


Clinical Research

The Diagnostic Accuracy of Clinical Tests in Identifying Structural Change in Achilles Tendinopathy

A Pilot Study

Wesley Matthews ,
Richard Ellis,
James Furness,
Evelyne Rathbone, and
Wayne Hing

Abstract: Background. With the development of a consensus agreement on the essential clinical diagnostic criteria for Achilles tendinopathy (AT), there is scope to investigate the diagnostic accuracy of clinical tests, their relationship to structural changes observed on ultrasound imaging (USI), and the potential role of USI in the clinical diagnosis of AT. Objectives. To evaluate the relationship between clinical tests and tendon structure via USI. Methods. A pilot cross-sectional study of 23 individuals (14 male, and 9 female) with unilateral, symptomatic AT were recruited from physiotherapy clinics. Assessment included subjective measures and patient-reported outcome measures (pain with loading, stiffness, self-reported function, Victorian Institute of Sport Assessment-Achilles [VISA-A], Pain Catastrophizing Scale, and 12-Item Short Form Survey). Objective tests included palpation, arc sign, Royal London Hospital Test, single-leg

heel raise, and hopping. USI subcategorized tendon structure as normal or abnormal. Diagnostic accuracy (positivity rates, sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV]), receiver operating characteristic analysis, area under the curve (AUC), and effect sizes were calculated. Results. Stiffness and pain on palpation demonstrated high positivity rates (82.6%). Hopping pain demonstrated high positivity rates (78.3%). Pain on palpation and pain during hopping had high sensitivity (0.94) and PPV (0.84) for detecting structural abnormalities. VISA-A and self-reported function had excellent AUC (0.89) for detecting structural change. No test distinguished the degree of structural change. Conclusions. Pain on palpation and

hopping pain are promising indicators of structural tendon pathology. Patient-reported outcome measures may aid in imaging decisions. Further studies are needed to validate findings.

“Pain on palpation and hopping pain are promising indicators of structural tendon pathology.”

Level of Evidence: Level IV

Keywords: tendinopathy; ultrasound; diagnosis; Achilles

Introduction

Achilles tendinopathy (AT) is defined as the clinical presentation of pain localized to the Achilles tendon, and the associated loss of function with tendon

DOI: 10.1177/19386400261434148. From Bond Institute of Health and Sport, Faculty of Health Sciences and Medicine, Bond University, Robina, Gold Coast, QLD, Australia (WM, JF, WH); Physical Activity, Sport, and Exercise Research Theme (PASER), Faculty of Health, Southern Cross University, Gold Coast, QLD, Australia (WM); 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 (RE); Department of Physiotherapy, School of Clinical Sciences, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand (RE); Faculty of Health Sciences and Medicine, Bond University, Gold Coast, QLD, Australia (ER). Address correspondence to: Wesley Matthews, Faculty of Health, Southern Cross University, Terminal Drive, Bilinga, Gold Coast, Queensland 4225, Australia; e-mail: wesley.matthews@scu.edu.au.

For reprints and permissions queries, please visit <http://www.sagepub.com/journalsPermissions.nav>.

Copyright © 2026 The Author(s)



loading activities.¹⁻³ AT is primarily diagnosed with a combination of patient history and clinical tests.⁴ However, there is a lack of consistency in the methods used for the clinical diagnosis of AT.⁵⁻⁷ Clinical tests are commonly compared with a reference standard, such as ultrasound imaging (USI),⁶ where the presence of structural change can be evaluated.^{8,9} Although USI has demonstrated both sensitivity and accuracy for detecting pathological structural tendon change,¹⁰⁻¹² it does not always correlate with clinical symptoms and function,¹³⁻¹⁵ with asymptomatic Achilles tendons demonstrating structural change.¹⁵ However, pathological change identified on USI has been shown to be a risk factor for developing symptoms in asymptomatic AT.^{16,17}

A recent scoping review identified the most common methods for clinically diagnosing AT.⁷ Similarly, a recent Delphi study provided consensus guidance for the clinical diagnosis of AT, including differential diagnosis.⁶ Consensus for essential diagnostic criteria included pain location, pain during activity, tests that provoke pain and palpation to assess pain.⁶ The diagnostic domains of “pain onset,” “morning or arising pain or perceived stiffness,” and “self-reported function” that were included in the scoping review,⁷ reached consensus agreement but were considered not essential to achieve a clinical diagnosis of AT.⁶ Notably, imaging tendon pathology reached consensus, although it was not essential for diagnosing AT if the essential diagnostic criteria were met.⁶ Of note, there was debate and discussion within the expert panel as to whether the stage of disease, the lack of a precise definition of normal imaging, the poor sensitivity of imaging modalities, and whether tendinopathy can be excluded in the presence of normal imaging was presented.⁶

To determine the validity of diagnostic tests, imaging is commonly used as the reference standard,⁶ with USI the preferred method for assessing tendon structure.^{18,19} However, a lack of consistency in the application of USI

methods has been suggested as a contributing factor in the uncertainty as to the relevance of USI in the diagnosis of AT.^{10,16,20} Standardization of the application of USI in the assessment of AT has been demonstrated as a reliable method for assessing tendon structural change,^{21,22} and may offer a more consistent reference standard for determining the diagnostic accuracy of clinical tests. Recently, a reliable method for staging AT based on USI features was developed,²¹ subclassifying AT as “Normal,” “Reactive/Early Dysrepair” (R/ED), and “Late Dysrepair/Degenerative” (LD/D). This provides a unified method for describing the degree of structural change identified on USI. It has been proposed that identifying subclassifications of AT may allow clinicians to tailor treatment programs to the appropriate stage of tendinopathy and more accurately monitor progress.^{1,2,13}

With the development of a consensus agreement on the essential clinical diagnostic criteria to diagnose AT,⁶ and a consistent method for describing the degree of structural change on USI,²¹ there is scope to investigate the diagnostic accuracy of clinical tests, their relationship to structural change observed on USI, and the potential role of USI in the clinical diagnosis of AT. Thus, the primary aim of this pilot cross-sectional study was to explore methodological issues in assessing the diagnostic accuracy of previously identified⁷ clinical tests in AT, as determined using USI as the reference standard, in participants with symptomatic clinically diagnosed AT. Furthermore, this study aimed to provide guidance for future larger research studies in developing a more uniform method for diagnosing AT and defining the role of USI in both research and clinical practice.

Methods

The reporting of the pilot diagnostic accuracy study followed the Standards for Reporting Diagnostic Accuracy Studies checklist.²³

Participants

A convenience sample of participants was recruited between March 2022 and March 2023 from local physiotherapy clinics on the Gold Coast, Australia. All potential participants were seeking treatment for symptomatic AT, with a diagnosis of AT determined by their treating physiotherapist, not involved in this study. Potential participants were screened by the lead researcher (WM) to ensure they met the inclusion criteria (Table 1). They were provided with an explanatory statement and gave consent to participate in the study. The same physiotherapist (WM) completed both the clinical assessment and USI assessment. The clinical assessment process is provided in Figure 1. Ethical approval for the study was provided by Bond University Human Research Council (WM00028).

The index tests for this study were extracted from a recent scoping review.⁷ Table 2 provides an overview of the subjective assessment, patient-reported outcome measures (PROMs), and objective index tests used.

Subjective Assessment

The subjective assessment included questions on the location of pain, pain with tendon loading and tendon stiffness.

Patient-Reported Outcome Measures

Victorian Institute of Sport Assessment—Achilles (VISA-A). The VISA-A is a valid and reliable tool that has been widely used to assess severity and disability for people with AT.^{7,24-27} The VISA-A contains eight questions covering the domains of pain (questions 1-3), function (questions 4-6), and activity (questions 7-8).²⁵ The maximum score is 100, with healthy subjects scoring a minimum of 96.²⁵

Pain Catastrophizing Scale. The PCS is a valid and reliable instrument used to measure catastrophic thinking related to pain,²⁸ is one of the more common instruments used to assess a person's

Table 1.

