The usual complaint was pain on activity and morning stiffness. Tenderness was found at the mid-tendon in 41 tendons, at the insertion in 12 tendons, and diffusely throughout in four tendons.

Abbreviations: US, ultrasound; MRI, magnetic resonance imaging; VISA-A, Victorian Institute of Sport assessment scale
No patient had a history of systemic inflammatory disorders, such as rheumatoid arthritis, or of hypercholesterolaemia.

Clinical severity
In this prospective study, we assessed clinical status at baseline, 12 months, and 24 months. At baseline, symptoms of Achilles tendon pain were categorised as present/absent; when symptoms were present, the severity was scored according to the Victorian Institute of Sport assessment scale (VISA-A). The scores can range from 0 to 100, where 100 represents no symptoms and perfect function. A change of 25 points is considered clinically significant. At the 12 month clinical follow up, we again recorded clinical history and repeated the VISA assessment. The two year follow up was by telephone. We used “sporting success” and “symptomatic benefit” as outcome measures at that time. Sporting success refers to return to sporting activities at the previous level, and symptomatic benefit refers to self reported improvement compared with baseline.

US examination
US was performed at baseline and 12 months, but not at 24 months. Real time US was performed by one of two radiologists using a high resolution 12 MHz array scanner (Advanced Technology Laboratories 5000, Bothell, Washington, USA). The radiologists were blinded to the clinical findings or other imaging findings. US was performed on the same day as (n = 22) or within one week of (n = 23) the clinical examination. Patients were positioned prone, with their feet hanging over the end of the scanning table in a relaxed position. An acoustic stand off pad or a synthetic gel spacer was not necessary. Sonograms were obtained in the sagittal (longitudinal) plane of the entire length of both tendons, as well as the transverse plane. Particular care was taken to ensure that the scan plane was parallel to the tendon fibres to avoid acoustic fibre anisotropy. Tendon thickness was measured by the maximum anteroposterior diameter in a transverse scan at a neutral position of the talocrural joint. A tendon with a diameter greater than 6 mm was defined as thickened.

A sonographic abnormality was defined as either one or more hypoechoic and/or hyperechoic (typically reported in the clinical setting as calcific) areas evident in both the longitudinal and transverse scans, or an area of thickening with or without hypoechoic areas. We measured hypoechoic areas using electronic calipers. Length was measured on the sagittal image, whereas width (mediolateral dimension) and height (anteroposterior dimension) were measured on the axial image. The approximate volume of each hypoechoic lesion was calculated using the product of length, width, and height. The tendons were also graded using the scheme of Archambault et al., in which grade 1 represents the normal tendon, grade 2 the thickened tendon with a homogeneous echotexture, and grade 3 the tendon with hypoechoic areas, or calcifications, within the tendon with or without thickening.

Retrocalcaneal bursitis was diagnosed if the anteroposterior diameter of the bursa was more than 2 mm. Both colour and power Doppler interrogation were used in all patients for both the entire tendon and the paratenon. The instrument settings for power and colour Doppler examination were optimised for slow flow (velocity and volume). Doppler results were categorised as either normal or abnormal, as there have been no reports of a reliable system for grading vascularity. Pertendinous fluid collections were also recorded.

MRI examination
MRI was performed using a 1.5 T Signa scanner (General Electric, Milwaukee, Wisconsin, USA) on the first 25 consecutive patients who enrolled in the study. Sixteen patients had unilateral and 18 bilateral symptoms totalling 34 symptomatic tendons and 16 asymptomatic tendons. MRI was performed within two weeks of the US and clinical examinations. With the patient supine, we obtained multiple sagittal and axial sequences of both ankles using a quadrature head coil to image both ankles simultaneously. The following sequences were used: for T1 weighted sagittal and axial spin echo imaging, repetition time (TR) = 500 milliseconds, echo time (TE) = 14 milliseconds, section thickness = 3 mm with no interslice gap, field of view (FOV) = 12 cm, matrix = 256 × 256, number of excitations = 2, and imaging time = 4 minutes 24 seconds; for sagittal fast short tau inversion recovery (FSTIR), TR = 4000 milliseconds, TE (effective) = 32 milliseconds, inversion time = 150 milliseconds, FOV = 16 cm × 16 cm, section thickness = 3 mm with no gap, matrix = 256 × 192, number of excitations = 3, and imaging time = 5 minutes 36 seconds; for two dimensional T2* weighted sagittal gradient recalled echo (GRE), TR = 800 milliseconds, TE = 30 milliseconds, flip angle = 70°, FOV = 12.2 cm, section thickness = 3 mm with no gap, matrix = 256 × 256, number of excitations = 1.5, and imaging time = 5 minutes 10 seconds. In-plane resolution was 0.5 × 0.5 mm for T1 weighted scans and 0.6 × 0.8 mm for STIR sequences. Images were read by two radiologists simultaneously in a blinded fashion; disagreement was resolved by consensus.

