ORIGINAL ARTICLE

Defining Enthesitis in Spondyloarthritis by Ultrasound: Results of a Delphi Process and of a Reliability Reading Exercise

L. TERSLEV,[1] E. NAREDO,[2] A. IAGNOCCO,[3] P. V. BALINT,[4] R. J. WAKEFIELD,[5] P. AEGERTER,[6] S. Z. AYDIN,[7] A. BACHTA,[8] H. B. HAMMER,[9] G. A. W. BRUYN,[10] E. FILIPPUCCI,[11] F. GANDJBAKHCH,[12] P. MANDL,[13] C. PINEDA,[14] W. A. SCHMIDT,[15] AND M. A. D'AGOSTINO,[6] ON BEHALF OF THE OUTCOME MEASURES IN RHEUMATOLOGY ULTRASOUND TASK FORCE

Objective. To standardize ultrasound (US) in enthesitis.
Methods. An initial Delphi exercise was undertaken to define US-detected enthesitis and its core components. These definitions were subsequently tested on static images taken from spondyloarthritis patients in order to evaluate their reliability.
Results. Excellent agreement (>80%) was obtained for including hypoechogenicity, increased thickness of the tendon insertion, calcifications, enthesophytes, erosions, and Doppler activity as core elementary lesions of US-detected enthesitis. US definitions were subsequently obtained for each elementary component. On static images, the intraobserver reliability showed a high degree of variability for the detection of elementary lesions, with kappa coefficients ranging from 0.13–1. The interobserver kappa values were variable, with the lowest kappa coefficient for enthesophytes (0.24) and the highest coefficient for Doppler activity at the enthesis (0.63).
Conclusion. This is the first consensus-based US definition of enthesitis and its elementary components and the first step performed to ensure a higher degree of homogeneity and comparability of results between studies and in daily clinical work.

INTRODUCTION

One of the key pathologic features of the group of disorders known as the spondyloarthritides (SpA) is enthesitis, which is defined as inflammation of the insertion of ten-dons, ligaments, and capsules into bone. Although the pathogenesis has not been fully elucidated, immunologic and biomechanical factors appear to be important in the

[1]L. Terslev, MD, PhD: Copenhagen University Hospital at Glostrup, Copenhagen, Denmark; [2]E. Naredo, MD: Hospital General Universitario Gregorio Marañón, Madrid, Spain; [3]A. Iagnocco, MD: Sapienza Università di Roma, Rome, Italy; [4]P. V. Balint, MD, PhD: National Institute of Rheu-matology and Physiotherapy, Budapest, Hungary; [5]R. J. Wakefield, MD: University of Leeds, Leeds, UK; [6]P. Aegerter, MD, PhD, M. A. D'Agostino, MD, PhD: Université Versailles-Saint Quentin, Paris, and AP-HP, Hôpital Ambroise Paré, Boulogne-Billancourt, France; [7]S. Z. Aydin, MD: Istanbul Medeniyet University, Goztepe Training and Research Hospital, Istanbul, Turkey; [8]A. Bachta, MD, PhD: Military Medical Institute, Warsaw, Poland; [9]H. B. Hammer, MD, PhD: Diakonhjemmet Hospital, Oslo, Norway; [10]G. A. W. Bruyn, MD, PhD: MC Groep, Lelystad, The Netherlands; [11]E. Filippucci, MD, PhD: Università Politecnica delle Marche, Jesi, Ancona, Italy; [12]F. Gandjbakhch, MD: Uni-versité Paris 6-Pierre et Marie Curie, Hôpital La Pitié Salpetrière, AP-HP, Paris, France; [13]P. Mandl, MD, PhD: University of Vienna, Vienna, Austria; [14]C. Pineda, MD: National Institute of Rehabilitation, Mexico City, Mexico;

[15]W. A. Schmidt, MD: Immanuel Krankenhaus Berlin, Medical Centre for Rheumatology, Berlin-Buch, Berlin, Ger-many.

Address correspondence to M. A. D'Agostino, MD, PhD, Université Versailles-Saint Quentin, Service de Rhumatolo-gie, Hôpital Ambroise Paré, 9, Avenue Charles de Gaulle, 92100 Boulogne-Billancourt, France. E-mail: maria-antonietta.dagostino@apr.ap-hop-paris.fr.

