



## Diagnostic accuracy of clinical tests for diagnosis of ankle syndesmosis injury: a systematic review

Amy D Sman, Claire E Hiller and Kathryn M Refshauge

*Br J Sports Med* 2013 47: 620-628 originally published online December 6, 2012

doi: 10.1136/bjsports-2012-091702

---

Updated information and services can be found at:

<http://bjsm.bmj.com/content/47/10/620.full.html>

---

*These include:*

### References

This article cites 37 articles, 15 of which can be accessed free at:  
<http://bjsm.bmj.com/content/47/10/620.full.html#ref-list-1>

Article cited in:

<http://bjsm.bmj.com/content/47/10/620.full.html#related-urls>

### Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

### Topic Collections

Articles on similar topics can be found in the following collections

[Editor's choice](#) (171 articles)

---

### Notes

---

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>



Editor's choice  
Scan to access more  
free content

Faculty of Health Sciences,  
Discipline of Physiotherapy,  
University of Sydney, Lidcombe,  
New South Wales, Australia

#### Correspondence to

Amy D Sman,  
Cumberland Campus C42,  
University of Sydney,  
PO Box 170, Lidcombe,  
NSW 1825, Australia;  
[asma6205@uni.sydney.edu.au](mailto:asma6205@uni.sydney.edu.au)

Received 3 September 2012  
Accepted 9 November 2012  
Published Online First  
6 December 2012

# Diagnostic accuracy of clinical tests for diagnosis of ankle syndesmosis injury: a systematic review

Amy D Sman, Claire E Hiller, Kathryn M Refshauge

## ABSTRACT

**Objectives** To determine the value of clinical tests for accurate diagnosis of ankle syndesmosis injury.

**Design** Systematic review.

**Data sources** An electronic database search was conducted (to 6 August 2012) of databases such as: MEDLINE, CINAHL, EMBASE, PubMed and Cochrane Databases. References from identified articles were examined and seven authors of eligible studies were contacted for additional information.

**Study selection** Studies of any design, without language restriction, were included; however, systematic reviews were excluded. Eligible studies included participants with a suspected ankle syndesmosis injury but without fracture. Reliability studies compared one or more clinical tests and studies of test accuracy compared the clinical test with a reference standard.

**Results** The database search resulted in 114 full text articles which were assessed for eligibility. Three studies were included in the review and raw data of these studies were retrieved after contacting the authors. Eight clinical diagnostic tests were investigated; palpation of the tibiofibular ligaments, external rotation stress test, squeeze, Cotton, fibula translation, dorsiflexion range of motion (ROM) and anterior drawer tests. Two studies investigated diagnostic accuracy and both investigated the squeeze test by with conflicting results. Likelihood ratios (LR) ranging from LR+1.50 to LR−1.50 were found for other tests. High intra-rater reliability was found for the squeeze, Cotton, dorsiflexion ROM and external rotation tests (83–100% close agreement). Inter-rater reliability was good for the external rotation test ( $ICC_{2,1} > 0.70$ ). Fair-to-poor reliability was found for other tests.

**Conclusions** This is the first systematic review to investigate the reliability and accuracy of clinical tests for the diagnosis of ankle syndesmosis injury. Few studies were identified and our findings show that clinicians cannot rely on a single test to identify ankle syndesmosis injury with certainty. Additional diagnostic tests, such as MRI, should be considered before making a final diagnosis of syndesmosis injury.

## INTRODUCTION

Ankle injuries are the most common injury among sporting populations<sup>1</sup> and between 1% and 20%<sup>2–7</sup> of these injuries involve the distal tibiofibular joint or ankle syndesmosis. Gerber *et al*<sup>2</sup> suggested that the true incidence in the general population is higher than that reported since syndesmosis injuries are probably under-diagnosed.

The ankle syndesmosis provides integrity to the tibia and fibula and the dense syndesmosis ligaments provide the main stability to the distal tibiofibular joint. The distal tibiofibular syndesmosis was

previously regarded as no more than an interosseous ligament,<sup>8</sup> but since 1977 the definition has been extended to include the interosseous membrane and ligament, together with the anterior inferior tibiofibular ligament (AITFL) and posterior inferior tibiofibular ligament (PITFL).<sup>8–10</sup>

Most ankle injury research has focused on the lateral ankle ligaments and inversion injuries, while research on injuries to the ankle syndesmosis, including diagnosis, is rare. However, several authors have reported that syndesmosis injury is difficult to evaluate and diagnose,<sup>3 11 12</sup> and is therefore often missed in the clinic,<sup>13</sup> but is thought to have a longer recovery time than other ankle injuries.<sup>4 14</sup> A lack of consensus on diagnostic criteria, and reliability and accuracy of the clinical diagnostic tests, can result in late or missed diagnosis of injury to the ankle syndesmosis. This could lead to an inappropriate and even harmful choice of treatment and might ultimately lead to chronic symptoms.

Injuries to the ankle syndesmosis are distinguished from other ankle ligamentous injuries based on the history, including mechanism of injury, and physical examination.<sup>15</sup> Several mechanisms of injury have been reported; however, dorsiflexion and external rotation with a firmly planted foot is most commonly described.<sup>4 7 12 16</sup> Dorsiflexion produces a slight widening between the tibia and fibula, as the fibula undergoes a small degree of lateral rotation because of the varying slope of the lateral surface of the body of the talus.<sup>9</sup> The classic feature of syndesmosis injury is palpatory tenderness over the anterior and posterior tibiofibular ligaments and between the tibia and fibula.<sup>15</sup> Patients may also have swelling just proximal to the ankle joint,<sup>15 17</sup> at or above the anterior tibiofibular ligament<sup>4 18–20</sup> in the acute stage, although ankle syndesmosis injuries are not always prone to severe swelling, potentially because the damaged tissue is extracapsular.<sup>15</sup>

Several clinical diagnostic tests are used to assess ankle syndesmosis injury; however, their reliability and accuracy have not been extensively researched. Palpation of the tibiofibular ligaments,<sup>14 16 21</sup> the dorsiflexion with external rotation stress test<sup>4 22 23</sup> and the squeeze test<sup>7 23–25</sup> are the most commonly described tests, but the Cotton<sup>21 23 26</sup> and fibula translation<sup>22 26</sup> tests have also been described.

