Doppler ultrasound – a valid and reliable tool to assess spondyloarthritis

Sandra Falcão¹, Eugenio De Miguel¹, Concepción Castillo², Jaime C Branco³, Emilio Martín-Mola²

ABSTRACT

Enthesitis is the hallmark of spondyloarthritis and is observed in all subtypes. Namely, a wide information on spondyloarthritis abnormalities, including synovitis, bursitis, tendinitis, enthesitis and cortical bone abnormalities (erosions and enthesophytes), can be efficiently perceived by ultrasound power Doppler. Furthermore, several studies on imaging of enthesis showed that imaging techniques are better than clinical examination to detect pathology at asymptomatic enthesis. Vascularized enthesitis detected by ultrasound power Doppler appears to be a valuable diagnostic tool to confirm spondyloarthritis diagnosis. This article focuses on the validity and reliability of ultrasound enthesitis assessment in the management of spondyloarthritis patients.

Keywords: Ultrasound; Power Doppler; Enthesitis; Spondyloarthritis

BACKGROUND

There is an increasing interest among rheumatologists for using ultrasonography (US) as an investigation and management tool for musculoskeletal disorders¹. This imaging modality has not only a number of advantages, such as low cost, good accessibility, and ability to dynamic real-time assessment of multiple joints in relatively short period of time, but also the ability to detect and monitor bone and joint soft tissue inflammation and its structural sequelae². A growing body of evidence in many rheumatic conditions, such as rheumatoid arthritis (RA) and ankylosing spondylitis (AS), suggests that US is a high sensitive and non-invasive tool in the detection and monitoring of early synovitis, bone erosions, tenosynovitis and enthesitis³⁴. The early detection of inflammatory joint pathologies would ideally allow clinicians to initiate relevant therapies in order to prevent destruction of bone and joint soft tissue, and, subsequently, improve morbidity and long-term outcome.

Spondyloarthritis (SpA) is a group of disorders that are characterized by inflammatory involvement of the enthesis and the adjacent bone³. Enthesitis is regarded as the primary lesion and is observed in all SpA subtypes, and may sometimes be present several years as an isolated clinical manifestation⁶⁻⁸. Encouraging data suggest that US enthesis scores could be used as a valid tool for SpA assessment⁹,¹⁰. Chronic inflammatory low back pain and radiographic changes involving the sacroiliac joints are key diagnostic features for AS according to the modified New York (MNY) criteria established in 1984¹¹. Although the MNY criteria have been widely used in both clinical and research settings the absence of both radiographic sacroiliitis and impaired spinal mobility at early stages of the disease have contributed to a long delay (5-10 years) in AS diagnosis in many patients¹². Recently the Assessment of SpondyloArthritis International Society (ASAS) developed new classification criteria for axial SpA that demonstrated much better specificity compared to the European Spondylarthropathy Study Group (ESSG) criteria modified for Magnetic Resonance Image (MRI) (sensitivity 85.1%, specificity 65.1%) and slightly superior to the modified Amor criteria (sensitivity 82.9%, specificity 77.5%)¹³. Nevertheless the sensitivity of ASAS classification criteria for peripheral SpA was slightly worse compared to the registered for axial SpA ASAS criteria (77.8% versus 82.9%)¹⁴. Despite these encouraging data, several studies on imaging of enthesis showed that imaging techniques, such as MRI or US are superior to clinical examination, and frequently pathology at asymptomatic enthesis might only be detected by imaging te-
Furthermore, it has been suggested that US might be superior to MRI for detecting early signs of enthesopathy\(^{16,17}\).

Despite the promising results about US examination of the enthesis in SpA, further research should be prompt to assess its specific role in diagnosis and follow up of disease course. This article focuses on the validity and reliability of US enthesitis assessment in the management of SpA patients.

**VALIDITY OF ULTRASOUND ENTHESITIS ASSESSMENT IN SPONDYLOARTHRITIS**

US provide a widely information on SpA abnormalities, including synovitis, bursitis, tendinosis, enthesitis and cortical bone abnormalities (erosions and enthesophytes).