Significance & Innovations

- Excellent agreement was obtained to include hypoechogenicity, increased thickness of the tendon insertion, calcifications, enthesophytes, erosions, and Doppler activity as elementary lesions of enthesitis.
- No agreement was obtained to include bursitis or tendinitis as elementary components of ultrasound-detected enthesitis.

development and persistence of local inflammation at these sites. Sometimes, and especially in longstanding disease, this inflammation can overcome the anatomic limit of the enthesis and include adjacent structures, such as the tendon body, peritenon, and bursa.

However, the involvement of entheseal-related structures is not always specific to SpA and may be observed in other noninflammatory and traumatic conditions with enthesis involvement (1–4). Therefore, to differentiate the origin of the pathologic process, the term enthesopathy is commonly used to describe a metabolic or degenerative/traumatic cause, whereas the term enthesitis refers to an inflammatory cause, such as in SpA (1–4). However, the 2 terminologies have often been used interchangeably in studies published in the literature, which has led a group of researchers to develop the concept of the enthesis organ, despite the origin (5).

The assessment of enthesitis has conventionally been performed for several years through clinical examination (i.e., presence of tenderness and swelling) and radiographs for the presence of bony changes, such as bone erosions and heel spurs (enthesophytes). Both evaluations, however, are known to lack accuracy (3,6). With the realization of the importance of an early, accurate diagnosis of inflammatory conditions like SpA, and with the development of new therapeutic options, new imaging modalities have been investigated to determine whether they improve entheseal disease assessment. In this respect, ultrasound (US) has been shown to be promising because, unlike clinical examination, it not only allows direct visualization of the enthesis and entheseal-related structures (3–8), but also it is increasingly available in the clinical setting.

Although several published studies have highlighted the role of US in the assessment of entheseal inflammation in SpA, a lack of consensus on which elementary structures of the enthesis should be examined and how to define any abnormality found has remained. This question was raised at the first Outcome Measures in Rheumatology (OMERACT) US Specialist Interest Group meeting held in Asimolar, California, in 2004. At this meeting, for the first time, an international group of US experts agreed on a preliminary definition of US-related enthesopathy to standardize the entity for future evaluation. The proposed definition was broad and included several soft tissue and bone elements and was defined 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" (9). In addition, there was a lack of consensus on which lesions best defined active (and therefore potentially reversible) disease and which were more permanent structural consequences. As a result, this led to a number of studies proposing different US scoring systems (3,4,10,11). This lack of consensus was highlighted in a recent systematic literature review (12) that showed nonuniformity both in the definitions being applied and on the technical parameters and scanning methods used. In order to improve the use of US in the evaluation of SpA-related enthesitis, an OMERACT Task Force subgroup was formed. The aim of this subgroup was to standardize the US definition and detection of SpA-related enthesitis. To this effect, the group proceeded in a 2-step manner, the details of which are described in this article. The first step aimed to develop consensus-derived definitions for US-defined enthesitis and agree on which elementary components should be included using a Delphi process among rheumatologists with an interest in SpA-related enthesitis. The second step aimed at evaluating the reliability of the definition of enthesitis and of each elementary lesion defining US enthesis by using static images.

MATERIALS AND METHODS

Study design. *First step: consensus process on the definition of enthesitis and elementary lesions.* We undertook a 2-round Delphi exercise using a written questionnaire sent by e-mail to 26 rheumatologists trained in musculoskeletal US from 13 countries (Denmark, France, Germany, Hungary, Italy, Mexico, The Netherlands, Norway, Poland, Spain, Turkey, the UK, and the US). The participants were selected because of their declared interest in participating in the OMERACT US Task Force on enthesitis and their previous publications in the field. See Appendix A for additional members of the OMERACT US Task Force.

The construction of the questionnaire was made using data collected from a recent systematic literature review (12), data from preceding studies performed by some members of the group (3,4,13), and information obtained from a preliminary meeting of the members of the Task Force listing the areas of interest. This collated information resulted in 14 statements grouped under the 4 following subheadings: 1) definition of a normal US enthesis and other related anatomic structures, 2) elementary lesions to be included in both B-mode and Doppler assessment of enthesitis, 3) US definition of individual elementary lesions, and 4) lesions reflecting inflammation and damage. In particular for subheadings 2 and 3, the suggested elementary lesions were change in the echogenicity (at the enthesis, tendon, and bursa); increased thickness (at the enthesis, tendon, and bursa); calcifications (at the enthesis, tendon, and bursa); and enthesophytes, erosions, cortical irregularities, and Doppler signal (at the enthesis, tendon, and bursa). For the suggested elementary lesions above, a

definition of each elementary component was specified in a subsequent question.

