

The Utility of Clinical Measures for the Diagnosis of Achilles Tendon Injuries: A Systematic Review With Meta-Analysis

Michael Reiman, DPT, LAT, ATC, SCS; Ciara Burgi, SPT, CSCS; Eileen Strube, SPT, LAT, ATC, CSCS; Kevin Prue, SPT, CSCS; Keaton Ray, SPT, ATC, CSCS; Amanda Elliott, SPT; Adam Goode, DPT, PhD, PT

Duke University Medical Center, Durham, NC

Objective: To summarize and evaluate the current diagnostic accuracy of clinical measures used to diagnose Achilles tendon injuries.

Data Sources: A literature search of MEDLINE, CINAHL, and EMBASE databases was conducted with key words related to diagnostic accuracy and Achilles tendon injuries.

Study Selection: Original research articles investigating Achilles tendon injuries against an acceptable reference standard were included.

Data Extraction: Three studies met the inclusion criteria. Quality assessment was conducted using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. DerSimonian-Laird random-effects models were used to pool sensitivity (SN), specificity (SP), and diagnostic odds ratios with their 95% confidence intervals (CIs).

Data Synthesis: The SN and negative likelihood ratio (–LR) values for Achilles tendon rupture measures ranged from 0.73 (95% CI = 0.65, 0.81) and 0.30 (95% CI = 0.23, 0.40) to 0.96 (95% CI = 0.93, 0.99) and 0.04 (95% CI = 0.02, 0.10), respectively, whereas SP and positive likelihood ratio (+LR) values ranged from 0.85 (95% CI = 0.72, 0.98) and 6.29 (95% CI = 2.33, 19.96) to 0.93 (95% CI = 0.84, 1.00) and 13.71 (95% CI = 3.54, 51.24), respectively, with the highest SN and SP both

reported in the calf-squeeze test. The SN and –LR values for Achilles tendinopathy measures ranged from 0.03 (95% CI = 0.00, 0.08) and 0.97 (95% CI = not reported) to 0.89 (95% CI = 0.75, 0.98) and 0.19 (95% CI = not reported), whereas SP and +LR values ranged from 0.58 (95% CI = 0.38, 0.77) and 2.12 (95% CI = not reported) to 1.00 (95% CI = 1.00, 1.00) and infinity, respectively, with the highest SN and SP reported for morning stiffness and palpation for crepitus. Pooled analyses demonstrated similar diagnostic properties in all 3 clinical measures (arc sign, palpation, and Royal London Hospital test), with SN and –LR ranging from 0.42 (95% CI = 0.23, 0.62) and 0.68 (95% CI = 0.50, 0.93), respectively, for the arc sign, to 0.64 (95% CI = 0.44, 0.81) and 0.48 (95% CI = 0.29, 0.80), respectively, for palpation. Pooled SP and +LR ranged from 0.81 (95% CI = 0.65, 0.91) and 3.15 (95% CI = 1.61, 6.18), respectively, for palpation, to 0.88 (95% CI = 0.74, 0.96) SP for the arc sign and 3.84 (95% CI = 1.69, 8.73) +LR for the Royal London Hospital test.

Conclusions: Most clinical measures for Achilles tendon injury have greater diagnostic than screening capability.

Key Words: diagnostic accuracy, sensitivity, specificity

Key Points

- Clinical measures of Achilles tendon injuries have not been well investigated.
- For Achilles tendon injuries, currently available clinical measures are stronger in their diagnostic than screening properties.
- The squeeze test was a useful diagnostic measure for an Achilles tendon tear, although we recommend caution because this finding was demonstrated in only 1 study with a high risk of bias.

The Achilles tendon is the largest and most frequently torn tendon in the human body.^{1–3} Achilles tendon injuries are among the most common sport-related injuries.^{3–8} The accurate diagnosis of an Achilles tendon injury, such as Achilles tendinopathy and, to a lesser degree, Achilles tendon tear, is not always clear and straightforward.^{1,5,9–12} The differential diagnosis of an Achilles tendon injury includes but is not limited to retrocalcaneal bursitis, os trigonum, tarsal tunnel syndrome, posterior tibialis tendon rupture, arthritic conditions, plantar fasciitis, and stress fracture.⁹

Sensitivity (SN) and specificity (SP) are accuracy properties used for both screening and diagnostic accuracy tests. The closer the SN is to 100% in the presence of a test with a negative result, the stronger the ability of that clinical measure to rule out the potential for a particular diagnosis. The closer the SP is to 100% in the presence of a test with a positive result, the stronger the ability of that clinical measure to rule in the potential for a particular diagnosis.

Diagnostic ultrasound and magnetic resonance imaging (MRI) have traditionally been considered the criterion reference standards to diagnose Achilles tendon injuries.^{1,2,9,10} However, this testing can be costly and may

Table 1. Article Search Strategy^a

Step	Search Terms	Boolean Operator	PubMed	CINAHL
1	<i>Achilles tendon</i>	OR	7004	1504
2	<i>Sprains and strains</i> OR <i>tendon injuries</i> OR <i>tendinopathy</i> OR <i>athletic injuries</i> OR <i>leg injuries</i> OR <i>rupture</i> OR <i>soft tissue injuries</i> OR <i>tears and lacerations</i>	OR	21 248	45 047
3	<i>Diagnosis</i> OR <i>physical therapy</i> OR <i>diagnostic tests</i> OR <i>physical examination</i>	OR	8 185 483	975 779
4	<i>Posterior cruciate ligament</i> OR <i>anterior cruciate ligament</i> OR <i>osteoarthritis</i> [MeSH]	OR	52 330	16 145
5	1, 2, 3	AND	2259	515
6	5, 4	NOT	1512	482

^a The number of duplicates was 557, and the number of articles for the title screen was 1437.

not result in accurate diagnosis.^{12–15} Because of limited evidence support, the American Academy of Orthopaedic Surgeons' clinical practice guidelines recommendation was inconclusive regarding the routine use of MRI for diagnosing acute Achilles tendon tears.^{15–17} In addition, the McKinsey Global Institute¹⁸ reported that diagnostic imaging from both MRI and computed tomography scans contribute to \$26.5 billion in unnecessary health care costs annually.