The Outcome Measures in Rheumatology Clinical Trials (OMERACT) defines enthesopathy as an “abnormally hypoechoic (loss of normal fibrillar architecture) and/or thickened tendon or ligament at its bony attachment (may occasionally contain hyperechoic foci consistent with calcification), seen in 2 perpendicular planes that may exhibit Doppler signal and/or bony changes including enthesophytes, erosions, or irregularity” (Figure 1)\(^{18}\). Although grey-scale US depict enthesis structural damage, it seems that the combination with power Doppler increases diagnostic accuracy for SpA. The first description was made by D’Agostino and colleagues who demonstrated that US, in grey-scale combined with power Doppler, allowed

**FIGURE 1.** Ultrasonographic images of Achilles enthesopathy: (A) Longitudinal view: thickening, loss of fibrillar pattern, bone proliferation and erosions. (B) Transversal view: loss of fibrillar pattern and erosions. (C) Longitudinal view: thickening, loss of fibrillar pattern, bursa, erosions and Doppler signal. (D) Transversal view: loss of fibrillar pattern, erosions and Doppler signal
the detection of abnormal vascularization at enthesis insertion into cortical bone profile in the majority of SpA patients, but not in mechanical back pain or RA patients. This study included 164 SpA patients and 64 controls (thirty four with mechanical back pain and thirty with RA). Abnormal US findings consistent with at least one enthesis were observed in 98% and 52% of SpA and control patients, respectively. US enthesis was most commonly distributed in the distal portion of the lower limbs, irrespective of SpA subtype and of skeletal distribution of clinical symptoms, and none of abnormal enthesis in control patients showed vascularization, compared with 81% of enthesis in SpA patients. More recently, the same author reported in a prospective single-center cohort study with 118 patients that power Doppler US detection of at least one vascularized enthesis provided good predictive value for diagnosing SpA (sensitivity 76.5%, specificity 81.3%, positive likelihood ratio 4.1, OR 14.1, p<0.0001).

In recent years, several US scores for enthesis SpA structural damage assessment have been developed. Balsam and colleagues developed the GUESS (Glasgow Ultrasound Enthesitis Scoring System). In this index an enthesis score is formulated from the detection of bursitis, structure thickness, bony erosion, and enthesophyte (bony spur) on US examination of both lower limbs at five enthesal sites (superior pole and inferior pole of patella, tibial tuberosity, Achilles tendon, and plantar aponeurosis). The authors concluded that US is better than clinical examination in the detection of enthesis in patients with SpA. Alcalde and colleagues have tested an index (Sonographic Enthesal Index - SEI) that evaluates five enthesal regions from both lower limbs. Hypoechogeticity, increased tendon thickness, peritendinous oedema and bursitis were considered as signs of acute injury (SEI-A); and insertional bone erosions, intratendinous calcifications, decreased thickness and tears as chronic lesions (SEI-C). Each variable was scored in absent (0) or present (1) and total SEI was the sum of SEI-A and SEI-C (maximum 76 points). SEI correlated with lower Shober’s test, but not with other SpA activity or severity parameters. These studies represent the landmark of US enthesis attempt for structural damage assessment, but they did not include power Doppler signal on US lesion evaluation. Based on previous data, power Doppler signal is highly sensitive for detection of small vessels, is significantly correlated with clinical examination, and enables the distinction between inflammatory enthesis and enthesal lesions of purely mechanical origin. In this regard, de Miguel and colleagues elaborated a 136-point US-based scoring, examining elemental structural damage, including power Doppler signal, in twelve enthesal areas: proximal planter fascia, distal Achilles tendon, distal and proximal patellar ligament, distal quadriceps and brachial triceps tendon. Enthesis thickness, structure, calcifications/cortical bone proliferation, erosions, bursa, and power Doppler were scored in cortical bone profile, tendon, and bursa of 25 SpA patients and 29 healthy controls. After logistic regression analysis of the core of three of the elementary lesions were overestimated: calcifications (0-3), Doppler (0 or 3) and erosion (0 or 3); while scoring tendon thickness, structure and bursa were classified as 0 or 1 (absence/presence). The established US score of ≥18 was the best cut-off point for differentiation between cases and controls; and demonstrated a sensitivity, specificity, positive and negative likelihood ratios of 83.3%, 82.8%, 4.8%, and 0.2%, respectively, for the diagnosis of SpA regardless of the presence of other clinical manifestations. Kiris and colleagues evaluated the relationship between power Doppler US SpA enthesis assessment and the Maastricht Ankylosing Spondylitis Enthesis Index. This study illustrated some important points that should be considered. Firstly, clinical examination of enthesis seemed to be sensitive, except for Achilles tendon assessment where the US power Doppler was significantly more accurate. Secondly, clinical and sono- graphic results were discordant for three regions of the thirteen explored enthesis (first costochondral joint, seventh costochondral joint and iliac crest) where tenderness of enthesis occurred without ultrasonographically proven enthesis. Results were probably related with the chosen US acoustic window. Thirdly, in agreement with other data, pain or tenderness of enthesis is related to local increased vascularity easily detected by power Doppler, and thus the value of a uniform system for grading enthesis should be properly adapted to assess its role in diagnosis and follow up of disease course.

