Bond University

Research Repository

Staging achilles tendinopathy using ultrasound imaging: The development and investigation of a new ultrasound imaging criteria based on the continuum model of tendon pathology

Matthews, Wesley; Ellis, Richard; Furness, James W.; Rathbone, Evelyne; Hing, Wayne

Staging achilles tendinopathy using ultrasound imaging: the development and investigation of a new ultrasound imaging criteria based on the continuum model of tendon pathology

Wesley Matthews [iD] ,[1] Richard Ellis,[2,3] James W Furness,[1] Evelyne Rathbone,[4] Wayne Hing[1]

ABSTRACT

Aim To develop a standardised ultrasound imaging (USI)-based criteria for the diagnosis of tendinopathy that aligns with the continuum model of tendon pathology. Secondary aims were to assess both the intra-rater and inter-rater reliability of the criteria.

Methods A criteria was developed following a face validity assessment and a total of 31 Achilles tendon ultrasound images were analysed. Intra-rater and inter-rater reliability were assessed for overall tendinopathy stage (normal, reactive/early dysrepair or late dysrepair/degenerative) as well as for individual parameters (thickness, echogenicity and vascularity). Quadratic weighted kappa (k_w) was used to report on reliability.

Results Intra-rater reliability was 'substantial' for overall tendinopathy staging (k_w rater A; 0.77, 95% CI 0.59 to 0.94, rater B; 0.70, 95% CI 0.52 to 0.89) and ranged from 'substantial' to 'almost perfect' for thickness (k_w rater A; 0.75, 95% CI 0.59 to 0.90, rater B; 0.84, 95% CI 0.71 to 0.98), echogenicity (k_w rater A; 0.78, 95% CI 0.62 to 0.95, rater B; 0.73, 95% CI 0.58 to 0.89) and vascularity (k_w rater A; 0.86, 95% CI 0.74 to 0.98, rater B; 0.89, 95% CI 0.79 to 0.99). Inter-rater reliability ranged from 'substantial' to 'almost perfect' for overall tendinopathy staging (k_w round 1; 0.75, 95% CI 0.58 to 0.91, round 2; 0.81, 95% CI 0.63 to 0.99), thickness (k_w round 1; 0.65, 95% CI 0.48 to 0.83, round 2; 0.77, 95% CI 0.60 to 0.93), echogenicity (k_w round 1; 0.70, 95% CI 0.54 to 0.85, round 2; 0.76, 95% CI 0.58 to 0.94) and vascularity (k_w round 1; 0.89, 95% CI 0.79 to 0.99, round 2; 0.86, 95% CI 0.74 to 0.98). Inter-rater reliability increased from 'substantial' in round 1 (k_w 0.75, 95% CI 0.58 to 0.91) to 'almost perfect' in round 2 (0.81, 95% CI 0.63 to 0.99).

Conclusion Intra-rater and inter-rater reliability were 'substantial' to 'almost perfect' when utilising an USI-based criteria to diagnose Achilles tendinopathy. This is the first study to use the continuum model of tendon pathology to develop an USI-based criteria to diagnose tendinopathy.

What are the new findings

► A standardised criteria for the diagnosis of tendinopathy aligning with the continuum model of tendon pathology is presented.
► The criteria is reliable in staging Achilles tendinopathy.
► The proposed criteria aligns with accepted clinical terminology used to describe tendinopathy.

INTRODUCTION

Tendinopathy refers to persistent tendon pain and dysfunction which is related to mechanical loading.[1] Tendinopathy accounts for 30%–50% of all overuse injuries,[2] with Achilles tendinopathy being among the most common.[3–6] While there are multiple models to describe the pathogenesis of tendinopathy,[7–10] the continuum model of tendon pathology proposed by Cook and Purdam in 2009,[9] and updated in 2016,[11] is widely used to clinically describe and diagnose tendinopathy.[3 9 11 12] The continuum model of tendon pathology used both imaging and clinical features to characterise the different stages of tendinopathy.[9 11] The continuum model proposed three key stages of tendon pathology; reactive tendinopathy, tendon dysrepair and degenerative tendinopathy.[9] Although the model is described in three distinct stages, it is acknowledged that tendon pathology occurs on a continuum, with continuity between stages.[9 11]

A clinical diagnosis of tendinopathy is primarily derived from the patient history and clinical tests.[11 13–21] Clinical tests have been shown to be sensitive for detecting tendinopathy, and they are not specific in identifying pathological change when compared with imaging[22–24] and do not allow clinicians to stage their patient according to the continuum model of tendon pathology.[11] Conversely, ultrasound imaging (USI) has been shown to be both accurate and sensitive for detecting

pathological structural change within tendons[25–27] but does not always correlate with pain and dysfunction.[11 28 29] Although reviews have demonstrated both an association and dissociation between tendon structure, and function and pain,[20 21] structural changes identified on USI can be considered a risk factor for the development of symptomatic tendinopathy.[11 29–31] The relative risk (RR) for developing pain in asymptomatic (clinical tests negative with structural changes on imaging) Achilles tendinopathy has been reported as 5.45–7.33.[30 31]

While USI is useful for detecting tendon structural change, it is considered an operator-dependent modality, and standardised practices for both capturing and assessing ultrasound images can increase inter-rater reliability when assessing neovascularity, structural change, calcification (k 0.46–0.96) and tendon thickness (ICC 0.85–0.98) for tendinopathy in both the upper limb[32] and lower limb.[33] Although the standardisation of criteria for assessing tendon structural change improves reliability, there is significant heterogeneity of USI-based diagnostic criteria.[31 34] This, in conjunction with the cross-sectional design of many imaging studies,[30] and lack of a clinical gold standard with which to compare images,[29] contributes to the uncertainty of the relevance of USI in the clinical diagnosis of tendinopathy.[30]

When assessing current methods for diagnosing tendinopathy using USI-based criteria, a recent systematic review[31] reported that the most common parameters used to measure tendon pathology include tendon thickness, echogenicity and vascularity. Furthermore, while the stages of tendinopathy are distinguished by specific imaging features,[9 11] no studies used the continuum model of tendon pathology to stage tendinopathy.[31] Additionally, it was discovered that the use of a combination of the three most commonly reported parameters demonstrated a higher risk (RR 6.49) for developing symptomatic tendinopathy when compared with using two parameters (RR 3.66).[31] Therefore, it was suggested that future USI-based criteria use a combination of three parameters (tendon thickness, echogenicity and vascularity) in addition to an ordinal-based USI criteria that better aligns with the continuum model of tendon pathology ('normal', 'reactive/early dysrepair' or 'late dysrepair/degenerative').[31]