The 14 statements were distributed among the participants who were asked to rate their level of agreement or disagreement for each statement according to a 1–5 Likert scale, where 1 = strongly disagree and 5 = strongly agree. Space for additional free comments was also included at the end of each statement. The participants were asked to respond within 1 month. The group answers from the first round questionnaire were summarized with mean scores by the coordinator (MAD) and resent with a revised questionnaire to the same group (round 2). The questionnaire is available in Supplementary Appendix A (available in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/acr.22191/abstract).

Second step (reliability exercise: collection of US images representative of the agreed elementary lesions for enthesitis). The questionnaire respondents were asked to collect US images of entheses from their daily practice that were considered either normal or demonstrated entheseal changes consistent with SpA and that represented the components agreed on in the Delphi process. Each expert was asked to collect at least 1 US image of each individual elementary component presented in the longitudinal plane. The anatomic sites analyzed were as follows: Achilles tendon insertion, plantar fascia, patellar ligament insertion (proximal and distal insertions), quadriceps tendon insertion, and lateral epicondyle tendon insertion. After a collection period of 2 months, the images were sent by e-mail to the coordinator of the study (MAD). The 119 collected images were uploaded to a web site that was specifically designed for evaluating reliability and that has already been used for a similar purpose (13).

Second step (reliability exercise: consensus web exercise). By using the designed web site, the collected images were displayed and the participants were asked to read each image and determine the presence/absence of each elementary component and to state whether these were compatible with the final diagnosis of enthesitis (yes/no). A randomly selected group of images were displayed twice to evaluate the intrareader reliability.

Statistical analysis. In the Delphi process, agreement was assessed on the following 2 levels: 1) does the expert respondent agree with the issue under consideration, and 2) does each expert respondent agree with the opinion of other experts on a certain issue (consensus element). The 1–5 Likert scale used was graded as follows: 1 indicated "definitely no/definitely not important," 2 indicated "probably no/probably not important," 3 indicated "no opinion," 4 indicated "probably yes/probably important," and 5 indicated "definitely yes/essential." The results were expressed as the cumulative percentage of respondents scoring an item either 4 (probably yes/probably important) or 5 (definitely yes/essential; total cumulative agreement). Group agreement with the issue under consideration was defined as total cumulative agreement ≥80% after the second Delphi round. Only sentences having reached a score of 70% in the first round were included in the second round. If both of these parameters were satisfied (i.e., agreement of at least 70% in the first round and of at least 80% in the second round), we considered that the group had reached a consensus and that the category was defined as appropriate. Only the statements satisfying these requirements were used for finally defining the 4 areas of interest.

Intra- and interobserver reliability were calculated using the standard kappa coefficient. Intraobserver reliability was assessed by Cohen's kappa. Interobserver reliability was studied by calculating the mean kappa on all pairs (i.e., Light's kappa) (14). Kappa coefficients were interpreted according to Landis and Koch (15). Kappa values of 0−0.20 were considered poor, 0.20−0.40 fair, 0.40−0.60 moderate, 0.60−0.80 good, and 0.80−1 excellent (15). The percentage of observed agreement (i.e., percentage of observations that obtained the same score) and prevalence of the observed lesions were also calculated.

RESULTS

Delphi exercise. The overall response rate from the first Delphi round was 92% (24 of 26 rheumatologists). Nonresponders to the first questionnaire were not included in the second round. After the second Delphi exercise, the total cumulative agreement scores for each of the 4 areas of interest outlined above were 89%, 87%, 84%, and 84%, respectively. The total cumulative agreement (%) after both rounds for each of the 14 statements is shown in Table 1. As shown in Table 1, consensus was reached after the first round for areas of interest 1 and 2, whereas areas of interest 3 and 4 were more controversial. According to these results, the final definitions are outlined below.

First area of interest: definition of a normal US enthesis and other normal anatomic structures (i.e., tendon and bursa). A normal enthesis was defined as an insertion of tendon, ligament, and capsule into bone with regular margin and with the same US appearance and thickness as the corresponding tendons or ligament (fibrillar echotexture or homogenous linear echotexture) and capsule (hyperechoic band). A normal bursa was defined as a thin hypoechoic layer surrounded by a hyperechoic line. It was agreed that a normal bursa is visible only in 2 of the considered sites: 1) the patellar tendon insertion on the tibia tuberosity (i.e., infrapatellar bursa) and 2) at the calcaneal level, at the insertion of the Achilles tendon (retrocalcaneal bursa). It was also agreed that, at the calcaneal level, the bursa may contain an iso- or hyperechoic structure corresponding to the echogenicity of the fat pad. The panelists therefore agreed to maintain separate definitions for a normal enthesis, normal tendon, and bursa.

