Cochrane Library

Cochrane Database of Systematic Reviews

Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement (Review)

Hanchard NCA, Lenza M, Handoll HHG, Takwoingi Y

Hanchard NCA, Lenza M, Handoll HHG, Takwoingi Y.
Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement.
Cochrane Database of Systematic Reviews 2013, Issue 4. Art. No.: CD007427.
DOI: 10.1002/14651858.CD007427.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement

Nigel CA Hanchard[1], Mário Lenza[2], Helen HG Handoll[1], Yemisi Takwoingi[3]

[1]Health and Social Care Institute, Teesside University, Middlesbrough, UK. [2]Orthopaedic and Trauma Department, Hospital Israelita Albert Einstein, São Paulo, Brazil. [3]Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham, UK

Contact: Nigel CA Hanchard, Health and Social Care Institute, Teesside University, Middlesbrough, Tees Valley, TS1 3BA, UK. n.hanchard@tees.ac.uk.

Editorial group: Cochrane Bone, Joint and Muscle Trauma Group.
Publication status and date: Edited (no change to conclusions), published in Issue 9, 2014.

Citation: Hanchard NCA, Lenza M, Handoll HHG, Takwoingi Y. Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *Cochrane Database of Systematic Reviews* 2013, Issue 4. Art. No.: CD007427. DOI: 10.1002/14651858.CD007427.pub2.

ABSTRACT

Background

Impingement is a common cause of shoulder pain. Impingement mechanisms may occur subacromially (under the coraco-acromial arch) or internally (within the shoulder joint), and a number of secondary pathologies may be associated. These include subacromial-subdeltoid bursitis (inflammation of the subacromial portion of the bursa, the subdeltoid portion, or both), tendinopathy or tears affecting the rotator cuff or the long head of biceps tendon, and glenoid labral damage. Accurate diagnosis based on physical tests would facilitate early optimisation of the clinical management approach. Most people with shoulder pain are diagnosed and managed in the primary care setting.

Objectives

To evaluate the diagnostic accuracy of physical tests for shoulder impingements (subacromial or internal) or local lesions of bursa, rotator cuff or labrum that may accompany impingement, in people whose symptoms and/or history suggest any of these disorders.

Search methods

We searched electronic databases for primary studies in two stages. In the first stage, we searched MEDLINE, EMBASE, CINAHL, AMED and DARE (all from inception to November 2005). In the second stage, we searched MEDLINE, EMBASE and AMED (2005 to 15 February 2010). Searches were delimited to articles written in English.

Selection criteria

We considered for inclusion diagnostic test accuracy studies that directly compared the accuracy of one or more physical index tests for shoulder impingement against a reference test in any clinical setting. We considered diagnostic test accuracy studies with cross-sectional or cohort designs (retrospective or prospective), case-control studies and randomised controlled trials.

Data collection and analysis

Two pairs of review authors independently performed study selection, assessed the study quality using QUADAS, and extracted data onto a purpose-designed form, noting patient characteristics (including care setting), study design, index tests and reference standard, and the diagnostic 2 x 2 table. We presented information on sensitivities and specificities with 95% confidence intervals (95% CI) for the index tests. Meta-analysis was not performed.

Main results

We included 33 studies involving 4002 shoulders in 3852 patients. Although 28 studies were prospective, study quality was still generally poor. Mainly reflecting the use of surgery as a reference test in most studies, all but two studies were judged as not meeting the criteria for having a representative spectrum of patients. However, even these two studies only partly recruited from primary care.

The target conditions assessed in the 33 studies were grouped under five main categories: subacromial or internal impingement, rotator cuff tendinopathy or tears, long head of biceps tendinopathy or tears, glenoid labral lesions and multiple undifferentiated target conditions. The majority of studies used arthroscopic surgery as the reference standard. Eight studies utilised reference standards which were potentially applicable to primary care (local anaesthesia, one study; ultrasound, three studies) or the hospital outpatient setting (magnetic resonance imaging, four studies). One study used a variety of reference standards, some applicable to primary care or the hospital outpatient setting. In two of these studies the reference standard used was acceptable for identifying the target condition, but in six it was only partially so. The studies evaluated numerous standard, modified, or combination index tests and 14 novel index tests. There were 170 target condition/index test combinations, but only six instances of any index test being performed and interpreted similarly in two studies. Only two studies of a modified empty can test for full thickness tear of the rotator cuff, and two studies of a modified anterior slide test for type II superior labrum anterior to posterior (SLAP) lesions, were clinically homogenous. Due to the limited number of studies, meta-analyses were considered inappropriate. Sensitivity and specificity estimates from each study are presented on forest plots for the 170 target condition/index test combinations grouped according to target condition.

Authors' conclusions

There is insufficient evidence upon which to base selection of physical tests for shoulder impingements, and local lesions of bursa, tendon or labrum that may accompany impingement, in primary care. The large body of literature revealed extreme diversity in the performance and interpretation of tests, which hinders synthesis of the evidence and/or clinical applicability.

PLAIN LANGUAGE SUMMARY

Physical tests for shoulder impingement in primary care

Impingement (or pinching) of soft-tissues in or around the shoulder is a common cause of pain and is often linked to tissue damage in and around the joint. If doctors and therapists could identify impingement and associated damage using simple, physical tests, it would help them to inform on the best treatment approach at an early stage. We were particularly interested in the primary (community) care setting, because this is where most shoulder pain is diagnosed and managed. We reviewed original research papers for evidence on the accuracy of physical tests for shoulder impingement or associated damage, in people whose symptoms and/or history suggest any of these disorders. To find the research papers, we searched the main electronic databases of medical and allied literature up to 2010. Two review authors screened assessed the quality of each research paper and extracted important information. If multiple research papers reported using the same test for the same condition, we intended to combine their results to gain a more precise estimate of the test's accuracy. We included 33 research papers. These related to studies of 4002 shoulders in 3852 patients. None of the studies exclusively looked at patients from primary care, though two recruited some of their patients from primary care. The majority of studies used arthroscopic surgery as the reference standard. There were 170 different target condition/index test combinations but only six instances where the same test was used in the same way, and for the same reason, in two studies. For this reason combining results was not appropriate. We concluded that there is insufficient evidence upon which to base selection of physical tests for shoulder impingement, and potentially related conditions, in primary care.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of results table

Setting	Most people with shoulder pain symptomatic of impingements and related pathologies are diagnosed and managed in the primary care setting.
Index tests	Physical tests used single or in combination to identify shoulder impingement and related pathologies.
Reference standard	While a definitive reference standard is lacking, surgery, whether open or arthroscopic, is generally regarded as the nest available. Non-invasive contenders include ultrasound and magnetic resonance imaging (MRI).
Importance	Accurate diagnosis using readily applied, convenient, low-cost physical tests would enable appropriate and well-timed management of these common causes of shoulder pain.
Studies	Index were 33 studies including 4002 shoulders in 3852 patients. These incorporated numerous standard, modified or combinations of index tests and 14 novel index tests.
Quality concerns	Although 28 studies were prospective, study quality was generally poor. All but two studies failed to meet the criteria for having a representative spectrum of patients.
Data analysis	The studies tested 170 target condition/index test combinations, with only six instances of any index test being performed and interpreted similarly in two studies. Meta-analysis of the latter was inappropriate, however.

Target condition	Subcategory of target condition, if applicable	Studies	Shoulders/patients	Tests or variants evaluated
Subacromial and internal impingement	Subacromial impingement	5	361/356	13
	subacromial *versus* Internal impingement	1	110/110	1
	Internal impingement	0	0	0
Rotator cuff tendinopathy or tears	Non-specific disease of the 'rotator cuff'	5	466/466	17
	Specific diseases of the 'rotator cuff'	5	503/503	15
	Non-specific disease of the 'posterosuperior rotator cuff'	2	220/220	4
	Specific disease of the 'posterosuperior rotator cuff'	2	166/157	3
	Non-specific disease of supraspinatus	4	792/678	11
	Specific disease of supraspinatus	6	887/870	18
	Disease of infraspinatus	3	719/605	5
	Non-specific disease of subscapularis	5	887/773	10
	Specific disease of subscapularis	3	145/136	10
LHB tendinopathy or tears		3	660/557	10
Glenoid labral lesions	Non-specific labral lesions	4	364/364	5

	Non-specific SLAP lesions	3	222/221	15
	Type II-IV SLAP lesions	2	315/307	5
	Type II SLAP lesions	3	405/405	18
Multiple, undifferentiated target conditions	LHB/labral pathology; LHB/SLAP lesions; SA-SD bursitis/bursal-side degeneration of supraspinatus; and SIS/rotator cuff tendinitis or tear.	4	201/200	10

BACKGROUND

Target condition being diagnosed

Shoulder pain and dysfunction are common in the general population. A systematic review reported point prevalences for shoulder pain ranging from 7% to 26% with some indication that prevalence increases with age (Luime 2004a). Data from the US National Ambulatory Medical Care Survey (NAMCS) 1993 to 2000 indicate that one per cent of all office visits to physicians are for shoulder pain, and that a quarter of these visits are to primary care physicians (Wofford 2005). Moreover, shoulder pain has little tendency to resolve quickly or completely; according to a Dutch study, one half of all sufferers still report problems a year after their initial consultation (Van der Heijden 1997).

Shoulder pain and dysfunction may result from various aetiologies and pathologies. A common cause is impingement (pinching), which causes 'catching' or aching pain without appreciable joint stiffness, and which has a number of subtypes.

Impingement was originally characterised by Neer and Welsh (Neer 1977) as pinching of the soft-tissue structures between the humerus (upper arm bone) and the bone-and-ligament coraco-acromial arch of the scapula (shoulder blade) on movement. These structures include the contents of the so-called subacromial outlet: the 'rotator cuff' of muscles and tendons that surrounds the shoulder joint and the large lubricating sac (the subacromial bursa) that overlies it; and also the biceps tendon, which arches over the humerus, deep to the rotator cuff and within the shoulder joint itself. Neer 1977 proposed a continuum of impingement severity, from irritation of the bursa and cuff (normally due to overuse, and reversible by conservative management) to full thickness tears of the cuff. It has since been theorised that any abnormal reduction in the subacromial outlet's volume (e.g. by bone shape, soft-tissue thickening, posture or minor joint instability) may predispose to, contribute to, perpetuate or aggravate this train of events (discussed by Hanchard 2004).

It is increasingly recognised that other forms of impingement exist which, in distinction from subacromial outlet impingement, involve pinching of intra-articular (internal joint) structures at the extremes of movement. The socket's rim (the glenoid rim), its fibrocartilage extension (the glenoid labrum), and the deep surface of the rotator cuff are all at risk from this internal impingement mechanism, which may be subcategorised as anterosuperior or posterosuperior glenoid impingement (respectively affecting the front and back of the shoulder joint). It is unclear to what extent internal impingement is limited to athletes, and whether instability is a prerequisite (Jobe 1996).

Sometimes, primary partial thickness tears occur inside the substance of the rotator cuff, possibly due to internal shear stress (Fukuda 2003). Such tears also have the potential to cause impingement pain (Fukuda 2003; Uchiyama 2010).

Index test(s)

When a person presents with a history and symptoms suggestive of shoulder impingement, the clinician performs a series of physical (non-invasive) tests that aim to establish the diagnosis, and inform treatment and prognosis. Such tests may include the 'painful arc' test, intended to identify impingement in general terms (Cyriax 1982); tests to identify subacromial impingement (e.g. Neer 1977) or internal impingement (e.g. Meister 2004); tests to differentiate subacromial from internal impingement (Zaslav 2001); tests to diagnose rotator cuff involvement, including tears (e.g. Gerber 1991a; Gerber 1996; Hertel 1996a), or biceps tendon involvement (e.g. Yergason 1931); or tests to diagnose glenoid labrum tears (e.g. Kim 2001; Liu 1996b; O'Brien 1998a). These tests are described in Table 1, and include tests that were identified in studies included in this review. See Table 2 for explanations of terms used in Table 1 and elsewhere. Sometimes, local anaesthetic is injected into or around the subacromial bursa on the premise that negation of a previously positive (painful) physical test for subacromial outlet impingement will confirm and localise the diagnosis (Neer 1977). While not encompassing local anaesthesia *per se*, we will consider it for inclusion in this review when it is used in this special adjunctive mode. (Some studies of diagnostic accuracy may use local anaesthesia as a reference test rather than an index test, as considered below.)

The attraction of physical tests is that they can be used at any stage in the patient's care pathway and in any setting. They are non-invasive (apart from optional, adjunctive local anaesthesia), convenient, quick, and yield immediate results. Their aim of replicating pain or functional deficits lends them implicit relevance to patients' symptoms whereas, by contrast, lesions detected by imaging or at open surgery may actually be asymptomatic (Dinnes 2003; MacDonald 2000a; Milgrom 1995; Sher 1995). Furthermore, they involve no cost additional to that of a clinical consultation.

Physical tests involve clinical and interpretative skills, and results have been shown to differ with testers' expertise (Hanchard 2005). This has implications for the generalisation of results relating to test performance from individual studies. Given this, we will summarise data on variability in test results reported by the included studies, whether this is between individuals, across settings, or both.

Alternative test(s)

Other tests, usually conducted subsequently and in secondary care settings by specialists, include ultrasonography, arthrography, bursography, magnetic resonance imaging (MRI) and magnetic resonance arthrography (MRA). Those considered as potential reference standards for this review are described in Table 3. Some of these tests are invasive and none is completely valid (Dinnes 2003). Specifically, the generally accepted gold standard of diagnosis, direct observation at open or arthroscopic ('keyhole') surgery (Table 3), is not completely valid because tears within the substance of the rotator cuff are not directly visible (Fukuda 2003) and conversely, visible tears may be asymptomatic (Dinnes 2003; MacDonald 2000a; Milgrom 1995; Sher 1995). Surgery carries a risk of complications (Blumenthal 2003; Boardman 1999; Borgeat 2001), and is not applicable in the primary care setting where the majority of consultations and treatment prescriptions occur. Moreover, approximately 70% of patients with shoulder impingement respond to conservative treatment (Morrison 1997a) and so those having surgery cannot be considered representative (spectrum bias).