The continuum model of tendon pathology provides a framework for describing the pathogenesis of tendinopathy and associated clinical and imaging features. The staging of tendon pathology may be beneficial for clinicians to target treatment according to the tendon structure.[9 11] Tendons in the 'reactive/early dysrepair' stage may have the ability to regain normal structure with appropriate management[35] and treatments aimed at inhibiting tendon cell response to prevent further tendon structural change may be prioritised.[11] Similarly, treatments that may aggravate symptoms and facilitate further tendon structural breakdown (eg, heavy-loaded eccentrics, intra-tendinous injections, etc) may be avoided.[11] Conversely, in 'degenerative' tendinopathy there is little

capacity for the tendon structure to change, and tendon structure does not relate to clinical outcome.[36] Thus, treatment should be aimed at improving the capacity of the tendon to handle variable loads rather than to stimulate a healing response in the degenerative portions of the tendon.[11]

Given the complex relationship between tendon structure, dysfunction and pain, there is scope to explore whether the development a USI-based criteria that aligns with the continuum model of tendon pathology is feasible and may help integrate clinical and imaging findings, allowing for more targeted treatment.[18 31 34] Therefore, the primary aim of this study was to develop a standardised USI-based criteria for the diagnosis of tendinopathy that aligns with the continuum model of tendon pathology. Following the development of the criteria, the secondary aims were to assess both the intra-rater and inter-rater reliability of the criteria.

MATERIALS AND METHODS
Study design
This study followed the protocols for diagnostic procedures in manual and musculoskeletal medicine.[37] It consisted of three distinct phases: (1) development of USI-based criteria, (2) sonographer education session and (3) inter-rater and intra-rater reliability study (figure 1). The study design established a methodological model for enhancing imaging procedures, assessment and consistency, and reducing rater subjectivity and recall bias. The aim of the criteria development phase was to establish a criteria that was evidence-based and clinically relevant. The reliability phase was separated into three key parts in order to ensure randomisation and blinding, and to reduce recall bias.

Phase 1: criteria development
Following a systematic review,[31] an initial criteria was proposed that incorporated both the continuum model of tendon pathology[9 11] and common parameters used to measure tendinopathy on USI.[31] Additionally, as identified in previous studies,[18 31 34] an ordinal scale was used that included measures of tendon thickness, echogenicity and vascularisation. The criteria were then sent to an expert panel (two musculoskeletal physiotherapists trained in USI, two radiologists and one sonographer) to undergo face validity assessment for feedback and recommendations regarding clinical relevance, design and ease of use. Subsequently, the criteria were further refined to reflect the expert panel feedback (figure 2).

Phase 2: Sonographer training and education
A 2-hour training session was conducted with the primary researcher (WM) along with two sonographers (raters A and B). The sonographers were provided with the USI criteria (figure 2), and an education sheet defining each USI parameter along with example images to guide the education session (table 1). The sonographers were educated on the overall criteria, as well as on
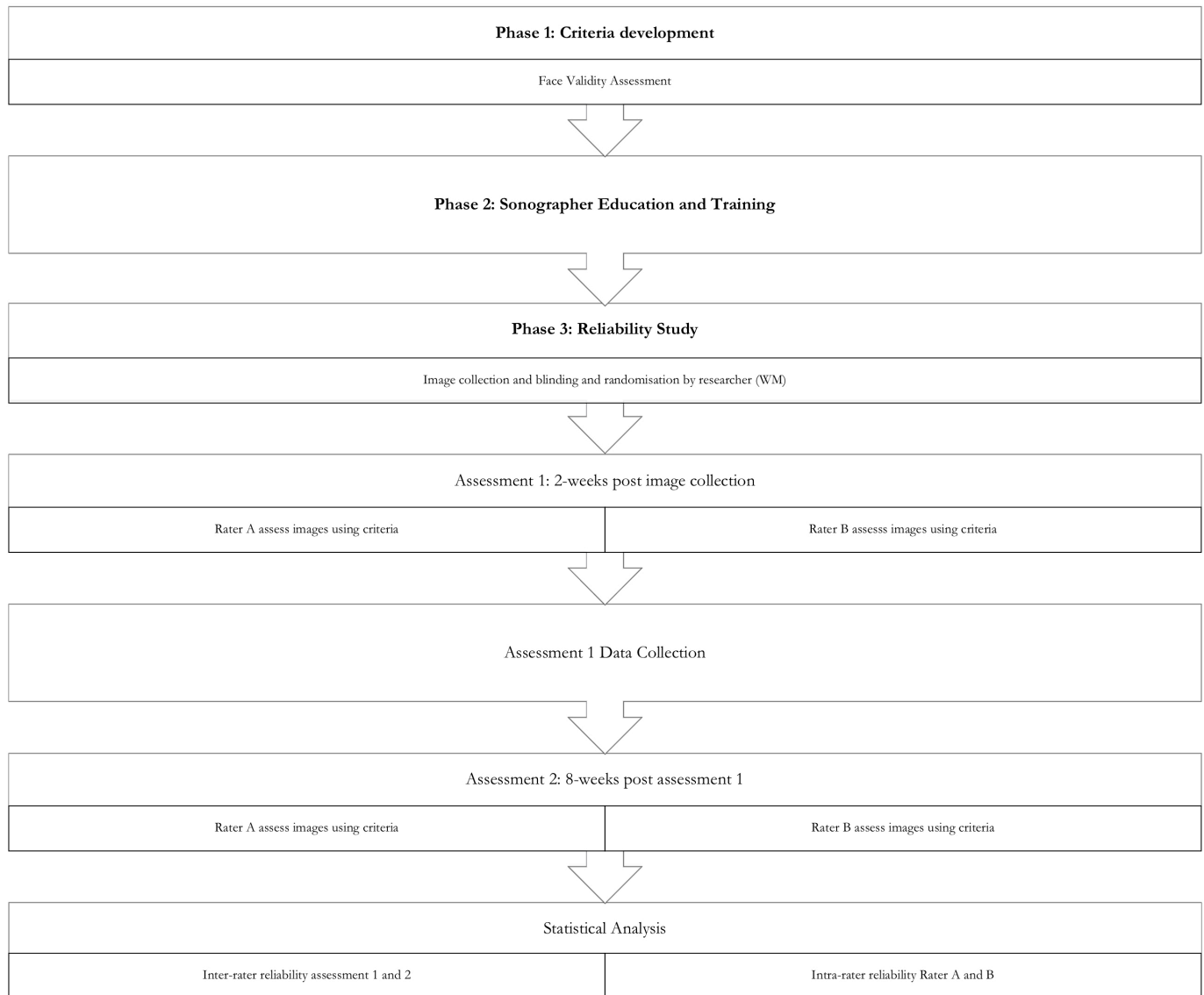
Figure 1 Flow chart of overall study design.