Second area of interest: elementary lesions in both B mode and Doppler to be included in the definition of US enthesis. Good to high agreement (>80%) was reached for hypoechogenicity, increased thickness of the tendon insertion, calcifications, enthesophytes, erosions, and Doppler activity as elementary lesions of enthesitis (Figure 1). Only moderate agreement (76%) was reached for cortical irregularities, defined as a loss of the normal regular bone contour without any clear sign of enthesophyte and/or erosion, and therefore these lesions were not

Table 1. Results of the Delphi exercise*			
Statements	**Round 1 cumulative agreement, %**	**Round 2 cumulative agreement, %**	**Category appropriated?**
1a. Hypoechogenicity of the insertion into the bone of tendon/ligament/capsule (i.e., enthesis), after correcting anisotropy artifact	92	100	Yes
1b. Hypoechogenicity could be defined as a lack of the homogeneous fibrillar pattern with loss of the tightly packed echogenic lines			
2. Increased thickness of tendon/ligament/capsule insertion into the bone, as compared to the body of tendon/ligament/capsule, with or without (please specify in the optional comment) blurring of the tendon/ligament/capsule margins	78	84	Yes
3a. Calcifications detected at the tendon insertion into the bone (i.e., enthesis)	72	82	Yes
3b. Calcifications could be defined as hyperechoic foci consistent with calcific deposits, with or without acoustic shadow, seen in 2 perpendicular planes, detected at the tendon insertion into the bone (i.e., enthesis)			
4a. Enthesophytes detected at the insertion of enthesis into the bone	92	100	Yes
4b. Enthesophyte could be defined as a step up of bony prominence at the end of the normal bone contour, seen in 2 perpendicular planes, with or without acoustic shadow			
5a. Erosions detected at the insertion of the enthesis into the bone	100	100	Yes
5b. Erosion could be defined as a cortical breakage with a step down contour defect, seen in 2 perpendicular planes, at the enthesis insertion			
6a. Cortical irregularities detected at the insertion of the enthesis into the bone	72	76	No
6b. Irregularity could be defined as a loss of the normal regular bone contour, without any clear sign of enthesophyte and/or erosion			
7. Presence of a Doppler signal at the cortical enthesis insertion, different from reflecting surface artifact or nutrition vessel signal, with or without (please specify in the optional comment) cortical irregularities, erosions, or enthesophytes	76	92	Yes
8. Presence of a Doppler signal at the tendon with or without (please specify in the optional comment) Doppler signal at enthesis insertion into the bone	58	–	No
9. Do you think that all of the US elementary lesions mentioned above should be used to define US enthesitis?	72	94	Yes
10. Do you agree to distinguish between active/inactive US aspects of enthesitis according to the presence of specific elementary lesions?	93	100	Yes
11. Do you agree to define hypoechogenicity, thickening (with or without mixed echotexture, and with or without blurring of the tendon margins), and Doppler signal as signs of acute or active US enthesitis, and erosions, enthesophytes, calcification, and cortical irregularities as signs of chronic and/or inactive US enthesitis?	72	78	No
12. Do you agree to include the increased of thickness of tendon insertion in the US chronic lesion, if not associated with clear loss of echogenicity?	56	–	No
13. Do you agree to include bursitis, if associated with other abnormalities of enthesitis insertion, as a sign of US chronic lesions?	48	–	No
14. Do you agree to include bursitis, if associated with other abnormalities of enthesitis insertion, as a sign of US acute lesions?	30	–	No

* The individual questions from the Delphi rounds and the cumulative agreement for each round are shown. US = ultrasonographic.

included. Similarly, even less agreement (<60%) was obtained to include bursitis or tendinitis as elementary components of enthesitis.