The reference tests are also affected by clinical and interpretation skills. Varying degrees of 'operator dependence' apply to the imaging techniques, among which ultrasonography is the most susceptible. Surgery is also operator dependent; evaluations using videotaped arthroscopies have demonstrated disappointing agreement between surgeons as to the presence, absence and

extent of pathology (Mohtadi 2004). As with the index tests (above), we will therefore summarise data reported by the included studies on the variability of the alternative reference tests.

Rationale

In a systematic review of interventions for shoulder pain, Green et al (Green 2003) observed that diverse and often conflicting diagnostic labelling hampered interpretation of the literature. Our review should help in this regard. In addition, timely diagnosis of impingement and the underlying structural deficits should enable rationalisation of patients' diagnostic pathways, as well as informing their management and prognosis.

At the inception of this review, we identified two relevant systematic reviews in this area. Dinnes et al (Dinnes 2003) reviewed diagnostic tests for shoulder pain due to soft tissue disorders, including cohort studies of physical tests, ultrasound, MRI or MRA in patients suspected of having soft tissue disorders (search date October 2001). Though they reported inclusion of 'clinical impingement syndrome', Dinnes et al's primary emphasis was on the detection of rotator cuff tears. Tests for disorders of the glenoid labrum were specifically excluded. Conversely, a systematic review by Luime et al (Luime 2004b) concentrated on clinical diagnostic studies of tests for glenoid labral tears and shoulder joint instability (reported search dates '2001' for CINAHL and EMBASE, and '2003' for MEDLINE). Our own review, as well as conducting an updated search for studies of clinical examination, extends the definition of shoulder impingement, as described above. The mutually distinct nature of tests for impingement and instability (despite the potential interrelationships between the two conditions) has enabled the review to focus on the former. Our review also differs from the others in placing emphasis on the primary care setting (while not excluding secondary or tertiary care) as most people with shoulder pain are diagnosed and managed in this setting (Broadhurst 2004). From the primary care perspective, patients studied at a later stage in the referral pathway or undergoing more than minimally invasive reference tests are not representative, and this issue of applicability is explicit in the quality assessment of included studies.

OBJECTIVES

To evaluate the diagnostic accuracy of physical tests, applied singly or in combination, for shoulder impingements (subacromial or internal) or local lesions of bursa, rotator cuff or labrum that may accompany impingement, in people whose symptoms and/or history suggest any of these disorders.

We also examined the physical tests according to whether they were intended to:

- identify impingement in general (or differentiate it from other causes of shoulder pain, e.g. 'frozen shoulder')
- subcategorise impingement as subacromial outlet impingement (impingement under the acromion process) or internal impingement (impingement within the shoulder joint)
- diagnose lesions of bursa, tendon or glenoid labrum that may be associated with impingement
- form part of a diagnostic package or process and, if so, according to the stages at which they may apply.

Investigation of sources of heterogeneity

We planned to investigate the following potential sources of heterogeneity.

- Study population: older general population; young athletic population; other well defined groups e.g. wheelchair users or swimmers (*see* the Differences between protocol and review)
- Stage of clinical care: primary (generally in the community setting), secondary (referral following preliminary screening) or tertiary (referral to a specialist centre)
- Study design: cross sectional (or cohort) versus case-control; retrospective versus prospective design
- Type of reference test. This will vary according to the target condition and setting, but generally surgery versus non-invasive imaging will be considered (*see*Table 3)
- Aspects of study conduct, specifically: blinding and reporting of uninterpretable or intermediate results.

METHODS

Criteria for considering studies for this review

Types of studies

We considered diagnostic test accuracy studies that directly compared the accuracy of one or more physical index tests for shoulder impingement against a reference test. We considered diagnostic test accuracy studies with cross-sectional or cohort designs (retrospective or prospective), case-control studies and randomised controlled trials. In particular, we noted whether the cases and controls in case-control studies were highly selected or acceptably representative of the patient population normally tested by the index test(s). We considered, but decided against, excluding cohort studies with an excessively long period between the index and reference test. We defined this as a period that, on average, equals or exceeds the reported mean duration of symptoms, or one month (whichever is shorter). We excluded studies that were reported only in abstract form.

Participants

Patients of any age and in any clinical setting with pain, dysfunction or both suspected to be due to shoulder impingement of any type (*see* Target conditions), whether subacromial, internal or secondary to rotator cuff disease, and with or without rotator cuff tears. Excluded were studies evaluating physical (index) tests under anaesthesia, or intra- or post-operatively. We also excluded studies that focused solely on pain due to acromioclavicular joint (ACJ) disorders; or that focused primarily on shoulder joint instability, fracture, acute or recurrent shoulder dislocation, or systemic disease (e.g. rheumatoid disease). Subsequent to the protocol we excluded studies with highly selected populations, such as overhead throwing athletes.

After evaluation of a patient's history, physical tests are normally the first stage in the diagnosis of shoulder impingement. However, the applicability of one physical test may be conditional upon the result of another (e.g. Zaslav 2001), and this was taken into account.

Index tests

Physical tests used singly or in combination to identify shoulder impingement, such as the painful arc test (Cyriax 1982); to classify

shoulder impingements, e.g. Neer's test (Neer 1977; Neer 1983), the modified relocation test (Hamner 2000), the internal rotation resistance strength test (Zaslav 2001); or to diagnose localised conditions that may accompany impingement, e.g. Yergason's test (Yergason 1931), the lift off test (Gerber 1991a; Gerber 1996; Hertel 1996a), the crank test (Liu 1996b), the active compression test (O'Brien 1998a) and the biceps load II test (Kim 2001) (*see* Table 1).

Ideally, articles for inclusion should have described a physical test, or reference a source that did so, in sufficient detail to enable its replication, and clearly indicate what constituted a positive index test result. Those that did not were included only if they provided sufficient information to be of clinical value. Studies reporting the collective diagnostic accuracy of a series of tests were considered, providing each component, and its manner of inclusion, were adequately described. Generic terms such as 'physical examination', as used to denote an unspecified combination of physical tests, led to exclusion unless further details were obtained from authors.

Target conditions

Subacromial or internal impingement of the shoulder and the localised conditions that may accompany these classifications, namely bursitis, rotator cuff tears, glenoid labrum tears, and inflammation or rupture of the biceps tendon.

Instability may underlie impingement, but tests of instability were only included if they were intended to demonstrate associated impingement pain, as in the modified relocation test (Hamner 2000), as opposed to instability *per se*. Similarly, tests for ACJ disorders were only included if, like the active compression test (O'Brien 1998a), they had a component intended to reproduce impingement pain.

Reference standards

In the absence of a definitive reference standard, surgery, whether open or arthroscopic, is generally regarded as the best available. We additionally considered ultrasound, which may be conducted in the primary care setting, and magnetic resonance imaging, magnetic resonance arthrography, subacromial local anaesthesia, arthrography and bursography, all of which may have more general applicability than surgery. These additional 'reference' tests are defined in Table 3. Their validity varies according to context, and are discussed case by case (*see* Table 3).

Search methods for identification of studies

Electronic searches

The search for studies was carried out in two stages (up to November 2005; 2005 to February 2010)

In the first stage, we searched MEDLINE (1966 to 14 November 2005), EMBASE (1974 to 14 November 2005), CINAHL (1982 to 14 November 2005) and AMED (Allied and Complementary Medicine Database) (1985 to 14 November 2005). We developed a sensitive search strategy (Appendix 1) as recommended in Chapter 5 and Appendix 5.4 of the Handbook (de Vet 2005). We also searched DARE (Database of Abstracts of Reviews of Effectiveness, *The Cochrane Library*) (1995 to 14 November 2005). While we recognise the potential association between language restriction and selection bias, pragmatic considerations required that the searches were restricted to articles written in the English language.

In the second stage, we searched MEDLINE, EMBASE and AMED (CINAHL had been removed to a separate search platform) from 2005 to 15 February 2010 (Appendix 2).

Searching other resources

We checked the reference lists of all relevant retrieved articles of primary diagnostic studies and systematic reviews.

Data collection and analysis

Selection of studies

Assisted by a pro-forma stating the review inclusion criteria, two review authors (NH and HH) independently screened the results of the electronic searches for the first batch (up to November 2005); and one review author (HH) screened the results of the second batch. Throughout, benefit of doubt was given for the assessment of study eligibility. After obtaining full text articles, two pairs of review authors (NH and HH; NH and ML) independently performed study selection. Disagreements were resolved by discussion between three review authors (NH, HH and ML).

Data extraction and management

We designed a review-specific data collection form (Whiting 2005a) and piloted it on three studies of diagnostic accuracy that focused on physical tests for shoulder instability (a condition outside the scope of the present review). Pairs of review authors (NH and HH; NH and ML) independently extracted all key study and participant information and data from the included studies without masking of trial authors and other identifying information. All disagreements were resolved by consensus.

We extracted the diagnostic 2 x 2 table data (number of true positives, false positives, false negatives, and true negatives) from the publications. If these were not available we attempted to reconstruct the 2 x 2 table(s) from summary estimates (Whiting 2005b). Studies presenting insufficient data for construction of 2 x 2 tables were excluded from the review.

We contacted authors mainly in regard to the availability of trial reports and more rarely identification of index tests and where there were minor and isolated discrepancies impeding the construction of 2 x 2 tables.

Discrepancies in 2 x 2 tables due to rounding errors were a common finding. A rule was devised whereby data were considered for inclusion only where the discrepancies in the back-calculated 2 x 2 table did not exceed 10% in any cells. Studies with multiple discrepant analyses were excluded. Where incorrectly reported summary statistics (borderline discrepancies in sensitivity or specificity not attributable to rounding error; or positive predictive value, negative predictive value or accuracy) were identified in included studies, this was highlighted as a cause for concern.

Assessment of methodological quality

At the same time as data collection, pairs of review authors (NH and HH; NH and ML) independently assessed study quality using all items of the QUADAS form (Whiting 2003), tailored to the review. Prior to the protocol, we had already undertaken a preliminary piloting exercise to establish a coding manual setting out review-

specific criteria (*see* Appendix 3). Disagreements were resolved by consensus.

Statistical analysis and data synthesis

For each index test, we plotted the observed sensitivities and specificities (with 95% confidence intervals) on forest plots for visual examination of variation in test accuracy across studies.

We planned to perform meta-analysis using hierarchical models if adequate data were available. However, due to the limited number of studies included for each test, meta-analysis was not possible and so descriptive analyses were undertaken.

Investigations of heterogeneity

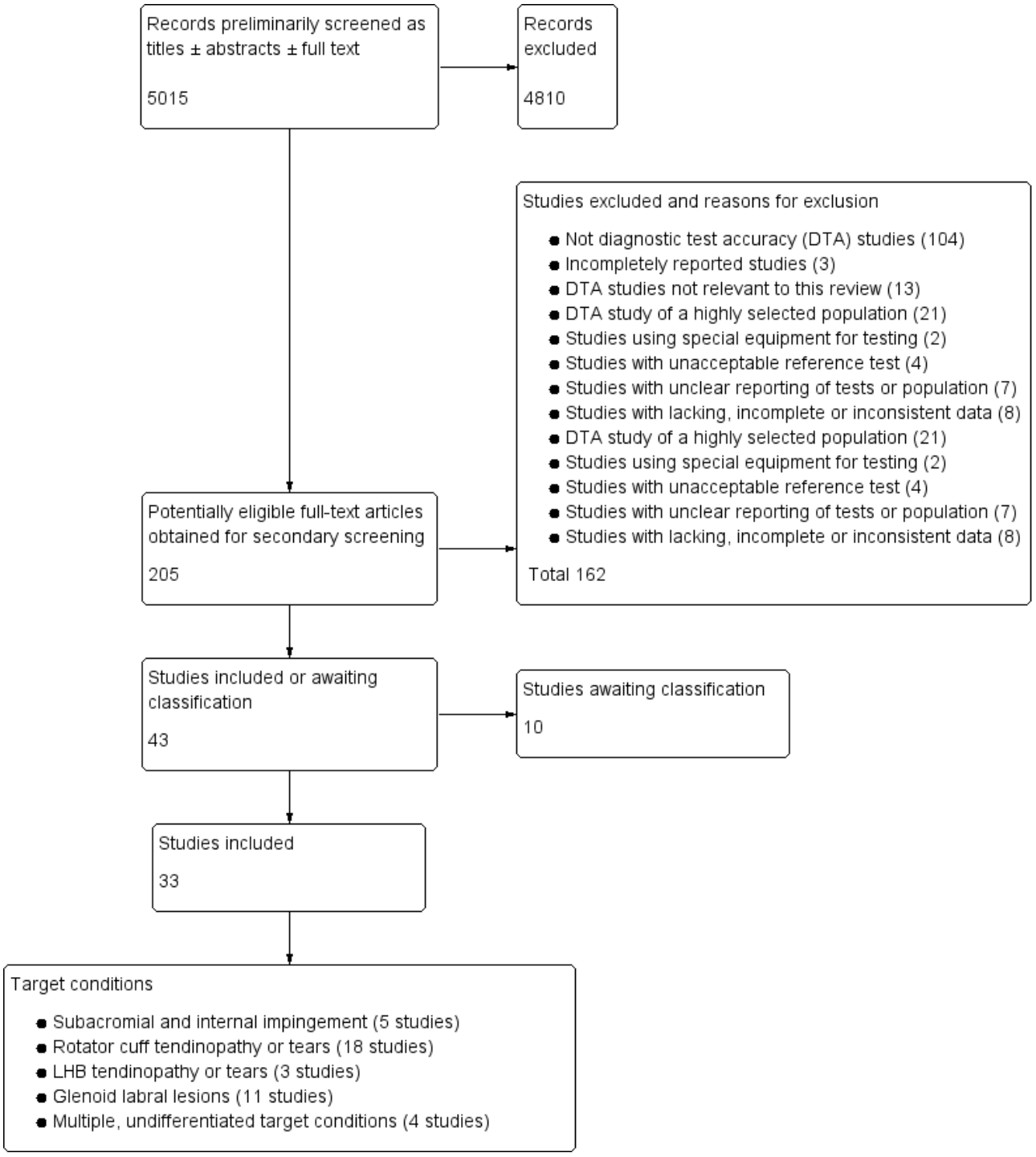
We planned to use meta-regression (by adding covariates to the hierarchical models) or subgroup analyses to explore the effect of potential sources of heterogeneity, such as the type of reference standard, on sensitivity and specificity. However, due to the limited number of studies available for each test, this was not possible.