each individual parameter's key features at each of the three stages of tendinopathy. Once the criteria had been explained, additional archived ultrasound images that were not included in the retrospective data collection were used to review and discuss until there was a consensus agreement on the appropriate stage of tendinopathy. A key feature of the continuum model of tendon pathology is the possibility of tendons demonstrating structural changes from different stages of tendon pathology, therefore the likelihood of tendons displaying features from different stages of tendon pathology in each individual parameter was discussed. It was determined that in these instances, the sonographer would need to determine the overall stage of tendinopathy based on their clinical experience and the available data. Once the overall criteria and each individual parameter had been defined, discussed and consensus agreement reached, the sonographers discussed the best way to record the data from

the USI analysis and a scoring table was then drafted and agreed on.

Phase 3: reliability study
Image collection

A retrospective search for suitable ultrasound images of the Achilles tendon, collected as part of usual patient care and stored on a local radiology clinic database between January 2018 and May 2019, was conducted. Gatekeeper consent and approval was obtained prior to commencing the database search. Ultrasound images were deemed eligible for inclusion if they met the following criteria: adults (either gender) aged over 18 years; presenting with Achilles tendon pain; referred for imaging as part of usual care by health practitioner; images contained measurements of tendon thickness and vascularity (colour/Power Doppler). General exclusion criteria were history of surgery; history of systemic inflammatory disease; injection used as an intervention; ultrasound
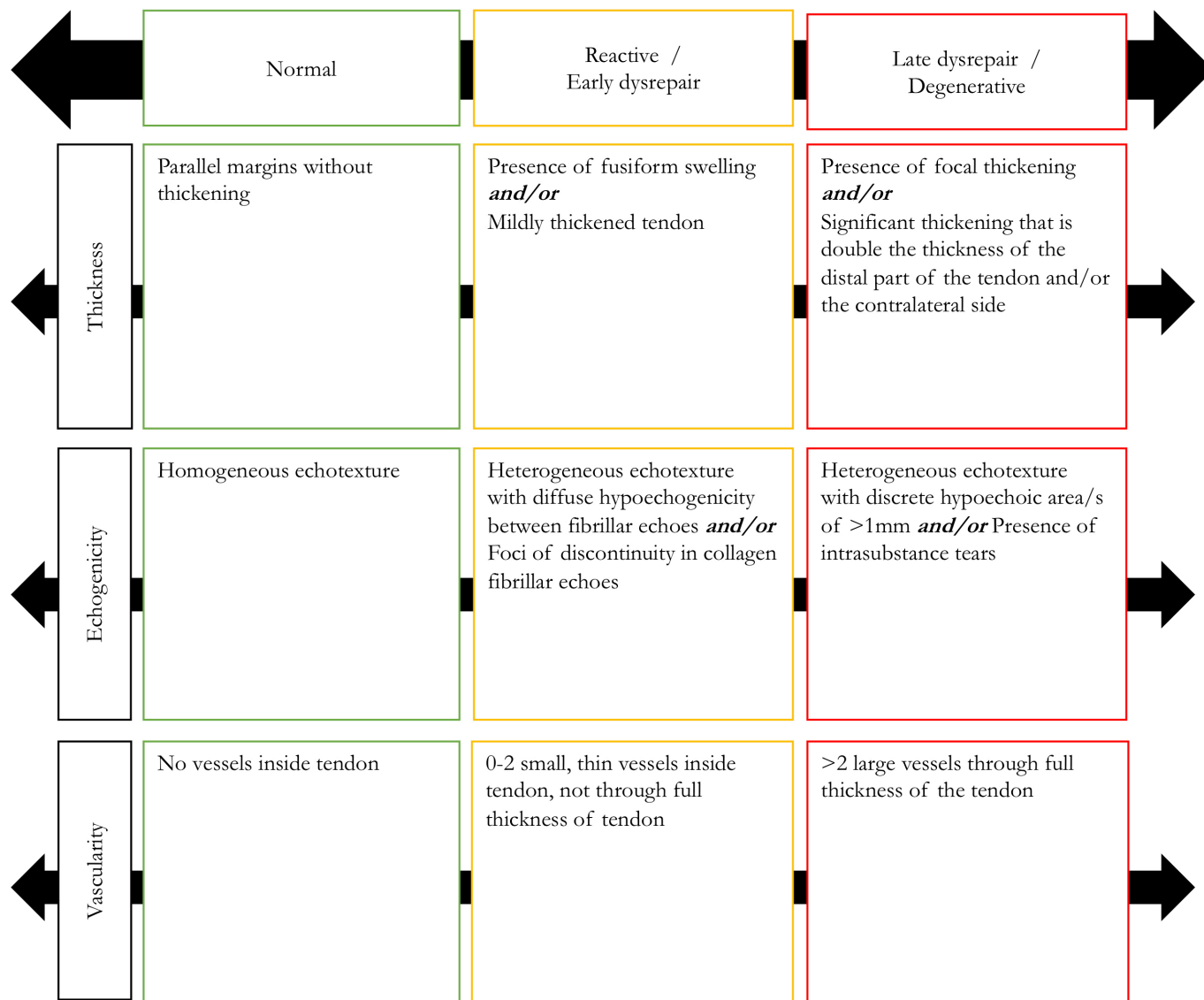
Figure 2 US-based criteria staging tendinopathy using the continuum model of tendon pathology.

diagnosed tendon tear or rupture. Images were extracted in JPEG format by the primary researcher (WM) prior to being de-identified.