The clinical diagnostic tests aim to reproduce symptoms by applying stress to the syndesmosis. The dorsiflexion with external rotation test is thought to reproduce pain over the ankle syndesmosis ligaments by mimicking the commonly described mechanism of injury. The squeeze test involves mid-calf compression of the tibia and fibula which is thought to cause separation at the distal tibiofibular joint, in turn increasing tension in

**To cite:** Sman AD, Hiller CE, Refshauge KM. *Br J Sports Med* 2013;**47**: 620–628.

the remaining syndesmosis ligament fibres resulting in pain at the ankle.<sup>27</sup> Biomechanical analysis has confirmed separation at the distal tibiofibular joint when the calf was compressed.<sup>24</sup>

The dense syndesmosis ligaments afford only a small degree of movement between the bones<sup>8</sup> in healthy ankles, and the Cotton test and fibula translation test are used for the diagnosis of ankle syndesmosis injury because they are thought to produce increased movement compared to the opposite ankle. With the Cotton test, medial and lateral forces are applied to the talus with the ankle in the neutral position.<sup>15 26</sup> For the fibula translation test, the fibula is translated from anterior to posterior on the tibia.<sup>15</sup> These tests are positive when increased mediolateral (Cotton) or anteroposterior (fibula translation) movement is felt compared to the opposite ankle.<sup>15 26</sup>

In addition to clinical examination, several surgical and radiographic tools have been investigated for their diagnostic value for ankle syndesmosis injury. The most valuable technique to accurately diagnose ankle syndesmosis injury is arthroscopy.<sup>22 28 29</sup> However, it is expensive and invasive and therefore is usually only performed to diagnose and treat the more severe grades of injury. That is, it is used to enable treatment and not simply as a diagnostic tool.<sup>27</sup>

Of the radiographic tools, MRI has been found to have the highest specificity and sensitivity, similar to that of arthroscopy.<sup>29–32</sup> However, the costs for such examinations are still relatively high. x-Ray is widely used, especially to rule out fractures, but the reliability of x-ray to diagnose injury to the ankle syndesmosis is questionable.<sup>27 33</sup> CT, however, has been shown to be more sensitive than x-ray for detecting ankle syndesmosis injury,<sup>34</sup> but has a high radiation dose. Therefore, although MRI would be the preferred diagnostic tool for this injury, x-ray and CT are also often used in clinical practice.

An accurate and reliable diagnostic protocol for diagnostic tests of ankle syndesmosis injury should facilitate timely and accurate diagnosis, enabling reduced use of expensive diagnostic tools, and an earlier start of appropriate treatment. Therefore, the aim of this systematic review was to determine whether clinical diagnostic tests accurately and reliably diagnose ankle syndesmosis injury.

## METHODS

### Identification of selected studies

#### Search strategy

Studies were retrieved from searches of the following databases from earliest record to 6 August 2012: MEDLINE, EMBASE, CINAHL, PsycINFO, PEDro, AMED, SPORTDiscus, Scopus, Web of Science, PubMed and Cochrane Register of Clinical Trials. The search strategy was developed for MEDLINE and modified for use in other databases (figure 1). The reference lists from identified articles were also examined for additional relevant references. Seven authors were contacted to obtain additional information.<sup>16 22 25 26 30 35 36</sup> We received five responses<sup>16 22 26 30 36</sup> and data from three studies<sup>16 26 36</sup> that met the inclusion criteria were provided.

#### Eligibility criteria

The criteria for the selection of eligible studies were developed in accordance with the handbook for systematic reviews of diagnostic test accuracy from the Cochrane collaboration<sup>37</sup> and by Devillé *et al.*<sup>38</sup> Studies were selected based on the following eligibility criteria: any trial design; participants who sustained an ankle injury with suspicion of an ankle syndesmosis injury without fracture. There was no language restriction. Studies examining the accuracy of clinical tests needed to compare the

```
Syndesmosis.mp OR Syndesmosis injury.mp OR
  syndesmosis ankle injury.mp
OR
  • Syndesmo*.mp
  • High ankle*.mp OR high ankle sprain*.mp OR Ankle
  sprain*.mp OR Ankle*.mp
  • Exp Ankle/pa,ra (pathology, radiography)
  • Exp ankle injuries/ di (diagnosis)
  • Athletic injur*.mp OR Sport injur*.mp
AND
  • Diagno*.mp
AND
  • Clinical test*.mp OR Manual test*.mp OR Physical
  test*.mp OR physical exam*.mp OR Exp physical
  examination/cl,mt,st (classification, methods,
  standards)
AND
  • x-rays/di,mt,st (diagnosis,methods,standards) OR
  diagnostic imagi*.mp OR tomograph*.mp OR x-ray
  computed*.mp OR magnetic resonance imagin*.mp*
  OR radiograph*.mp OR MRI*.mp OR CT*.mp OR
  Arthroscopy*.mp OR Arthrosc*.mp OR
  Ultrasound*.mp
AND
  • valid*.mp OR accura*.mp OR sensitiv*.mp OR
  specific*.mp OR reliab*.mp
```

Figure 1 Keyword search MEDLINE.

clinical test with a reference standard (eg, MRI, arthroscopy) and present relevant statistical values (eg, OR, sensitivity and specificity values). Studies investigating reliability needed to investigate one or more clinical tests and present relevant statistical values (eg, ICCs, Cohen's  $\kappa$ ). Systematic reviews were excluded; however, the reference lists from these systematic reviews were examined for any additional relevant references. A complete description of the eligibility criteria is shown in figure 2.

### Study selection

From the results of the initial search, titles and abstracts were screened independently by two of the authors, using the predetermined inclusion and exclusion criteria and full text articles were retrieved. Full text articles were further screened independently by two of the authors for inclusion. Differences between the reviewers were resolved by consensus. Where consensus was not reached, a third author adjudicated.

### Data extraction

Data extracted included study design, sample size and participant characteristics, clinical tests used, diagnostic criteria used for the clinical tests, the individual results of the clinical test(s) and the reference standard used. Data were extracted by two independent reviewers using a standard form developed for the review. Differences of opinion between reviewers were resolved by consensus. Where consensus was not reached, a third author adjudicated. A description of the examination protocols for the clinical tests used in each included study can be found in table 1.