Clinical measures, such as subjective reports of pain and stiffness, and objective clinical tests commonly described (palpation of a gap in the Achilles tendon, calf-squeeze test, palpation for tenderness of the Achilles tendon, and the arc sign) are being used more commonly by and are increasingly accessible to practicing clinicians for assistance with the diagnosis of Achilles tendon dysfunction. In fact, recent findings² suggested that the use of a comprehensive clinical examination incorporating such measures outperformed MRI with respect to diagnostic accuracy for Achilles tendon rupture. Additionally, delayed treatment for Achilles tendon dysfunction leads to poorer outcomes.^{19–24} Therefore, if these clinical measures are truly comparable with diagnostic imaging, the cost-effectiveness of the standard imaging techniques for detecting Achilles tendon dysfunction may be questioned. The clinical efficacy of clinical measures in relation to commonly accepted diagnostic imaging methods has not been determined.

To our knowledge, a comprehensive review of the diagnostic accuracy of clinical measures (both subjective [eg, reports of pain and stiffness] and objective [eg, palpation for tenderness and a gap in the tendon, arc sign, calf-squeeze test, single-legged heel raise, and hop test]) for Achilles tendon injury does not exist. Clinicians examining and treating patients with Achilles tendon conditions need a clear understanding of the utility of these various clinical measures. Therefore, the purpose of our study was to systematically evaluate the diagnostic accuracy of clinical measures used to diagnose Achilles tendon injuries.

METHODS

Data Sources

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines during the searching and reporting phases of this review. The PRISMA statement includes a 27-item checklist that is designed to be used as a basis for reporting systematic reviews of randomized trials,²⁵ but it can also be applied to multiple forms of research methods.²⁶

A systematic, computerized search of the literature in the MEDLINE, CINAHL, and EMBASE databases was conducted April 1, 2013. The MESH search terms for MEDLINE are listed in Table 1, with limits for English language and humans. Three reviewers (C.B., E.S., K.R.) independently performed the search. Because computerized search results for diagnostic accuracy data frequently omit many relevant articles,²⁷ the reference lists of all selected publications were checked to retrieve relevant publications that were not identified in the computerized search. The gray literature, which included publications, posters, abstracts, and conference proceedings, was also hand searched. Two reviewers (C.B., E.S.) searched the reference lists and gray literature. To identify relevant articles, titles and abstracts of all identified citations were independently screened. Full-text articles were retrieved if the abstract provided insufficient information to establish eligibility or if the article passed the first eligibility screening.

Selection Criteria

Articles examining clinical measures for Achilles tendon conditions were eligible if they met all of the following criteria: (1) included individuals with Achilles tendon pain, (2) included at least 1 clinical Achilles tendon examination measure, (3) used an acceptable reference standard, (4) reported the results in sufficient detail to allow reconstruction of contingency tables, and (5) were written in English.

An article was excluded if (1) the condition was associated with an injury located elsewhere (eg, knee joint) that referred pain to the Achilles tendon region, (2) insufficient detail was provided to calculate diagnostic accuracy, (3) the clinical measures were performed under any form of anesthesia or on cadavers, (4) specialized instrumentation not readily available to all clinicians was used, and (5) clinical measures were performed on infants or toddlers.

All criteria were independently applied by 2 reviewers (K.P., K.R.) to the full text of the articles that passed the first eligibility screening. An independent reviewer (M.R.) verified inclusion of all articles in the review. Disagreements among the reviewers were discussed and resolved during a consensus meeting (K.P., K.R., M.R.).

Risk of Bias and Quality Assessment

Two reviewers (C.B., E.S.) independently reviewed each full-text article and scored it with the Quality Assessment of Diagnostic Accuracy Studies 2 scores (QUADAS-2)

tool.²⁸ Disagreements among the reviewers were discussed and resolved during a consensus meeting and then by an independent third reviewer (M.R.). The QUADAS-2 is a quality-assessment tool composed of 4 domains: patient selection, index test, reference standard, and flow and timing. The risk of bias is assessed in each domain; the first 3 domains are also assessed for applicability with a *low*, *high*, or *unclear* rating. *Applicability* in the QUADAS-2 refers to whether certain aspects of an individual study match or do not match the review question. Unlike the QUADAS-1, the QUADAS-2 does not use a comprehensive quality score; instead, it uses an overall judgment of *low*, *high*, or *unclear* risk. An overall risk rating of *low risk of bias* or *low concern regarding applicability* requires the study to be ranked as *low* on all relevant domains. A *high* or *unclear* rating in 1 or more domains may require that the study be rated as *at risk of bias* or as having *concerns regarding applicability*.

Data Extraction and Analysis

Three reviewers (M.R., K.P., K.R.) independently extracted information and data regarding study population, setting, special test performance, injury, diagnostic reference standard, and numbers of true positives, false positives, false negatives, and true negatives for calculation of SN, SP, positive likelihood ratio (+LR), and negative likelihood ratio (–LR) when not provided. *Sensitivity* is defined as the percentage of people who test positive for a specific condition among a group of people who have the condition. *Specificity* is the percentage of people whose test results are negative for a specific condition among a group of people who do not have the condition. A +LR is the ratio of a positive test result in people with the condition to a positive test result in people without the condition. A +LR identifies the strength of a test in determining the presence of a finding and is calculated by the formula $SN/(1 - SP)$. A –LR is the ratio of a negative test result in people with the condition to a negative test result in people without the condition and it is calculated by the formula $(1 - SN)/SP$. The higher the +LR and lower the –LR, the more the posttest probability is altered. The diagnostic odds ratio (DOR) is a single indicator, independent of prevalence, that represents the ratio of the odds of positivity in those with disease relative to the odds or positivity in those without disease. The values for the DOR range from 0, indicating no test discrimination, to infinity, with higher scores indicating better discrimination.²⁹ Posttest probability can be altered to a minimal degree with +LRs of 1 to 2 or –LRs of 0.5 to 1, to a small degree with +LRs of 2 to 5 or –LRs of 0.2 to 0.5, to a moderate degree with +LRs of 5 to 10 and –LRs of 0.1 to 0.2, and to a large and almost conclusive degree with +LRs greater than 10 and –LRs less than 0.1. *Pretest probability* is defined as the probability of the target condition before a diagnostic test result is known. It represents the probability that a specific patient, with a specific past history, presenting to a specific clinical setting, with a specific symptom complex, has a specific condition.³⁰