Image analysis

Retrospective analysis was performed of all USI images. In total, 31 Achilles tendon images were assessed. Images were digitally stored on an external hard drive in JPEG format and included both pathological and normal Achilles tendons. Once images were collected, one investigator (WM) independently de-identified and randomised the images. The images were then sent to two sonographers (raters A and B), who were blinded to the original diagnosis, for classification according to the USI criteria (figure 2). Each rater individually assessed each image to determine the tendinopathy stage.
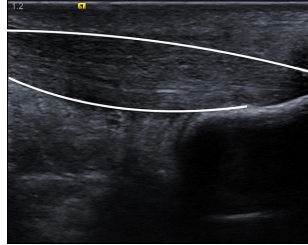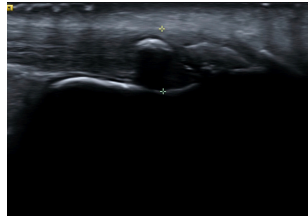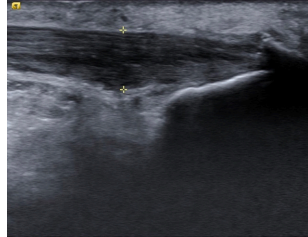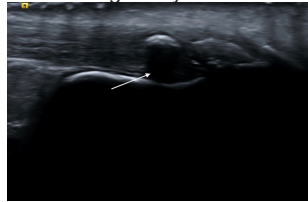
The same images were then re-randomised and sent to the same two sonographers 8 weeks later for re-assessment using the same criteria. Sonographers were again blinded to their original diagnosis and tendinopathy stage. Memory recall bias has been shown to effect intra-rater reliability studies, yet the ideal duration between assessments is unknown,[38] with 'washout' periods ranging from 24 hours to 4 weeks,[39–41] therefore this time frame was doubled to reduce the chance of recall bias.

Statistical analysis
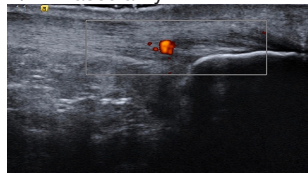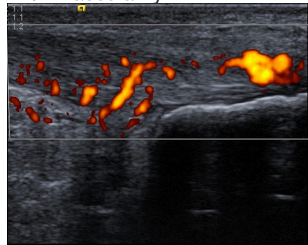
Participant characteristics were reported using mean (SD) and frequency. Statistical analysis was performed following the methods of previous reliability studies.[32 33] Quadratic-weighted Cohen's κ (k_w) with 95% CI was used to calculate intra-rater and inter-rater reliability for all ordinal variables (tendon thickness, vascularity and echogenicity) and overall classification (normal, reactive/early dysrepair or late dysrepair/degenerative). Agreement was interpreted as 'poor' ($k_w \leq 0.00$), 'slight' (k_w 0.01–0.20), 'fair' (k_w 0.21–0.40), 'moderate' (k_w 0.41–0.60), 'substantial' (k_w 0.61–0.80) and 'almost perfect' (k_w 0.81–1.00).[42] A contingency table was produced to display the frequencies and percentages of agreement

Table 1 Sonographer education session using example images

	R/ED	LD/D	Example images
Thickness			
Thickness was measured in millimetre at the thickest point of the tendon. Overall tendon shape and thickness were analysed and compared with the criteria	Presence of fusiform swelling and/or mildly thickened tendon	Presence of focal thickening Significant thickening that is double the thickness of the distal part of the tendon and/or the contralateral side	R/ED thickness LD/D thickness
Echogenicity			
Echogenicity was described as the overall appearance of collagen fascicles on USI. The main differences between stages were determined to be the distribution of fibrillar changes and size of hypoechoic areas	Heterogeneous echotexture with diffuse hypoechogenicity between fibrillar echoes and/or foci of discontinuity in collagen fibrillar echoes	Heterogeneous echotexture with discrete hypoechoic area/s of >1 mm and/or presence of intrasubstance tears	R/ED echogenicity LD/D echogenicity
Vascularity			

Continued

Table 1 Continued

	R/ED	LD/D	Example images
Vascularity was described using the number, size and location of blood vessels throughout the tendon	0–2 small, thin vessels inside tendon, not through full thickness of tendon	>2 large vessels through full thickness of the tendon	R/ED vascularity LD/D vascularity

LD/D, late dysrepair/degenerative; R/ED, reactive/early dysrepair; USI, ultrasound imaging.

between the two raters. The observed proportions of agreement (P_o) and the corresponding 95% CI for each category and the overall classification were also reported. All statistical analyses were performed using the SPSS software package (IBM SPSS Statistics for Macintosh, V.26.0, IBM) and VassarStats statistical software.[43] A sample size of 30 or greater is recommended for reliability studies,[44] thus we used 31 participants.

Patient and public involvement

It was not possible to involve patients in the design, or conduct, or reporting, or dissemination of our research. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS

Participant characteristics are presented in table 2. Of the 31 included participants, ultrasound images for 15 left and 16 right Achilles tendons were assessed. Pooled data for intra-rater reliability were measured, with both raters A and B demonstrating 'substantial' agreement (k_w 0.70–0.77) for the overall staging of tendinopathy (table 3). The intra-rater reliability of individual ultrasound parameters ranged from 'substantial' to 'almost perfect' for

both raters. In addition to intra-rater reliability, inter-rater reliability was measured for both rounds of image assessment (table 4). Pooled data for overall inter-rater agreement for the first round of assessment were determined to be 'substantial' (k_w 0.75, 95% CI 0.58 to 0.91), with this increasing to 'almost perfect' (k_w 0.81, 95% CI 0.63 to 0.99) in the second round. For both rounds of assessment, individual parameter inter-rater reliability ranged from 'substantial' to 'almost perfect'.