Third area of interest: US definition of elementary lesions. Hypoechogenicity was defined as a lack of the homogeneous fibrillar pattern with loss of the tightly

packed echogenic lines after correcting for anisotropy. Increased thickness of the enthesis was defined as increased thickness of the tendon/ligament/capsule insertion into the bone, as compared to the body of the tendon/ligament/capsule, with or without blurring of the tendon/ligament/capsule margins. Enthesophytes were defined as a step up
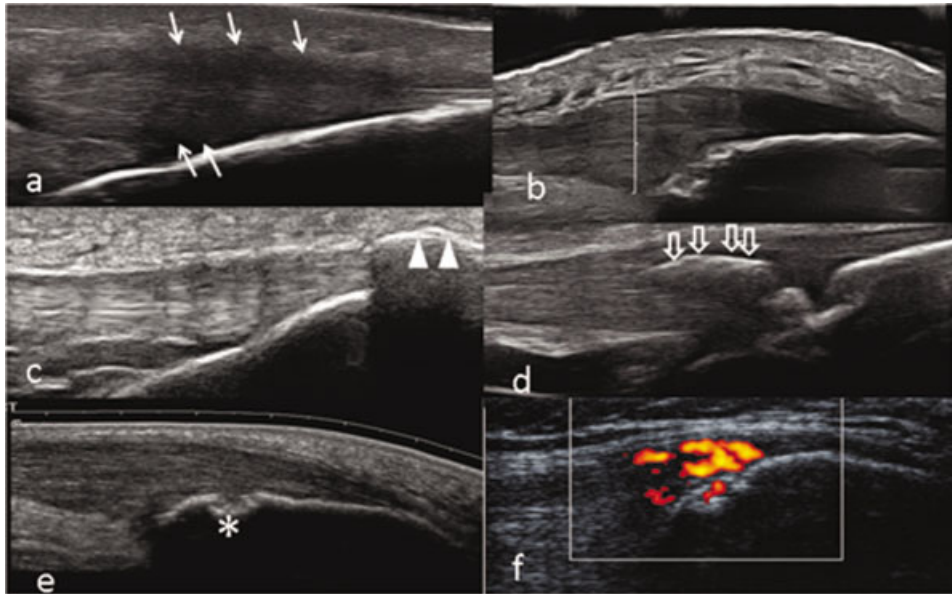
Figure 1. The elementary components are as follows: **a**, hypoechogenicity (**white arrows** indicate increased thickness with blurring of the tendon margins); **b**, increased thickness of tendon insertion (**white line**); **c**, enthesophyte (the step up of the bony prominence at the end of the normal bone contour is marked with **white arrowheads**); **d**, calcifications (the hyperechoic focus consistent with calcific deposit is marked by **open arrows**); **e**, bone erosion at the enthesis marked with an **asterisk**; and **f**, Doppler at enthesis <2 mm from the bone insertion.

of bony prominence at the end of the normal bone contour, seen in 2 perpendicular planes, with or without acoustic shadow. Calcifications were defined as hyperechoic (bright) foci consistent with calcific deposits, with or without acoustic shadow, seen in 2 perpendicular planes, detected at the tendon insertion into the bone (i.e., enthesis). Erosion was defined as a cortical breakage with a step down contour defect, seen in 2 perpendicular planes, at the insertion of the enthesis to the bone, according to the OMERACT definition (9). The Doppler signal at the enthesis was defined as Doppler activity approximately <2 mm near the bony cortex. The Doppler signal must be at the enthesis, different from reflecting surface artifact or nutrition vessel signal, with or without cortical irregularities, erosions, or enthesophytes.

Fourth area of interest: elementary lesions reflecting inflammation and structural damage. Excellent agreement (93%) was obtained for separating signs of active inflammation from signs of structural damage. A definition was therefore proposed that included hypoechogenicity, thickening, and Doppler signal as signs of inflammation (and therefore of acute/active US enthesitis) and included erosions, enthesophytes, calcification, and cortical irregularities as signs of structural damage (and therefore of chronic/inactive US enthesitis). There was, however, poor agreement (<60%) for differentiation between acute and chronic enthesitis, and therefore this differentiation could not be made.

Reliability of reading images. Of the 24 rheumatologists, 15 (63%) collected 119 images of the elementary

components from the plantar fascia, Achilles tendon, lateral epicondyle, quadriceps, and patellar tendon (both insertions). The intraobserver reliability was performed on 31 of the 119 images. The equipment used for the image collections differed from institution to institution and comprised the following US machines: General Electric Logic 9 and Logic 5, Siemens Antares, Phillips HDI 5000, and Esaote MyLab 70 XVG and MyLab 60.