Meta-analysis

Studies were statistically pooled when ≥ 2 studies examined the same index test and diagnosis with the same

reference standard. DerSimonian-Laird³¹ random-effects models, which consider both between-studies and within-study heterogeneity, were used to produce summary estimates of SN, SP, +LR, –LR, and DORs. An I^2 value of $>50\%$ and Cochrane-Q P value of $<.10$ were the criteria to indicate significant between-studies heterogeneity of SN and SP and likelihood ratios, respectfully. We did not formally test publication bias because of the low power of the tests with limited included studies.³² No significant threshold effects were found using Spearman correlation coefficients. When a cell was empty, we added 0.5 to all 4 cells as suggested by Cox.³³ All analyses were conducted in Meta-DiSc (version 1.4; Informer Technologies, Inc, Dallas, TX) by 1 author (A.G.), who was blinded to the results of the search, inclusion and exclusion criteria, and study quality.³⁴

RESULTS

Selection of Studies

The systematic search through MEDLINE, CINAHL, and EMBASE netted 1512 abstracts, and 3 additional papers were identified through an extensive hand search. In total, 1437 titles were initially retained after duplicates were removed. Abstract and full-text review reduced the number of acceptable papers to 3,^{1,9,10} and the 2 articles^{1,9} investigating Achilles tendinopathy qualified for meta-analysis (Figure 1). This review included 219 participants across the 3 studies and the investigation of 14 clinical measures. The sample sizes of the studies were 174,¹ 14,¹⁰ and 21.⁹ Two of the studies^{1,10} investigated physical examination measures only, 1 for Achilles tendon tear¹ and the other for Achilles tendinopathy,¹⁰ whereas the third study⁹ examined both subjective and physical examination measures (Table 2).

Quality Scores

The individual items for the risk of bias and applicability concerns are provided in Table 3. Patient selection and index test use were common reasons for the high risk of bias and applicability concerns. The κ value between testers for the overall bias score using QUADAS-2 was 0.59 (95% confidence interval [CI] = 0.20, 0.98), with this point estimate reflecting moderate agreement.³⁵ No disagreement persisted between reviewers after the consensus meeting.

Results of Individual Diagnostic Clinical Measures

Three studies met the inclusion criteria: 1 study for Achilles tendon rupture¹ and 2 studies for Achilles tendinopathy.^{9,10} Two studies were classified as high bias^{1,10} and 1 as low bias⁹ per QUADAS-2. The characteristics of each study included in the review are listed in Table 2. The diagnostic accuracy of clinical measures for Achilles tendon injuries is given in Table 4, the pooled diagnostic properties of the clinical measures for Achilles tendon conditions are shown in Table 5, and the studied clinical measures are described in Table 6.

Achilles Tendon Rupture. One study qualified for inclusion based on the diagnosis of Achilles tendon rupture.¹ This study was classified as high bias per QUADAS-2. Four clinical measures for this diagnosis

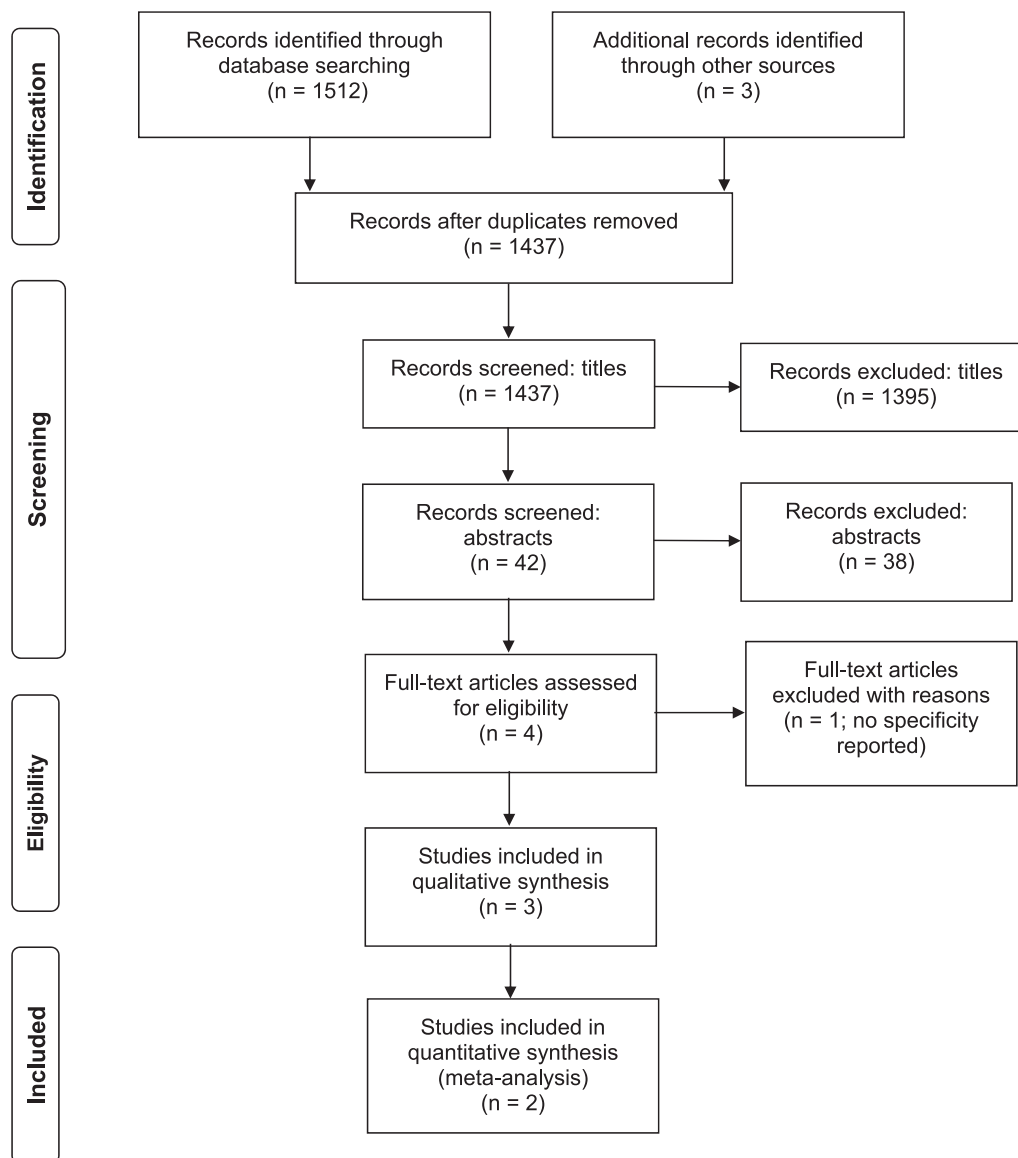


Figure. Flow diagram for study inclusion.