Participant Inclusion and Exclusion Criteria.

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> • Clinical diagnosis of Achilles tendinopathy made by treating physiotherapist • Over 18 years of age • Current Achilles tendon symptoms 	<ul style="list-style-type: none"> • Previous Achilles tendon tear or rupture confirmed by ultrasound imaging • Previous Achilles tendon surgery • Previous injection into the Achilles tendon • Diagnosed with inflammatory disease (e.g., Rheumatoid Arthritis)

pain experience with AT,^{7,24} and has been used as a method to assess the psychological effect of AT.²⁹ Participants are asked to indicate the degree to which they have specific thoughts and feelings when experiencing pain, using a scale from 0 (not at all) to 4 (all the time). A total score is yielded (ranging from 0 to 52), with a score of 30 or higher indicating a clinically relevant level of catastrophising.²⁸

12-Item Short Form Survey. The 12-Item Short Form Survey (SF-12) is a questionnaire commonly used to assess quality of life in people with AT.^{7,24,30} It contains 12 questions that assess eight key domains: physical limitations due to health difficulties, social limitations due to physical or emotional difficulties, limitations in activities in daily living due to physical difficulties, physical pain, general mental health, limitations in activities in daily living due to emotional difficulties, vitality, and general health perceptions.³¹ Scoring was completed using version 1.0 of the SF-12.³²

Objective Assessment

Objective assessment of the Achilles tendon was conducted after USI using commonly reported measures.^{7,24,33-35} Tests were performed in a standardized order: palpation, the arc sign, and Royal London Hospital Test (RLHT), followed by clinician demonstration and verbal instruction for the single-leg heel raise (SLHR) and hopping tests. Each test was

completed on the non-affected limb first, with rest periods to allow symptoms to return to baseline before proceeding.

Palpation. Tendon palpation (pain, swelling, and thickness) is the most commonly utilized clinical test to diagnose AT.⁷ Palpation has previously demonstrated moderate specificity (73%-85%) and sensitivity (58%-84%) when compared with USI in the diagnosis of AT.^{33,34} Palpation was performed as previously reported.^{7,33,34}

The arc sign. The arc sign has been shown to demonstrate low to moderate sensitivity (25%-53%) and high specificity (83%-100%) for assessing AT when compared with USI.^{33,34} The arc sign was performed as previously reported.^{7,33,34} If intratendinous swelling was not present, the thickest portion of the tendon was palpated.

Royal London Hospital Test. The RLHT has previously demonstrated moderate sensitivity (51%-54%) and high specificity (91%-93%) when compared with USI in assessing AT.^{33,34} The RLHT was performed as previously reported.^{7,33,34}

Single-leg heel raise. The SLHR is commonly utilized in the diagnosis of AT to assess pain during tendon loading and functional capacity.^{7,24,33} It has demonstrated low sensitivity (22%) and high specificity (93%) compared with USI.³³ The test was performed as previously described,^{8,36} on both the

affected and non-affected legs, and ceased when full range could no longer be achieved or pain prevented continuation. Pain was recorded using a Visual Analogue Scale (VAS), with clinicians documenting total repetitions, repetitions before pain onset, and pain test completion.⁷

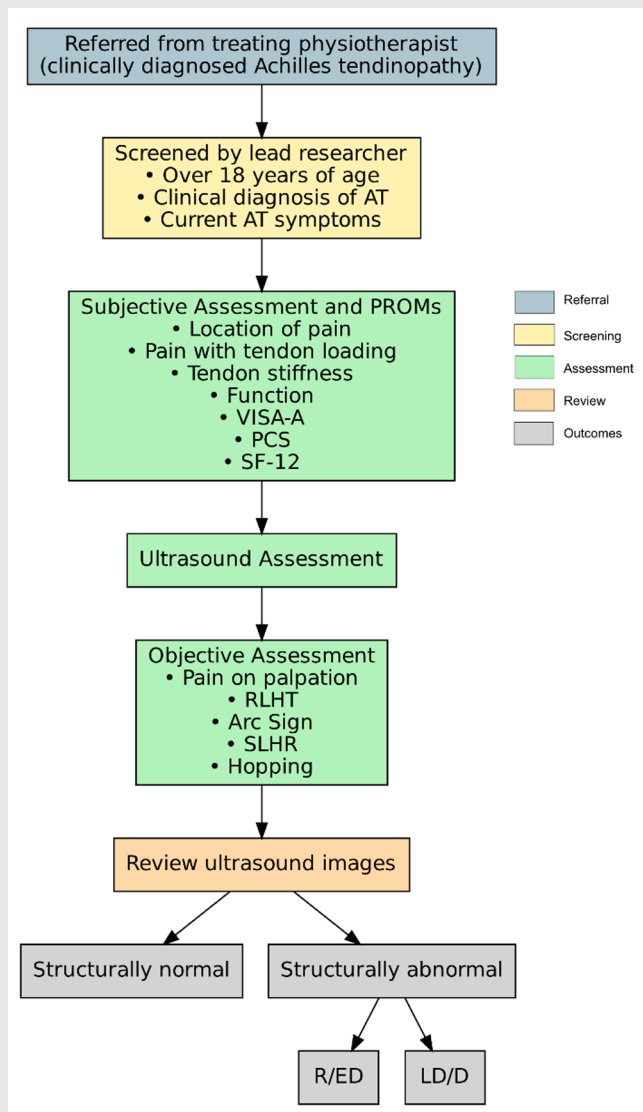
Hopping. Hopping was utilized to assess tendon pain and functional capacity under high loads.^{7,24,33} It has previously demonstrated moderate sensitivity (43%) and high specificity (87%) compared with USI for AT diagnosis.³³ The test was performed as described by Silbernagel et al,^{36,37} with participants hopping rhythmically on the affected then non-affected leg. Testing ceased, consecutive hops were no longer possible due to pain or after 50 repetitions. Previous studies stopped at 25 repetitions,³⁶ however, this was extended to account for fitness variability. Pain was recorded using a VAS, and clinicians documented total repetitions, repetitions before pain onset, and pain test completion.

Ultrasound Assessment

USI was utilized as the reference standard for assessing the accuracy of the index tests. USI was completed by the same researcher (WM) who completed the subjective assessment, PROMs, and objective assessment. USI was completed using a Phillips Lumify L12-4 Linear Array Transducer

Figure 1.

Clinical assessment process.



Abbreviations: AT, Achilles tendinopathy; PROMs, patient-reported outcome measures; VISA-A, Victorian Institute of Sport Assessment—Achilles; PCS, Pain Catastrophizing Scale; SF-12, 12-Item Short Form Survey; RLHT, Royal London Hospital Test; SLHR, single-leg heel raise; R/ED, reactive/early dysrepair; LD/D, late dysrepair/degenerative.

(Koninklijke Philips N.V., Amstelplein 2, 1096 BC Amsterdam, The Netherlands). All participants were assessed according to the European Society of Musculoskeletal Radiology guidelines for the ankle.³⁸ The ultrasound images captured during the examination were not assessed at this

time. Images were assessed following a 48-hour wash-out period to minimize clinician bias. As demonstrated in Table 3, the images were assessed according to the previously developed criteria²¹ and sub-grouped as structurally normal or structurally abnormal (R/ED and LD/D).