Recent data suggest that power Doppler US can also be efficiently used in diagnosis of early SpA. Using the Madrid Sonography Enthesis Index (MASEI) as a model de Miguel and colleagues developed a cross-sectional, blinded and controlled study including 113 early SpA patients. A cut-off point of MASEI ≥18 achieved a positive likelihood ratio of 4.26, sensitivity and specificity of 87.26% and 84.21%, respective-
ly, for SpA diagnosis. Furthermore, it seems that enthesis are early affected in SpA, and the incidence of involvement is higher in men and independent of SpA subtypes, HLA-B27 status or presentation pattern (axial, peripheral or mixed forms)\(^3\).

The therapeutic follow up of structural damage represents another potential goal for power Doppler US use in SpA. The introduction of targeting therapies, in particular tumor necrosis factor (TNF)-alpha blocking drugs, has seen unprecedented success in the treatment of signs and symptoms of SpA\(^2\), but current radiographic follow up data suggest that these drugs do not affect the bone proliferation process\(^2\). This apparent lack of structural stopping effect is in sharp contrast to what is seen for the erosive joint destruction in RA and represents a differential finding in SpA\(^4\). Although evidence supports the apparent inefficacy of anti-TNF drugs on bone formation in SpA pathophysiologic process, it seems that other entheseal abnormalities detected by US, such as morphologic abnormalities (tendon hypoechogenicity and/or thickening), power Doppler signal, and bursitis are responsive to these drugs\(^2\). Underlying the relevance of knowledge on entheseal pathophysiologic events in SpA; and supporting the challenge of introduce the US enthesis study in SpA to improve the objective knowledge about disease activity, structural permanent damage, and its relation with present and future disease assessment tools.

**RELIABILITY OF ULTRASOUND ENTHESIS ASSESSMENT IN SPONDYLOARTHRITIS**

It is a consensual fact that one of the major disadvantages of musculoskeletal US is operator dependency. The images generated are mainly qualitative and agreement as to be reached by different observers as to the presence or absence of pathological signs of disease. If quantitative measurements are required, then intra- and inter-observer errors become more important. Few studies have evaluated the overall reliability of US in rheumatology, most of them have been accomplished in RA patients or regarded joint examination\(^8\). Despite promising results, prior to the implementation of US as a valid method for detecting and monitoring the disease process in SpA patients, reliability assessment is crucial.