The size of any intratendinous lesion was measured and the approximate volume of the lesion calculated as a product of the length (craniocaudal dimension on longitudinal plane), width (mediolateral dimension on axial plane), and height (anteroposterior dimension on axial plane). Only foci that showed increased signal on FSTIR were designated as lesions, as normal tendons contain high signal foci on T1 and T2. Lesions were graded from 1 to 3 on a similar manner to the US grading. Grade 1 represents a normal tendon, grade 2 a thickened tendon with homogeneous signal, and grade 3 describes intratendinous signal intensity change. The paratenon was not evaluated.

Data analysis
Statistical analysis was performed using the Statistical Package for the Social Sciences software (SPSS) for Windows (version 7.0). Patient characteristics were analysed using descriptive data. Diagnostic data were analysed on a case by case basis. Grades of severity assessed clinically, by US and MRI were compared using Pearson's correlation coefficient and Spearman's rank correlation for non-parametric data. The relation between clinical and imaging findings was analysed.
using a $\chi^2$ analysis with $2 \times 3$ contingency tables. The relation between severity grades and possible confounding factors was assessed using Pearson’s correlation. Follow up data were analysed using a $\chi^2$ analysis with $2 \times 3$ contingency tables.

**RESULTS**

**US**

US identified abnormal morphology in 37 of the 57 (65%) symptomatic tendons and normal morphology in 19 of the 28 (68%) asymptomatic tendons. Thus, US had a sensitivity of 0.80 and specificity of 0.49 in diagnosing tendinopathy compared with the clinical yardstick. The positive predictive value was 0.65 and negative predictive value 0.68. Table 1 shows the US results for the 50 tendons that were also examined by MRI. Note that US detected an abnormal signal in only two of 16 symptomatic tendons diagnosed clinically as having insertional tendinopathy. Thus, in tendons with midsubstance tendinopathy, US abnormality was present in 35 of 45 (78%), which corresponds with a specificity of 0.66 and a positive predictive accuracy of 0.78.

Neither colour nor power Doppler interrogation enhanced the accuracy of US. Of the 39 tendons with increased colour flow, 21 were thickened or nodular on clinical examination ($r = 0.17; p<0.01$). Positive colour flow correlated with age (Spearman’s $r = 0.56; p<0.01$) but not with presence of symptoms ($r = 0.111; p = 0.74$) or sports participation ($r = 1.05; p = 0.59$).

The total volume of hypoechoic lesions correlated with the US grade of severity ($r = 0.87; p<0.01$), and both of these measures correlated with the VISA-A score (Spearman’s $r = -0.33; p<0.01$ for volume of lesions and $-0.34; p<0.01$ for US classification).

**MRI**

MRI identified abnormal morphology in 19 of the 34 (56%) symptomatic tendons and normal morphology in 15 of 16 asymptomatic tendons (table 1). Thus, MRI had a sensitivity of 0.95, a specificity of 0.50, a positive predictive value of 0.56, and a negative predictive value of 0.94 in diagnosing tendinopathy compared with the clinical yardstick. The false negative cases were clinically milder than the true positive cases (Student’s t test; $p<0.01$), and were less likely to be thickened.

The volume of intratendinous high signal intensity lesions on MRI correlated with the grading for MRI (Spearman’s $r = 0.74; p<0.01$). This volume also correlated significantly with the clinical severity (Spearman’s $r = -0.42; p<0.01$) and the volume of hypoechoic lesions on US (Spearman’s $r = 0.67; p<0.05$).