Statistical Analysis

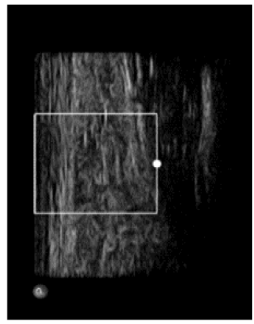
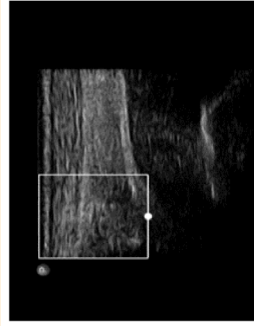
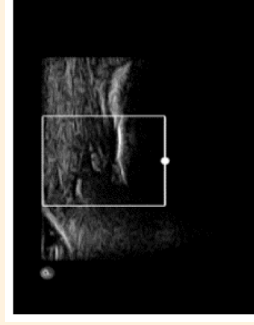
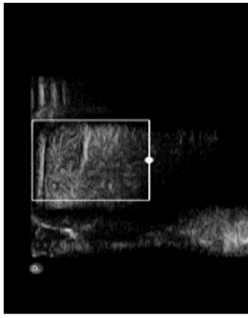
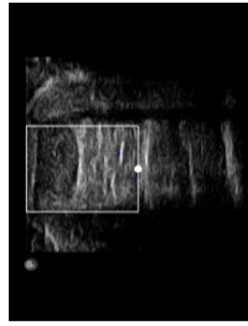
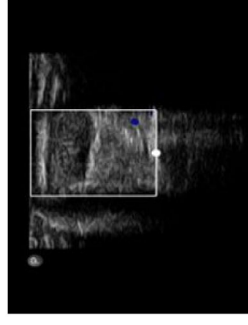
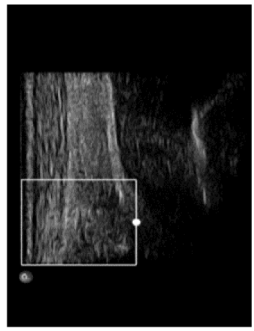
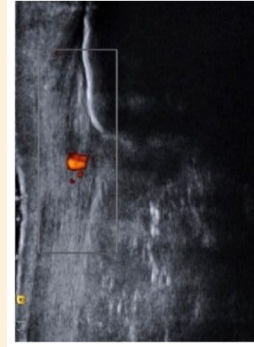
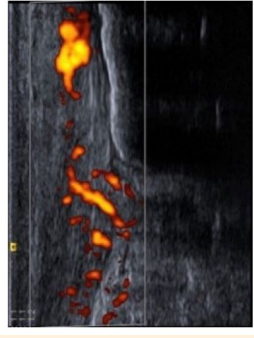
General patient characteristics were reported using mean (*SD*) or median (interquartile range, *IQR*) for continuous variables, depending on data distribution. Normality was checked using a combination of histograms, normal Q-Q plots, and the Shapiro-Wilk test. Data was pooled for structurally abnormal tendons. Variables, such as AT stiffness, AT loading stiffness, pain with SLHR, and pain with hopping were dichotomized as present or not present to facilitate statistical analysis. Positivity rates for these categorical variables were calculated with 95% confidence intervals (CIs) using the Clopper-Pearson exact method, and the exact binomial test was used to determine statistical significance. Diagnostic performance metrics including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), were derived from 2×2 contingency tables with 95% CI calculated using Clopper-Pearson exact intervals; Fisher's exact test was used to evaluate statistical significance for these comparisons. Continuous variables were analyzed using receiver operating characteristic (ROC). The area under the curve (AUC) was estimated with bootstrapped 95% CI to assess discriminative ability. Youden's index was calculated to identify optimal cutoff values for continuous variables, with a value >0.80 considered clinically significant.³⁹ Effect sizes were reported as odds ratios (ORs) with 95% CI and Fisher's exact test for categorical variables, and Hedges' *g* with bootstrapped 95% CI for continuous variables. Exact tests and bootstrapping were used due to the small sample size ($N = 23$) and non-normal distribution of some variables, ensuring more reliable estimates and valid inference without relying on parametric assumptions. Results were deemed to be statistically significant at the $P < .05$ level. Statistical analysis was completed using R (Version 4.5.1)⁴⁰ and IBM SPSS Statistics (IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, New York: IBM Corp.).

Table 2.
Index Tests for Achilles Tendinopathy.

	Definition	Result
Subjective assessment		
Self-reported location of pain	Patient asked "Can you point out where you get your pain"	Pain located 2 to 6 cm above the calcaneal insertion (midportion) or pain located in the distal 2 cm of the Achilles tendon (insertion)
Self-reported pain with tendon loading	Patient reported intensity of pain using the VAS while performing a single-leg heel raise and hop test	Patient reported increased pain on the VAS with the single-leg heel raise and hop test
Self-reported tendon stiffness or pain over a specified time	Clinician asks about pain and stiffness over specified timeframes (eg, morning, night, 24 hours, with activity)	Patient reported morning stiffness or pain or with tendon loading that may "warm-up"
PROMs		
Self-reported overall rating of Achilles tendon function	Clinician asks "Can you rate your Achilles tendon where 100% represents no problems and 0% is the worst-case scenario"	Patient reported level of function where the maximum score is 100
VISA-A	The VISA-A is a valid and reliable tool to assess severity and disability in AT	The maximum score is 100, with healthy subjects scoring a minimum of 96
PCS	The PCS is a valid and reliable instrument used to measure catastrophic thinking related to pain	A score of 30 or higher indicating a clinically relevant level of catastrophizing
SF-12	The SF-12 is a self-reported outcome measure assessing the impact of health on an individual's everyday life	The SF-12 creates two summary scores, mental health and physical health and compared with the population average score of 50
Objective assessment		
Palpation	Clinician palpates the whole length of the tendon in a proximal to distal direction	Patient reported pain located 2 to 6 cm above the calcaneal insertion (midportion) or distal 2 cm of Achilles tendon (insertion), with or without subjective opinion of tendon thickening or swelling
Single-leg heel raise	Patient rising up on to toes and lowering back down in a controlled manner	Clinician recorded number of completed single-leg heel raises on each leg
Hopping	Participant hopping on the spot	Clinician recorded number of completed hops on each leg up to 50 consecutive hops
The Royal London Hospital Test	Clinician palpating the tendon for any local tenderness with the ankle either in neutral position or in slight plantarflexion. Patient actively dorsiflexes and the portion of the tendon found to be tender is palpated again	Patient reported pain on palpation reduces significantly or disappear with maximum dorsiflexion
The arc sign	Clinician identifying intratendinous swelling in the tendon. The patient actively dorsiflexes and plantarflexes the ankle joint observing the movement of the swelling between the malleoli	The intratendinous swelling moves relative to the malleoli with the Achilles tendon during the ankle movement

Abbreviations: PROMs, patient-reported outcome measures; AT, Achilles tendinopathy; VAS, Visual Analogue Scale; VISA-A, Victorian Institute of Sport Assessment—Achilles; PCS, Pain Catastrophizing Scale; SF-12, 12-Item Short Form Survey.

Table 3. Ultrasound Imaging Criteria for Diagnosing and Sub-Grouping Achilles Tendinopathy.

		Structurally abnormal	
		Reactive/early dysrepair	Late dysrepair/degenerative
Thickness	Structurally normal Parallel margins without thickening 	Presence of fusiform swelling and/or mildly thickened tendon 	Presence of focal thickening and/or significant thickening that is double the thickness of the distal part of the tendon and/or the contralateral side 
Echogenicity	Homogenous echotexture 	Heterogeneous echotexture with diffuse hypoechoogenicity 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 
Vascularity	No vessels inside tendon 	0-2 small, thin vessels inside tendon, not through full thickness of tendon 	> 2 large vessels through full thickness of the tendon 

Results

General characteristics are presented in Table 4. All 23 participants (14 male and 9 female) had unilateral AT. USI assessment determined 6 participants had structurally normal AT, with 17 participants demonstrating structurally abnormal AT (12 R/ED and 5 LD/D). There were some differences in the number of repetitions between the affected and non-affected sides in the hopping test; however, the differences were not statistically significant. Overall, participants with abnormal AT demonstrated a lower number of median repetitions, 15.0 (IQR = 5.0-40.0), with participants with both R/ED AT (18.5, IQR = 5.0-37.8) and LD/D AT (6.0, IQR = 3.0-50.0) demonstrating lower repetitions compared with their non-affected limb (R/ED = 50.0 [48.0-50.0]; LD/D = 48.0 [40.0-50.0]).