The contingency table demonstrates the frequencies and proportions of agreement for each category in the criteria (table 5). Reactive/early dysrepair tendinopathy was the most commonly diagnosed stage of tendinopathy by both rater A (round 1, 54.8%; round 2, 61.3%) and rater B (round 1, 41.9%; round 2, 67.7%). Late dysrepair/degenerative tendinopathy was diagnosed by rater A in 19.4% of participants in round 1 and 22.6% of participants in round 2. Rater B diagnosed late dysrepair/degenerative tendinopathy 32.3% of participants in round 1 and 22.6% of participants in round 2. The overall observed proportion of agreements (P_o) for the staging of tendinopathy using USI-based criteria were

Table 2 Participant characteristics (n=31)

Mean age years (SD)	**60.9 (18.2)**
Sex (M/F)	15/16
Tendon (RAT/LAT)	15/16

F, female; LAT, left Achilles tendon; M, male; RAT, right Achilles tendon.

Table 3 Intra-rater reliability of tendinopathy classification criteria and individual parameters

	Rater A k_w (95% CI)	Rater B k_w (95% CI)
Overall agreement	0.77 (0.59 to 0.94)	0.70 (0.52 to 0.89)
Thickness	0.75 (0.59 to 0.90)	0.84 (0.71 to 0.98)
Echogenicity	0.78 (0.62 to 0.95)	0.73 (0.58 to 0.89)
Vascularity	0.86 (0.74 to 0.98)	0.89 (0.79 to 0.99)

k_w, weighted kappa.

Table 4 Inter-rater reliability of tendinopathy classification criteria and individual parameters

	Round 1 k_w (95% CI)	Round 2 k_w (95% CI)
Overall agreement	0.75 (0.58 to 0.91)	0.81 (0.63 to 0.99)
Thickness	0.65 (0.48 to 0.83)	0.77 (0.60 to 0.93)
Echogenicity	0.70 (0.54 to 0.85)	0.76 (0.58 to 0.94)
Vascularity	0.89 (0.79 to 0.99)	0.86 (0.74 to 0.98)

k_w, weighted kappa.

0.74 (95% CI 0.55 to 0.87) in round 1 and 0.87 (95% CI 0.69 to 0.95) in round 2.

DISCUSSION

The intra-rater and inter-rater reliability ranged from 'substantial' to 'almost perfect' for identifying individual features (tendon thickness, echogenicity and vascularity) and determining overall staging of tendinopathy, using standardised criteria. Inter-rater agreement increased from round 1 (k_w 0.75, 95% CI 0.58 to 0.91) to round 2 (k_w 0.81, 95% CI 0.63 to 0.99), which may indicate that as raters became more acquainted with the criteria there was a learning effect which resulted in improved reliability.

Tendon thickness

Tendon thickness demonstrated the largest variation in intra-rater reliability, with rater A demonstrating 'substantial' agreement (k_w 0.75) and rater B demonstrating 'almost perfect' agreement (k_w 0.84). Similarly, there was variation in inter-rater reliability, with tendon thickness demonstrating the lowest agreement in round 1 (k_w 0.65) and having had the largest improvement in agreement in round 2 (k_w 0.77). While other studies have demonstrated good reliability of tendon thickness using continuous measures,[33 45–47] the current study is the first to report reliability of tendon thickness measured

in a categorical manner. The decision to use categorical criteria followed the face validity assessment, where the expert panel highlighted inaccuracies when measuring small changes in tendon thickness, and difficulty using either the distal part of the same tendon or contralateral tendon as a reference point due to the frequent presence of asymptomatic tendon thickening.

Previous studies have demonstrated 'almost perfect' intra-rater and inter-rater reliability when using the intraclass correlation coefficient (ICC) to assess Achilles tendon thickness using continuous variables.[33 45] Although there is difficulty in comparing ICC and kappa statistics numerically, the grading of the results allows comparisons to be made. Hence, intra-rater reliability of these previous studies was comparable with the 'almost perfect' reliability of rater B and better than 'substantial' agreement achieved by rater A.[33 45] Similarly, previous studies[33 45] have demonstrated 'almost perfect' inter-rater reliability when using the ICC compared with the 'substantial' agreement demonstrated in this study. This variation in both intra-rater and inter-rater reliability is in line with a recent systematic review in which both the intra-rater agreement (ICC 0.65–0.94) and inter-rater agreement (ICC 0.65–0.84) ranged from 'good' to 'almost perfect'.[48]

Echogenicity

The scoring of echogenicity using the proposed criteria demonstrated 'substantial' intra-rater and inter-rater reliability. Intra-rater agreement (rater A, k_w 0.78; rater B, k_w 0.73) was consistent with results reported by Sunding et al,[33] in which intra-rater agreement for the measurement of tendon echogenicity ranged from k 0.54 to 0.84. Inter-rater agreement (round 1, k_w 0.70; round 2, k_w 0.76) for the current study was higher than those reported in previous Achilles tendon studies (k 0.07–0.58).[33 49] When comparing the current results with those of other studies that have measured echogenicity in different tendons,

Table 5 Contingency table of frequencies and percentages of agreement between raters A and B for the ultrasound-based diagnosis of tendinopathy using standardised criteria

	Rater B: n (%)				
Rater A: n (%)	Normal	R/ED	LD/D	Total (%)	P_o (95% CI)
Round 1					
Normal	6 (19.4)	2 (6.5)	0 (0.0)	8 (25.8)	0.60 (0.27 to 0.86)
R/ED	2 (6.5)	11 (35.5)	4 (12.9)	17 (54.8)	0.57 (0.33 to 0.78)
LD/D	0 (0.0)	0 (0.0)	6 (19.4)	6 (19.4)	0.60 (0.27 to 0.86)
Total	8 (25.8)	13 (41.9)	10 (32.3)	31 (100.0)	0.74 (0.55 to 0.87)
Round 2					
Normal	3 (9.7)	2 (6.5)	0 (0.0)	5 (16.1)	0.60 (0.17 to 0.92)
R/ED	0 (0.0)	18 (58.1)	1 (3.2)	19 (61.3)	0.81 (0.58 to 0.94)
LD/D	0 (0.0)	1 (3.2)	6 (19.4)	7 (22.6)	0.75 (0.35 to 0.95)
Total	3 (9.7)	21 (67.7)	7 (22.6)	31 (100.0)	0.87 (0.69 to 0.95)

LD/D, late dysrepair/degenerative; P_o, observed proportion of agreement; R/ED, reactive/early dysrepair.

intra-rater agreement was lower than those reported for the patella tendon (k 0.78–0.87)[33] and similar to those reported for the supraspinatus tendon (k 0.76–0.79).[50] However, the proposed criteria used within the current study demonstrated better inter-rater agreement than that previously reported for the patella tendon (k 0.52–0.60)[33] and supraspinatus tendon (k 0.51–60),[32] and it is comparable with using sonoelastography to assess echogenicity for the supraspinatus tendon (k 0.71–0.81).[50] Agreement on structural changes within a tendon is considered difficult, with previous studies using an ordinal scale ('normal', 'mild abnormality', 'severe abnormality') that stages tendinopathy according to subtle changes that may be difficult to distinguish from anisotropy.[32 33] By utilising criteria that bases diagnosis on the distribution of structural changes ('diffuse' or 'focal' echogenicity), as suggested by the continuum model of tendon pathology,[9 11] rather than just the presence of structural change, reliability may be increased.