Inclusion criteria	
<i>Design</i>	
<ul style="list-style-type: none"> <li>All cross-sectional, delayed cross-sectional, cohort type accuracy studies, case-control type accuracy studies, longitudinal, prospective, retrospective were included in the search</li> <li>Full text studies in peer-reviewed journals</li> <li>Any sample size</li> </ul>	
<i>Participants</i>	
<ul style="list-style-type: none"> <li>Human participants of any age, gender, ethnic group, geographic region or with any given level of care</li> <li>Ankle pain after injury</li> <li>Suspicion of Syndesmosis injury</li> </ul>	
<i>Comparisons</i>	
<ul style="list-style-type: none"> <li>Papers investigating intra-rater or inter-rater reliability of one or more clinical tests</li> <li>Papers comparing clinical diagnostic tests to a gold standard (arthroscopy) or a reference standard (MRI, CT, X-ray, Ultrasound) to investigate diagnostic accuracy.</li> </ul>	
<i>Outcome measures</i>	
<ul style="list-style-type: none"> <li>Studies reporting sensitivity, specificity, odds ratios, likelihood ratios and/or receiver operating curves (ROC) for validity studies</li> <li>Studies report ICCs, Pearson's r, Spearman's p, Cohen's Kappa, Fleiss' Kappy, or CCCs for reliability studies</li> </ul>	
<i>Exclusion criteria</i>	
<ul style="list-style-type: none"> <li>Systematic review</li> <li>Studies including participants with fractures</li> <li>Cadaver studies</li> </ul>	

**Figure 2** Eligibility criteria systematic review.

### Quality assessment

Methodological quality of the included articles was assessed using the recommended Quality Assessment for Diagnostic Accuracy studies (QUADAS) tool.<sup>39 40</sup> The tool consists of 14 items rated as either 'yes', 'no', 'unclear' or 'N/A'. Nine items relate to bias (item 3–7, 10–12 and 14), three to the quality of reporting (items 8, 9 and 13) and two to variability (items 1 and 2). Methodological quality was assessed by two independent reviewers.

### SYNTHESIS OF RESULTS

Sensitivity, specificity and likelihood ratios (LR), with 95% CIs, were calculated. Diagnostic accuracy was calculated using the

method as described by Portney and Watkins (true positive + true negative)/total number of cases.<sup>41</sup> LR are considered to be both clinically useful statistics<sup>42 43</sup> and the best indices of diagnostic accuracy.<sup>44</sup> Guidelines described by Jaeschke *et al*<sup>45</sup> were used to interpret the LR. A likelihood ratio between 0.5 and 2 is considered a very small or irrelevant shift, 0.2–0.5 or 2–5 is a small shift, 0.1–0.2 or 5–10 is a moderate shift and <1 or >10 is a large, often conclusive shift.<sup>45</sup> Intra-rater reliability could not be calculated with ICCs due to insufficient variability between scores. Therefore, intra-reliability results were analysed by percentage agreement between sessions for the same rater. Inter-rater reliability scores were calculated with ICC<sub>(2,1)</sub> and 95% CIs. Interpretation of the ICC values was according to the guidelines by Fleiss.<sup>46</sup>

### RESULTS

#### Description of included studies

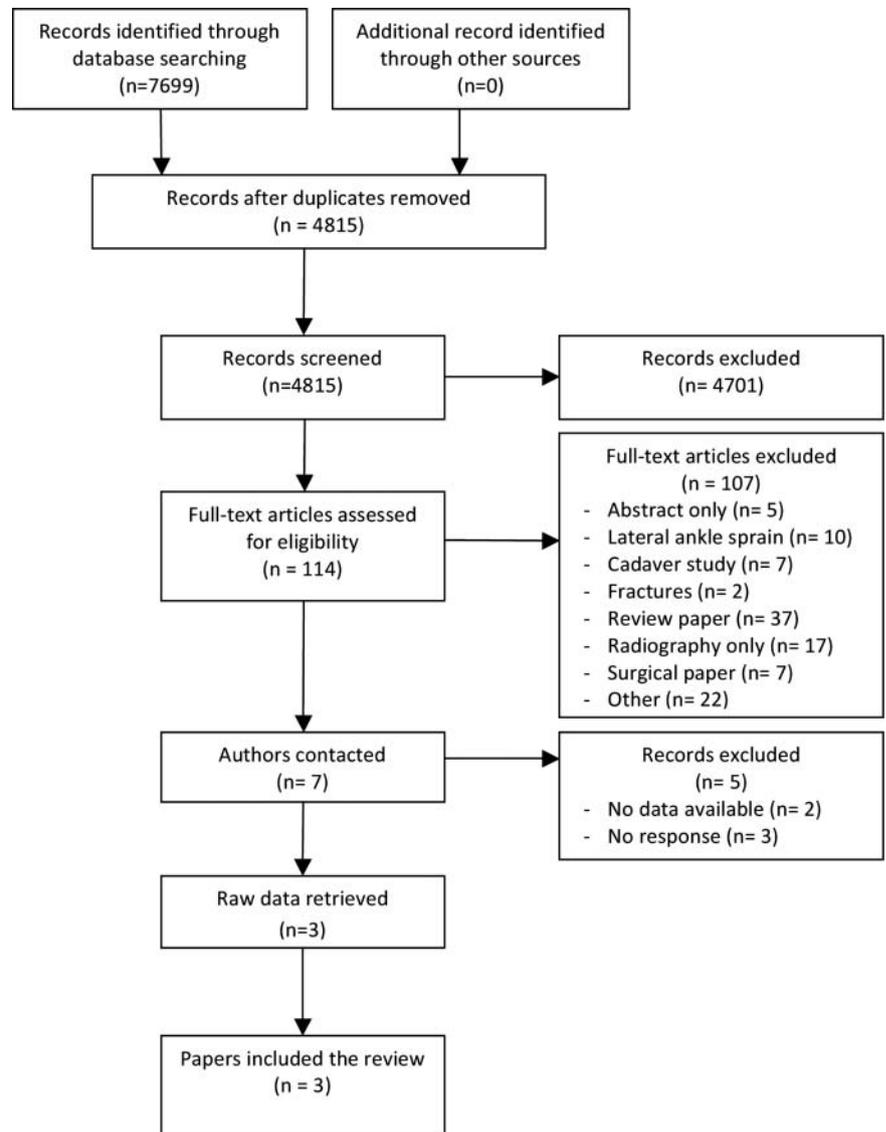
The initial electronic database search resulted in a total of 7699 articles, leaving 4815 articles after the removal of duplicates. Following the title and abstract screening, 114 were selected for possible inclusion in the review and full text articles were retrieved. Following the screening of the full text, seven articles met the inclusion criteria<sup>16 22 25 26 30 35 36</sup> (figure 2). Raw data from three articles were available and included in the systematic review for further analysis.<sup>16 26 36</sup> Two studies were included for the analysis of accuracy of the clinical tests.<sup>16 26</sup> Nussbaum *et al*<sup>16</sup> compared findings from three clinical tests with the reference standard of radiography in 60 participants and Beumer *et al*<sup>26</sup> compared findings from six clinical tests with arthroscopy in three participants. Two studies were included for the reliability analysis.<sup>26 36</sup> Beumer *et al*<sup>26</sup> reviewed six clinical tests on 12 participants with seven examiners and Alonso *et al*<sup>36</sup> reviewed four tests on 53 participants with two examiners. A flow diagram of the search history and selection process is presented in figure 3.