Patient-Reported Outcome Measures

Overall, participants with clinically diagnosed AT recorded lower self-reported function (66.2 [SD = 25.3]) and lower VISA-A scores (61.9 [SD = 19.1]) (Table 5). While not statistically significant, there were larger differences in self-reported function between participants with structurally normal AT (88.3 [7.5]) and structurally abnormal AT (58.4 [SD = 24.8]). Similarly, VISA-A scores were lower in participants with structurally abnormal AT (55.4 [SD = 16.2]) and structurally normal AT (80.2 [SD = 14.8]).

Objective Assessment

The positivity rates for self-reported stiffness, stiffness with activity, pain on palpation, the arc sign, the RLHT, self-reported pain with SLHR and self-reported pain with hopping were calculated. Four tests demonstrated statistical significance (Figure 2): 3 tests—self-reported stiffness, stiffness

with activity, and pain on palpation—each showed a positivity rate of 82.6% (95% CI = 0.61-0.95; $P = .003$) while 78.3% of participants reported pain with hopping (95% CI = 0.56-0.92; $P = .011$).

Diagnostic Accuracy

Patient-reported outcome measures. The ROC analysis in Figure 3 evaluated the ability of PROMs to discriminate between the presence or absence of structural change on USI. The VISA-A demonstrated the highest accuracy, with an AUC of 0.92 (95% CI = 0.79-1.00), followed by self-reported function with an AUC of 0.89 (95% CI = 0.76-1.00). Youden's index identified optimal thresholds of 60.5 for the VISA-A and 77.5 for self-reported function. The PCS demonstrated moderate accuracy in detecting the presence of structural change (AUC = 0.77, 95% CI = 0.50-1.00); however, no participant scored higher than 30, which would indicate a clinically relevant level of catastrophising.²⁸

Objective tests. Figure 4 demonstrates the diagnostic accuracy of tests to differentiate between AT without structural change on USI (normal) and with structural change on USI (R/ED or LD/D). Of the 7 tests assessed, only hopping pain ($P < .01$) and pain on palpation ($P = .04$) demonstrated statistical significance. Hopping pain demonstrated high sensitivity (0.94; 95% CI = 0.71-1.00), and high PPV (0.89; 95% CI = 0.65-0.99), with pain on palpation also demonstrating high sensitivity (0.94; 95% CI = 0.71-1.00) and high PPV (0.84; 95% CI = 0.60-0.97). Self-reported stiffness and stiffness with activity demonstrated high sensitivity (stiffness = 0.88, 95% CI = 0.64-0.99; stiffness with activity, 0.82, 95% CI = 0.57-0.96); however, results were not statistically significant. The arc sign (1.00; 95%

CI = 0.54-1.00), RLHT (0.83; 95% CI = 0.36-1.00) and SLHR pain (0.83; 95% CI = 0.36-1.00) demonstrate high specificity and PPV (arc test = 1.00, 95% CI = 0.59-1.00; RLHT = 0.91, 95% CI = 0.59-1.00; SLHR pain = 0.92, 95% CI = 0.62-1.00); however, none were statistically significant.

Effect Size

Given the pilot study's low sample size ($N = 23$), effect sizes were calculated to evaluate associations between the clinical tests, PROMs and USI-identified structural change. Figure 5 summarizes the ORs of the categorical variables. Pain on palpation and hopping demonstrated the strongest associations with structural change, with statistically significant results ($P = .04$ and $P < .01$, respectively). The RLHT and SLHR pain demonstrated elevated ORs, but the associations were not statistically significant. The arc sign produced an undefined OR due to sparse data. Confidence intervals were wide across all tests, reflecting the limited sample size and should be interpreted with caution. All of the PROMs demonstrated a medium to very large effect size (Figure 6).

Discussion

This pilot study examined methodological challenges in evaluating the diagnostic accuracy of clinical tests for AT, specifically in distinguishing between AT with and without structural changes. All participants were clinically diagnosed with AT by their treating physiotherapist, without strict criteria, reflecting the lack of consistency in AT diagnosis.^{5-7,33} Recent consensus identified essential diagnostic elements of AT, with pain location, pain during activity, tests that provoke pain, and palpation to assess pain considered essential, while measures, such as pain onset, stiffness, and self-reported function, were considered supportive but

Table 4.

General Characteristics of Included Participants.

	Total (<i>N</i> = 23)	Normal (<i>n</i> = 6; 26.1%)	USI diagnosis		
			Abnormal (<i>n</i> = 17; 73.9%)	R/ED (<i>n</i> = 12; 52.2%)	LD/D (<i>n</i> = 5; 21.7%)
Age (years), mean (SD)	48.9 (13.6)	36.8 (13.3)	53.1 (11.3)	54.1 (11.8)	50.8 (10.7)
Sex, <i>n</i> (%)					
Male	14 (60.9)	2 (8.7)	12 (52.2)	7 (30.4)	5 (21.7)
Female	9 (39.1)	4 (17.4)	5 (21.7)	5 (21.7)	0 (0.0)
Affected tendon, <i>n</i> (%)					
Left	11 (47.8)	2 (8.7)	9 (39.1)	7 (30.4)	2 (8.7)
Right	12 (52.2)	4 (17.4)	8 (34.8)	5 (21.7)	3 (13.0)
Previous Achilles tendinopathy, <i>n</i> (%)					
No	14 (60.9)	5 (21.7)	9 (39.1)	6 (26.1)	3 (13.0)
Yes	9 (39.1)	1 (4.3)	8 (34.8)	6 (26.1)	2 (8.7)
Location of symptoms, <i>n</i> (%)					
Insertion	6 (26.1)	2 (8.7)	4 (17.4)	1 (4.3)	3 (13.0)
Midportion	17 (73.9)	4 (17.4)	13 (56.5)	11 (47.8)	2 (8.7)
Duration of symptoms, <i>n</i> (%)					
<1 month	2 (8.7)	1 (4.3)	1 (4.3)	0 (0.0)	1 (4.3)
1-4 months	7 (30.4)	0 (0.0)	7 (30.4)	6 (26.1)	1 (4.3)
>4 months	14 (60.9)	5 (21.7)	9 (39.1)	6 (26.1)	3 (13.0)
SLHR repetitions, mean (<i>SD</i>)					
Non-affected	21.0 (5.9)	23.0 (6.4)	20.2 (5.7)	20.0 (5.7)	20.8 (6.3)
Affected	18.8 (9.2)	21.3 (6.3)	17.9 (10.0)	18.5 (9.7)	16.6 (12.0)
Hopping repetitions, median (IQR)					
Non-affected	50.0 (41.0-50.0)	50.0 (41.0-50.0)	50.0 (42.0-50.0)	50.0 (48.0-50.0)	48.0 (40.0-50.0)
Affected	32.0 (5.5-50.0)	50.0 (40.3-50.0)	15.0 (5.0-40.0)	18.5 (5.0 – 37.75)	6.0 (3.0-50.0)

Abbreviations: USI, ultrasound imaging; R/ED, reactive/early dysrepair; LD/D, late dysrepair/degenerative; *n*, number; %, percentage (percentages under USI groups are of the grand total for the variable); *SD*, standard deviation; AT, Achilles tendon.

not essential.⁶ A scoping review⁷ further recommended incorporating PROMs addressing disability, psychological factors and quality of life, aligned with

the core health-related domains of tendon pathology.²⁴

The results of this pilot study largely align with recent consensus

recommendations.⁶ Pain during activity was assessed through self-reported pain during Achilles tendon loading tests (SLHR and hopping), though it can also

Table 5.