(palpation, calf-squeeze test, Matles test, and Copeland test; Table 4) were included. The SN values for Achilles tendon rupture measures ranged from 0.73 (95% CI = 0.65, 0.81) to 0.96 (95% CI = 0.93, 0.99), whereas SP values ranged from 0.85 (95% CI = 0.72, 0.98) to 0.93 (95% CI = 0.84, 1.00), with the highest SN and SP both reported in the calf-squeeze test. Palpation for a gap in the Achilles tendon had an SN of 0.73 (95% CI = 0.65, 0.81), an SP of 0.89 (95% CI = 0.71, 0.97), a +LR of 6.64 (95% CI = 2.32, 19.91), and a -LR of 0.30 (95% CI = 0.23, 0.40). When the gap test result was positive, the posttest probability of a diagnosis of an Achilles tendon tear for that patient was altered to a moderate degree. When the palpation test result was negative, the posttest probability of such a diagnosis was only altered to a small degree. The calf-squeeze test had an SN of 0.96 (95% CI = 0.93, 0.99), an SP of 0.93 (95% CI = 0.75, 0.99), a +LR of 13.71 (95% CI = 3.54, 51.24), and a -LR of 0.04 (95% CI = 0.02, 0.10). Therefore, when this test result was positive, the posttest probability of an Achilles tendon tear was altered to a large and almost

conclusive degree. Additionally, when this test result was negative, it also altered the posttest probability of not having an Achilles tendon tear to the same degree. The reported diagnostic values for the Matles test included an SN of 0.88 (95% CI = 0.78, 0.94) and an SP of 0.85 (95% CI = 0.66, 0.95), which resulted in a +LR of 6.29 (95% CI = 2.33, 19.96) and a -LR of 0.14 (95% CI = 0.07, 0.25). A positive result on the Matles test altered the posttest probability of a diagnosis of an Achilles tendon tear to a moderate degree; a negative result altered the probability of a diagnosis of not having an Achilles tendon tear to the same degree. The author reported only the SN value (0.78; 95% CI = 0.49, 0.94) for the Copeland test; therefore, +LR and -LR were not quantifiable.¹⁰ Thus, we could not compute the posttest probability of ruling out the diagnosis of an Achilles tendon tear for the Copeland test.

Achilles Tendinopathy. Two studies qualified for inclusion based on the diagnosis of Achilles tendinopathy.^{9,10} One study¹⁰ was ranked as high bias, whereas the other study⁹ was ranked as low bias. Ten clinical measures (arc sign,

Table 2. Summary of Studies Included in the Review

Measure, Source, and Year	Participants	Symptom Duration	Clinical Measures Investigated
Achilles tendon tear			
Maffulli, ¹ 1998	174 individuals (25 females) with unilateral complete subcutaneous Achilles tendon tears Treated conservatively: mean age = 48.1 ± 13.2 y, sex NR, n = 22 Treated with percutaneous surgical repair: mean age = 42.5 ± 11.6 y, sex NR, n = 19 Treated with open surgical repair: mean age = 38.4 ± 10.5 y, sex NR, n = 133	Day of injury (n = 98), 13 d (n = 45), 47 d (n = 14), 1428 d (n = 10)	Tendon-gap palpation (n = 133 patients, 28 controls) Calf-squeeze test (n = 133 patients, 28 controls) Matles test (n = 77 patients, 28 controls) Copeland test (n = 44 patients, 28 controls)
Achilles tendinopathy			
Maffulli et al, ¹⁰ 2003	10 male athletes with Achilles tendinopathy (mean age = 28.5 ± 6.8 y) and 14 controls (mean age = 27.1 ± 7.4 y)	NR	Palpation Arc sign Royal London Hospital test Combined tests (all of the above)
Hutchison et al, ⁹ 2013	21 patients: 10 with known chronic midbody Achilles tendinopathy (mean age = 48 y, 4 females); 6 with known posterior ankle pain, not due to Achilles tendinopathy (mean age = 51 y, 3 females); 5 with no posterior ankle pain or tendinopathy (mean age = 54 y, 5 females)	NR	Subjective measures • Self-report of pain • Self-report of morning stiffness Palpation tests • Tendon thickening • Crepitus • Palpation • Royal London Hospital test • Arc sign Tendon-loading tests • Stretch on passive dorsiflexion with knee joint flexed • Single-legged heel raise • Hop test

Abbreviation: NR, not reported.

palpation, Royal London Hospital test, self-report of pain, self-report of morning stiffness, palpation of tendon thickening, palpation of crepitus, stretch on passive dorsiflexion with knee joint in flexion, single-legged heel raise, and hop test; Table 4) for Achilles tendinopathy were included. Three of these measures qualified for meta-analysis: the arc sign, palpation, and the Royal London Hospital test (Table 5). The individual SN values for Achilles tendinopathy measures ranged from 0.03 (95% CI = 0.00, 0.08; palpation for crepitus)⁹ to 0.89 (95% CI = 0.75, 0.98; self-report of morning stiffness),⁹ whereas the individual SP values ranged from 0.58 (95% CI = 0.38, 0.77; self-report of morning stiffness)⁹ to 1.00 (95% CI = 1.00, 1.00; palpation of crepitus and arc sign).⁹ Composite testing (combining palpation, the arc sign, and the Royal London Hospital test)¹⁰ did not improve the diagnostic accuracy of the test and in fact was less diagnostically accurate (0.59 [95% CI = 0.47, 0.74]; SP = 0.83 [95% CI = 0.76, 0.89]; +LR 3.47 [CI not reported (NR)]; and -LR = 0.29 [CI NR]). For the subjective measures, self-reports of

pain (SN = 0.78 [95% CI = 0.58, 0.94]) and morning stiffness (SN = 0.89 [95% CI = 0.75, 0.98]) demonstrated a stronger screening value than diagnostic value (SP range, 0.58–0.77).⁹ Both of these measures, with +LRs/-LRs of 3.39/0.29 and 2.12/0.19, respectively, are only able to alter posttest probability to a small to moderate degree. Palpation for tendon thickening (SP = 0.90 [95% CI = 0.83, 0.97]; +LR = 5.9 [CI NR]) and crepitus (SP = 1.0 [95% CI = 1.00, 1.00]; +LR = infinity) function more strongly as diagnostic than as screening measures (SN range = 0.03–0.59). With +LRs of 5.9 and infinity, respectively, these measures will alter posttest probability for this diagnosis to a moderate to large, and almost conclusive, degree. All of the tendon-loading measures were much stronger as diagnostic measures (SP range = 0.87–0.93; +LR = 1.0–3.31) than as screening (SN range = 0.13–0.43; -LR = 1.0–0.66) measures. The shift in posttest probability for a diagnosis of Achilles tendinopathy is still altered only between a very small degree (passive dorsiflexion) and a small degree (single-legged heel raise and hop tests).