A total of eight clinical diagnostic tests were investigated, of which six tests have previously been described in the literature as commonly used tests for the diagnosis of ankle syndesmosis

**Table 1** Protocol used for clinical tests in the included studies

Test	Authors	Protocol	Positive findings
Squeeze test	Nussbaum <i>et al</i> <sup>16</sup> Alonso <i>et al</i> <sup>36</sup> Beumer <i>et al</i> <sup>26</sup>	Patient sitting over side of the bed. Compression of fibula to the tibia above the midpoint of the calf	Pain over area of the syndesmosis ligaments
Dorsiflexion with compression test	Alonso <i>et al</i> <sup>36</sup>	Patient standing and actively dorsiflexing the ankle. Performed once unassisted and once with the therapist applying a manual compressive force to the malleoli	Significant increase in ankle range of motion when compression added, or decreased pain at end of range when compression added
Dorsiflexion range of motion	Beumer <i>et al</i> <sup>26</sup>	Application of passive dorsiflexion	Reduced dorsiflexion range of motion compared with contralateral ankle
External rotation test	Alonso <i>et al</i> <sup>36</sup> Beumer <i>et al</i> <sup>26</sup>	Participant sitting over side of the bed. Application of a passive external rotation stress to affected foot and ankle with the knee at 90° and ankle in neutral	Pain over the syndesmosis ligaments
Palpation	Nussbaum <i>et al</i> <sup>16</sup> Alonso <i>et al</i> <sup>36</sup>	Performed with the ankle passively and maximally dorsiflexed in external rotation on a fixed lower leg	Reproduction of pain along the syndesmosis
Cotton test	Nussbaum <i>et al</i> <sup>16</sup> Alonso <i>et al</i> <sup>36</sup>	Palpation over anterior tibiofibular ligament	Report of pain after pressing directly over the ligament
Fibula translation	Beumer <i>et al</i> <sup>26</sup>	Distal tibia stabilised and lateral force applied to the foot	Increased lateral translation of the talus from medial to lateral compared with contralateral side
Anterior drawer	Beumer <i>et al</i> <sup>26</sup>	Apply anterior–posterior translation of the fibula	Anteroposterior displacement of the fibula is greater than contralateral side
	Beumer <i>et al</i> <sup>26</sup>	Not described	Used as a differential diagnosis between lateral collateral instability and syndesmosis instability

**Figure 3** Flow of studies through the review.



injury. These are the palpation of the tibiofibular ligaments,<sup>14 16 21</sup> the external rotation stress test,<sup>4 22 23</sup> the squeeze test,<sup>7 23–25</sup> the Cotton test<sup>21 23 26</sup> and the fibula translation test.<sup>22 26</sup> Beumer *et al*<sup>26</sup> also investigated the diagnostic accuracy of reduced dorsiflexion range of motion (ROM) and the anterior drawer test. However, although these tests are used as part of clinical examination, their use for the diagnosis of ankle syndesmosis injury has not been previously researched or described.

### Quality assessment

Moderate quality was found for the three included studies.<sup>16 26 36</sup> All three studies adequately addressed the majority of items on bias, variability and quality of reporting where applicable to the studies. Of the nine items relating to eliminating bias, Beumer *et al*<sup>26</sup> addressed six items, whereas Nussbaum *et al*<sup>16</sup> addressed five. Variability accounted for two items and was not addressed by Beumer *et al*;<sup>26</sup> however, Nussbaum *et al*<sup>16</sup> addressed both items. The final three items related to quality of reporting; Beumer *et al*<sup>26</sup> addressed two of the three, and Nussbaum *et al*<sup>16</sup> addressed all items. For Alonso *et al*,<sup>36</sup> however, only 6/14 items applied to the study and ‘yes’ was scored on all of these (table 2).

### Diagnostic accuracy of clinical tests

Of the two studies<sup>16 26</sup> that investigated diagnostic accuracy (table 3) only the squeeze test was investigated in both studies; however, conflicting results were found. Nussbaum *et al*<sup>16</sup> reported a positive likelihood ratio of 2.68 (95% CI 1.93 to 3.73) whereas Beumer *et al*<sup>26</sup> showed a positive likelihood ratio of 0.67 (95% CI 0.39 to 1.15) and a negative likelihood ratio of 3.00 (95% CI 0.44 to 20.32). LR for the anterior drawer, Cotton test, dorsiflexion ROM, external rotation and fibula translation tests ranged from LR+1.50 to LR–1.50. Diagnostic accuracy, LR and sensitivity and specificity values could not be calculated for the external rotation with dorsiflexion test and AITFL tenderness (table 4).

### Reliability of clinical tests

#### Intra-rater reliability

One study<sup>26</sup> was retrieved with data that enabled the calculation of intra-rater reliability (table 5). We were not able to calculate ICCs due to insufficient variability between scores and the results are therefore presented as per cent agreement. Intra-rater reliability for the squeeze, Cotton, dorsiflexion ROM and the external rotation tests ranged from 83% to 100% agreement between the two test occasions. Intra-rater reliability for fibula

**Table 2** Quality assessment—QUADAS tool

Author	Was the spectrum of patients representative of the patients who will receive the test in practice?	Were selection criteria clearly described?	Was the reference standard likely to correctly classify the target condition?	Was the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two sets?	Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?	Did the patients receive the same reference standard regardless of the index test result?	Was the reference standard independent of the index test? (ie the index test did not form part of the reference standard)	Was the execution of the index test described in sufficient detail to permit its replication?	Was the execution of the reference test described in sufficient detail to permit its replication?	Were the index test results interpreted without knowledge of the results of the reference standard?	Were the reference standard results interpreted without knowledge of the results of the index test?	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Were un-interpretable /intermediate test results reported?	Were withdrawals from the study explained?
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Beumer <i>et al</i> <sup>26</sup>	No	No	Yes	Yes	No	No	N/A	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes
Nussbaum <i>et al</i> <sup>16</sup>	Yes	Yes	No	Yes	Yes	Yes	N/A	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes
Alonsoet <i>al</i> <sup>36</sup> *	Yes	Yes	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	Yes	Yes	Yes

\*QUADAS tool appropriate for reliability, not for validity. Therefore no reference test was required (items 3–7, 9–11).  
N/A, the item is not applicable to this study.