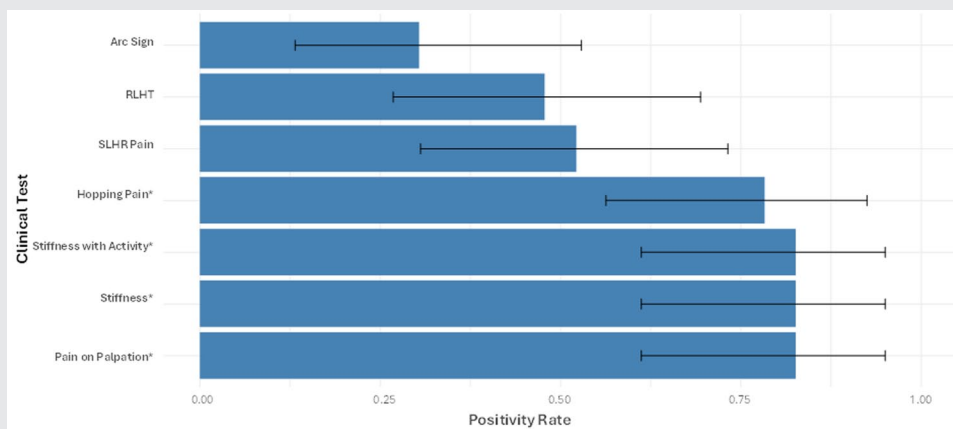
Patient-Reported Outcome Measures.

	USI diagnosis				
	Total (N = 23)	Normal (n = 6)	Abnormal (n = 17)	R/ED (n = 12)	LD/D (n = 5)
Self-reported function, mean (SD)	66.2 (25.3)	88.3 (7.5)	58.4 (24.8)	56.5 (27.5)	63.0 (18.6)
VISA-A, mean (SD)	61.9 (19.1)	80.2 (14.8)	55.4 (16.2)	55.3 (19.2)	55.8 (6.2)
PCS, mean (SD)	7.7 (6.1)	4.0 (4.5)	9.0 (6.1)	8.6 (6.8)	9.6 (4.6)
SF-12 physical, mean (SD)	44.8 (9.1)	48.5 (6.5)	43.5 (9.7)	43.0 (9.8)	44.6 (10.3)
SF-12 mental, mean (SD)	55.6 (6.4)	58.2 (2.7)	54.7 (7.2)	54.3 (6.8)	55.5 (8.8)

Abbreviations: USI, ultrasound imaging; R/ED, reactive/early dysrepair; LD/D, late dysrepair/degenerative; n, number; SD, standard deviation.

Figure 2.

Positivity rates of clinical tests for the presence of Achilles tendinopathy, accompanied by 95% CI (as determined by the exact binomial test, a rate is statistically significant if the 95% CI do not include 0.5).



RLHT, Royal London Hospital Test; SLHR, single-leg heel raise; CI, confidence interval. *, statistically significant.

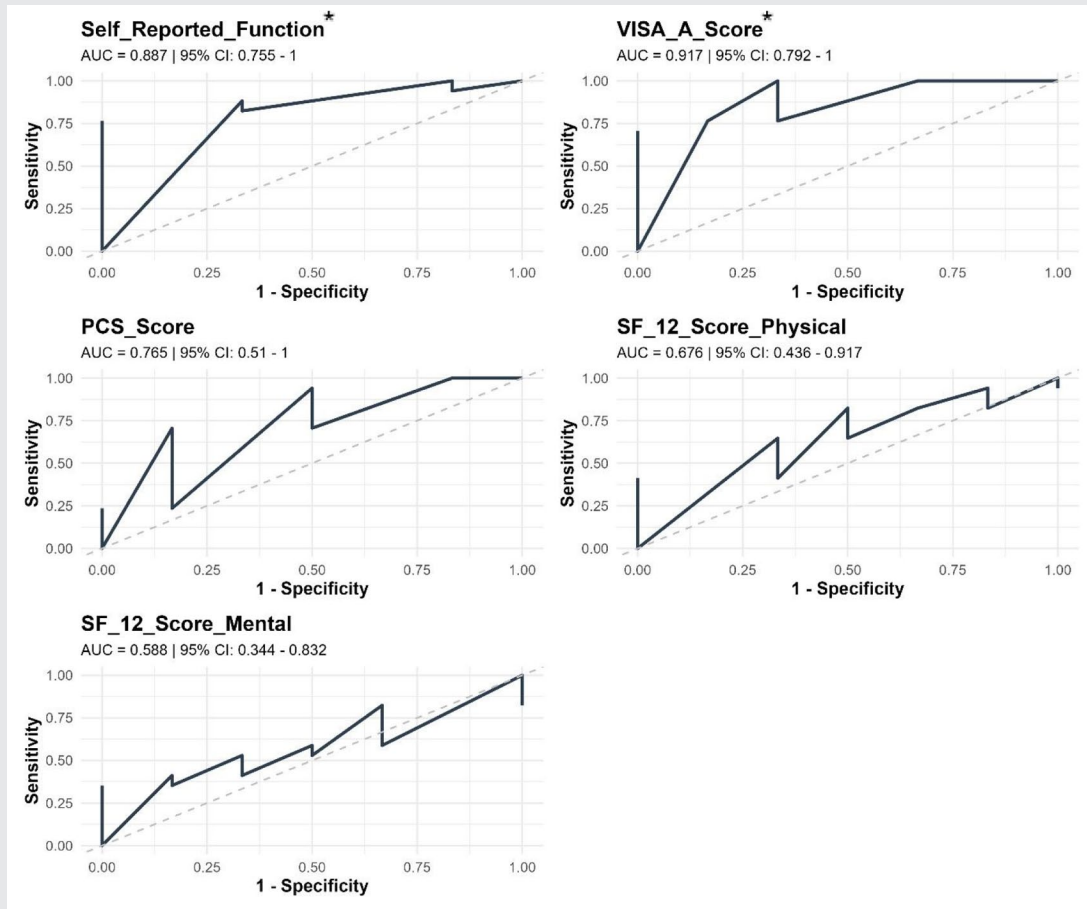
be captured via subjective history.⁶ Pain on palpation demonstrated the high positivity rate (82.6%), followed by hopping (78.3%), both reinforcing the prevalence of pain with tendon loading and palpation in symptomatic AT.⁶ In contrast, perceived stiffness, while reaching consensus, was not deemed

essential for diagnosis.⁶ Our results suggest similar positivity rates for self-reported stiffness and stiffness during activity as for palpation, indicating that tendon stiffness may warrant consideration in the diagnosis of AT. The AT-specific tests, such as the arc sign and RLHT, which consensus excluded from

essential diagnostic criteria,⁶ demonstrated the lowest positivity rates in this study, supporting their limited diagnostic value. Although imaging reached consensus as supportive rather than essential for AT diagnosis,⁶ debate persists regarding its role, particularly when

Figure 3.

Receiver operating characteristic (ROC) curves differentiating between Achilles tendinopathy with structural change on ultrasound imaging using patient-reported outcome measures.



AUC, area under the curve; VISA-A, Victorian Institute of Sport Assessment-Achilles; PCS, Pain Catastrophizing Scale; SF-12, short form 12; CI, confidence intervals; *, statistically significant.

imaging appears normal, uncertainties surrounding disease staging, defining “normal,” and the limited sensitivity of imaging modalities.⁶ Inconsistent application of USI has been cited as a contributor to this uncertainty.^{10,16,20} However, standardized USI protocols have demonstrated reliability in assessing tendon structural change,^{21,22} and may offer a consistent reference standard for evaluating the diagnostic accuracy of clinical tests. Therefore, this study examined the diagnostic accuracy of clinical tests for detecting