RESULTS

Results of the search

We screened 3127 records from the first stage of the search and 1888 from the second stage (*see* Appendix 2). We obtained over 400 full text articles, some (numbers not fully documented) prompted by our scrutiny of references lists of reviews and primary studies. Of the 205 potentially eligible studies, 162 were excluded, 10 await classification, and 33 were included. The study flow diagram is shown in Figure 1.

Figure 1. Flow diagram.



Included studies

The Characteristics of included studies table gives details of the 33 studies, which evaluated a total of 4002 shoulders in 3852 patients. Apart from five studies (Castoldi 2009; Itoi 2006; Norwood 1989; Oh 2008 and Schlechter 2009), all were prospective. Fourteen (42%) were conducted in the USA. The remainder took place in Canada (Holtby 2004a; Holtby 2004b; MacDonald 2000; Razmjou 2004), South Korea (Kim 2001; Kim 2006; Kim 2007b; Oh 2008), Italy (Castoldi 2009; Gumina 2008; Iagnocco 2003), Denmark (Frost 1999; Suder 1994), Japan (Itoi 1999; Itoi 2006), Spain (Naredo 2002), Switzerland (Hertel 1996), Turkey (Calis 2000) and the UK (Miller 2008b).

Most of the studies were set in secondary or tertiary care, and only a few used reference standards that would be applicable to primary care, the intended focus of this review, or to the hospital outpatient setting. These were Calis 2000 (local anaesthesia, MRI); Iagnocco 2003, Miller 2008b, Naredo 2002 (ultrasonography); Frost

1999, Itoi 1999, Kim 2006 (MRI); and O'Brien 1998 (radiography and MRI, but also arthroscopic and open surgery, in various unspecified combinations). Apart from O'Brien 1998, previously mentioned, four studies (Castoldi 2009; Hertel 1996; Razmjou 2004; Speer 1994) used a mixture of arthroscopic and open surgery. The remainder, comprising 20 (61%) studies, used arthroscopic surgery alone.

Studies were grouped according to their target condition (*see* Table 4).

Subacromial impingement

Five studies (Calis 2000; Gumina 2008; Iagnocco 2003; MacDonald 2000; Naredo 2002) evaluated tests for subacromial impingement explicitly, or SA-SD bursitis, which we considered synonymous, on a total of 889 shoulders in 781 patients (*see* Table 5 for overview). One of these studies, Calis 2000, evaluated tests not only for subacromial bursitis but also, using dynamic ultrasonography as a reference standard, for subacromial impingement as an observable event in real time.

Internal impingement

No studies evaluated tests for internal impingement in isolation.

Subacromial versus internal impingement

One study (Zaslav 2001) of 110 shoulders in 110 patients evaluated a test to differentiate subacromial from internal impingement (*see* Table 5 for overview).

Rotator cuff tendinopathy or tears

Eighteen studies (Barth 2006; Calis 2000; Castoldi 2009; Frost 1999; Gumina 2008; Hertel 1996; Holtby 2004b; Iagnocco 2003; Itoi 1999; Itoi 2006; Kim 2006; MacDonald 2000; Miller 2008b; Naredo 2002; Norwood 1989; Speer 1994; Suder 1994; Wolf 2001) evaluated tests for rotator cuff tendinopathy or tears on a total of 2477 shoulders in 2337 patients (*see* Table 6 for overview).

LHB (long head of biceps) tendon tendinopathy or tears

Three studies (Iagnocco 2003; Kibler 2009; Naredo 2002) evaluated tests for LHB tendon tendinopathy or tears on a total of 660 shoulders in 557 patients (*see* Table 7 for overview).

Glenoid labral lesions

Eleven studies (Guanche 2003; Kibler 2009; Kim 2001; Kim 2007b; Liu 1996b; O'Brien 1998; Oh 2008; Parentis 2006; Schlechter 2009; Stetson 2002; Suder 1994) evaluated tests for glenoid labral lesions on a total of 1245 shoulders in 1236 patients (*see* Table 8 for overview).

Multiple, undifferentiated target conditions

Four studies evaluated tests for multiple, undifferentiated target conditions. These were Bennett 1998 (LHB/labral pathology; 46 shoulders in 45 patients), Holtby 2004a (LHB/SLAP lesions; 50 shoulders in 50 patients), Michener 2009 (SA-SD bursitis/bursal-side degeneration of supraspinatus; 55 shoulders in 55 patients) and Razmjou 2004 (subacromial impingement syndrome/rotator cuff tendinitis or tear; 50 shoulders in 50 patients) (*see* Table 9 for overview).

Excluded studies

The reasons for excluding, usually from inspection of the full text article, 162 studies are given in the Characteristics of excluded studies. Table 10 shows the trials grouped by their primary reason. The main and often listed as the sole reason for exclusion was that the study was not a diagnostic test accuracy study. Of the 104 studies for which this was the case, 11 were systematic reviews without reporting of results from an associated primary study. Three studies were not reported in full and it appears unlikely that this will ever be the case. The rest were confirmed diagnostic test accuracy studies. Of these, five were not of physical tests and eight did not study a target condition of this review. Twenty-one trials studied a highly selected population, either in terms of a high risk population (e.g. overhead throwing athletes as in Hamner 2000), previous injury (e.g. anterior shoulder dislocation), 100% prevalence of a condition by intent (e.g. all had SLAP lesions in Berg 1998) or a highly selected population by exclusion of key conditions (e.g. Liu 1996a). Two studies were excluded because special equipment (a hand held dynamometer) was used and four studies were excluded because the reference test was unacceptable (e.g. MRI was used as a reference standard for impingement in Silva 2008). In seven studies there was unclear reporting of physical tests, testing and/or the population. Lastly, eight studies were excluded because of the lack, incompleteness or gross inconsistency of reported data.

Studies awaiting classification

Ten studies await classification. The reasons are given in the Characteristics of studies awaiting classification. The reports for eight of these studies (Gill 2007; Jia 2008; Jia 2009; Kim 2003a; Kim 2003b; Kim 2004a; McFarland 2002; Park 2005), which apparently draw on the same clinical database, demonstrate substantial threats to validity. Verification is especially warranted in view of these studies' large patient numbers and potentially influential nature. The remaining two studies (Kelly 2010; Nanda 2008) presented insufficient data for adequate analysis.

Methodological quality of included studies

The methodological quality tables in the Characteristics of included studies give details of the results of the methodological quality assessment based on the 14 items of the QUADAS tool and using the coding manual set out in Appendix 3. In these tables, the results are expressed in terms of the methodological quality ('high', 'low' or 'unclear'). relating respectively to the review author's judgements ('yes', 'no', 'unclear'). Figure 2 summarises the judgements on each of the 14 methodological quality items for each included study.

Figure 2. Methodological quality summary: review authors' judgements about each methodological quality item for each included study. Coding: + = 'Yes'; - = 'No'; ? = unclear.

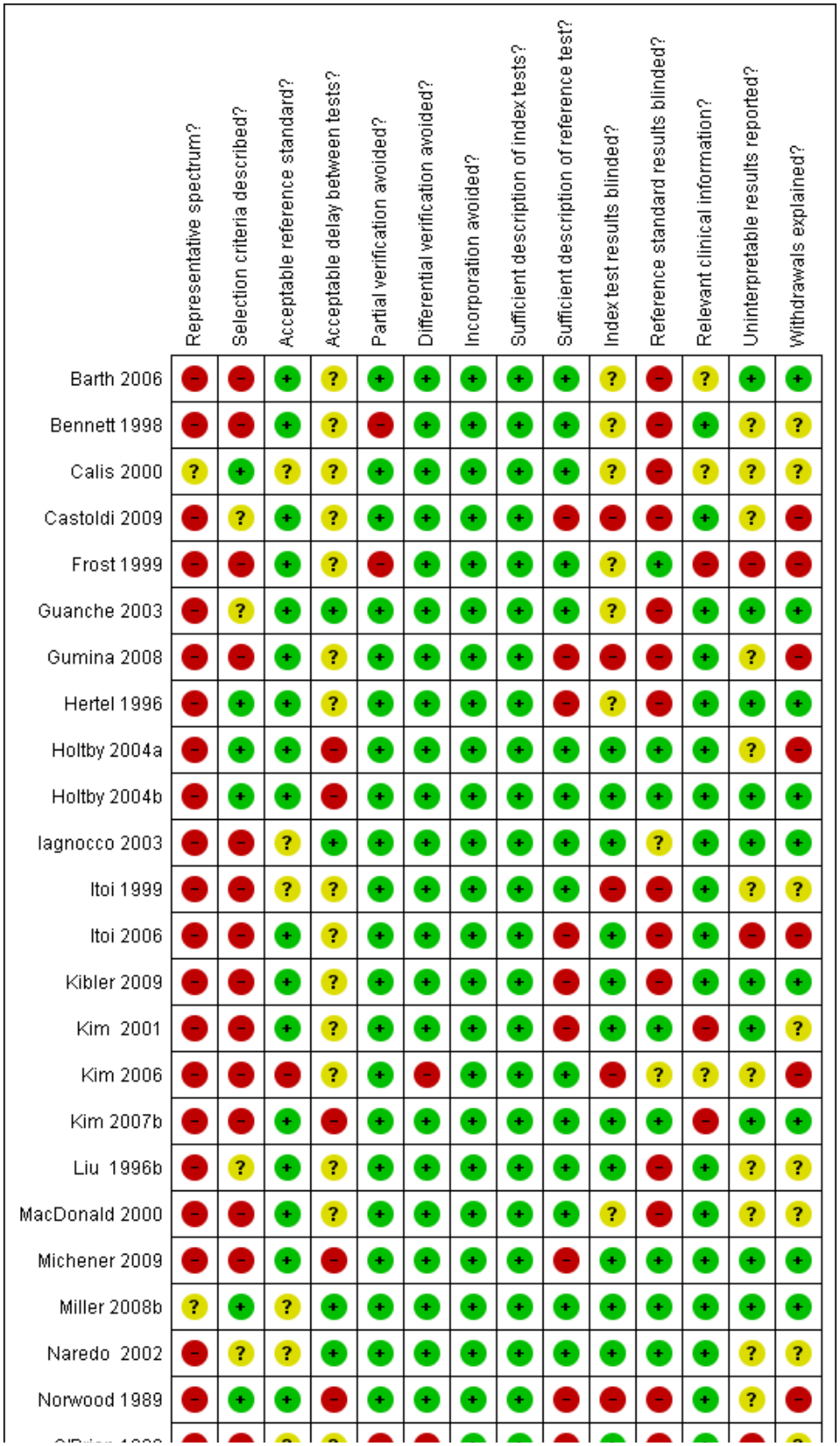


Figure 2. (Continued)