**Table 3** Summary of included studies investigating diagnostic accuracy

Study	Participants	Inclusion–exclusion criteria	Procedures	Tests investigated	Reference standard used
Nussbaum <i>et al</i> <sup>16</sup>	N=60 (51M, 9F) NCAA Division I-A athletes with syndesmosis sprain 34 of 60 injuries in football, 9 lacrosse, 8 soccer, 2 rowing and 1 each in wrestling, gymnastics, swimming, track, field hockey, cheerleading, basketball Right ankle 28 Left ankle 32	<ul style="list-style-type: none"> <li>▶ History and physical examination consistent with an acute ankle sprain</li> <li>▶ Functional disability by failure on a single leg hop test</li> <li>▶ Evaluated and treated within 24 h of injury</li> <li>▶ No history of prior syndesmosis ankle sprain</li> </ul>	Physical examination and standard radiographs including AP, internal rotation mortise and lateral views Examined within 24 h of injury by certified athletic trainer. Physician reported on radiographs	Squeeze test Ligament palpation External rotation stress test with passive dorsiflexion	Radiographs
Beumer <i>et al</i> <sup>26</sup>	N=12 3 Participants with suspected chronic syndesmosis injury underwent arthroscopy 9 healthy subjects with asymptomatic ankles	Not defined	Participants sat behind a curtain with exposure only of the lower legs Both legs were examined twice in a different order by 7 examiners (4 orthopaedic surgeons, 3 orthopaedic registrars) Participants did not speak and indicated pain by tapping on a wooden board then pointing to the place where pain was felt Following day, 3 patients with suspected chronic syndesmosis injury underwent arthroscopy	Squeeze test Cotton test Fibula translation External rotation stress test Anterior drawer Dorsiflexion range of motion	Arthroscopy

translation and anterior drawer test showed a greater variability ranging from 46% to 92% agreement between the two test occasions (table 6).

#### Inter-rater reliability

Inter-rater reliability was calculated from the raw data of two studies (table 5).<sup>26 36</sup> Good reliability was found for the

external rotation test in both studies ( $ICC_{2,1}$  0.732<sup>26</sup> and  $ICC_{2,1}$  0.744<sup>36</sup>). Fair reliability was found for the squeeze test ( $ICC_{2,1}$  0.461<sup>26</sup> and 0.490<sup>36</sup>) and ligament palpation ( $ICC_{2,1}$  0.490<sup>36</sup>). Poor reliability was found for the anterior drawer test ( $ICC_{2,1}$  0.059<sup>36</sup>), the Cotton test ( $ICC_{2,1}$  0.156<sup>36</sup>) and the fibula translation test ( $ICC_{2,1}$  0.282<sup>36</sup>) for the single study that evaluated these tests (table 7).

**Table 4** Diagnostic accuracy of clinical tests, diagnostic accuracy, sensitivity, specificity, likelihood ratios (LR) (95% CI),  $\chi^2$  and p value

Author	N	Test ↓	True positive	False positive	False negative	True negative	Diagnostic accuracy (%) <sup>*</sup>	Sensitivity	Specificity	LR+	LR–	$\chi^2$	p Value
Beumer <i>et al</i> <sup>26†</sup>	21‡	Anterior drawer	5	4	9	3	38.1	0.36 (0.16 to 0.61)	0.43 (0.16 to 0.75)	0.63 (0.24 to 1.62)	1.50 (0.59 to 3.84)	0.219	0.640
		Cotton test	4	2	10	5	42.9	0.29 (0.12 to 0.55)	0.71 (0.36 to 0.92)	1.00 (0.24 to 4.20)	1.00 (0.56 to 1.78)	0.000	1.000
		Dorsiflexion	7	3	7	4	52.4	0.50 (0.27 to 0.73)	0.57 (0.25 to 0.84)	1.17 (0.43 to 3.18)	0.88 (0.38 to 2.00)	0.000	1.000
		External rotation	7	7	7	0	33.3	0.50 (0.27 to 0.73)	0.00 (0 to 0.35)	0.50 (0.30 to 0.84)	N/A	3.241	0.072
		Fibula translation	9	3	5	4	61.9	0.64 (0.39 to 0.84)	0.57 (0.25 to 0.84)	1.50 (0.59 to 3.84)	0.63 (0.24 to 1.62)	0.219	0.640
		Squeeze test	8	6	6	1	42.9	0.57 (0.33 to 0.79)	0.14 (0.03 to 0.51)	0.67 (0.39 to 1.15)	3.00 (0.44 to 20.32)	0.670	0.413
Nussbaum <i>et al</i> <sup>16§</sup>	60	Squeeze test	1	22	0	37	63.3	1.00 (0.21 to 1)	0.63 (0.50 to 0.74)	2.68 (1.93 to 3.73)	N/A	0.059	0.809
		External rotation¶	1	59	0	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		AITFL tenderness	1	59	0	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A

<sup>\*</sup>(True positive + true negative)/number of cases, as described by Portney and Watkins.<sup>41</sup>

<sup>†</sup>Arthroscopy was used as the reference standard.

<sup>‡</sup>Three people had arthroscopy and were clinically diagnosed by seven (blinded) assessors.

<sup>§</sup>X-Ray was used as the reference standard.

<sup>¶</sup>External rotation test performed in maximal passive dorsiflexion.

AITFL, anterior inferior tibiofibular ligament N/A, results were not able to be calculated.