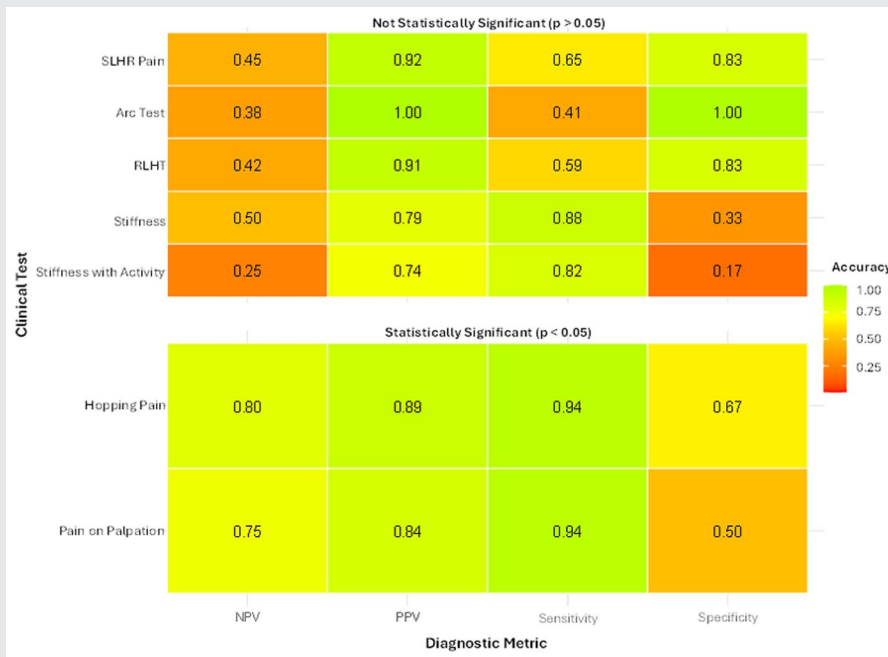
USI-confirmed structural change. Pain on palpation and pain during hopping both demonstrated high sensitivity, high PPV, and large ORs, suggesting these tests not only serve as essential diagnostic criteria for AT⁶ but may also indicate structural changes. However, their significant positivity rates in participants without structural change highlight the need for additional assessments before recommending USI. While not statistically significant, SLHR demonstrated high specificity and elevated ORs, indicating pain during

SLHR testing may warrant imaging. Self-reported stiffness, while not considered essential diagnostic criteria,⁶ also demonstrated high sensitivity for structural change and may be a useful adjunct when combined with pain-provoking tests, such as pain on palpation and pain with hopping. Overall, stiffness may help guide decisions on USI rather than serve as a standalone diagnostic marker.

Although PROMs are not considered essential for diagnosing AT,⁶ they have

Figure 4.

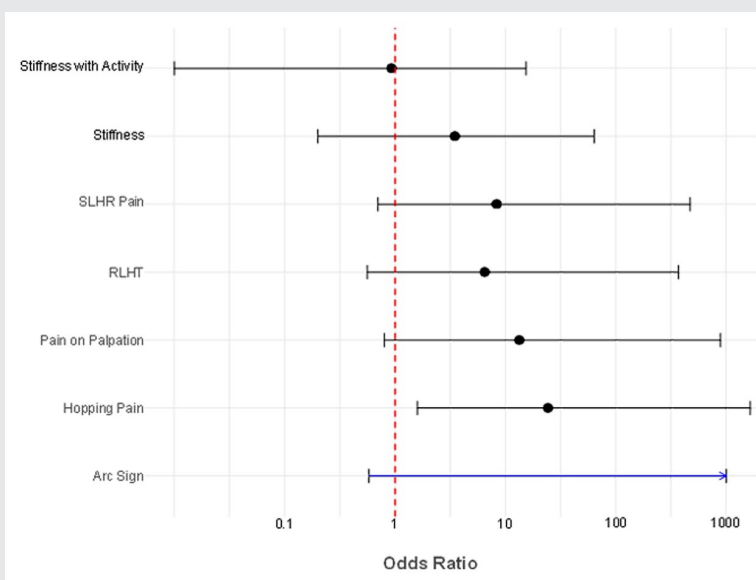
Heatmap showing diagnostic accuracy metrics for 7 clinical tests to detect structural changes identified on ultrasound imaging.



Tests are grouped by statistical significance. Color gradient reflects metric value (green = higher accuracy; orange/red = lower accuracy). Abbreviations: SLHR, single-leg heel raise; RLHT, Royal London Hospital Test; NPV, negative predictive value; PPV, positive predictive value.

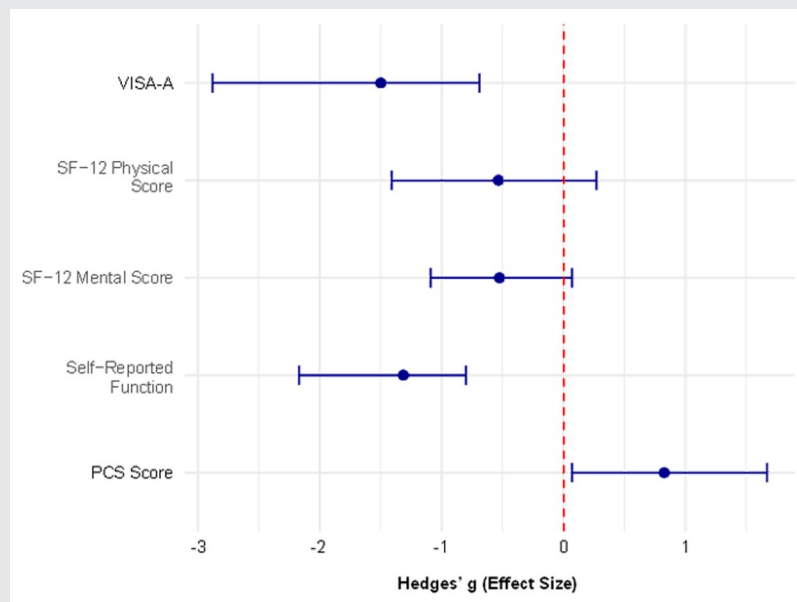
Figure 5.

Effect size of clinical tests measured by odds ratios.



Abbreviations: RLHT, Royal London Hospital Test; SLHR, single-leg heel raise; CI, confidence interval.

been proposed as useful adjuncts for AT diagnosis and management.^{7,35} Self-reported function reached consensus agreement for the diagnosis of AT, but was not considered essential.⁶ This study identified that participants with structurally normal AT reported higher mean scores for self-reported function (88.3) and VISA-A (80.2) compared with those with structural changes (58.4 and 55.4, respectively). Receiver operating characteristic analysis demonstrated clinically significant AUC values for both measures in detecting USI-confirmed structural change, with thresholds of 77.5 for self-reported function and 60.5 for the VISA-A. Effect size analysis further supports their potential as adjuncts for guiding imaging decisions. Incorporating PROMs alongside pain-provoking tests may help clinicians identify patients most likely to exhibit structural changes on USI, enabling more

Figure 6.Effect size of patient-reported outcome measures assessed with Hedges' *g*.

Magnitude guidelines: 0.2 = small; 0.5 = medium; 0.8 = large. Values > 1.0 suggest very large effect. Abbreviations: PROMs, patient-reported outcome measures; VISA-A, Victorian Institute of Sport Assessment-Achilles; PCS, Pain Catastrophizing Scale; SF-12, short form 12; CI, confidence intervals.

targeted imaging and efficient resource use.

The secondary aim of this study was to inform future research by promoting a more standardized approach to diagnosing AT and clarifying the role of USI in both research and clinical settings. While the essential diagnostic criteria established by consensus⁶ align with these findings, this study adds insight into non-essential tests and PROMs that may help determine when USI is warranted (Figure 7). These results provide a foundation for developing a structured decision-making model that integrates pain-provoking tests, self-reported stiffness, and PROM thresholds to guide imaging decisions. Future research should validate this model through a larger,

adequately powered diagnostic accuracy study, enabling refinement of test combinations, threshold values, and predictive algorithms.