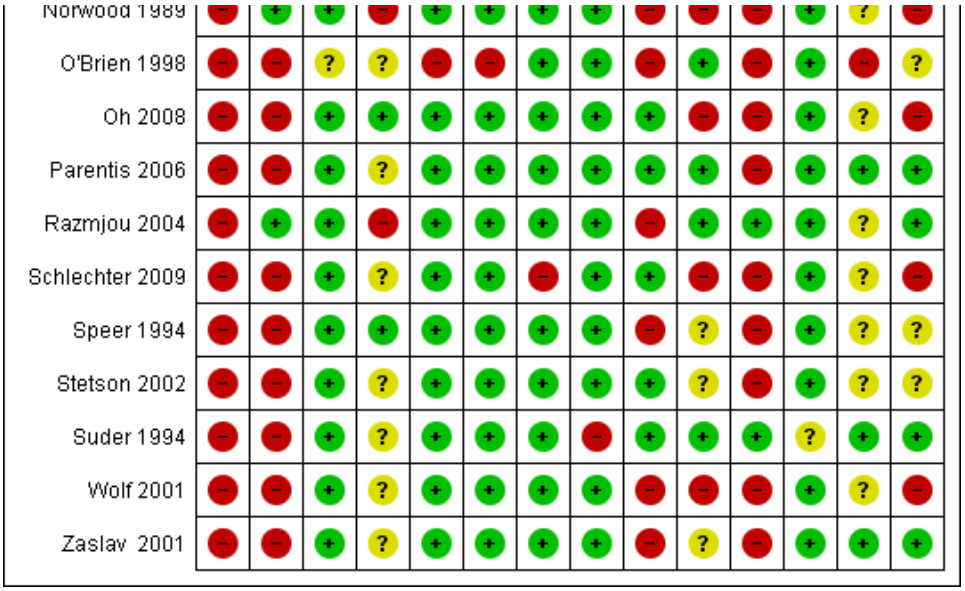
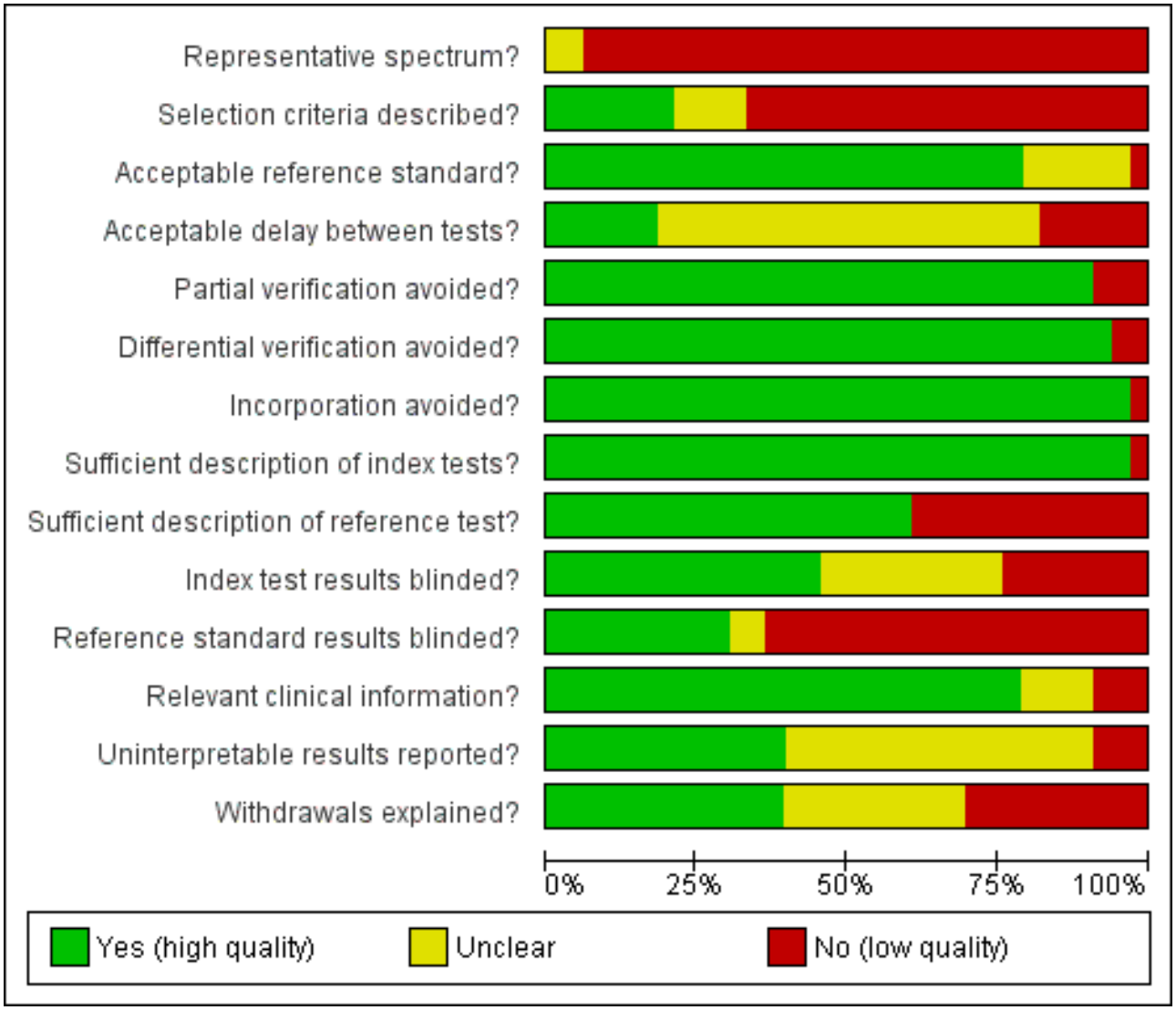


Figure 3 presents a graph showing the percentages of each of the three judgements for each quality item across all included studies. Some observations on these are given below.

Representative spectrum?: Mainly reflecting the use of surgery as a reference test, it is noteworthy that all but two studies (Calis 2000; Miller 2008b) were judged as not meeting the criteria for having a representative spectrum of patients. The reference tests in Calis 2000 were subacromial local anaesthetic injection and MRI, and ultrasound in Miller 2008b. However, the setting was mixed primary, secondary and tertiary in Calis 2000 and probably secondary in Miller 2008b and thus both were rated as having unclear risk of spectrum bias.

Selection criteria described?: The patient selection criteria were clearly described in seven studies and described but insufficiently clearly so in four others. The majority of studies (22 studies), however, gave a very limited description of their selection criteria.

Figure 3. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies.



Acceptable reference standard?: While 26 studies were considered to have an acceptable reference standard, the reference tests were sub-optimal (meriting an 'unclear' rating) in six studies (Calis 2000; Iagnocco 2003; Itoi 1999; Miller 2008b; Naredo 2002; O'Brien 1998). While not covered in our coding scheme for this item, Kim 2006 was given a 'No' rating for this item, which reflected the very unsatisfactory nature of the application of the two reference tests in this study.

Acceptable delay between tests?: Six studies met the criteria for an acceptable time period between the performance of the index and reference tests but in the majority (21) of studies there was insufficient information to judge this. The interval between the index and reference test was inappropriately long in six studies (Holtby 2004a; Holtby 2004b; Kim 2007b; Michener 2009; Norwood 1989; Razmjou 2004) putting them at risk of disease progression bias.

Partial verification avoided?: This was avoided in all studies except three studies (Bennett 1998; Frost 1999; O'Brien 1998) that were considered at high risk of partial verification bias.

Differential verification bias avoided?: Just two studies (Kim 2006; O'Brien 1998) were considered to be a high risk of differential verification bias.

Incorporation avoided?: Schlechter 2009 alone was considered to be a high risk of incorporation bias. In Schlechter 2009, the pathology found at surgery was "matched [to] the history, clinical presentation and symptoms".

Sufficient description of index tests?: All except Suder 1994 were judged as giving sufficient details to permit replication of the index test(s). Lack of sufficient details of the index tests generally resulted in exclusion of studies but we judged that the two tests, which were neither referenced nor described, in Suder 1994 were almost certainly the Neer's sign and Neer's test.

Sufficient description of reference test?: It was considered that sufficient details were given to replicate the reference test(s) in 20 studies but not in the other 13 studies, where often very limited or no detail was provided.

Index test results blinded?: There was a clear statement of blinding of the index tests in 15 studies and sufficient indication to merit an 'unclear' risk of test review bias in 10 prospective studies where the index test(s) clearly preceded the reference test. The other eight studies were judged to be an high risk of test review bias.

Reference tests blinded?: There was a clear statement of blinding of the reference test in 10 studies and sufficient indication that the reference test had been conducted independently in a further two studies (thus given an 'unclear' risk of bias rating). However, the lack of blinding in the remaining 21 studies meant that there was high risk of diagnostic review bias as foreknowledge of the index test result(s) may have influenced the interpretation of the reference test results.

Relevant clinical information?: Appropriate demographical and historical data were judged as being available when index tests were being interpreted in 26 studies; there being insufficient information to judge this in a further four studies. Such data were not available for interpreting the index tests in three studies (Frost 1999; Kim 2001; Kim 2007b).

Uninterpretable results reported?: Based on an assessment of the study design, recruitment and participant flow, 13 studies were judged to have fulfilled the criteria for this item and a further 17 studies may have done so but provided insufficient information to be certain. We judged that three studies (Frost 1999; Itoi 1999; O'Brien 1998) were at high risk of bias for this item.

Withdrawals explained?: Again, based on an assessment of the study design, recruitment and participant flow, 13 studies were judged to have had no withdrawals or to have accounted for these. In 10 studies there was insufficient information to be certain of this and in the remaining 10 studies, withdrawals were possible but either not reported or considered.

Findings

Overview of analyses and target condition/index test combinations

The complexity of the evidence is illustrated by the large number (170) of target condition/index test combinations. These were grouped by five main target conditions: subacromial or internal impingement; rotator cuff tendinopathy or tear; tendinopathy of the long head of biceps; glenoid labral lesions; and undifferentiated target conditions. The five main target conditions, which are also shown in Table 4, are exploded in Table 5 to Table 9. There were numerous standard, modified (see below) or combination index tests, and 14 novel index tests (tests being evaluated for the first time in the report in question, and originated by the authors of the report). The latter included the internal rotation resistance strength test for differentiating subacromial from internal impingement; active abduction, the drop sign, the external rotation lag sign, the Gum-Turn test and the internal rotation lag sign for rotator cuff tears; the upper cut test for LHB (long head of biceps) or labral lesions; and the active compression test, the biceps load II test, the crank test, modified dynamic labral shear, the passive compression test and the passive distraction test for labral lesions.

Subacromial and internal impingement (five studies)

The sensitivity and specificity estimates from each study for the tests of subacromial and internal impingement are shown in forest plots in Figure 4.

Figure 4. Subacromial and internal impingement

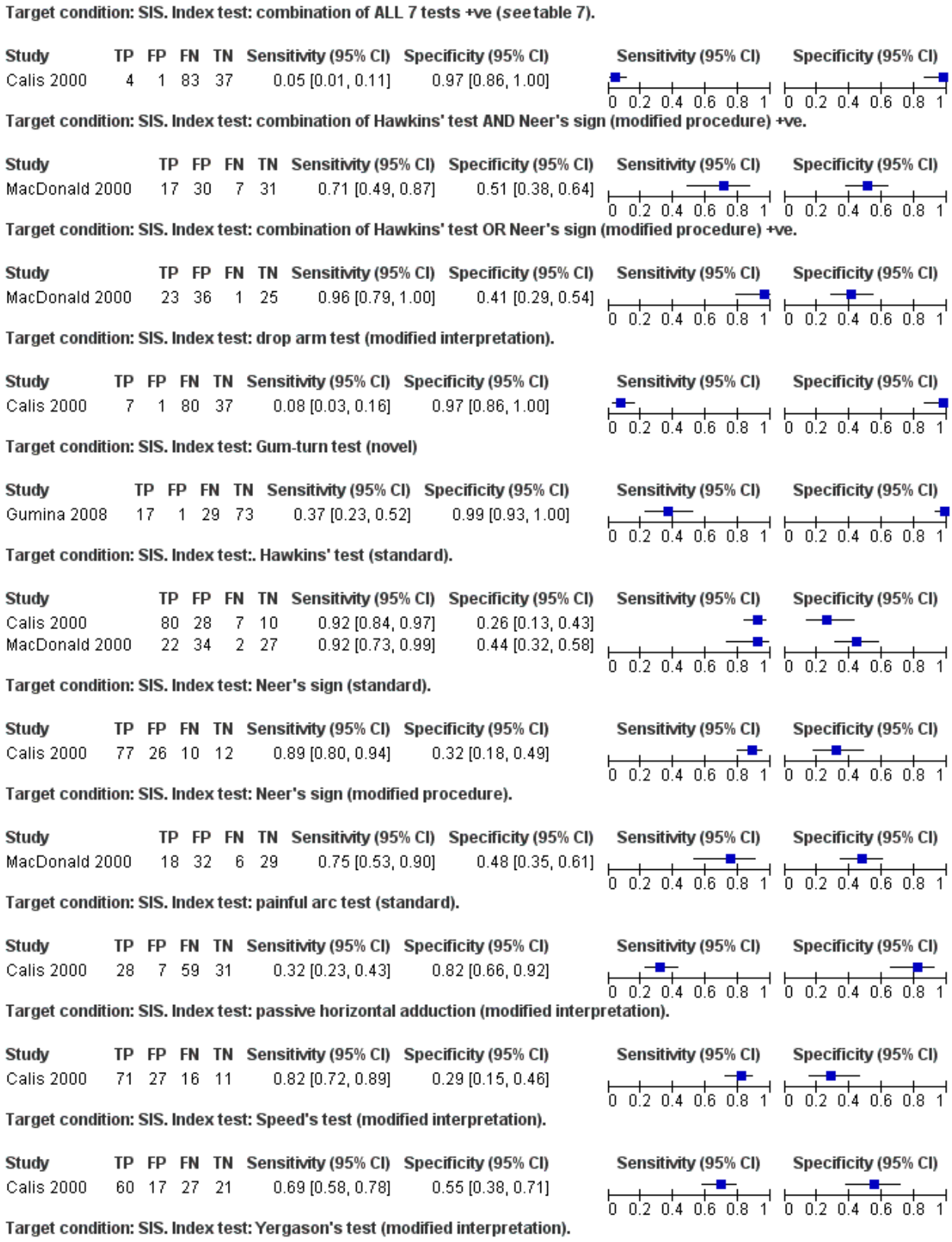
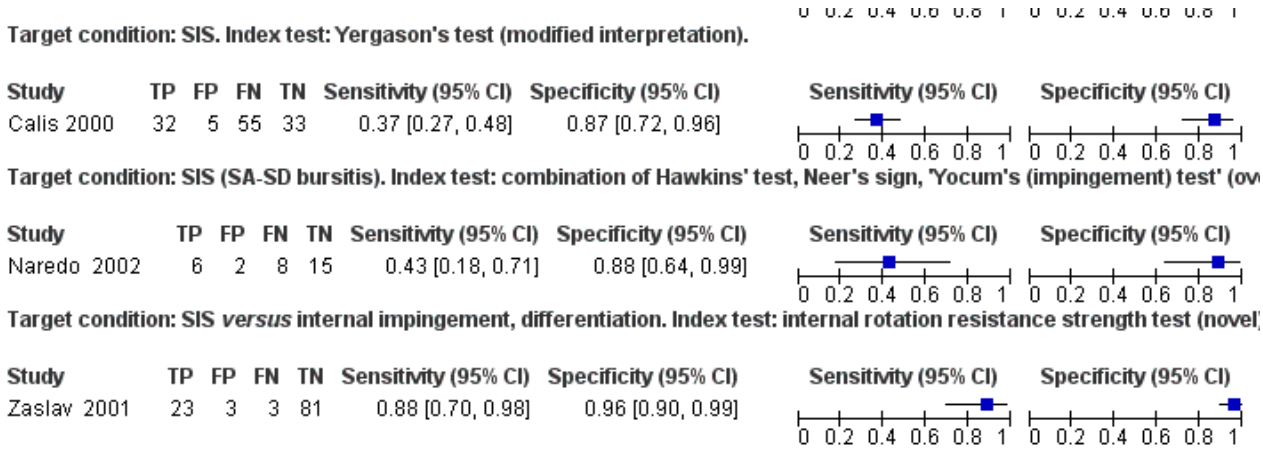