**Correlation between US and MRI**

Among the 34 symptomatic tendons, both US and MRI were correctly positive in 18 tendons (53%) and falsely negative in 10 tendons (29%) (table 1). The VISA-A scores of the patients with false negative imaging were significantly better—that is, representing less pathology—than those with positive imaging ($p<0.05$). Of 11 asymptomatic tendons, seven were correctly identified as normal on both imaging modalities.

**Clinical outcome**

One year after baseline measurement, 36 of the 45 patients were available for follow up. There were 45 symptomatic and 25 asymptomatic tendons (two tendons were excluded from follow up because they had had symptoms before being “asymptomatic” at baseline). One tendon that had been asymptomatic at baseline had become painful. Of the 45 symptomatic tendons, 14 had improved significantly (>25 points improvement) on the VISA-A scale, and 31 remained stable. None had deteriorated significantly. The baseline clinical VISA-A score correlated with the one year VISA-A score (Pearson’s $r = 0.56; p<0.01$) and the two year VISA-A score (Pearson’s $r = 0.21; p<0.05$).

Two years after baseline measurement, 32 of the original 45 patients were available for follow up, and all but two of them had returned to sport at their previous level. One of the two who had not returned to sport felt she was significantly improved. Thus, 94% of patients available for follow up achieved sporting success (as defined) and 97% obtained symptomatic benefit. The mean (95% confidence interval) VISA score at the two year follow up was 91.8 (88.2 to 95.4).

**Imaging findings and 12 month clinical outcome**

US grade of severity measured at baseline was not significantly associated with outcome at one year ($r^2 = 5.45; p = 0.25$). Similarly, less severe US abnormality at follow up was not associated with clinical improvement at that time ($r^2 = 0.89; p = 0.64$; table 2). Reduced area of abnormal signal on US did not correlate with improved VISA score ($r = -0.10, p = 0.59$).

MRI grade of severity at baseline was significantly associated with VISA-A score at 12 month follow up ($r^2 = 0.54; p = 0.02$). MRI was not performed at follow up, and therefore change in MRI findings could not be compared with clinical follow up.

**DISCUSSION**

**Clinical utility of US and MRI in patients with Achilles tendon symptoms**

In patients with Achilles tendon disorders ranging in clinical severity, we found that US and MRI had similar accuracy.
when clinical criteria were used as the diagnostic yardstick (see below). These findings, in patients whose symptoms did not warrant surgical intervention, complement those of Astrom et al., who found little difference in diagnostic accuracy between US and MRI for patients who had been referred for surgery. We note that US was generally normal in patients who were clinically diagnosed as having insertional tendinopathy.

The greater sensitivity of US (81%) and MRI (96%) in that study may reflect the greater severity of symptoms in those cases. Similarly Karjalainen et al. found a similar sensitivity for MRI of 94%, for patients referred for MRI by orthopaedic surgeons, which is very similar to our findings. These authors also reported a 19% false positive rate among asymptomatic tendons, which reinforces the caution needed in interpreting MRI findings. A limitation of our study was that we did not evaluate paratendinous changes using MRI. Some authors, but not others, have found these changes to be common and important in Achilles tendinopathy.

A recent laboratory experiment adds validity to our finding of false negative US imaging in abnormal tendons. In a cadaver study of surgically produced longitudinal tears of the posterior tibialis tendon, the sensitivity of US was 67%, and that of MRI was 77.5%. We note, however, that Karjalainen et al. reported high sensitivity and specificity and significant ability to predict outcome using specific MRI sequences including FLASH sequences in a large, prospective study.

In our study, the value of US was not enhanced by the addition of colour and power Doppler interrogation. Colour and power Doppler sonography can depict high volume flow in low velocity areas such as the musculoskeletal soft tissues. Because grey scale sonography is operator dependent, it was hoped that the addition of colour or power Doppler assessment would add objective evidence of pathology. These techniques have been reported in studies imaging the patellar tendon.