Vascularity

Vascularity represented the most reliable of the USI-based parameters measured, with both intra-rater and inter-rater agreement being 'almost perfect'. The intra-rater agreement (rater A, k_w 0.86; rater B, k_w 0.89) is improved than that reported by Sunding *et al*[33] where intra-rater agreement for Achilles tendon vascularity ranged from k 0.64 to 0.78. When compared with studies that have examined vascularity in other tendons, the current results align with those reported for the patella tendon (k 0.79–0.86).[33] Inter-rater agreement (round 1, k_w 0.89; round 2, k_w 0.86) in this study was also higher than that reported by Sunding *et al*[33] where inter-rater agreement for the measurement of tendon vascularity ranged from k 0.59 to 0.87. Similarly, the current results are higher than those reported for the patella tendon (k 0.45–0.76) and similar to those reported for the supraspinatus tendon (k 0.96–0.98).[32] The differences in reliability seen from the current study compared with those previously reported may be due to the design of the criteria used in the current study. While Sunding *et al*[33] used a scale based on subjective interpretations of the quantity of blood vessels present ('mild', 'moderate' and 'severe'), the current study used more objective measures such as number and size of blood vessels present. Similarly, Ingwersen *et al*[32] used a more objective measure of the amount of Doppler activity present, expressed as a percentage of the region of interest (<25%, 25%–50%, >50%), hence the similarity of results to the current study.

The continuum model of tendon pathology

Overall, the proposed criteria demonstrated 'substantial' to 'almost perfect' intra-rater and inter-rater reliability for staging Achilles tendinopathy using the continuum model of tendon pathology. This study aligns with others in demonstrating that by using a standardised USI-based criteria to stage tendon pathology, both intra-rater and inter-rater reliability are increased.[32 33 51] While many previous studies have staged tendinopathy using a variety of different criteria,[25 30 32 35 52–59] no study has developed an USI-based criteria that utilises the continuum model of tendon pathology.[31 34] This is the first study to propose an USI criteria that is aligned with the continuum model of tendon pathology which is widely accepted and utilised to diagnose tendinopathy based on specific clinical and imaging findings.[9 11] Although the role of imaging in the diagnosis of tendinopathy is debated,[11 13 17 54 60 61] imaging may provide additional information that may help stage tendinopathy and better direct treatment.[11] Furthermore, there is scope to develop reliable and valid criteria that use accepted terminology to describe pathological tendon changes.[1]

Clinical relevance

While the proposed criteria demonstrates 'substantial' to 'almost perfect' intra-rater and inter-rater reliability for staging Achilles tendinopathy using the continuum model of tendon pathology, it is important to note that the diagnosis of tendinopathy is multifactorial.[1 13] The role of USI in the diagnosis of tendinopathy remains challenging due to a lack of clinical gold standard to compare images.[29] Therefore, while USI can be utilised to assist in the diagnosis of tendinopathy,[1 13] the clinical diagnosis of tendinopathy does not require the presence of structural changes as shown on imaging.[1] Rather, by standardising the method with which tendon pathological change is reported and aligning terminology with current conceptual models of tendon pathology, larger multicentre studies can be conducted to further investigate the clinical relevance of USI in the diagnosis and management of tendon pathology and they will allow for greater transferability between studies.

Limitations

The main limitation in this study is the retrospective design of the study. This did not allow for control of imaging procedure, machine, sonographer or imaging settings. While standardised imaging procedures, settings and machines have been shown to increase reliability,[32 33 46 51] there is a question as to whether there is transferability to the clinical setting, where images are captured using a variety of machines and imaging procedures.[32] Standardised protocols can make reliability appear artificially high when compared with the clinical setting,[32] however if reliability is low with a standardised protocol, reliability can be assumed to be poor and less relevant for the clinical setting.[32] Additionally, included images were static in nature which may affect the external validity and reliability of the criteria as clinically USI is considered a dynamic investigation.[62] Further study limitations include the small sample size, and assessment of only the Achilles tendon, which may make generalisation of the results difficult. While a larger sample size will give greater confidence in results, a sample size of 30–50 has been suggested as sufficient to determine clinical relevance.[44] Although only the Achilles tendon was assessed in this

study, other studies have shown that standardised criteria are reliable in assessing tendon pathological changes in the patella and supraspinatus tendons.[32 63 64]

CONCLUSION

Intra-rater and inter-rater reliability were 'substantial' to 'almost perfect' when utilising an USI-based criteria to diagnose Achilles tendinopathy based on the continuum model of tendon pathology. This is the first study to use the continuum model of tendon pathology to develop an USI-based criteria to diagnose tendinopathy. The proposed USI criteria aligns with accepted clinical terminology used to describe tendinopathy and provides an objective criteria to reliably stage tendinopathy. Future research should test both the intra-rater and inter-rater reliability of the criteria in a prospective manner within the clinical setting and in other tendons. Additionally, future research should explore the correlation of clinical assessment with the proposed criteria.