Figure 4. (Continued)



Four studies evaluated 13 standard, modified or combination tests for subacromial impingement. The standard tests were Hawkins' test, Neer's sign and the painful arc test. The modified tests were Neer's sign, passive horizontal adduction, Speed's test and Yergason's test. There were three combination tests, which comprised: all of seven specific tests (*see* Calis 2000 in Table 5); Hawkins' test or Neer's sign; and Hawkins' test and Neer's sign. The sensitivity estimates ranged from 5% (95% CI 1% to 11%) for the combination of seven tests to 96% (95% CI 79% to 100%) for the combination of Hawkins' test or Neer's sign. The specificity estimates ranged from 26% (95% CI 13% to 43%) for the standard Hawkins' test in Calis 2000 to 99% (95% CI 93% to 100%) for the Gum-Turn test. Only one test was performed and interpreted similarly in two studies. This was the standard Hawkins' test, but different and possibly incomparable reference standards were used (Calis 2000; MacDonald 2000).

One study evaluated the novel external rotation resistance strength test to differentiate subacromial from internal impingement and gave a sensitivity of 88% (95% CI 70% to 98%) and specificity of 96% (95% CI 90% to 99%).

No study evaluated any test for internal impingement.

Rotator cuff tendinopathy or tears (18 studies)

Non-specific disease of the 'rotator cuff' (five studies)

The sensitivity and specificity estimates from each study for the tests of non-specific disease of the 'rotator cuff' are shown in forest plots in Figure 5.

Figure 5. Rotator cuff tendinopathy or tears - non-specific disease of the 'rotator cuff'

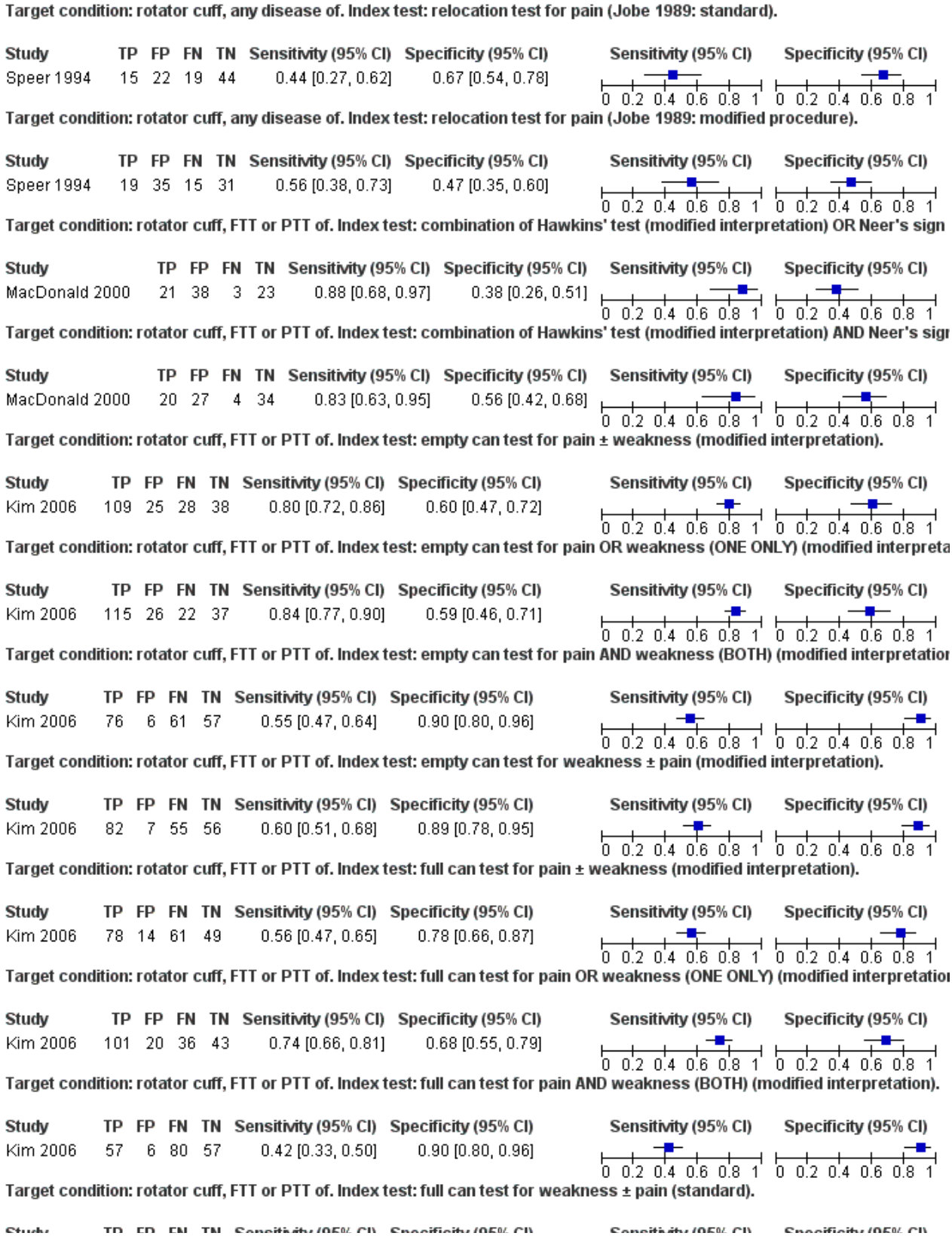
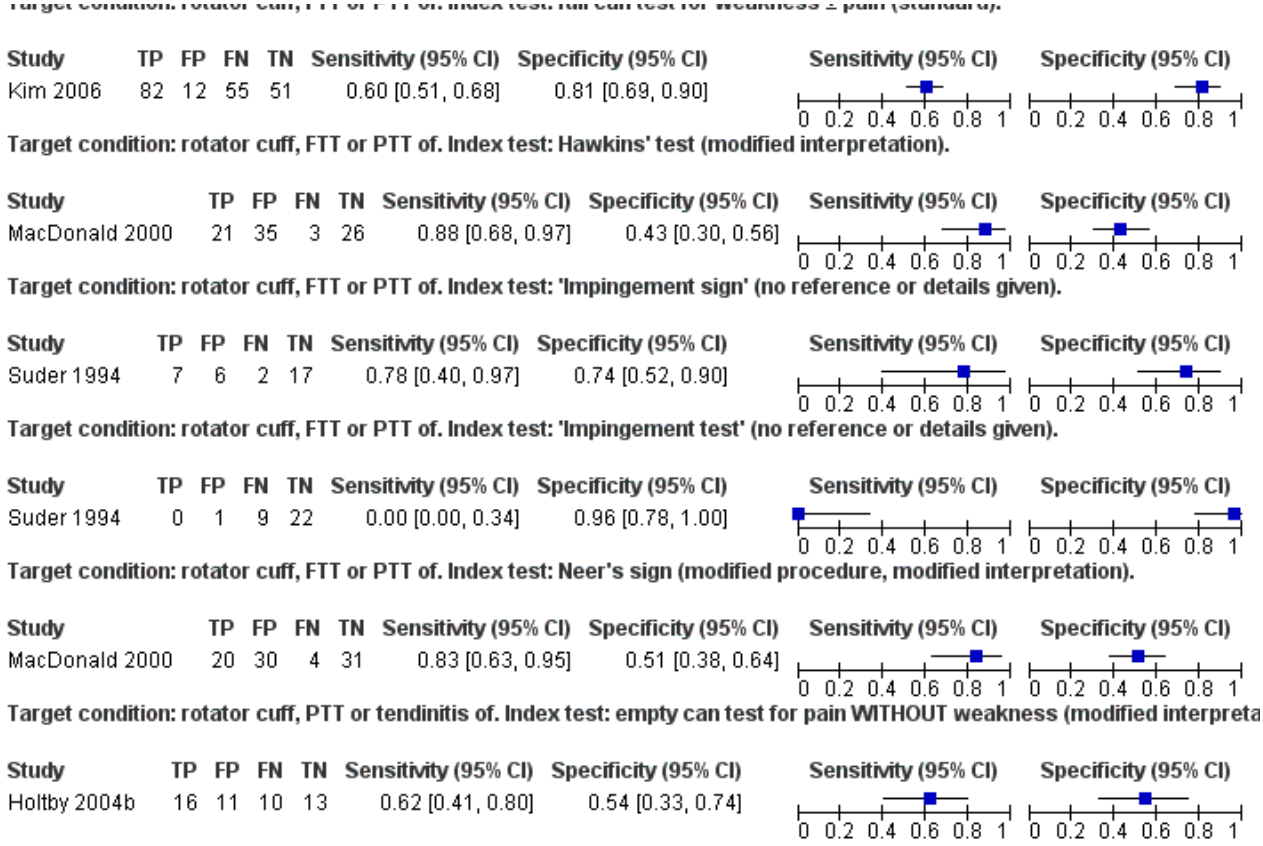


Figure 5. (Continued)



One study evaluated two tests for diseases of the 'rotator cuff', without attempting to discriminate between these diseases. The tests were a standard (Jobe 1989) and a modified relocation test for pain with sensitivities of 44% (95% CI 27% to 62%) and 56% (95% CI 38% to 73%), and specificities of 67% (95% CI 54% to 78%) and 47% (95% CI 35% to 60%) respectively.

Three studies evaluated 14 standard, modified or combination tests for tears of the 'rotator cuff' without attempting to discriminate between full thickness tears (FTT) and PTT. The same test was not performed and interpreted similarly in any two studies. The standard test was the full can test. The modified tests included four variants of the empty can test, three variants of the full can test, Hawkins' test and Neer's sign. There were two combination tests: Hawkins' test or Neer's sign; and Hawkins' test and Neer's sign. Two tests, an 'impingement sign' and an 'impingement test' were insufficiently defined to be categorised. The sensitivities ranged from 0% (95% CI 0% to 34%) for the undefined 'impingement test' to 88% for a modified Hawkins' test and the combination of Hawkins' test or Neer's sign (95% CI 68% to 97% in both instances). The specificities ranged from 38% (95% CI 26% to 51%) for the combination of Hawkins' test or Neer's sign to 96% (95% CI 78% to 100%) for the undefined 'impingement test'.

One study evaluated a modified empty can test for partial thickness tears (PTT) or tendinitis of the 'rotator cuff' without attempting to discriminate between these diseases and gave a sensitivity of 62% (95% CI 41% to 80%) and specificity of 54% (95% CI 33% to 74%).

Specific diseases of the 'rotator cuff' (five studies)

The sensitivity and specificity estimates from each study for the tests of specific diseases of the 'rotator cuff' are shown in forest plots in Figure 6.