The prevalence of the detected elementary lesions, the observed agreement, and the kappa coefficients concerning intra- and interobserver reliability on static images are shown in Tables 2 and 3, respectively. For the intraobserver reliability (Table 2), the highest mean prevalence on static images was seen for enthesophytes (range 61.3–100%), while the lowest mean prevalence was observed for bursitis (range 5–55.4%). The observed agreement was highest for Doppler activity at the enthesis (range 0.79–1) and for global enthesitis (range 0.84–1) and lowest for calcifications (range 0.40–1). The agreement on all elementary lesions was generally found to be in the same range. There was, however, a high degree of variability for the detection of elementary lesions, with kappa coefficients ranging from 0.14–1. The kappa values for global enthesitis were better than those for the morphologic elementary components (range 0.25–1).

The results of the interobserver reliability (Table 3) showed that the detected prevalence of elementary lesions seen on static images was quite variable among the participating sonographers, being lowest for bursitis (22.7%) and highest for enthesophytes (73.6%). The observed agreement was highest for global enthesitis (0.93) and low-

Table 2. Results of intraobserver reliability of ultrasonographic elementary components and global definition of enthesitis*

	Enthesophytes, yes/no	Hypoechogenicity of enthesis, yes/no	Thickened enthesis, yes/no	Erosions, yes/no	Power Doppler inside enthesis, yes/no	Calcifications, yes/no	Bursitis, yes/no	Global enthesitis, yes/no
Mean prevalence, range of %	61.3–100	56–96.4	26.7–51.6	13.3–51.6	26.7–53.6	12–67.9	5–55.4	6.5–16.7
Observed agreement, range	0.65–1	0.68–1	0.50–1	0.64–1	0.79–1	0.40–1	0.68–1	0.84–1
Cohen's κ, range	0.20–0.84	0.24–1	0.13–1	0.14–1	0.57–1	0.23–1	0.19–1	0.25–1

* The range of Cohen's kappa values among observers are shown for every single component of enthesitis agreed on in the Delphi exercise. Also shown is the kappa value for the observer's opinions on whether enthesitis was present or not.

est for hypoechogenicity of the enthesis (0.69). Overall, the interobserver agreement on all elementary lesions was generally in the same range as that observed for the intraobserver results. The interobserver kappa values were variable, with the lowest kappa value for enthesophytes (0.24) and the highest for Doppler activity at the enthesis (0.63) (Table 3).

DISCUSSION

Defining enthesitis by US is quite a difficult challenge because of the numerous anatomic structures that can be involved in the inflammatory process. A recent systematic literature review highlighted a great variability in the definition of enthesitis in SpA applied in US studies since the first publication in 1985 (12) and, in particular, definitions of its constituent elementary components. The literature review underlined the necessity for a more refined definition of enthesitis and its components. The Delphi and web exercises therefore focused solely on obtaining agreement about which elementary components should be included in the US definition of enthesitis and subsequently on the ability to agree on these definitions on static images. Although all images involved originated from SpA patients, further studies are needed to highlight what relates to the presence of enthesitis in SpA (i.e., the inflammatory disease) independently of other possible factors (i.e., biomechanical, traumatic, or metabolic) and other diseases.

In order to ensure a consensus, we conducted the present Delphi exercise and were able to obtain a good to high agreement (>80%) on the inclusion of hypoechogenicity and increased thickness of the tendon insertion and enthesophytes, calcifications, erosions, and Doppler signal at the enthesis ≤2 mm near the bony cortex as elementary components of enthesitis, corresponding to what has already been proposed in the literature (16–18). In their definitions, both calcification and enthesophyte findings with and without acoustic shadowing were included, and it is worth noting that, although in most cases both pathologies give acoustic shadowing, this may not necessarily be the case for small abnormalities (19–21). No agreement was obtained, however, for the inclusion of bursitis, tendinitis, or cortical irregularities as elementary lesions of enthesitis, although their inclusion has been suggested by others (3,22–24). The obtained consensus is a major step toward ensuring homogeneity in future studies. Although high agreement was present for the elementary components, it was not possible to obtain agreement on how to differentiate between acute and chronic enthesis involvement, which may partly be related to the lack of histopathologic data, and further studies are needed in this area to clarify the possible differences.