Author affiliations
[1]Bond Institute of Health and Sport, Bond University Faculty of Health Sciences and Medicine, Gold Coast, Queensland, Australia
[2]Active Living and Rehabilitation: Aotearoa New Zealand, Health and Rehabilitation Research Institute, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand
[3]Department of Physiotherapy, School of Clinical Sciences, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand
[4]Bond University Faculty of Health Sciences and Medicine, Gold Coast, Queensland, Australia

ORCID iD
Wesley Matthews http://orcid.org/0000-0002-2281-0464

REFERENCES

1 Scott A, Squier K, Alfredson H, *et al*. Icon 2019: international scientific tendinopathy symposium consensus: clinical terminology. *Br J Sports Med* 2020;54:260–2.
2 Scott A, Ashe MC. Common tendinopathies in the upper and lower extremities. *Curr Sports Med Rep* 2006;5:233–41.
3 McCreesh K, Lewis J. Continuum model of tendon pathology—where are we now? *Int J Exp Pathol* 2013;94:242–7.
4 Rees JD, Maffulli N, Cook J. Management of tendinopathy. *Am J Sports Med* 2009;37:1855–67.
5 Cassel M, Baur H, Hirschmüller A, *et al*. Prevalence of Achilles and patellar tendinopathy and their association to intratendinous changes in adolescent athletes. *Scand J Med Sci Sports* 2015;25:e310–8.
6 Hopkins C, Fu S-C, Chua E, *et al*. Critical review on the socio-economic impact of tendinopathy. *Asia Pac J Sports Med Arthrosc Rehabil Technol* 2016;4:9–20.
7 Abate M, Silbernagel KG, Siljeholm C, *et al*. Pathogenesis of tendinopathies: inflammation or degeneration? *Arthritis Res Ther* 2009;11:235–35.
8 Arnoczky SP, Lavagnino M, Egerbacher M. The mechanobiological aetiopathogenesis of tendinopathy: is it the over-stimulation or the under-stimulation of tendon cells? *Int J Exp Pathol* 2007;88:217–26.
9 Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *Br J Sports Med* 2009;43:409–16.
10 Fu S-C, Rolf C, Cheuk Y-C, *et al*. Deciphering the pathogenesis of tendinopathy: a three-stages process. *Sports Med Arthrosc Rehabil Ther Technol* 2010;2:30.
11 Cook JL, Rio E, Purdam CR, *et al*. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research? *Br J Sports Med* 2016;50:1187–91.
12 Rees JD, Stride M, Scott A. Tendons—time to revisit inflammation. *Br J Sports Med* 2014;48:1553–7.
13 Scott A, Docking S, Vicenzino B, *et al*. Sports and exercise-related tendinopathies: a review of selected topical issues by participants of the second International scientific tendinopathy symposium (ISTS) Vancouver 2012. *Br J Sports Med* 2013;47:536–44.
14 Coombes BK, Bisset L, Vicenzino B. Management of lateral elbow tendinopathy: one size does not fit all. *J Orthop Sports Phys Ther* 2015;45:938–49.
15 Lewis J, McCreesh K, Roy J-S, *et al*. Rotator cuff tendinopathy: navigating the diagnosis-management conundrum. *J Orthop Sports Phys Ther* 2015;45:923–37.
16 Lewis J. Rotator cuff related shoulder pain: assessment, management and uncertainties. *Man Ther* 2016;23:57–68.
17 Malliaras P, Cook J, Purdam C, *et al*. Patellar tendinopathy: clinical diagnosis, load management, and advice for challenging case presentations. *J Orthop Sports Phys Ther* 2015;45:887–98.
18 Scase E, Purdam C, Cook J. Matching tendinopathy stage with efficacious intervention. *Soundeffects* 2011;4:14–18.
19 Maffulli N, Khan KM, Puddu G. Overuse tendon conditions: time to change a confusing terminology. *Arthroscopy* 1998;14:840–3.
20 Littlewood C, Malliaras P, Bateman M, *et al*. The central nervous system-—an additional consideration in 'rotator cuff tendinopathy' and a potential basis for understanding response to loaded therapeutic exercise. *Man Ther* 2013;18:468–72.
21 Coombes BK, Bisset L, Vicenzino B. A new integrative model of lateral epicondylalgia. *Br J Sports Med* 2009;43:252–8.
22 Cook JL, Khan KM, Kiss ZS, *et al*. Reproducibility and clinical utility of tendon palpation to detect patellar tendinopathy in young basketball players. Victorian Institute of sport tendon Study Group. *Br J Sports Med* 2001;35:65.
23 Grimaldi A, Mellor R, Nicolson P, *et al*. Utility of clinical tests to diagnose MRI-confirmed gluteal tendinopathy in patients presenting with lateral hip pain. *Br J Sports Med* 2017;51:519–24.
24 Hutchison A-M, Evans R, Bodger O, *et al*. What is the best clinical test for Achilles tendinopathy? *Foot Ankle Surg* 2013;19:112–7.
25 Khan KM, Forster BB, Robinson J, *et al*. Are ultrasound and magnetic resonance imaging of value in assessment of Achilles tendon disorders? A two year prospective study. *Br J Sports Med* 2003;37:149–53.
26 Warden SJ, Kiss ZS, Malara FA, *et al*. Comparative accuracy of magnetic resonance imaging and ultrasonography in confirming clinically diagnosed patellar tendinopathy. *Am J Sports Med* 2007;35:427–36.
27 Westacott DJ, Minns JI, Foguet P. The diagnostic accuracy of magnetic resonance imaging and ultrasonography in gluteal tendon tears—a systematic review. *Hip Int* 2011;21:637–45.
28 Rio E, Moseley L, Purdam C, *et al*. The pain of tendinopathy: physiological or pathophysiological? *Sports Med* 2014;44:9–23.

29 Docking SI, Ooi CC, Connell D. Tendinopathy: is imaging telling us the entire story? *J Orthop Sports Phys Ther* 2015;45:842–52.

30 McAuliffe S, McCreesh K, Culloty F, *et al*. Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? A systematic review and meta-analysis. *Br J Sports Med* 2016;50:1516–23.

31 Matthews W, Ellis R, Furness J, *et al*. Classification of tendon matrix change using ultrasound imaging: a systematic review and meta-analysis. *Ultrasound Med Biol* 2018;44:2059–80.

32 Ingwersen KG, Hjarbaek J, Eshoej H, *et al*. Ultrasound assessment for grading structural tendon changes in supraspinatus tendinopathy: an inter-rater reliability study. *BMJ Open* 2016;6:e011746.

33 Sunding K, Fahlström M, Werner S, *et al*. Evaluation of Achilles and patellar tendinopathy with greyscale ultrasound and colour Doppler: using a four-grade scale. *Knee Surg Sports Traumatol Arthrosc* 2016;24:1988–96.