Figure 6. Rotator cuff tendinopathy or tears - specific disease of the 'rotator cuff'

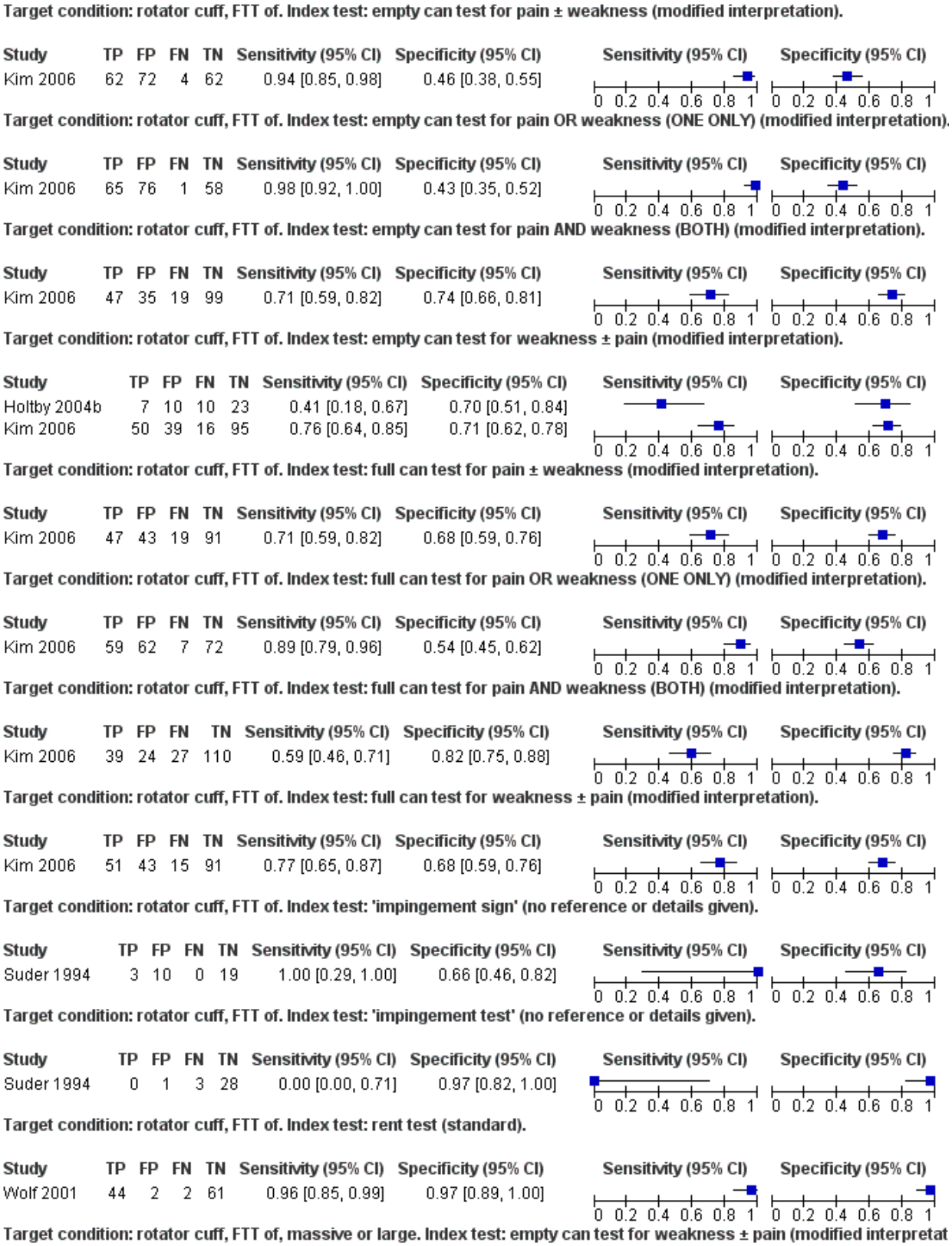
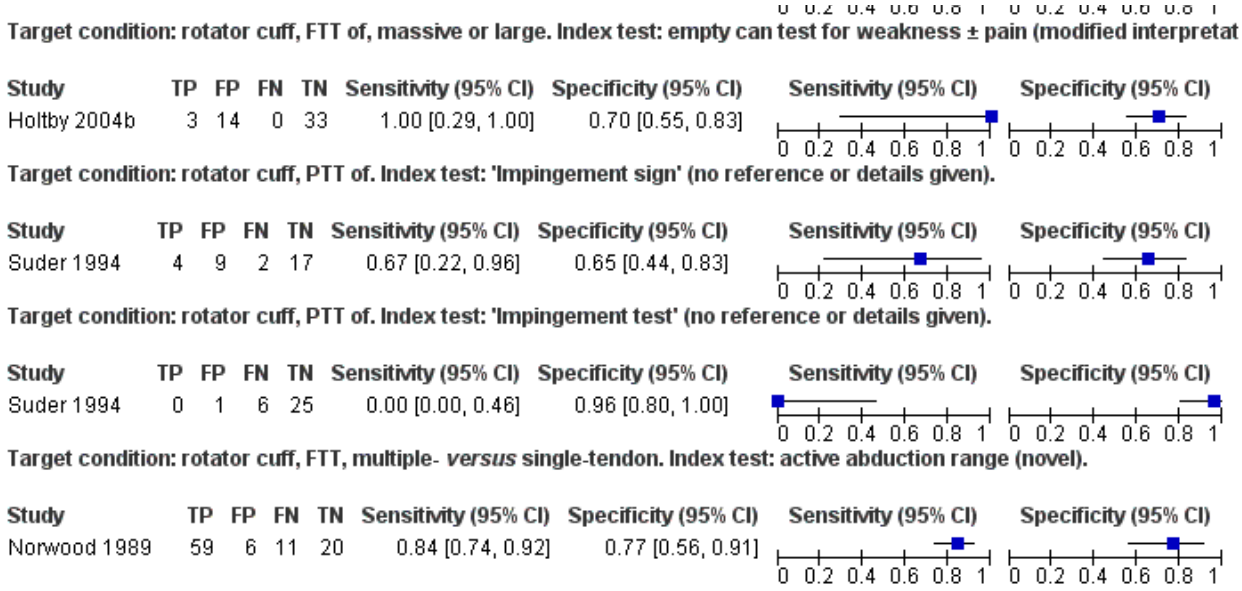


Figure 6. (Continued)



Four studies evaluated one or more of 11 standard or modified tests for FTT of the 'rotator cuff'. The standard test was the rent test. The modified tests were the empty can test (four variants) and the full can test (four variants). Two tests, an 'impingement sign' and an 'impingement test' were insufficiently defined to be categorised. The sensitivity estimates ranged from 0% (95% CI 0% to 71%) for the undefined 'impingement test' to 100% (95% CI 29% to 100%) for the undefined 'impingement sign'. The specificity estimates ranged from 43% (95% CI 35% to 52%) for a variant of the empty can test to 97% for an undefined 'impingement test' and the rent test (95% CI 82% to 100% and 89% to 100% respectively). There was one instance of a test being performed and interpreted similarly in two studies. This was a modified empty can test (Holtby 2004b; Kim 2006).

One study evaluated a modified empty can test for massive or large FTT of the 'rotator cuff' with a sensitivity of 100% (95% CI 29% to 100%) and a specificity of 70% (95% CI 55% to 83%).

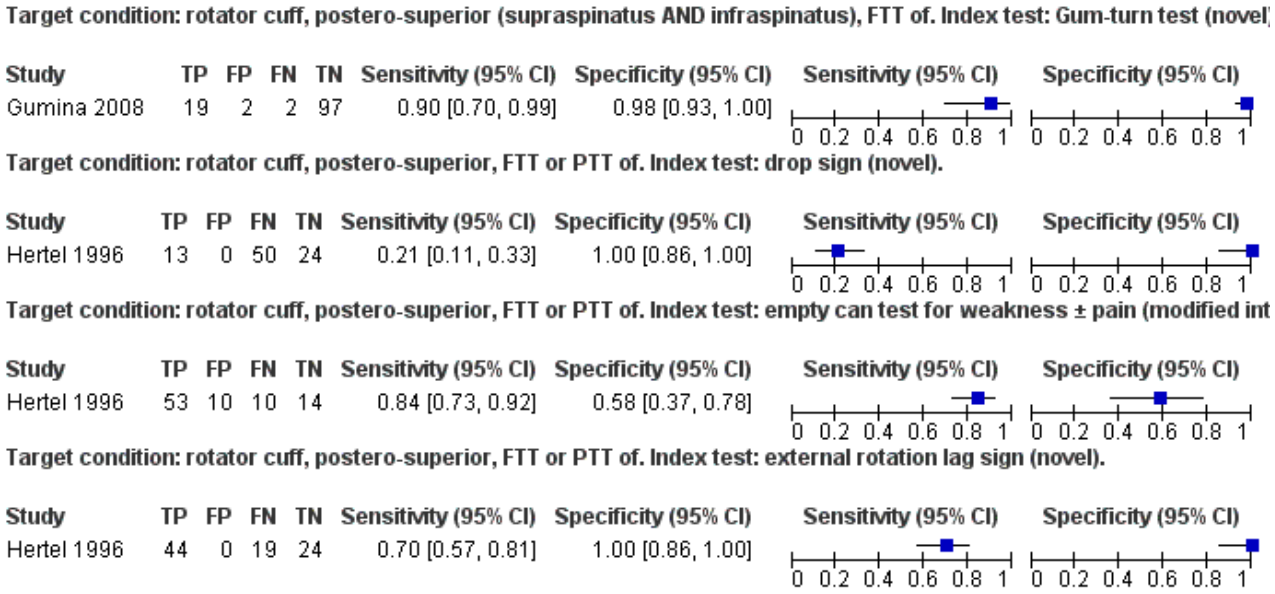
One study evaluated two tests for PTT of the 'rotator cuff'. These were an undefined 'impingement sign' and an undefined 'impingement test'. The sensitivity estimates were 67% (95% CI 22% to 96%) for the undefined 'impingement sign' and 0% (95% CI 0% to 36%). The specificity estimates were 65% (95% CI 44% to 83%) and 96% (95% CI 80% to 100%) respectively.

One study evaluated a novel active abduction range test to discriminate between single- and multiple-tendon FTT of the 'rotator cuff' with a sensitivity of 84% (95% CI 74% to 92%) and a specificity of 77% (95% CI 56% to 91%).

Non-specific disease of the 'posterosuperior rotator cuff' (two studies)

The sensitivity and specificity estimates from each study for the tests of non-specific disease of the 'posterosuperior rotator cuff' are shown in forest plots in Figure 7.

Figure 7. Rotator cuff tendinopathy or tears: non-specific disease of the 'posterosuperior rotator cuff'



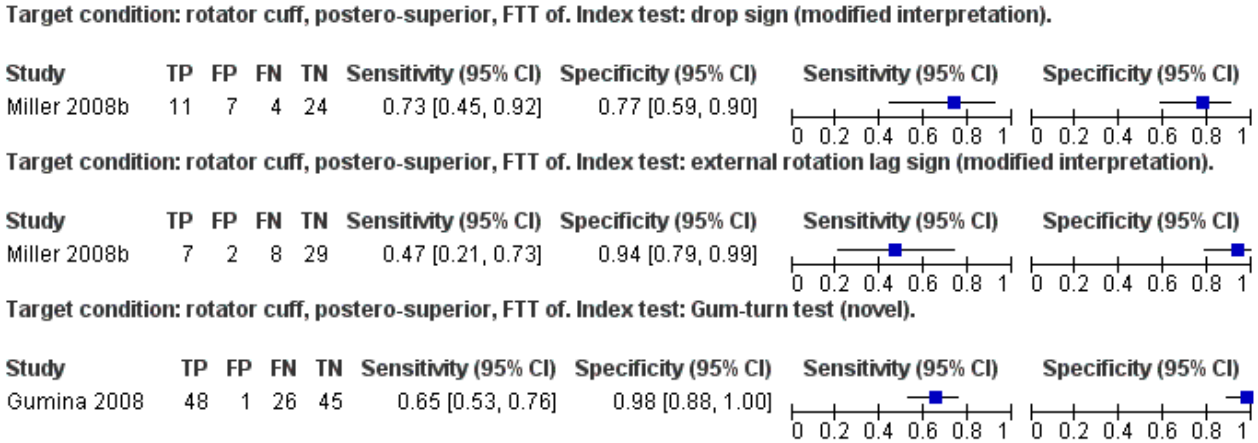
One study evaluated the novel Gum-Turn test for non-specific disease of the 'posterosuperior rotator cuff' (affecting supraspinatus AND infraspinatus) with a sensitivity of 90% (95% CI 70% to 99%) and a specificity of 98% (95% CI 93% to 100%).

One study evaluated three novel or modified tests for tears of the 'posterosuperior rotator cuff' without attempting to discriminate between FTT and PTT. The novel tests were the drop sign and the external rotation lag sign. The modified test was the empty can test. The sensitivity estimates ranged from 21% (95% CI 11% to 33%) for the novel drop sign to 84% (95% CI 73% to 92%) for the modified empty can test. The specificity estimates ranged from 58% (95% CI 37% to 78%) for the modified empty can test to 100% for the novel drop sign and novel external rotation lag sign (95% CI 86% to 100% in both instances).

Specific diseases of the 'posterosuperior rotator cuff' (two studies)

The sensitivity and specificity estimates from each study for the tests of specific diseases of the 'posterosuperior rotator cuff' are shown in forest plots in Figure 8 .

Figure 8. Rotator cuff tendinopathy or tears: specific disease of the 'posterosuperior rotator cuff'.