Limitations

The main limitation of this study was the small sample size, limiting the ability of statistical tests to detect small or meaningful differences that may be present in larger sample sizes, and needing to interpret the results with caution. Although prior research⁴¹ suggests a sample size of 65 is needed for statistical significance, this pilot study, primarily aimed to evaluate the diagnostic utility of clinical tests for AT and explore the role of USI to inform a larger study. Standardizing clinical test

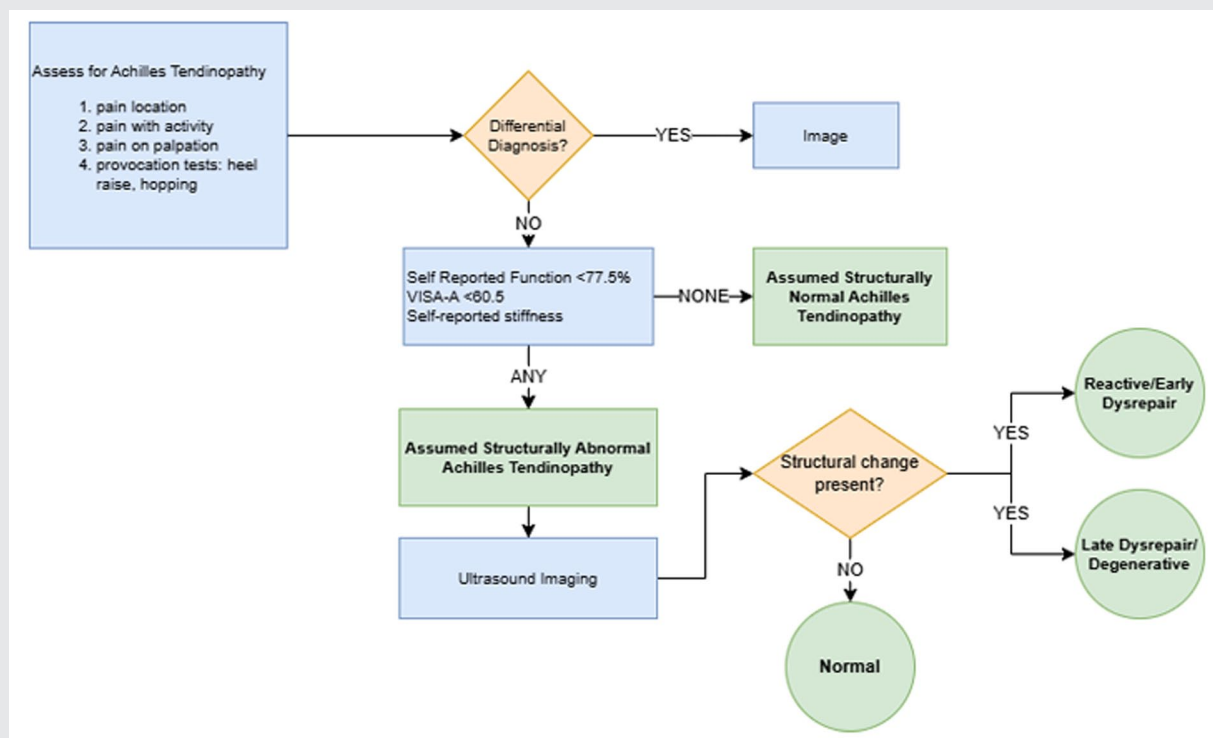
methods is essential for determining their relevance, yet considerable variability exists.^{5-7,33,34} Silbernagel et al³⁶ described the SLHR protocol for endurance and hopping used in this study, demonstrating excellent reliability (intraclass correlation coefficient = 0.76-0.94). Standardization of SLHR testing parameters, including body position, height of raise, tempo, and termination criteria, has been recommended,⁴²⁻⁴⁴ but consensus on its purpose, optimal parameters, adequate outcome measurements, or the appropriate normative values remains lacking.^{45,46} Similarly, hopping has been proposed as a method for assessing muscle-tendon elastic function,^{46,47} however, methods vary with recommendations to use objective measures of hopping height and distance to quantify AT-related function.⁴⁶

Conclusion

Pain on palpation, pain during hopping, and self-reported stiffness exhibited the highest rates of positivity among individuals with symptomatic AT. Notably, pain on palpation and pain with hopping may serve as useful clinical indicators of underlying structural changes detectable via USI. The diagnostic value of these findings appears to be enhanced when combined with self-reported stiffness, functional status, and VISA-A scores. To validate these preliminary findings, larger-scale studies are warranted to evaluate the diagnostic accuracy of clinical tests and their association with tendon structural pathology in both symptomatic and asymptomatic populations. Future research should incorporate USI as the reference standard to clarify its role in the clinical diagnostic process and its utility in identifying potential differential diagnoses.

Figure 7.

Clinical decision flow chart to determine the role of ultrasound imaging in the diagnosis of Achilles tendinopathy.



Abbreviation: VISA-A, Victorian Institute of Sport Assessment—Achilles.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an Australian Government Research Training Program Scholarship.

Ethical Approval and Informed Consents

Ethical approval for this study was provided by the Bond University Human Research Ethics Council (WM00028). Informed consent was obtained from all participants both verbally and in writing.

Trial Registration

Not applicable, because this article does not contain any clinical trials.

Data Availability Statement

All data produced in the present study are available upon reasonable request to the corresponding author.

ORCID iD

Wesley Matthews  <https://orcid.org/0000-0002-2281-0464> 

References

- de Vos RJ, van der Vlist AC, Winters M, van der Giesen F, Weir A. Diagnosing Achilles tendinopathy is like delicious spaghetti carbonara: it is all about key ingredients, but not all chefs use the same recipe. *Br J Sports Med.* 2021;55(5):247-248. doi:10.1136/bjsports-2020-102863.
- Millar NL, Silbernagel KG, Thorborg K, et al. Tendinopathy. *Nat Rev Dis Primers.* 2021;7(1):1. doi:10.1038/s41572-020-00234-1.
- Maffulli N, Longo UG, Kadakia A, Spiezia F. Achilles tendinopathy. *Foot Ankle Surg.* 2020;26(3):240-249. doi:10.1016/j.fas.2019.03.009.
- Traweger A, Scott A, Kjaer M, et al. Achilles tendinopathy. *Nat Rev Dis Primers.* 2025;11(1):20. doi:10.1038/s41572-025-00602-9.
- Maffulli N, Oliva F, Loppini M, Aicale R, Spiezia F, King JB. The Royal London Hospital Test for the clinical diagnosis of patellar tendinopathy. *Muscles Ligaments Tendons J.* 2017;7(2):315-322. doi:10.11138/mltj/2017.7.2.315.
- Malliaras P, Gravare Silbernagel K, de Vos RJ, et al. Diagnostic domains, differential diagnosis and conditions requiring further medical attention that are considered important in the assessment for Achilles tendinopathy: a Delphi consensus study.