**Table 5** Summary of included studies investigating reliability

Study	Participants	Inclusion–exclusion criteria	Study design	Tests investigated
Alonso <i>et al</i> <sup>26</sup>	N=53 (38M, 15F) Age: 24.3±8.5 years (12–52) Time between injury and test: 34.2±125 days, median 5 60.4% tested within 7 days of injury All involved in a ball sport	Excluded if: ▶ Open wound in the area of the lower leg and ankle ▶ Fracture of the tibia or fibula ▶ Ankle pain of non-mechanical origin ▶ Any condition that might be exacerbated by test procedures	9 physiotherapists, 2 clinics Experience in sports injuries 1–11 years with average of 5 years Training session for the 4 tests First rater was participant's therapist. Second rater blinded to results. Results independently documented Tests performed at first consultation during routine examination	Squeeze test Dorsiflexion with compression test External rotation test Ligament palpation
Beumer <i>et al</i> <sup>26</sup>	N=12 3 with suspected chronic syndesmosis injury 9 healthy subjects with asymptomatic ankles	Not defined	Participants sat behind a curtain with exposure only of the lower legs Both legs were examined twice in a different order by 7 examiners (4 orthopaedic surgeons, 3 orthopaedic registrars) Participants did not speak and indicated pain by tapping on a wooden board then pointing to the place where pain was felt Injury diagnosis based on medical history, physical examination and diagnostic imaging	Squeeze test Cotton test Fibula translation External rotation stress test Anterior drawer Dorsiflexion range of motion

## DISCUSSION

This systematic review found that the diagnostic sensitivity and specificity, and inter-rater reliability, of clinical tests to identify ankle syndesmosis injury were very low although the intra-rater reliability was adequate. The clinical tests reviewed were the anterior drawer, Cotton, dorsiflexion, external rotation, fibula translation and squeeze tests. However, the findings are based on two articles that investigated diagnostic accuracy and two articles that investigated the reliability of clinical tests used. Overall eight clinical tests and two reference standards were used in the diagnosis of a total number of 125 patients.

### Diagnostic accuracy of clinical tests

Diagnostic accuracy for all evaluated clinical tests was very low. LR were all close to 1, indicating that the tests are unhelpful in the clinic. The squeeze test was the only test with a clinically important result.<sup>16 45</sup> The conclusion that clinical tests are not accurate should be carefully interpreted. The aim of the study conducted by Nussbaum *et al*<sup>16</sup> was not to investigate diagnostic accuracy, but to evaluate time missed from sport in an attempt to better define the injury. Therefore, the use of x-rays was in relation to clinical pragmatism not for use as a reference standard. Nussbaum *et al*<sup>16</sup> did report a significant relationship ( $p=0.03$ ) between a positive squeeze test and a longer time lost from competition. This indicates that although the squeeze test might not be very useful as a diagnostic tool, it could be useful as a prognostic tool.

Additionally, the external rotation test in passive dorsiflexion and palpatory tenderness along the AITFL as investigated by Nussbaum *et al*<sup>16</sup> were excluded from the analysis because there was no variability among the measures. The x-rays, which were used as the reference standard by Nussbaum *et al*<sup>16</sup> were negative in 59 of the 60 cases, while AITFL tenderness and the external rotation in passive dorsiflexion test was positive for all 60 cases.

Beumer *et al*,<sup>26</sup> however, used arthroscopy as the reference standard which is considered as a very valuable tool to accurately diagnose ankle syndesmosis injury. Unfortunately arthroscopy was only performed in three participants because the remaining nine participants were healthy controls.

Despite our extensive search of the literature we did not find any diagnostic accuracy studies using MRI as a reference standard, even though the ability of MRI to accurately detect ankle syndesmosis injuries is similar to that of arthroscopy.<sup>29–32</sup> This could be due to the fact that most studies were conducted prior to knowledge of the excellent diagnostic accuracy of MRI. Future investigation of diagnostic accuracy of clinical tests with the use of MRI as a reference standard would be recommended.

### Reliability of the clinical tests

The reliability of the clinical tests included in this systematic review varied among tests, but were consistent between the two studies included in our review. Intra-rater reliability was investigated in only one study,<sup>26</sup> which had such limited

**Table 6** Intra-rater reliability in percentages\*<sup>26</sup>

Rater ↓/test →	Squeeze test	Cotton test	Fibula translation	External rotation	Anterior drawer	Dorsiflexion
A	91.7	91.7	87.5	91.7	70.8	95.8
B	91.7	87.5	87.5	100	91.7	100
C	91.7	83.3	87.5	95.8	87.5	95.8
D	91.7	95.8	70.8	95.8	45.8	100
E	87.5	100	87.5	95.8	87.5	87.5
F	87.5	91.7†	91.7	95.8	87.5	83.3
G	87.5	87.5	75.0	87.5	66.7	87.5

\*Percentage of agreement between session 1 and session 2 measured in 12 patients (24 legs).

†Data from one leg in a single participant from one of the raters during one session was missing.

**Table 7** Inter-rater reliability, ICC<sub>2,1</sub> (95% CI)

Author./test →	Anterior drawer	Cotton test	Dorsiflexion	External rotation	Fibula translation	Squeeze test	Palpation
Beumer <i>et al</i> <sup>26*</sup>	0.059 (-0.19 to 0.205)	0.156 (0.042 to 0.346)	0.355 (0.203 to 0.558)	0.732 (0.599 to 0.851)	0.282 (0.141 to 0.486)	0.461 (0.300 to 0.652)	
Alonso <i>et al</i> <sup>36</sup>			0.326 (0.061 to 0.547)	0.744 (0.596 to 0.844)		0.490 (0.259 to 0.669)	0.490 (0.259 to 0.669)

\*Two sessions were performed and the results of the first session are described here as this reflects when the tests are used in the clinic.

variability between test results that ICCs could not be calculated. However, the per cent close agreement was high for the squeeze, Cotton, external rotation and dorsiflexion tests (83–100% agreement).

Inter-rater reliability was investigated in two studies, finding good reliability for the external rotation stress test and fair-to-poor reliability for the squeeze, dorsiflexion, Cotton, anterior drawer and fibula translation tests. Both studies used similar clinical test examination protocols with experienced examiners who were appropriately blinded to the results. The main limiting factor was that only data from the three tests described were available from the two studies.<sup>26 36</sup> In addition, of these tests, only the squeeze and external rotation test had common protocols. The protocol differed between studies for the dorsiflexion test: Beumer *et al*<sup>26</sup> interpreted findings based on limited ROM whereas Alonso *et al*<sup>36</sup> performed a dorsiflexion lunge with compression test. As seen in the results of this systematic review, there are many different ways to perform these clinical tests and few have been validated. We recommend that studies use trained examiners to ensure standardised performance of the investigated tests and consistent interpretation of findings.