In this investigation of Achilles tendons, we found that all tendons with positive colour flow and power Doppler also had positive findings on the grey scale. This finding contradicts speculation that power Doppler may be more suitable because it depends on scan angle. Until the histological cause of positive colour and power Doppler sonography is identified, the mechanism for increases in flow remains speculative. However, we read with great interest that Swedish researchers suggest an association between neovascularisation in tendinopathy and pain. These authors successfully treated abnormal vessels visualised by colour Doppler with sclerosis.

We used the volume of intratendinous abnormalities seen on US to order imaging. Although this has also been used by Movin et al., the reproducibility and validity of this measurement has not been assessed. Despite this limitation, imaging grade of severity correlated well with clinical severity. However, no additional information was obtained that was not evident clinically.

Determination of gold standard for clinicoradiological tendon research

Although imaging has often been used as the yardstick in radiological studies, we note that Kannus and Jozsa found histopathological abnormalities in 160 of 445 previously healthy cadaver Achilles tendons. Furthermore, many authors have described abnormal signals in MR images of asymptomatic, very active athletes. Thus, it can be argued that careful clinical assessment may provide a relevant yardstick against which to compare the value of imaging. Imaging and prediction of clinical outcome

This study was performed in a university based referral sports medicine practice staffed by sports medicine doctors experienced in treating tendinopathies. Thus, it is possible that the participants in the study reflect a “filtered” patient group. Nevertheless, none were referred for surgery either initially or within the two year follow up period. In this population, US imaging was unable to differentiate between cases that would improve and those that would worsen. Despite normal US imaging, some patients were worse at follow up. This is in contrast with the studies of Mathieson et al. and Nehrer et al., but more in keeping with the findings of Astrom et al. and Marcus et al. in the Achilles tendon. Several researchers who evaluated the patellar tendon in a similar fashion with both US and MRI describe a similar phenomenon—that is, symptoms do not mirror appearances at baseline, and US change over time is not closely associated with change in symptoms. From these findings, we infer that clinicians should not use imaging to monitor patient progress or determine readiness to return to sport. Our data suggest that clinical judgment is still required in this setting.

Severity of baseline MR was associated with 12 month clinical status by VISA score, and this is consistent with the results of Karjalainen et al., who reported a worse prognosis in patients with MRI proven insertional tendinosis or focal intratendinous lesions. Treatment was not controlled in either of these studies, and thus must be considered a potential confounding variable. If future studies closely control the treatment protocol prescribed for various pathologies, it would add strength to any finding of an association between imaging appearance and clinical outcome.

Some authors advocate the use of T1 weighted gradient echo sequences in the evaluation of the Achilles tendon, stating that the highest intrinsic signal in tendons is achieved with short TE gradient sequences. Although we did not use T1 weighted gradient echo sequences, we did use T2* weighted gradient echo sequences, with a longer TE of 30 milliseconds, but obtained at relatively high resolution (0.47 × 0.47 mm, slice thickness 3 mm).

We conclude that clinicians should exercise discretion in ordering imaging tests and in interpreting their findings in chronic Achilles tendon disorders, as both US and MRI commonly result in false positive and false negative diagnoses. As always, careful clinical correlation with imaging findings is essential. Selective use of imaging in patients who fail to respond clinically and who may have clinically underestimated tendon changes may be warranted. Although our findings suggest that imaging adds little information of use for expert sports medicine clinicians in diagnosing tendinopathy, it may be useful for inexperienced clinicians who are unsure of their diagnoses and unfamiliar with grading schemes such as VISA-A.

We note that most patients did well over a two year period, and that most patients who were still symptomatic at 12 months improved during the second year of observation. This is consistent with the long term follow up study of Paavola et al. reporting mainly good outcomes in patients with Achilles tendinopathy. Our study was not designed to evaluate various types of treatment.

With respect to outcome prediction, we found that both clinical baseline score and MRI severity were associated with 12 month clinical outcome. Baseline US appearance did not predict 12 month clinical outcome. The hypothesis that Achilles tendon MRI can contribute to improved outcome by enhancing diagnosis needs to be tested in further prospective studies while treatment is carefully controlled.
Assessment of Achilles tendon disorders

References