On the basis of the consensus agreement on which elementary components to include in the definition, the intra- and interreader reliability of these components were tested on static images in a web exercise and showed some de-

Table 3. Results of interobserver reliability of ultrasonographic elementary components and global definition of enthesitis*

	Enthesophytes, yes/no	Hypoechogenicity of enthesis, yes/no	Thickened enthesis, yes/no	Erosions, yes/no	Power Doppler inside enthesis, yes/no	Calcifications, yes/no	Bursitis, yes/no	Enthesitis, yes/no
Mean prevalence, %	73.6	61.8	39.9	34.6	30.8	29.2	22.7	8.9
Mean observed agreement	0.71	0.69	0.81	0.81	0.83	0.77	0.83	0.93
Mean Cohen's κ	0.24	0.34	0.60	0.50	0.63	0.45	0.52	0.61

* The Cohen's kappa values are shown for every single component of enthesitis agreed on in the Delphi exercise as well as the kappa value for the observers' opinions on whether enthesitis was present or not.

gree of variability. The global definition of enthesitis was shown to be more reliable than some of the individual single components. Although there was good to high agreement (>80%) on including hypoechogenicity and increased thickness of the enthesis, calcifications, enthesophytes, erosions, and Doppler activity as elementary lesions of enthesitis, there was low reliability on thickened enthesis, enthesophytes, and calcifications when testing the elementary components on static images, which is in accordance with a previous study (13). One reason for the low reliability may be that the evaluation was performed on static images and not during live scanning, in which case either optimization of images or more information obtained by changing the scan area may have assisted this. In addition, a lack of training in detecting each elementary component despite the theoretical agreement reached on the Delphi exercise could be another reason for the low reliability. In a previous study, Doppler activity at the enthesis was noted to be of great importance for detecting active enthesitis (13), but in the present Delphi exercise, it was not possible to obtain agreement on a definition of how to separate acute from chronic and active from inactive enthesitis changes, and more work is needed in this area. Although Doppler activity was an important finding, there was only agreement of its presence in the static images at the enthesis level in less than one-third of the images, but when Doppler activity was observed, it showed both high inter- and intraobserver reliability. This could be related to the definition of Doppler activity only being part of the enthesitis when found very close to the bony cortex (<2 mm), which means that Doppler activity further away or in part of the tendon should not be included to make sure not to include possible tendinitis. In addition, highly sensitive Doppler technology is necessary to detect a Doppler signal at the level of the enthesis. Again, training would be expected to increase the reliability even further. The implementation of new rules probably takes some time and further studies are needed on the reliability for detecting the single components of enthesitis, including testing these components in patients.

In conclusion, this US study represents the first agreement on US definitions and elementary lesions in SpA-related enthesitis. This is an important first step toward ensuring a higher degree of homogeneity between studies and daily clinical work, thereby facilitating the comparability of results.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. D'Agostino had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Terslev, Wakefield, Aegerter, D'Agostino.
Acquisition of data. Terslev, Iagnocco, Balint, Aydin, Bachta, Hammer, Bruyn, Filippucci, Gandjbakhch, Mandl, Pineda, Schmidt, D'Agostino.
Analysis and interpretation of data. Terslev, Naredo, Balint, Aegerter, D'Agostino.