Table 3. Risk of Bias of the Studies Included in the Review

Source, Year	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Maffulli, ¹ 1998	↓	↓	↑	↓	↑	↑	↑
Maffulli et al, ¹⁰ 2003	↓	↓	↑	↑	↓	↑	↑
Hutchison et al, ⁹ 2013	↑	↑	↑	↑	↑	↑	↑

Abbreviations: ↑, low risk; ↓, high risk.

Table 4. Summary of Studies Reporting on the Diagnostic Accuracy of Clinical Measures for Achilles Tendon Injuries

Source, y, and Measure	Sensitivity (95% CI)	Specificity (95% CI)	+Likelihood Ratio (95% CI)	-Likelihood Ratio (95% CI)	Reliability (κ)	Reference standard (Sensitivity/Specificity) ^a
Achilles tendon tear: Maffulli,¹ 1998						
Tendon-gap palpation (n = 133 patients, 28 controls)	0.73 (0.65, 0.81)	0.89 (0.71, 0.97)	6.64 (2.32, 19.91)	0.30 (0.23, 0.40)	NR	Surgery (gold standard)
Calif-squeeze test (n = 133 patients, 28 controls)	0.96 (0.93, 0.99)	0.93 (0.75, 0.99)	13.71 (3.54, 51.24)	0.04 (0.02, 0.10)		Control group: magnetic resonance imaging (0.91 ^c ;
Mattes test (n = 77 patients, 28 controls)	0.88 (0.78, 0.94)	0.85 (0.66, 0.95)	6.29 (2.33, 19.96)	0.14 (0.07, 0.25)		0.94/0.06) for complete
Copeland test (n = 44 patients, 28 controls)	0.78 (0.49, 0.94)	NR	NA	NA		tears, and US (0.50/0.81) ^{12,b}
						for proximal tendon injury
Achilles tendinopathy						
Maffulli et al, ¹⁰ 2003						
Palpation	0.58 (0.39, 0.75)	0.84 (0.75, 0.91)	3.63 (NR)	0.50 (NR)	0.27–0.72 (intrater)	US (0.50/0.81) ^{12,b}
Arc sign	0.53 (0.35, 0.70)	0.83 (0.72, 0.91)	3.06 (NR)	0.58 (NR)	0.72–0.85 (intrater)	94/99 ^{36,c} and histology
Royal London Hospital test	0.54 (0.35, 0.73)	0.91 (0.86, 0.95)	6.0 (NR)	0.51 (NR)	0.28–0.75 (intrater)	
Combined tests (all of the above)	0.59 (0.47, 0.74)	0.83 (0.76, 0.89)	3.47 (NR)	0.49 (NR)	0.55–0.72 (intrater)	
Hutchison et al, ^{9,10} 2013						
Subjective measures						
Self-report of pain	0.78 (0.58, 0.94)	0.77 (0.60, 0.91)	3.39 (NR)	0.29 (NR)	0.81 (intrater)	US
Self-report of morning stiffness	0.89 (0.75, 0.98)	0.58 (0.38, 0.77)	2.12 (NR)	0.19 (NR)	0.75 (intrater)	
					0.88 (intrater)	
					0.79 (intrater)	
Palpation tests						
Tendon thickening	0.59 (0.39, 0.78)	0.90 (0.83, 0.97)	5.90 (NR)	0.46 (NR)	0.70 (intrater)	
Crepitus	0.03 (0.00, 0.08)	1.0 (1.00, 1.00)	Infinity	0.97 (NR)	0.58 (intrater)	
Palpation	0.84 (0.68, 0.98)	0.73 (0.53, 0.92)	3.11 (NR)	0.22 (NR)	0.66 (intrater)	
Arc sign	0.25 (0.09, 0.44)	1.0 (1.00, 1.00)	Infinity	0.75 (NR)	-0.02 (intrater)	
Royal London Hospital test	0.51 (0.33, 0.70)	0.93 (0.83, 1.00)	7.29 (NR)	0.53 (NR)	0.96 (intrater)	
					0.74 (intrater)	
					0.81 (intrater)	
					0.77 (intrater)	
					0.70 (intrater)	
					0.37 (intrater)	
Tendon-loading tests						
Stretch on passive dorsiflexion with knee joint in flexion	0.13 (0.06, 0.20)	0.87 (0.77, 0.96)	1.00 (NR)	1.00 (NR)	0.43 (intrater)	
Single-legged heel raise	0.22 (0.08, 0.36)	0.93 (0.82, 1.00)	3.14 (NR)	0.95 (NR)	-0.14 (intrater)	
Hop test	0.43 (0.27, 0.59)	0.87 (0.74, 0.98)	3.31 (NR)	0.66 (NR)	0.58 (intrater)	
					0.26 (intrater)	
					0.54 (intrater)	
					0.56 (intrater)	

Abbreviations: CI, confidence interval; NA, not applicable; NR, not reported; US, ultrasonography.

^a Confidence intervals not reported.

^b Compared with reference standard of magnetic resonance imaging.

^c Compared with reference standard of clinical examination.

Table 5. Pooled Diagnostic Properties for the Diagnosis of Achilles Tendinopathy

Measure	No. of Studies (Sample Size)	Sensitivity (95% CI)	Specificity (95% CI)	+Likelihood Ratio (95% CI)	-Likelihood Ratio (95% CI)	Diagnostic Odds Ratio (95% CI)
Palpation	2 (n = 45) ^{9,10}	0.64 (0.44, 0.81)	0.81 (0.65, 0.91)	3.15 (1.61, 6.18)	0.48 (0.29, 0.80)	7.34 (2.40, 22.45)
Arc sign	2 (n = 45) ^{9,10}	0.42 (0.23, 0.62)	0.88 (0.74, 0.96)	3.24 (1.35, 7.81)	0.68 (0.50, 0.93)	5.49 (1.61, 18.71)
Royal London Hospital test	2 (n = 45) ^{9,10}	0.54 (0.34, 0.73)	0.86 (0.72, 0.95)	3.84 (1.69, 8.73)	0.54 (0.36, 0.81)	7.41 (2.31, 23.74)

Abbreviation: CI, confidence interval.