### Clinical implications

Although specific clinical tests have formed a critical part of the triage for diagnosis of ankle syndesmosis injury, the usefulness of these tests has not been adequately researched. Studies that reported the reliability or accuracy of clinical tests to diagnose ankle syndesmosis injury were limited, and therefore the diagnostic accuracy of the clinical diagnostic tests remains unclear. It is important to note that each test in this systematic review has been investigated and analysed for its individual ability to accurately diagnose ankle syndesmosis injury. Our findings show that clinicians cannot rely on a single test in isolation as no test has been found to be sufficient to identify ankle syndesmosis injury with certainty. It is likely that a combination of tests and inclusion of other elements such as symptoms and the patient's history might further assist in diagnosis. However, the studies included in this review had several limiting factors in methodology, including the choice of a reference standard and the results should therefore be interpreted with caution. None of the sample sizes of the three articles we retrieved, after contacting the authors, were appropriately powered.

The ability of clinical tests to diagnose ankle syndesmosis injury accurately without the necessity of imaging investigations would enable early start of appropriate management and potentially decrease the risk of persistence of problems. Importantly, it would limit costs for patients and the healthcare system.

### CONCLUSION

This is the first systematic review that has investigated reliability and accuracy of clinical tests for the diagnosis of ankle syndesmosis injury. There are clear gaps in the research about the reliability and diagnostic accuracy of the clinical tests and the usefulness of these clinical diagnostic tests remains uncertain. We recommend that future studies concentrate on using MRI or arthroscopy as a reference standard, recruit appropriate sample sizes and combine clinical tests with factors such as the patient's history. Clinicians should be aware of the limitations of the current clinical tests and conduct additional diagnostic tests, such as MRI or arthroscopy to make a final diagnosis of syndesmosis injury.

### What are the new findings

- ▶ This is the first systematic review that has investigated reliability and accuracy of clinical tests for the diagnosis of ankle syndesmosis injury. The clinical tests reviewed were the anterior drawer, Cotton, dorsiflexion, external rotation, fibula translation and squeeze tests.
- ▶ The findings showed that the diagnostic sensitivity and specificity were very low and only the squeeze test showed a clinically important result. Inter-rater reliability of clinical tests to identify ankle syndesmosis injury was investigated in two studies, finding only good reliability for the external rotation stress test. The intra-rater reliability was investigated in one study and was found to be adequate (83–100% agreement). However, these findings are based on two articles that investigated diagnostic accuracy and two articles that investigated the reliability of clinical tests used.
- ▶ Although specific clinical tests have formed a critical part of the triage for diagnosis of ankle syndesmosis injury, the usefulness of these tests has not been adequately researched. Studies that reported the reliability or accuracy of clinical tests to diagnose ankle syndesmosis injury were limited and had several limiting factors in methodology, including the choice of a reference standard and the results should therefore be interpreted with caution. None of the sample sizes of the three articles we retrieved, after contacting the authors, were appropriately powered.
- ▶ There are clear gaps in the research about the reliability and diagnostic accuracy of the clinical tests and the usefulness of these clinical diagnostic tests remains uncertain. We recommend that future studies concentrate on using MRI or arthroscopy as a reference standard, recruit appropriate sample sizes and combine clinical tests with factors such as the patient's history.

### How might it impact on clinical practice in the near future

- ▶ Our findings show that clinicians cannot rely on a single test in isolation as no test has been found to be sufficient to identify ankle syndesmosis injury with certainty. It is likely that a combination of tests and inclusion of other elements such as symptoms and the patient's history might further assist in diagnosis.
- ▶ Clinicians should be aware of the limitations of the current clinical tests and conduct additional diagnostic tests, such as MRI or arthroscopy to make a final diagnosis of syndesmosis injury.
- ▶ The ability of clinical tests to diagnose ankle syndesmosis injury accurately without the necessity of imaging investigations would enable early start of appropriate management and potentially decrease the risk of persistence of problems. Importantly, it would limit costs for patients and the healthcare system.

drafting and revising of the article critically for important intellectual content and the final approval of the version to be published. CEH contributed to the conception and design of the study, the collection and extraction of the data, the interpretation of the findings, the revising of the article critically for important intellectual content and the final approval of the version to be published. KMR contributed to the conception and design of the study, the interpretation of the findings, the revising of the article critically for important intellectual content and the final approval of the version to be published.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interest** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