REFERENCES

1. Balint PV, D'Agostino MA. Spondyloarthritis: a journey within and around the joint. Rheumatology (Oxford) 2012;51 Suppl 7:vii13–7.
2. Ruta S, Gutierrez M, Pena C, Garcia M, Arturi A, Filippucci E, et al. Prevalence of subclinical enthesopathy in patients with spondyloarthropathy: an ultrasound study. J Clin Rheumatol 2011;17:18–22.
3. Balint PV, Kane D, Wilson H, McInnes IB, Sturrock RD. Ultrasonography of entheseal insertions in the lower limb in spondyloarthropathy. Ann Rheum Dis 2002;61:905–10.
4. D'Agostino MA, Said-Nahal R, Hacquard-Bouder C, Brasseur JL, Dougados M, Breban M. Assessment of peripheral enthesitis in the spondylarthropathies by ultrasonography combined with power Doppler: a cross-sectional study. Arthritis Rheum 2003;48:523–33.
5. Benjamin M, McGonagle D. The enthesis organ concept and its relevance to spondyloarthropathies. Adv Exp Med Biol 2009;649:57–70.
6. Borman P, Koparal S, Babaoglu S, Bodur H. Ultrasound detection of entheseal insertions in the foot of patients with spondyloarthropathy. Clin Rheumatol 2006;25:373–7.
7. Galluzzo E, Lischi DM, Taglione E, Lombardini F, Pasero G, Perri G, et al. Sonographic analysis of the ankle in patients with psoriatic arthritis. Scand J Rheumatol 2000;29:52–5.
8. Lehtinen A, Taavitsainen M, Leirisalo-Repo M. Sonographic analysis of enthesopathy in the lower extremities of patients with spondylarthropathy. Clin Exp Rheumatol 1994;12:143–8.
9. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al, for the OMERACT 7 Special Interest Group. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2485–7.
10. Alcalde M, Acebes JC, Cruz M, Gonzalez-Hombrado L, Herrero-Beaumont G, Sanchez-Pernaute O. A sonographic enthesitic index of lower limbs is a valuable tool in the assessment of ankylosing spondylitis. Ann Rheum Dis 2007;66:1015–9.
11. De Miguel E, Cobo T, Munoz-Fernandez S, Naredo E, Uson J, Acebes JC, et al. Validity of enthesis ultrasound assessment in spondyloarthropathy. Ann Rheum Dis 2009;68:169–74.
12. Gandjbakhch F, Terslev L, Joshua F, Wakefield RJ, Naredo E, D'Agostino MA, for the OMERACT Ultrasound Task Force. Ultrasound in the evaluation of enthesitis: status and perspectives. Arthritis Res Ther 2011;13:R188.
13. D'Agostino MA, Aegerter P, Jousse-Joulin S, Chary-Valckenaere I, Lecoq B, Gaudin P, et al. How to evaluate and improve the reliability of power Doppler ultrasonography for assessing enthesitis in spondylarthritis. Arthritis Rheum 2009;61:61–9.
14. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull 1971;76:365–77.
15. Landis JR, Koch GG. An application of hierarchical κ-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977;33:363–74.
16. Falsetti P, Frediani B, Fioravanti A, Acciai C, Baldi F, Filippou G, et al. Sonographic study of calcaneal entheses in erosive osteoarthritis, nodal osteoarthritis, rheumatoid arthritis and psoriatic arthritis. Scand J Rheumatol 2003;32:229–34.
17. Iagnocco A, Riente L, Delle Sedie A, Filippucci E, Salaffi F, Meenagh G, et al. Ultrasound imaging for the rheumatologist: XXII. Achilles tendon involvement in spondyloarthritis: a multi-centre study using high frequency volumetric probe. Clin Exp Rheumatol 2009;27:547–51.
18. Aydin SZ, Bas E, Basci O, Filippucci E, Wakefield RJ, Celikel C, et al. Validation of ultrasound imaging for Achilles entheseal fibrocartilage in bovines and description of changes in humans with spondyloarthritis. Ann Rheum Dis 2010;69:2165–8.
19. Kim BK, Choi YS, Kwon HJ, Lee JS, Heo JJ, Han YJ, et al. Relationship between patterns of calcification in the thyroid nodules and histopathologic findings. Endocr J 2013;60:155–60.

20. Gufler H, Buitrago-Tellez CH, Madjar H, Allmann KH, Uhl M, Rohr-Reyes A. Ultrasound demonstration of mammographically detected microcalcifications. Acta Radiol 2000;41:217–21.
21. Yang WT, Suen M, Ahuja A, Metreweli C. In vivo demonstration of microcalcification in breast cancer using high resolution ultrasound. Br J Radiol 1997;70:685–90.
22. Kiris A, Kaya A, Ozgocmen S, Kocakoc E. Assessment of enthesitis in ankylosing spondylitis by power Doppler ultrasonography. Skeletal Radiol 2006;35:522–8.
23. Hatemi G, Fresko I, Tascilar K, Yazici H. Increased enthesopathy among Behçet's syndrome patients with acne and arthritis: an ultrasonography study. Arthritis Rheum 2008;58:1539–45.
24. Filippucci E, Aydin SZ, Karadag O, Salaffi F, Gutierrez M, Direskeneli H, et al. Reliability of high-resolution ultrasonography in the assessment of Achilles tendon enthesopathy in seronegative spondyloarthropathies. Ann Rheum Dis 2009; 68:1850–5.

APPENDIX A: MEMBERS OF THE OUTCOME MEASURES IN RHEUMATOLOGY ULTRASOUND TASK FORCE

Members of the Outcome Measures in Rheumatology Ultrasound Task Force are as follows: Marina Backhaus, Fred Joshua, David Bong, Isabelle Chary-Valckenaere, Paz Collado, Eugenio De Miguel, Jane E. Freeston, Walter Grassi, Marwin Gutierrez, Sandrine Jousse-Joulin, David Kane, Helen I. Keen, Damien Loeuille, Ingrid Möller, Wolfgang A. Schmidt, Marcin Szkudlarek, and Hans-Rudolf Ziswiler.