Meta-analysis

The diagnostic properties and total sample sizes of the 2 studies included in the meta-analysis are provided in Table 4. Both groups^{9,10} investigated palpation, the arc sign, and

the Royal London Hospital test; all tests were for Achilles tendinopathy. Pooled analyses demonstrated similar diagnostic properties in all 3 clinical measures, with SN ranging from 0.64 (95% CI = 0.44, 0.81) for palpation to 0.42 (95%

Table 6. Description of Clinical Measures

Measure	Description	Positive Finding
Achilles tendon tear		
Palpation of tendon gap ¹	Examiner gently palpates the course of the Achilles tendon, feeling for a gap in continuity of the tendon.	Tendon gap
Calf-squeeze test ¹	Examiner gently squeezes the patient's calf muscles with his or her hand.	Ankle remains still or only minimally plantar flexed
Matles test ¹	Patient flexes the knee to 90°. Positions of the ankles and feet are observed.	Foot on the affected side falls into neutral or dorsiflexion
Copeland test ¹	With patient prone, a sphygmomanometer cuff is placed around the middle of the calf and inflated to 100 mm Hg. Examiner then dorsiflexes the ankle.	Little or no pressure rise
Achilles tendinopathy		
Self-report of pain ⁹	Examiner asks patient, "Can you point out where you get pain?"	Patient points to 2–6 cm above insertion of the Achilles tendon on the calcaneus
Self-report of morning stiffness ⁹	Examiner asks patient, "How are your symptoms first thing in the morning when you get out of bed?"	Pain worse for the first few steps in the morning
Tendon thickening on palpation ⁹	Examiner palpates the Achilles tendon distally to proximally, 2–6 cm above insertion into the calcaneus, gently squeezing the tendon between the index finger and thumb, feeling for localized thickening.	Subjective tendon thickening
Crepitus on palpation ⁹	Examiner palpates the Achilles tendon 2–6 cm above insertion into the calcaneus, gently squeezing the tendon between the index finger and thumb, feeling for crepitation ("wet leather" sign) with passive ankle motion.	Subjective crepitation
Palpation (tenderness) ^{9,10}	Examiner gently palpates the course of the Achilles tendon proximally to distally, gently squeezing the tendon between the thumb and index finger.	Patient complains of tenderness
Arc sign ^{9,10}	Patient is asked to dorsiflex and plantar flex the ankle.	Area of tendon swelling evidenced by palpation (or 3 cm proximal to calcaneal insertion if no swelling is present) moves with ankle dorsiflexion and plantar flexion
Royal London Hospital test ^{9,10}	Once examiner has elicited local tenderness by palpating the tendon with the ankle in neutral position or slightly plantar flexed, patient is asked to actively dorsiflex and plantar flex the ankle. Tendon is again palpated with the ankle in end-range dorsiflexion and plantar flexion.	Tenderness present in ankle dorsiflexion
Stretch of passive dorsiflexion with knee joint in flexion ⁹	Patient places the affected leg forward (toes must face forward) and leans onto the leg until a stretch is felt in the tendon. Patient's heel must not lift off the ground.	Pain reported at extreme range of movement
Single-legged heel raise ⁹	Patient rises up on tiptoe (while standing on leg to be assessed) and lowers body back to the floor.	Pain on either the upward or downward movement
Hop test ⁹	Patient hops forward over a line marked on the floor, using the leg to be assessed.	Pain with testing

CI = 0.23, 0.62) for the arc sign. Pooled SP ranged from 0.81 (95% CI = 0.65, 0.91) for palpation to 0.88 (95% CI = 0.74, 0.96) for the arc sign. The DOR ranged from 5.49 (95% CI = 1.61, 18.71) for the arc sign to 7.41 (95% CI = 2.31, 23.74) for the Royal London Hospital test. No significant heterogeneity was found between studies.

DISCUSSION

To our knowledge, this is the first systematic review investigating the diagnostic accuracy of both subjective and objective orthopaedic clinical measures for Achilles tendon conditions. Only 3 studies met our inclusion criteria. Although all 14 examination measures affected posttest probability to some degree, we found inconsistencies in the study methods. As reported in previous systematic reviews of individual orthopaedic measures,³⁷⁻⁴⁴ the measures investigated in this review should not serve as the sole means of screening (ruling out) or diagnosing (ruling in) an Achilles tendon condition.

The diagnostic values of clinical measures for Achilles tendon tear have moderate to high diagnostic capability and are better used for diagnosis (SP and +LR values) than for screening (SN and -LR values).¹ The calf-squeeze test demonstrated the strongest SN, SP, +LR, and -LR, indicating that it is currently the best test to both screen for and confirm a diagnosis of Achilles tendon rupture. Both the +LR and -LR altered to a significant and almost conclusive degree the posttest probability of the diagnosis of Achilles tendon tear.³⁰ Palpation of a tendon gap demonstrated a moderate capacity (+LR 6.64) to confirm an Achilles tendon rupture but a moderate to low SN, as well as a -LR (0.30), which altered the posttest diagnosis probability to a small degree, indicating that this test should not be used in isolation for screening purposes. The reliability of these clinical measures was not reported. The only study¹ investigating these tests demonstrated a high risk of bias, as well as study design concerns that included nonstandardized use of the reference standard, reference standard results that were interpreted with knowledge of the index test results, 13 surgeons performing the tests without an assessment of intrarater or interrater reliability, and lack of routine use of intertester blinding. Each of these methodologic concerns could limit the clinical utility of the diagnostic accuracy values gleaned from this study. Strengths of this study included a large sample size (149 men and 25 women), use of a control group with possible other ankle injuries, and patient selection applicability. The patients in this study ranged from 38 to 48 years old.¹ Achilles tendon ruptures are most typically observed in men in the fourth to fifth decades of life,⁴⁵ making this study demographically applicable.