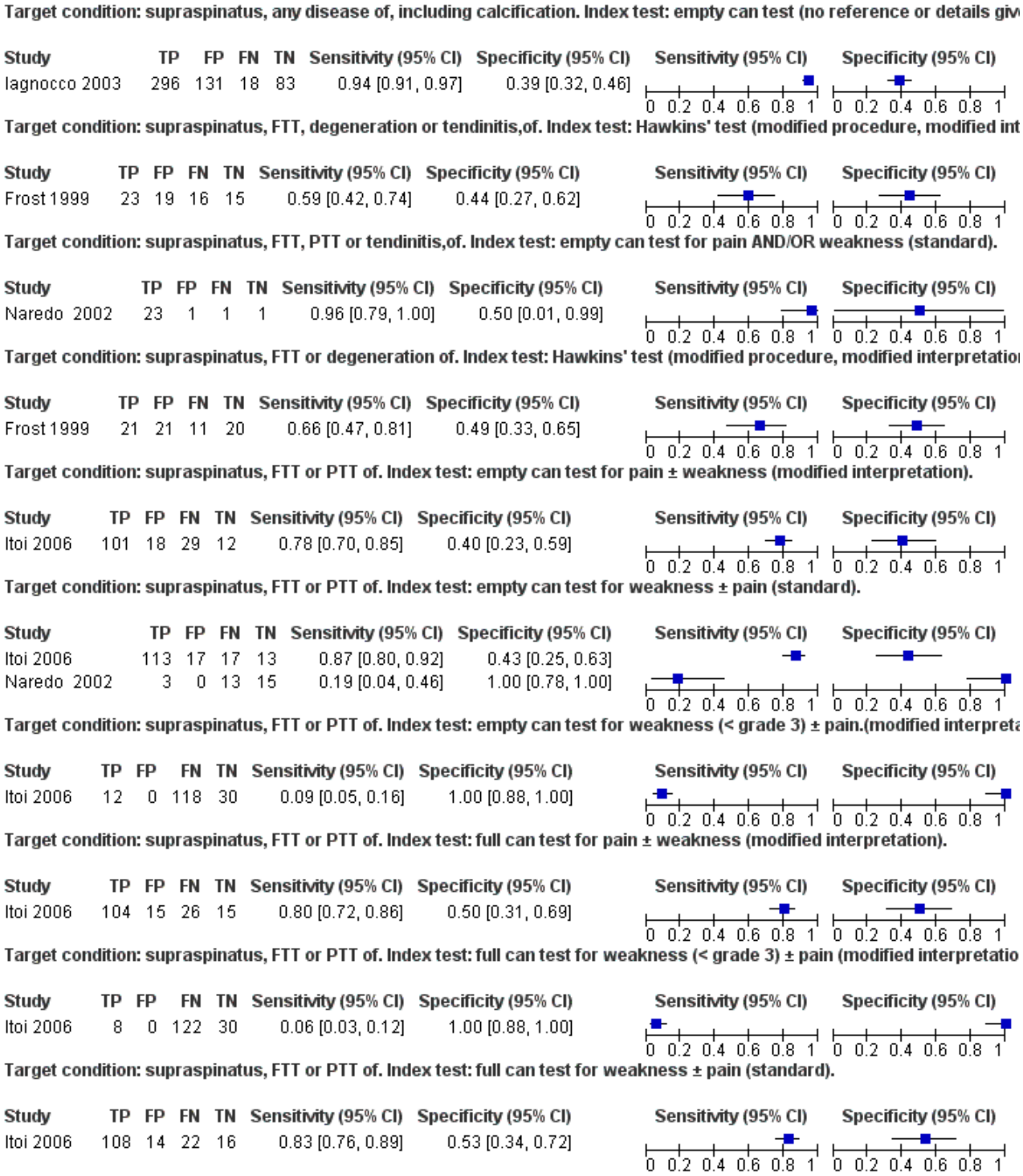


Two studies evaluated three novel or modified tests for FTT of the 'posterosuperior rotator cuff'. The novel test was the Gum-Turn test. The modified tests were the drop sign and the external rotation lag sign. The sensitivity estimates ranged from 47% (95% CI 21% to 73%) for the modified external rotation lag sign to 73% (95% CI 45% to 92%) for the modified drop sign. The specificity estimates ranged from 77% (95% CI 59% to 90%) for the modified drop sign to 98% (95% CI 88% to 100%) for the Gum-Turn test.

Non-specific disease of supraspinatus (four studies)

The sensitivity and specificity estimates from each study for the tests of non-specific disease of the supraspinatus are shown in forest plots in Figure 9.

Figure 9. Rotator cuff tendinopathy or tears: non-specific disease of supraspinatus



One study evaluated an undefined empty can test for diseases (calcification included) of supraspinatus without attempting to discriminate between these, with a sensitivity of 94% (95% CI 91% to 97%) and a specificity of 39% (95% CI 32% to 46%).

One study evaluated a modified Hawkins' test for FTT, degeneration or tendinitis of supraspinatus without attempting to discriminate between these, with a sensitivity of 59% (95% CI 42% to 74%) and a specificity of 44% (95% CI 27% to 62%).

One study evaluated the standard empty can test for FTT, PTT or tendinitis of supraspinatus without attempting to discriminate between these, with a sensitivity of 96% (95% CI 79% to 100%) and a specificity of 50% (95% CI 1% to 99%).

One study evaluated a modified Hawkins' test for FTT or degeneration of supraspinatus without attempting to discriminate between these, with a sensitivity of 66% (95% CI 47% to 81%) and a specificity estimate of 49% (95% CI 33% to 65%).

Two studies evaluated six standard or modified tests for FTT or PTT of supraspinatus, without attempting to discriminate between these diseases. The standard tests were the empty can test and the full can test. The modified tests were the empty can test (two variants) and the full can test (two variants). The estimates of sensitivity ranged from 6% (95% CI 3% to 12%) for a modified full can test to 87% (95% CI 80% to 92%) for a standard empty can test, and the specificity estimates ranged from 40% (95% CI 23% to 59%) for a modified empty can test to 100% for a standard empty can test, a modified empty can test, and a modified full can test (95% CI 78% to 100%; 88% to 100% and 88% to 100% respectively). One test was performed and interpreted similarly in both studies. This was the standard empty can test (Itoi 2006; Naredo 2002).

Specific diseases of supraspinatus (six studies)

The sensitivity and specificity estimates from each study for the tests of specific diseases of the supraspinatus are shown in forest plots in Figure 10.

Figure 10. Rotator cuff tendinopathy or tears: specific disease of supraspinatus.

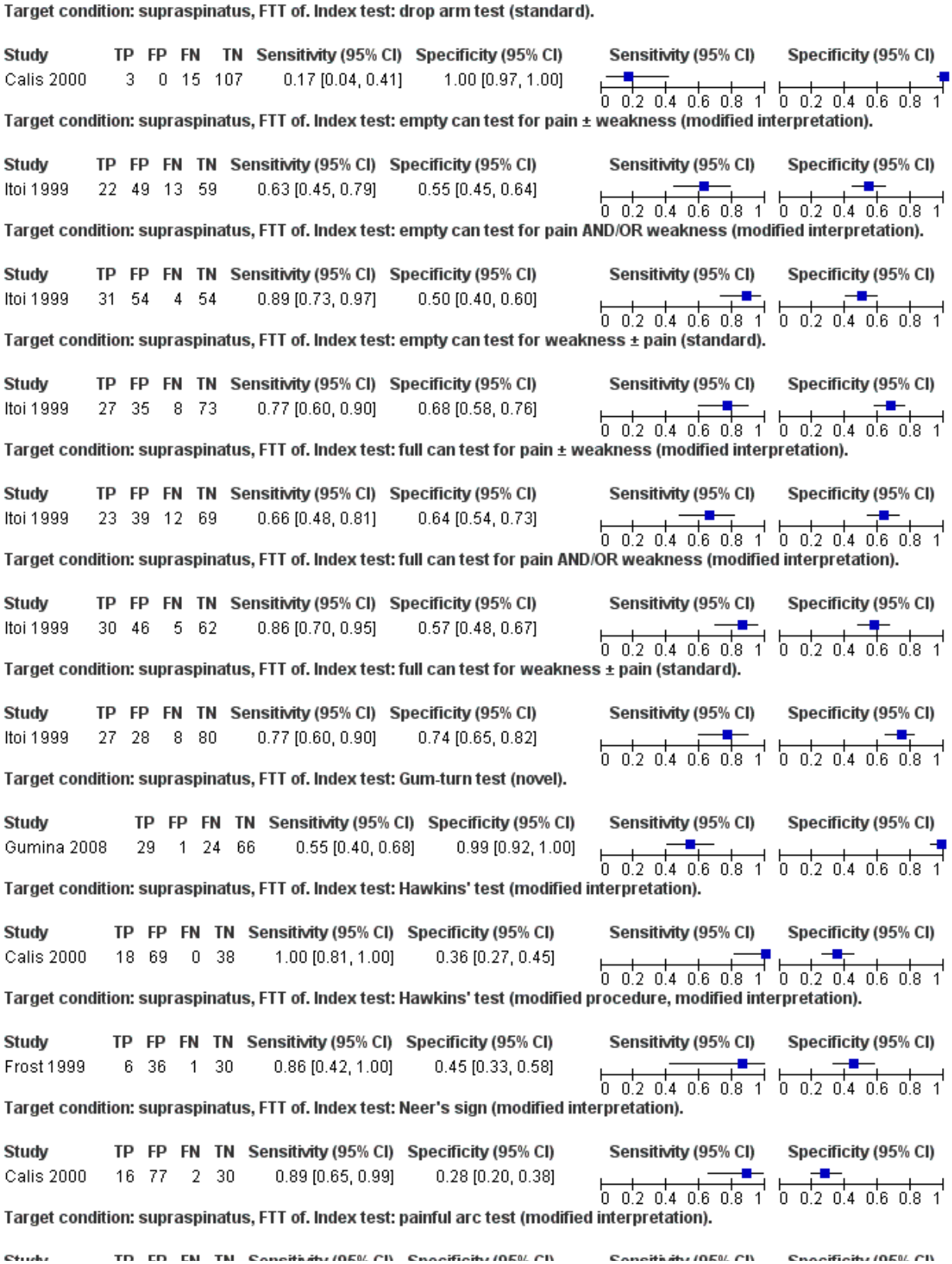
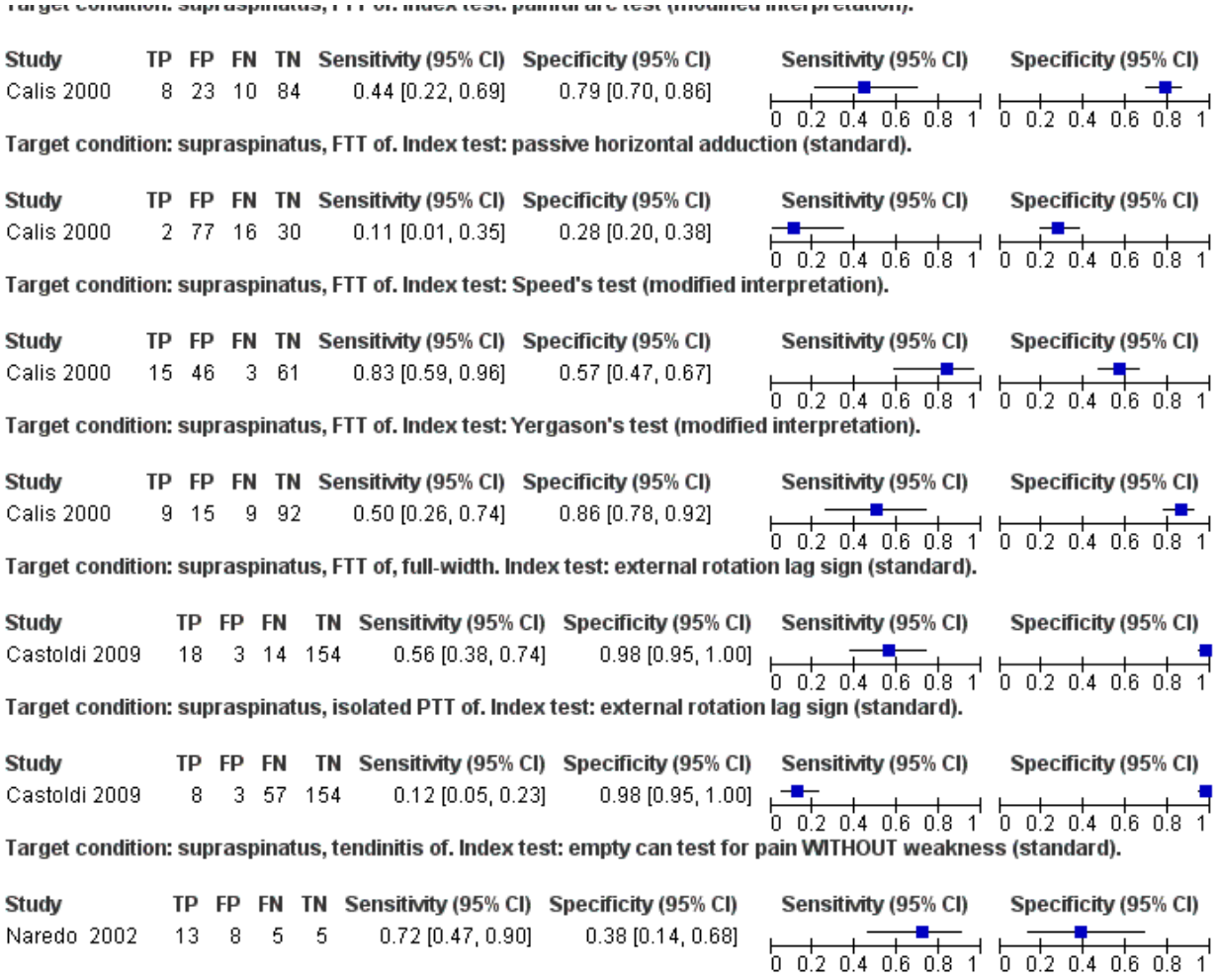


Figure 10. (Continued)



Four studies evaluated 15 novel, standard or modified tests for FTT of supraspinatus. There were no instances of the same test being performed and interpreted similarly in two or more studies. The novel test was the Gum-Turn test. The standard tests were the drop arm test, the empty can test and the full can test. The modified tests were the empty can test, the full can test and Hawkins' test (two variants each), and Neer's sign, the painful arc test, passive horizontal adduction, Speed's test and Yergason's test (one variant each). The sensitivity estimates ranged from 11% (95% CI 1% to 35%) for modified passive horizontal adduction to 100% (81% to 100%) for a modified Hawkins' test. The specificity estimates ranged from 28% (95% CI: 20% to 38%) for both the modified passive horizontal adduction and the modified Neer's sign to 100% (95% CI 97% to 100%) for the standard drop arm test.

One study evaluated the standard external rotation lag sign for full-width, FTT of supraspinatus, with a sensitivity of 56% (95% CI 38% to 74%) and a specificity of 98% (95% to 100%).