- Br J Sports Med.* 2025;59(13):891-901. doi:10.1136/bjsports-2024-109185.
7. Matthews W, Ellis R, Furness J, Hing WA. The clinical diagnosis of Achilles tendinopathy: a scoping review. *PeerJ.* 2021;9:e12166. doi:10.7717/peerj.12166.
 8. Silbernagel KG, Shelley K, Powell S, Varrecchia S. Extended field of view ultrasound imaging to evaluate Achilles tendon length and thickness: a reliability and validity study. *Muscles Ligaments Tendons J.* 2016;6(1):104-110. doi:10.11138/mltj/2016.6.1.104.
 9. Zellers JA, Bley BC, Pohlig RT, Alghamdi NH, Silbernagel KG. Frequency of pathology on diagnostic ultrasound and relationship to patient demographics in individuals with insertional Achilles tendinopathy. *Int J Sports Phys Ther.* 2019;14(5):761-769.
 10. 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(2):149-153. doi:10.1136/bjism.37.2.149.
 11. 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(6):637-645. doi:10.5301/HIP.2011.8759.
 12. Warden SJ, Kiss ZS, Malara FA, Ooi AB, Cook JL, Crossley KM. Comparative accuracy of magnetic resonance imaging and ultrasonography in confirming clinically diagnosed patellar tendinopathy. *Am J Sports Med.* 2007;35(3):427-436. doi:10.1177/0363546506294858.
 13. Cook JL, Rio E, Purdam CR, Docking SI. Revisiting the continuum model of tendon pathology: what is its merit in clinical practice and research. *Br J Sports Med.* 2016;50(19):1187-1191. doi:10.1136/bjsports-2015-095422.
 14. Rio E, Moseley L, Purdam C, et al. The pain of tendinopathy: physiological or pathophysiological? *Sports Med.* 2014;44(1):9-23. doi:10.1007/s40279-013-0096-z.
 15. Docking SI, Ooi CC, Connell D. Tendinopathy: is imaging telling us the entire story. *J Orthop Sports Phys Ther.* 2015;45(11):842-852. doi:10.2519/jospt.2015.5880.
 16. Matthews W, Ellis R, Furness J, Hing W. Classification of tendon matrix change using ultrasound imaging: a systematic review and meta-analysis. *Ultrasound Med Biol.* 2018;44(10):2059-2080. doi:10.1016/j.ultrasmedbio.2018.05.022.
 17. McAuliffe S, McCreesh K, Culloty F, Purtill H, O'Sullivan K. Can ultrasound imaging predict the development of Achilles and patellar tendinopathy? a systematic review and meta-analysis. *Br J Sports Med.* 2016;50(24):1516-1523. doi:10.1136/bjsports-2016-096288.
 18. Chimenti RL, Neville C, Houck J, Cuddeford T, Carreira D, Martin RL. Achilles pain, stiffness, and muscle power deficits: midportion Achilles tendinopathy revision - 2024. *J Orthop Sports Phys Ther.* 2024;54(12):CPG1-CPG32. doi:10.2519/jospt.2024.0302.
 19. de Vos RJ, van der Vlist AC, Zwerver J, et al. Dutch multidisciplinary guideline on Achilles tendinopathy. *Br J Sports Med.* 2021;55(20):1125-1134. doi:10.1136/bjsports-2020-103867.
 20. McAuliffe S, McCreesh K, Purtill H, O'Sullivan K. 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. doi:10.1016/j.ptsp.2016.12.002.
 21. Matthews W, Ellis R, Furness JW, Rathbone E, Hing W. Staging achilles tendinopathy using ultrasound imaging: the development and investigation of a new ultrasound imaging criteria based on the continuum model of tendon pathology. *BMJ Open Sport Exerc Med.* 2020;6(1):e000699. doi:10.1136/bmjsem-2019-000699.
 22. Sunding K, Fahlström M, Werner S, Forssblad M, Willberg L. Evaluation of Achilles and patellar tendinopathy with greyscale ultrasound and colour Doppler: using a four-grade scale. *Knee Surg Sports Traumatol Arthrosc.* 2016;24(6):1988-1996. doi:10.1007/s00167-014-3270-4.
 23. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ.* 2015;351:h5527. doi:10.1136/bmj.h5527.
 24. Vicenzino B, de Vos RJ, Alfredson H, et al. ICON 2019-International Scientific Tendinopathy Symposium consensus: there are nine core health-related domains for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. *Br J Sports Med.* 2020;54(8):444-451. doi:10.1136/bjsports-2019-100894.
 25. Grävare Silbernagel K, Malliaras P, de Vos RJ, et al. ICON 2020-International Scientific Tendinopathy Symposium consensus: a systematic review of outcome measures reported in clinical trials of Achilles tendinopathy. *Sports Med.* 2022;52(3):613-641. doi:10.1007/s40279-021-01588-6.
 26. Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *Br J Sports Med.* 2001;35(5):335-341. doi:10.1136/bjism.35.5.335.
 27. Paantjens MA, Helmhout PH, Backx FJG, Bakker EWP. Victorian Institute of Sport Assessment-Achilles thresholds for minimal important change and return to presymptom activity level in active soldiers with mid-portion Achilles tendinopathy. *BMJ Mil Health.* 2024;170(e2):e156-e160. doi:10.1136/military-2022-002326.
 28. Silbernagel KG, Thomee R, Karlsson J. Cross-cultural adaptation of the VISA-A questionnaire, an index of clinical severity for patients with Achilles tendinopathy, with reliability, validity and structure evaluations. *BMC Musculoskelet Disord.* 2005;6(1):12. doi:10.1186/1471-2474-6-12.
 29. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychol Assess.* 1995;7(4):524-532. doi:10.1037/1040-3590.7.4.524.
 30. Slagers AJ, van Veen E, Zwerver J, Geertzen JHB, Reininga IHF, van den Akker-Scheek I. Psychological factors during rehabilitation of patients with Achilles or patellar tendinopathy: a cross-sectional study. *Phys Ther Sport.* 2021;50:145-152. doi:10.1016/j.ptsp.2021.04.010.
 31. Hörterer H, Oppelt S, Böcker W, et al. Patient-reported outcomes of surgically treated insertional Achilles tendinopathy. *Foot Ankle Int.* 2021;42(12):1565-1569. doi:10.1177/10711007211023060.
 32. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34(3):220-233. doi:10.1097/00005650-199603000-00003.
 33. Hutchison AM, Evans R, Bodger O, et al. What is the best clinical test for Achilles tendinopathy. *Foot Ankle Surg.* 2013;19(2):112-117. doi:10.1016/j.fas.2012.12.006.
 34. OrthoToolKit . Free online SF-12 score calculator. <https://orthotoolkit.com/sf-12/>. Accessed July 16, 2023.
 35. Maffulli N, Kenward MG, Testa V, Capasso G, Regine R, King JB. Clinical diagnosis of Achilles tendinopathy with tendinosis. *Clin J Sport Med.* 2003;13(1):11-15. doi:10.1097/00042752-200301000-00003.

36. Silbernagel KG, Gustavsson A, Thomeé R, Karlsson J. Evaluation of lower leg function in patients with Achilles tendinopathy. *Knee Surg Sports Traumatol Arthrosc.* 2006;14(11):1207-1217. doi:10.1007/s00167-006-0150-6.
37. Silbernagel KG, Hanlon S, Sprague A. Current clinical concepts: conservative management of Achilles tendinopathy. *J Athl Train.* 2020;55(5):438-447. doi:10.4085/1062-6050-356-19.
38. Martinoli C. Musculoskeletal ultrasound: technical guidelines. *Insights Imaging.* 2010;1(3):99-141. doi:10.1007/s13244-010-0032-9.
39. Çorbacioğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: a guide to interpreting the area under the curve value. *Turk J Emerg Med.* 2023;23(4):195-198. doi:10.4103/tjem.tjem_182_23.
40. *RStudio: Integrated Development Environment for R.* Posit Software, PBC; 2025. <http://www.posit.co/>. Accessed November 6, 2025.
41. Grimaldi A, Mellor R, Nicolson P, Hodges P, Bennell K, Vicenzino B. Utility of clinical tests to diagnose MRI-confirmed gluteal tendinopathy in patients presenting with lateral hip pain. *Br J Sports Med.* 2017;51(6):519-524. doi:10.1136/bjsports-2016-096175.
42. Silbernagel KG, Thomeé R, Thomeé P, Karlsson J. Eccentric overload training for patients with chronic Achilles tendon pain: a randomised controlled study with reliability testing of the evaluation methods. *Scand J Med Sci Sports.* 2001;11(4):197-206. doi:10.1034/j.1600-0838.2001.110402.x.
43. Svantesson U, Osterberg U, Thomeé R, Grimby G. Muscle fatigue in a standing heel-rise test. *Scand J Rehabil Med.* 1998;30(2):67-72. doi:10.1080/003655098444156.
44. Haber M, Golan E, Azoulay L, Kahn SR, Shrier I. Reliability of a device measuring triceps surae muscle fatigability. *Br J Sports Med.* 2004;38(2):163-167. doi:10.1136/bjism.2002.002899.
45. Hébert-Losier K, Newsham-West RJ, Schneiders AG, Sullivan SJ. Raising the standards of the calf-raise test: a systematic review. *J Sci Med Sport.* 2009;12(6):594-602. doi:10.1016/j.jsams.2008.12.628.
46. McAuliffe S, Tabuena A, McCreesh K, et al. Altered strength profile in Achilles tendinopathy: a systematic review and meta-analysis. *J Athl Train.* 2019;54(8):889-900. doi:10.4085/1062-6050-43-18.
47. Fukushima S, Komi PV, Järvinen M, Miyashita M. In vivo Achilles tendon loading during jumping in humans. *Eur J Appl Physiol Occup Physiol.* 1995;71(5):453-458. doi:10.1007/BF00635880.