Most Achilles tendinopathy clinical measures demonstrated greater diagnostic than screening capabilities. A total of 5 of the 14 clinical tests demonstrated strong SP, including crepitus, the arc sign, the Royal London Hospital test, single-legged heel raise, and tendon thickening.^{9,10} Despite this relatively high SP, we suggest caution in applying any of these measures as individual clinical assessments, given that morning stiffness and palpation were the only tests with moderate SN. The +LR of these 5 tests ranged from 3.06 (alteration of posttest diagnosis probability to a small degree) to infinity (almost conclusive

posttest probability alteration). Considering the meta-analysis of 3 of these tests, +LRs ranging from 3.15 to 3.24 would again alter posttest diagnosis probability of Achilles tendinopathy only to a small degree. The reliability (κ) for measures investigating Achilles tendinopathy was determined for both intrarater and interrater reliability. Intrarater reliability ranged from poor to fair agreement (palpation [range = 0.27-0.72],¹⁰ arc sign [range = 0.55-0.72]¹⁰) to excellent agreement (Royal London Hospital test [range = 0.60-0.89],¹⁰ self-report of pain [0.81],⁹ self-report of morning stiffness [0.88],⁹ palpation [0.96],⁹ and arc sign [0.81]⁹). Interrater reliability ranged from poor to fair agreement (crepitus [0.02],⁹ Royal London Hospital test [0.37],⁹ stretch on passive dorsiflexion [0.14],⁹ and single-legged heel raise [0.26])⁹ to substantial to excellent agreement (palpation [range = 0.72-0.85],¹⁰ self-report of pain [0.75],⁹ self-report of morning stiffness [0.79],⁹ palpation [0.74],⁹ and arc sign [0.77]⁹). Although intrarater and interrater reliability were variable for the study by Maffulli et al,¹⁰ and the same measure in the 2 studies^{9,10} also showed variability, most of the clinical measures demonstrated substantial agreement.

We found no significant heterogeneity between studies for the measures qualifying for meta-analysis. The results of the study by Maffulli et al¹⁰ are nongeneralizable because of the distinct patient population (10 male athletes already on the waiting list at a special clinic for Achilles tendon surgical exploration). The study by Hutchison et al,⁹ the only study in this review demonstrating a low risk of bias, had methodologic concerns worthy of mention, including a limited number of participants (n = 21), risk of selection bias (1 of the 3 groups included colleagues of the authors), and the use of ultrasonography as a criterion reference. Although ultrasonography is considered an acceptable criterion reference, the limited investigation regarding its diagnostic accuracy compared with the longer-held standard of MRI is unconvincing (SN = 0.50 and SP = 0.81, respectively).¹² The most appropriate clinical reference standard for the diagnosis of Achilles tendinopathy may need further investigation.

Currently, copious debate exists regarding the most effective treatment for Achilles tendinopathy.⁴⁶⁻⁵² This discussion relies on the accurate determination of the existence of the condition. If future researchers focus on the diagnostic accuracy of the most appropriate clinical measures for Achilles tendon injuries and stratify patients within an applicable treatment-based classification, some of the current disparity may be eliminated.

Accordingly, the use of the clinical measures in this review should be tempered. The extent of their clinical utility would be strengthened with additional investigations, larger sample sizes, less bias, and consistent comparison with the strongest criterion reference(s) (Table 3). Reliance on the use of these measures as stand-alone screening or diagnostic measures is not suggested.

This study is not without limitations, including limiting the search strategy to only those articles written in English, selecting a study that did not investigate SP for the Copeland test,¹ not comparing patient-inclusion and -exclusion criteria across the studies, relying on the general diagnosis *Achilles tendon pathology*, and looking at single measures rather than clustered measures. The phrase *Achilles tendon pathology* does not differentiate tendinitis from tendinosis, for example.

Therefore, the reader may not be clear as to whether these measures are assessments of tendinitis versus tendinosis, although the study characteristics listed in Table 1 may be helpful. Clustering of findings includes using 2 or more clinical measures and statistically combining their diagnostic accuracy values to determine whether the combination of testing findings improves their accuracy. One study¹⁰ did cluster the findings of 3 individual measures.

CONCLUSIONS

In general, Achilles tendon tear measures have stronger diagnostic accuracy properties than do measures investigating Achilles tendinopathy. The calf-squeeze test has the strongest diagnostic properties of all measures investigated, with a +LR of 13.71 and a -LR of 0.04, giving it the ability to rule in or rule out an Achilles tendon tear to a large and almost conclusive degree. The diagnostic properties of the different measures for ruling in and ruling out Achilles tendinopathy are quite variable but overall demonstrate a stronger diagnostic than screening capability. Because only 1 study demonstrated low bias, further high-quality studies investigating these measures are suggested.

REFERENCES

1. Maffulli N. The clinical diagnosis of subcutaneous tear of the Achilles tendon: a prospective study in 174 patients. *Am J Sports Med.* 1998;26(2):266–270.
2. Garras DN, Raikin SM, Bhat SB, Taweel N, Karanjia H. MRI is unnecessary for diagnosing acute Achilles tendon ruptures: clinical diagnostic criteria. *Clin Orthop Relat Res.* 2012;470(8):2268–2273.
3. Jozsa L, Kvist M, Balint BJ, et al. The role of recreational sport activity in Achilles tendon rupture: a clinical, pathoanatomical, and sociological study of 292 cases. *Am J Sports Med.* 1989;17(3):338–343.
4. Suchak AA, Bostick G, Reid D, Blitz S, Jomha N. The incidence of Achilles tendon ruptures in Edmonton, Canada. *Foot Ankle Int.* 2005; 26(11):932–936.
5. Jarvinen TA, Kannus P, Maffulli N, Khan KM. Achilles tendon disorders: etiology and epidemiology. *Foot Ankle Clin.* 2005;10(2): 255–266.
6. Kannus P, Natri A. Etiology and pathophysiology of tendon ruptures in sports. *Scand J Med Sci Sports.* 1997;7(2):107–112.
7. Kvist M. Achilles tendon injuries in athletes. *Sports Med.* 1994; 18(3):173–201.
8. Blackburn L, Del Buono A, Maffulli N. Mid-portion Achilles tendinopathy: why painful? An evidence-based philosophy. *Knee Surg Sports Traumatol Arthrosc.* 2012;20(8):1653–1654.
9. 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.
10. 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.
11. Maffulli N, Longo UG, Maffulli GD, Khanna A, Denaro V. Achilles tendon ruptures in elite athletes. *Foot Ankle Int.* 2011;32(1):9–15.
12. Kayser R, Mahlfeld K, Heyde C, Tschauer C. Partial rupture of the proximal Achilles tendon: a differential diagnostic problem in ultrasound imaging. *Br J Sports Med.* 2005;39(11):838–842.
13. Kuwada GT. Surgical correlation of preoperative MRI findings of trauma to tendons and ligaments of the foot and ankle. *J Am Podiatr Med Assoc.* 2008;98(5):370–373.