Previous reports have already studied the intra and inter-observer reliability in US enthesis study, using a pragmatic methodology concerning the interpretation of static images\(^8\). More recent data have been published, with valid results on acquisition and interpretation of images using patients\(^10\). Furthermore, moderate to excellent intra and interobserver agreements were found, not only for most of the US elementary lesions indicative of enthesopathy\(^1,2\), but also for the conventional two-dimension and three-dimensional assessments of entheseal sonographic lesions\(^2\). Filippucci and colleagues obtained weighted kappa values estimating the inter and intraobserver agreements for soft tissue inflammation of 0.696 and 0.816, respectively and for tissue damage of 0.711 and 0.901, respectively. The levels of agreement were estimated using a dichotomous (presence/absence) and a semi-quantitative score system for assessment of Achilles tendon enthesis. The elementary Achilles entheseal US findings that characterized soft tissue inflammation were tendon hypoechogenicity, tendon thickening, entheseal hypoechogenicity, bursal effusion and power Doppler signal at tendon, entheseal or bursal level. Therefore lesions such as intratendinous or entheseal calcifications, enthesophytes, bone erosions or bone irregularities were considered as tissue damage. This study revealed a relatively poor agreement concerning two elementary lesions: bone irregularity and entheseal hypoechogenicity, with inter-observer unweight kappa values of 0.2 and 0.49, respectively. Reasons such as the affection of tendon hypoechogenicity by anisotropy or the difficulty to reach a consensus about which bone irregularities represent a pathologic state were pointed by the authors for the relatively low Kappa values obtained\(^3\). This study illustrated an important point: lower levels of agreement generally are in the context of interobserver image acquisition rather than intra-reader agreement level, which is generally more demanding and examiner-dependent.

Moreover, when evaluated the level of agreement between 2D and 3D of Achilles enthesis erosions two different authors achieved excellent kappa values of 0.85 and 0.86\(^2\). However, to improve the reliability of US to assess enthesis in SpA more data are needed concerning evaluation of inter-machine reliability\(^1\). In this regard we studied prospectively one hundred and ninety-two enthesis of nine SpA patients with two US equipments: Acuson-Antares Siemens (Medical Systems equipment) with a lineal 5-13 MHz probe (equipment 1; 9 patients 108 enthesis) and LOGIQ9 (GE Helhcare)
with a lineal 9-14 MHz probe (equipment 2; 7 patients 84 enthesis). The following entheseal sites were examined bilaterally in longitudinal and transversal plane: proximal plantar fascia, distal Achilles tendon, distal and proximal patellar ligament, distal quadriceps, and brachial triceps tendons. The obtained images were punctuated according MASEI score by two sonographers with different levels of experience in US (an expert US rheumatologist familiarized with MASEI score and a rheumatologist with training on US). Prior to the study consensus rules about enthesis scanning, the definition about abnormal findings, and the MASEI score system were achieved. The lectures were blinded to clinical data and carried out independently. Inter-reader, and inter-US equipment reliability were assessed according to the two-way, mixed-effect model (absolute agreement) and single-measure intraclass correlation coefficients (ICCs). A Kappa value less than 0.20 was considered poor, between 0.41 and 0.60 moderate, between 0.61 and 0.80 good, and between 0.81 and 1 excellent. They were determined by SPSS (version 10.5), and values of p<0.05 were considered significant. The inter-reader and inter-equipment agreements are reported in Table I and Table II, respectively.

The best of our knowledge, this was the first study to investigate the inter-equipment reliability for assessing enthesis in SpA. The low variability achieved could reflect the importance of specific training, and consensus rules that took place prior the study. The two sonographers had different levels of experience in performing MASEI scoring system, but they routinely search for morphologic and elementary lesions used in this score system in their daily practice. Such aspect would possibly make easier the improvement of skills that they were already very familiar with, and lowered the variability of a technique that includes several possible sources of discrepancy between sonographers, such as definition, detection, and scoring of lesions, and data acquisition. In this sense a standardized model for teaching sonographers with different levels of experience achieves an effective learning on enthesis US.

In summary, US seem to be a valid and reliable tool, even inter-equipment, to assess enthesis in SpA. Nevertheless, its ultimate utility in the core set of SpA diagnostic criteria and disease monitoring remains to be determined, and prior clinical application, learning and strict consensus guidelines are highly warranted.

CORRESPONDENCE TO
Sandra Falcão
Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental E.P.E.
Rheumatology Department
Rua da Junqueira nº126
1200 Lisbon, Portugal
E-mail: sfalcao76@gmail.com

REFERENCES