34 Ellis R, Manuel T. Identifying tendon matrix changes with ultrasound imaging: a literature review and proposed ultrasound criteria for staging for tendinopathy. *Physiotherapy* 2015;101:eS357.

35 Malliaras P, Purdam C, Maffulli N, *et al*. Temporal sequence of greyscale ultrasound changes and their relationship with neovascularity and pain in the patellar tendon. *Br J Sports Med* 2010;44:944–7.

36 Drew BT, Smith TO, Littlewood C, *et al*. Do structural changes (EG, collagen/matrix) explain the response to therapeutic exercises in tendinopathy: a systematic review. *Br J Sports Med* 2014;48:966–72.

37 Patijn J. Reproducibility and validity studies of diagnostic procedures in manual/musculoskeletal medicine protocol formats, 2nd edition.. *Journal of Orthopaedic Medicine* 2002;24:58–70.

38 Boone D, Halligan S, Mallett S, *et al*. Systematic review: bias in imaging studies—the effect of manipulating clinical context, recall bias and reporting intensity. *Eur Radiol* 2012;22:495–505.

39 Sarris I, Ioannou C, Dighe M, *et al*. Standardization of fetal ultrasound biometry measurements: improving the quality and consistency of measurements. *Ultrasound Obstet Gynecol* 2011;38:681–7.

40 Tudorache S, Cara M, Iliescu DG, *et al*. First trimester two- and four-dimensional cardiac scan: intra- and interobserver agreement, comparison between methods and benefits of color Doppler technique. *Ultrasound Obstet Gynecol* 2013;42:659–68.

41 Chen W, Lou EHM, Zhang PQ, *et al*. Reliability of assessing the coronal curvature of children with scoliosis by using ultrasound images. *J Child Orthop* 2013;7:521–9.

42 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.

43 Lowry R. *VassarStats: website for statistical computation*. New York, USA: Vassar College, 2019.

44 Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil* 2005;84:719–23.

45 Del Baño-Aledo ME, Martínez-Payá JJ, Ríos-Díaz J, *et al*. Ultrasound measures of tendon thickness: intra-rater, inter-rater and inter-machine reliability. *Muscles Ligaments Tendons J* 2017;7:192–9.

46 Poltawski L, Ali S, Jayaram V, *et al*. Reliability of sonographic assessment of tendinopathy in tennis elbow. *Skeletal Radiol* 2012;41:83–9.

47 Gellhorn AC, Carlson MJ. Inter-rater, intra-rater, and inter-machine reliability of quantitative ultrasound measurements of the patellar tendon. *Ultrasound Med Biol* 2013;39:791–6.

48 Mc Auliffe S, Mc Creesh K, Purtill H, *et al*. A systematic review of the reliability of diagnostic ultrasound imaging in measuring tendon size: is the error clinically acceptable? *Phys Ther Sport* 2017;26:52–63.

49 Archambault JM, Wiley JP, Bray RC, *et al*. Can sonography predict the outcome in patients with achillodynia? *J Clin Ultrasound* 1998;26:335–9.

50 Boesen AP, Boesen MI, Torp-Pedersen S, *et al*. Associations between abnormal ultrasound color Doppler measures and tendon pain symptoms in badminton players during a season: a prospective cohort study. *Am J Sports Med* 2012;40:548–55.

51 Comin J, Cook JL, Malliaras P, *et al*. The prevalence and clinical significance of sonographic tendon abnormalities in asymptomatic ballet dancers: a 24-month longitudinal study. *Br J Sports Med* 2013;47:89–92.

52 Cook JL, Khan KM, Kiss ZS, *et al*. Asymptomatic hypoechoic regions on patellar tendon ultrasound: a 4-year clinical and ultrasound follow-up of 46 tendons. *Scand J Med Sci Sports* 2001;11:321–7.

53 Cook JL, Khan KM, Kiss ZS, *et al*. Prospective imaging study of asymptomatic patellar tendinopathy in elite junior basketball players. *J Ultrasound Med* 2000;19:473–9.

54 de Jonge S, de Vos RJ, Van Schie HTM, *et al*. One-Year follow-up of a randomised controlled trial on added splinting to eccentric exercises in chronic midportion Achilles tendinopathy. *Br J Sports Med* 2010;44:673–7.

55 Khan KM, Cook JL, Kiss ZS, *et al*. Patellar tendon ultrasonography and jumper's knee in female basketball players: a longitudinal study. *Clin J Sport Med* 1997;7:199–206.

56 van Schie HTM, de Vos RJ, de Jonge S, *et al*. Ultrasonographic tissue characterisation of human Achilles tendons: quantification of tendon structure through a novel non-invasive approach. *Br J Sports Med* 2010;44:1153–9.

57 Docking S, Rosengarten S, Daffy J, *et al*. The role of ultrasound tissue characterisation in the management and prevention of tendinopathy in athletes. *J Sci Med Sport* 2013;16:e85–6.

58 Docking S, Rio E, Cook J. When pictures tell more than a thousand words: the clinical role of imaging for tendinopathy and emergence of UTC. *J Sci Med Sport* 2013;16:e85.

59 Docking SI, Rosengarten SD, Cook J. Achilles tendon structure improves on UTC imaging over a 5-month pre-season in elite Australian football players. *Scand J Med Sci Sports* 2016;26:557–63.

60 Docking SI, Cook J. Pathological tendons maintain sufficient aligned fibrillar structure on ultrasound tissue characterization (UTC). *Scand J Med Sci Sports* 2016;26:675–83.

61 Docking S, Bedi H, Jowett C, *et al*. Improvement in pain and function is not mediated by changes in tissue structure following surgery for Achilles tendinopathy. *J Sci Med Sport* 2015;19:e100.

62 Mapes-Gonnella T. The impact of education. *J Diagn Med Sonog* 2013;29:260–8.

63 Naredo E, Möller I, Moragues C, *et al*. Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. *Ann Rheum Dis* 2006;65:14–9.

64 Brage K, Hjarbaek J, Kjaer P, *et al*. Ultrasonic strain elastography for detecting abnormalities in the supraspinatus tendon: an intra- and inter-rater reliability study. *BMJ Open* 2019;9:e027725.