14. Kalebo P, Allenmark C, Peterson L, Sward L. Diagnostic value of ultrasonography in partial ruptures of the Achilles tendon. *Am J Sports Med.* 1992;20(4):378–381.
15. Kou J. AAOS clinical practice guideline: acute Achilles tendon rupture. *J Am Acad Orthop Surg.* 2010;18(8):511–513.
16. Chiodo CP, Glazebrook M, Bluman EM, et al. American Academy of Orthopaedic Surgeons clinical practice guideline on treatment of Achilles tendon rupture. *J Bone Joint Surg Am.* 2010;92(14):2466–2468.
17. Chiodo CP, Glazebrook M, Bluman EM, et al. Diagnosis and treatment of acute Achilles tendon rupture. *J Am Acad Orthop Surg.* 2010;18(8):503–510.
18. Angrisano C, Farrell D, Kocher B, Laboissiere M, Parker S. *Accounting for the Cost of Health Care in the United States.* New York, NY: McKinsey Global Institute; 2007.
19. Porter DA, Mannarino FP, Snead D, Gabel SJ, Ostrowski M. Primary repair without augmentation for early neglected Achilles tendon ruptures in the recreational athlete. *Foot Ankle Int.* 1997;18(9):557–564.
20. Myerson MS. Achilles tendon ruptures. *Instr Course Lect.* 1999;48: 219–230.
21. Gabel S, Manoli A 2nd. Neglected rupture of the Achilles tendon. *Foot Ankle Int.* 1994;15(9):512–517.
22. Den Hartog BD. Surgical strategies: delayed diagnosis or neglected Achilles' tendon ruptures. *Foot Ankle Int.* 2008;29(4):456–463.
23. Chiodo CP, Wilson MG. Current concepts review: acute ruptures of the Achilles tendon. *Foot Ankle Int.* 2006;27(4):305–313.
24. Rajasekar K, Gholve P, Faraj AA, Kosygan KP. A subjective outcome analysis of tendo-Achilles rupture. *J Foot Ankle Surg.* 2005; 44(1):32–36.
25. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg.* 2010;8(5):336–341.
26. Swartz MK. The PRISMA statement: a guideline for systematic reviews and meta-analyses. *J Pediatr Health Care.* 2011;25(1):1–2.
27. Doust JA, Pietrzak E, Sanders S, Glasziou PP. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *J Clin Epidemiol.* 2005;58(5):444–449.
28. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–536.
29. Glas AS, Lijmer JG, Prins MH, Bossel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol.* 2003;56(11):1129–1135.
30. Jaeschke R, Guyatt GH, Sackett DL; The Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test. B: What are the results and will they help me in caring for my patients? *JAMA.* 1994;271(9):703–707.
31. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177–188.
32. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol.* 2000;53(11):1119–1129.
33. Cox D. *The Analysis of Binary Data.* London, England: Methuen; 1970.
34. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol.* 2006;6:31.
35. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33(2):363–374.
36. De Zordo T, Chhem R, Smekal V, et al. Real-time sonoelastography: findings in patients with symptomatic achilles tendons and comparison to healthy volunteers. *Ultraschall Med.* 2010;31(4): 394–400.

37. Cook C, Mabry L, Reiman MP, Hegedus EJ. Best tests/clinical findings for screening and diagnosis of patellofemoral pain syndrome: a systematic review. *Physiotherapy*. 2012;98(2):93–100.
38. Reiman MP, Goode AP, Hegedus EJ, Cook CE, Wright AA. Diagnostic accuracy of clinical tests of the hip: a systematic review with meta-analysis. *Br J Sports Med*. 2013;47(14):893–902.
39. Wright AA, Wassinger CA, Frank M, Michener LA, Hegedus EJ. Diagnostic accuracy of scapular physical examination tests for shoulder disorders: a systematic review. *Br J Sports Med*. 2013;47(14):886–892.
40. Cook C, Hegedus E. Diagnostic utility of clinical tests for spinal dysfunction. *Man Ther*. 2011;16(1):21–25.
41. Hegedus EJ, Cook C, Hasselblad V, Goode A, McCrory DC. Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. *J Orthop Sports Phys Ther*. 2007;37(9):541–550.
42. Hegedus EJ, Goode AP, Cook CE, et al. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. *Br J Sports Med*. 2012;46(14):964–978.
43. Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utilities for assessing meniscal injury. *Clin Rehabil*. 2008;22(2):143–161.
44. Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utility for assessing superior labral anterior posterior lesions. *Am J Sports Med*. 2009;37(11):2252–2258.
45. Hess GW. Achilles tendon rupture: a review of etiology, population, anatomy, risk factors, and injury prevention. *Foot Ankle Spec*. 2010;3(1):29–32.
46. Silvan DM, Pazos FB, Valero AR, Justo XD, Jaen TF. The effectiveness of an intensive exercise programme in the functional recovery of a severe degenerative Achilles tendinopathy: is it enough by itself? A case report. *Br J Sports Med*. 2013;47(9):e2.
47. Alfredson H, Spang C, Forsgren S. Studies on Achilles tendinosis: bilateral recovery after unilateral surgery, and similar histopathological appearances bilaterally. *Br J Sports Med*. 2013;47(9):e2.
48. Alfredson H, Andersson G, Backman L, Bagge J, Danielson P, Forsgren S. US and Doppler-guided surgical treatment based on immunohistochemical findings in midportion Achilles tendinopathy shows good clinical results and fast return to activity. *Br J Sports Med*. 2013;47(9):e2.
49. Harmon K, Drezner J, Rao A. Platelet rich plasma for chronic tendinopathy. *Br J Sports Med*. 2013;47(9):e2.
50. Gurdezi S, Kohls-Gatzoulis J, Solan MC. Results of proximal medial gastrocnemius release for Achilles tendinopathy. *Foot Ankle Int*. 2013;34(10):1364–1369.
51. Murtaugh B, Ihm JM. Eccentric training for the treatment of tendinopathies. *Curr Sports Med Rep*. 2013;12(3):175–182.
52. Wiegerinck JI, Kerkhoffs GM, van Sterkenburg MN, Sierevelt IN, van Dijk CN. Treatment for insertional Achilles tendinopathy: a systematic review. *Knee Surg Sports Traumatol Arthrosc*. 2013;21(6):1345–1355.

Address correspondence to Michael Reiman, DPT, LAT, ATC, SCS, Duke University Medical Center, 2200 W Main Street, Durham, NC 27705. Address e-mail to michael.reiman@duke.edu.