### REFERENCES

- 1 Fong DT, Hong Y, Chan LK, *et al.* A systematic review on ankle injury and ankle sprain in sports. *Sports Med* 2007;37:73–94.
- 2 Gerber JP, Williams GN, Scoville CR, *et al.* Persistent disability associated with ankle sprains: a prospective examination of an athletic population. *Foot Ankle Int* 1998;19:653–60.
- 3 Brosky T, Nyland J, Nitz A, *et al.* The ankle ligaments: consideration of syndesmotom injury and implications for rehabilitation. *J Orthop Sports Phys Ther* 1995;21:197–205.
- 4 Boytim MJ, Fischer DA, Neumann L. Syndesmotom ankle sprains. *Am J Sports Med* 1991;19:294–8.
- 5 Fritschy D. An unusual ankle injury in top skiers. *Am J Sports Med* 1989;17:282–6.
- 6 Wright RW, Barile RJ, Surprenant DA, *et al.* Ankle syndesmosis sprains in National Hockey League players. *Am J Sports Med* 2004;32:1941–5.
- 7 Hopkinson WJ, St. Pierre P, Ryan JB, *et al.* Syndesmosis sprains of the ankle. *Foot Ankle* 1990;10:325–30.
- 8 Gray H. Arthrology. In: Peter L, Williams RW, eds. *Gray's anatomy*. 36 edn. Edinburgh: Churchill Livingstone, 1980: 421–2.
- 9 Gray H. Ankle and foot. In: Standing S, ed. *Gray's anatomy*. 40 edn. London: Churchill Livingstone, 2008:1429–63.
- 10 Levangie PK, Norkin CC. *The ankle and foot complex. Joint structure and function*. Ferny Creek, Victoria, Australia: MacLennan & Petty Pty Limited, 2001:369–70.
- 11 Taylor DC, Bassett Iii FH. Syndesmosis ankle sprains: diagnosing the injury and aiding recovery. *Phys Sportsmed* 1993;21:39–46.
- 12 Kjaer MKM, Magnusson P, Engebretsen L, *et al.* *Sports injury: regional considerations. Diagnosis and treatment. Textbook of sports medicine*. Oxford: Blackwell Science Ltd 2003:540–51.
- 13 Bloemers FW, Bakker FC. Acute ankle syndesmosis injury in athletes. *Eur J Trauma* 2006;32:350–6.
- 14 Taylor DC, Englehardt DL, Bassett Iii FH. Syndesmosis sprains of the ankle. The influence of heterotopic ossification. *Am J Sports Med* 1992;20:146–50.
- 15 Mulligan EP. Evaluation and management of ankle syndesmosis injuries. *Phys Ther Sport* 2011;12:57–69.
- 16 Nussbaum ED, Hosea TM, Sieler SD, *et al.* Prospective evaluation of syndesmotom ankle sprains without diastasis. *Am J Sports Med* 2001;29:31–5.
- 17 Fites B, Kunes J, Madaleno J, *et al.* Latent syndesmosis injuries in athletes. *Orthopedics* 2006;29:124–7.
- 18 Miller CD, Shelton WR, Barrett GR, *et al.* Deltoid and syndesmosis ligament injury of the ankle without fracture. *Am J Sports Med* 1995;23:746–50.
- 19 Karl EL, Wrazidlo W. Fresh rupture of the syndesmosis of the proximal ankle joint. Clinical significance and arthrographic diagnosis. *Der Unfallchirurg* 1987;90:92–6.
- 20 Frick H. The isolated tear of the tibio-fibular syndesmosis—mechanism, clinical observations, diagnosis and therapy. *Unfallheilkunde* 1978;81:542–5.
- 21 Mullins JF, Sallis JG. Recurrent sprain of the ankle joint with diastasis. *J Bone Joint Surg Br* 1958;40:270–3.
- 22 Ogilvie-Harris DJ, Reed SC. Disruption of the ankle syndesmosis: diagnosis and treatment by arthroscopic surgery. *Arthroscopy* 1994;10:561–8.
- 23 Magee D. *Lower leg, ankle and foot. Orthopedic Physical Assessment. 4 edn.* Philadelphia: Saunders, 2002:765–809.
- 24 Teitz CC, Harrington RM. A biomechanical analysis of the squeeze test for sprains of the syndesmotom ligaments of the ankle. *Foot Ankle Int* 1998;19:489–92.
- 25 Kiter E, Bozkurt M. The crossed-leg test for examination of ankle syndesmosis injuries. *Foot Ankle Int* 2005;26:187–8.
- 26 Beumer A, Swierstra BA, Mulder PGH. Clinical diagnosis of syndesmotom ankle instability: evaluation of stress tests behind the curtains. *Acta Orthop Scand* 2002;73:667–9.
- 27 Wataru Miyamoto MT. Management of chronic disruption of the distal tibiofibular syndesmosis. *World J Orthop* 2011;2:1–6.
- 28 Takao M, Ochi M, Naito K, *et al.* Arthroscopic diagnosis of tibiofibular syndesmosis disruption. *Arthroscopy* 2001;17:836–43.
- 29 Takao M, Ochi M, Oae K, *et al.* Diagnosis of a tear of the tibiofibular syndesmosis. The role of arthroscopy of the ankle. *J Bone Joint Surg Br* 2003;85:324–9.

**Contributors** ADS contributed to the conception and design of the study, the collection, extraction and analysis of the data and interpretation of the findings, the

- 30 Vogl TJ, Hochmuth K, Diebold T, *et al.* Magnetic resonance imaging in the diagnosis of acute injured distal tibiofibular syndesmosis. *Invest Radiol* 1997;32:401–9.
- 31 Oae K, Takao M, Naito K, *et al.* Injury of the tibiofibular syndesmosis: value of MR imaging for diagnosis. *Radiology* 2003;227:155–61.
- 32 Han SH, Lee JW, Kim S, *et al.* Chronic tibiofibular syndesmosis injury: the diagnostic efficiency of magnetic resonance imaging and comparative analysis of operative treatment. *Foot Ankle Int* 2007;28:336–42.
- 33 Nielson JH, Gardner MJ, Peterson MGE, *et al.* Radiographic measurements do not predict syndesmotic injury in Ankle Fractures: an MRI STUDY. *Clin Orthop Relat Res* 2005;436:216–21.
- 34 Ebraheim NA, Lu J, Yang H, *et al.* Radiographic and CT evaluation of tibiofibular syndesmotic diastasis: a cadaver study. *Foot Ankle Int* 1997;18:693–8.
- 35 Stricker PR, Spindler KP, Gautier KB. Prospective evaluation of history and physical examination: variables to determine radiography in acute ankle injuries. *Clin J Sport Med* 1998;8:209–14.
- 36 Alonso A, Khoury L, Adams R. Clinical tests for ankle syndesmosis injury: reliability and prediction of return to function. *J Orthop Sports Phys Ther* 1998;27:276–84.
- 37 Bossuyt PM, Leeflang MM. Chapter 6: Developing Criteria for Including Studies. *Handbook for Systematic Reviews of Diagnostic Test Accuracy*. The Cochrane Collaboration, 2008.
- 38 Devillé W, Buntinx F, Bouter L, *et al.* Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:1–13.
- 39 Whiting P, Rutjes AW, Reitsma JB, *et al.* The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
- 40 Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, *et al.* Chapter 9: Assessing methodological quality. In: Deeks JJP, Gatsonis C eds. *Cochrane handbook for systematic reviews of diagnostic test accuracy*. Birmingham: The Cochrane Collaboration, 2009.
- 41 Portney LG, Watkins MP. Chapter 27: statistical measures of validity. In: Cohen M, Kerian M, eds. *Foundations of clinical research: applications to practice*. 3rd edn. Upper Saddle River, NJ: Pearson Prentice Hall, 2009: 619–58.
- 42 Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *Br Med J* 2004;329:168–9.
- 43 Deeks JJ. Systematic reviews of evaluations of diagnostic and screening tests. *Br Med J* 2001;323:157–62.
- 44 Riddle DL, Stratford PW. Interpreting validity indexes for diagnostic tests: an illustration using the Berg balance test. *Phys Ther* 1999;79:939–48.
- 45 Jaeschke R, Guyatt GH, Sackett DL. 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? The Evidence-Based Medicine Working Group. *JAMA* 1994;271:703–7.
- 46 Fleiss J. *The design and analysis of clinical experiments*. New York: John Wiley & Sons, Inc., 1986: 1–32.