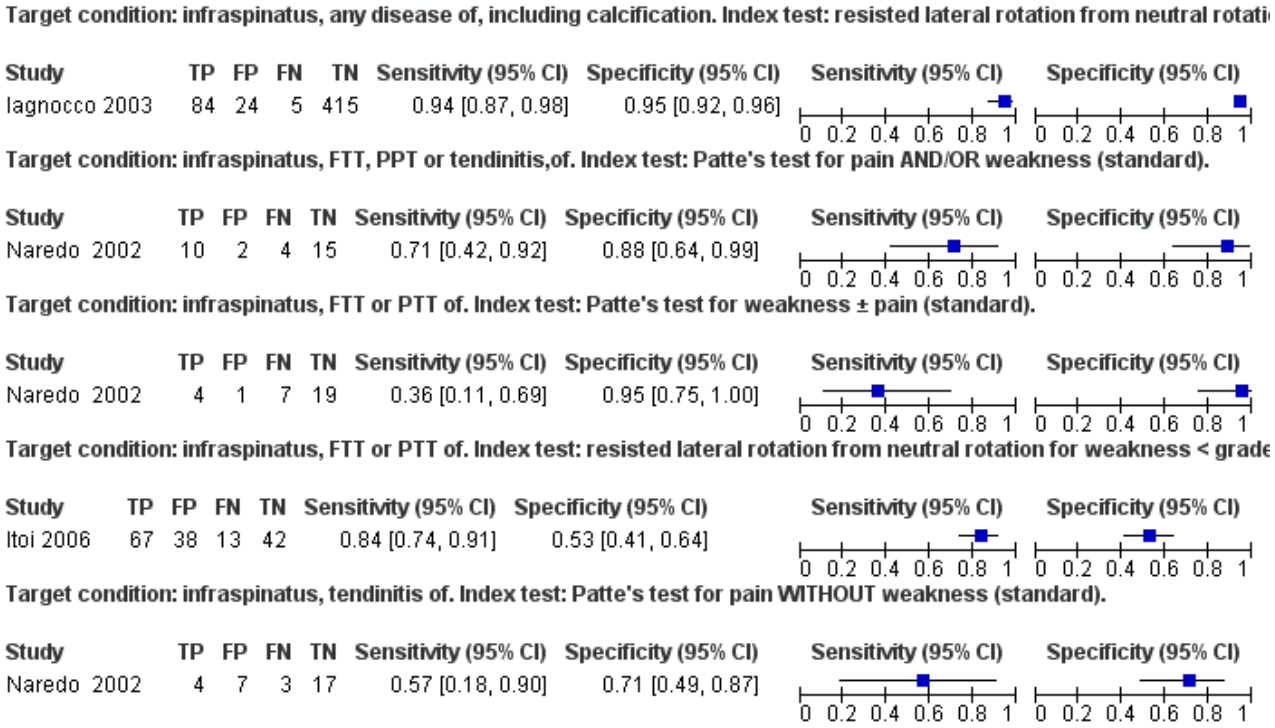
One study evaluated the standard external rotation lag sign for isolated PTT of supraspinatus, with a sensitivity of 12% (95% CI 5% to 23%) and a specificity of 98% (95% CI 95% to 100%).

One study evaluated the standard empty can test for tendinitis of supraspinatus, with a sensitivity of 72% (95% CI 47% to 90%) and a specificity of 38% (95% CI 14% to 68%).

Disease of infraspinatus (three studies)

The sensitivity and specificity estimates from each study for the tests of disease of infraspinatus are shown in forest plots in Figure 11.

Figure 11. Rotator cuff tendinopathy or tears: disease of infraspinatus



One study evaluated undefined resisted lateral rotation from neutral rotation for diseases of infraspinatus (calcification included) without attempting to discriminate between these, with a sensitivity of 94% (95% CI 87% to 98%) and a specificity of 95% (95% CI 92% to 96%).

One study evaluated the standard Patte's test for identifying and discriminating between tears and tendinitis of infraspinatus, with a sensitivity of 71% (95% CI 42% to 92%) and a specificity of 88% (95% CI 64% to 99%).

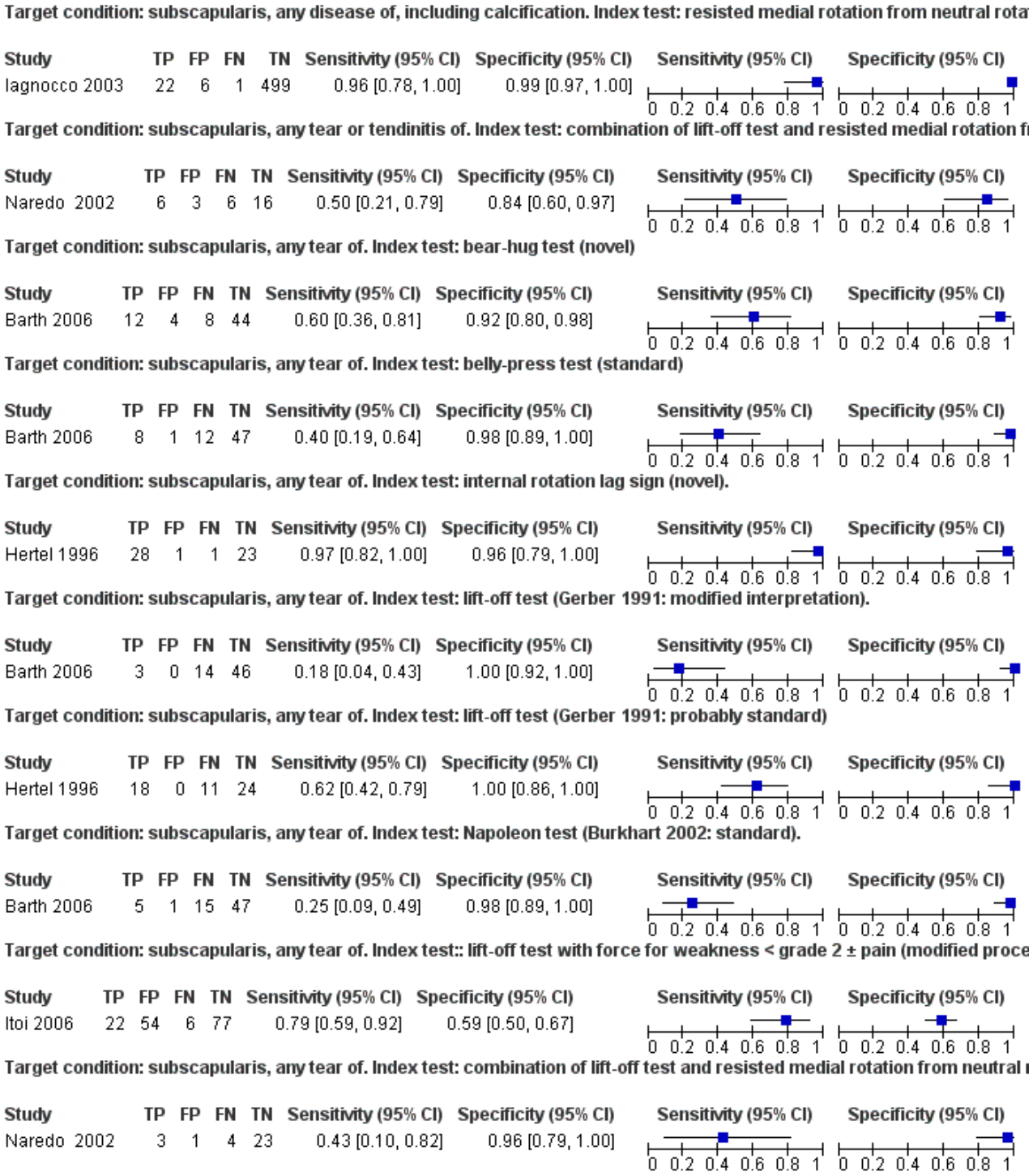
One study evaluated the standard Patte's test and another study evaluated modified resisted lateral rotation from neutral rotation for tears of infraspinatus. In neither case was there an attempt to differentiate between FTT and PTT. The sensitivity estimates were 36% (95% CI 11% to 69%) for the standard Patte's test and 84% (95% CI 74% to 91%) for resisted lateral rotation; and the specificity estimates were 95% (95% CI 75% to 100%) and 53% (95% CI 41 to 64%) respectively.

One study evaluated the standard Patte's test for infraspinatus tendinitis with a sensitivity of 57% (95% CI 18% to 90%) and a specificity of 71% (95% CI 49% to 87%).

Non-specific disease of subscapularis (five studies)

The sensitivity and specificity estimates from each study for the tests of non-specific disease of subscapularis are shown in forest plots in Figure 12.

Figure 12. Rotator cuff tendinopathy or tears: non-specific disease of subscapularis



One study evaluated undefined resisted medial rotation from neutral rotation for diseases of subscapularis (calcification included) without attempting to discriminate between these, with a sensitivity of 96% (95% CI 78% to 100%) and a specificity of 99% (95% CI 97% to 100%).

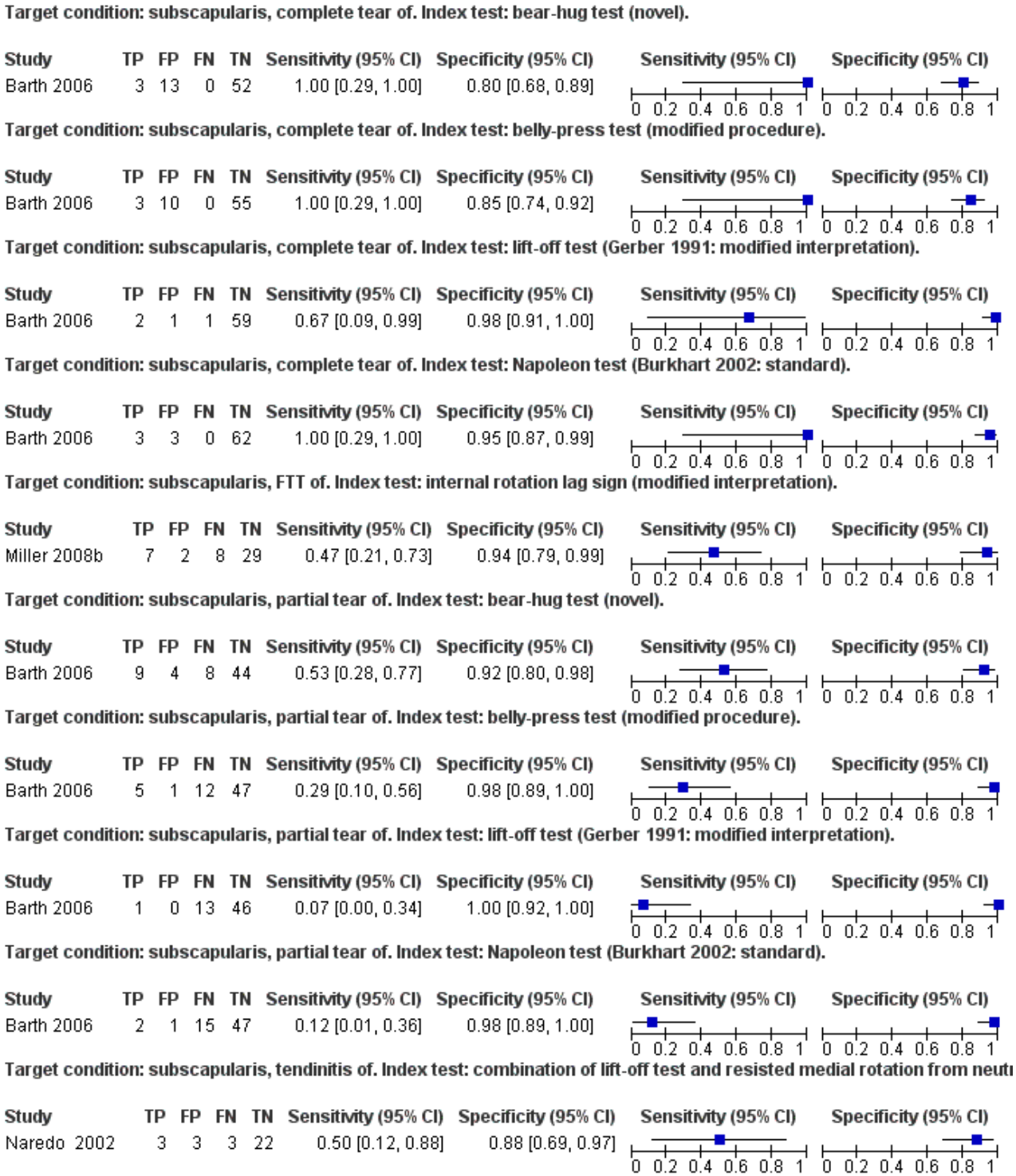
One study evaluated an incompletely defined combination of the lift-off test and resisted medial rotation from neutral rotation for any tear or tendinitis of subscapularis without attempting to differentiate between these diseases, with a sensitivity of 50% (95% CI 21% to 79%) and a specificity of 84% (95% CI 60% to 79%).

Four studies evaluated eight novel, standard, modified or combination tests for tears of subscapularis, without attempting to discriminate between types of tears. There were no instances of a test being performed and interpreted similarly in two or more studies. The novel tests were the bear-hug test and the internal rotation lag sign. The standard tests were the belly-press test, the lift-off test and the Napoleon test. The modified tests were the lift-off test and the lift-off with force. The combination test, which was incompletely defined, comprised the lift-off test and resisted medial rotation from neutral rotation. The sensitivity estimates ranged from 18% (95% CI 4% to 43%) for the modified lift-off test to 97% (95% CI 82% to 100%) for the internal rotation lag sign. The specificity estimates ranged from 59% (95% CI 50% to 67%) for the modified lift-off test with force to 100% for the standard lift-off test and the modified lift-off test (95% CI 86% to 100% and 92% to 100% respectively).

Specific diseases of subscapularis (three studies)

The sensitivity and specificity estimates from each study for the tests of specific diseases of subscapularis are shown in forest plots in Figure 13.

Figure 13. Rotator cuff tendinopathy or tears: specific disease of subscapularis



One study evaluated four novel, standard or modified tests for a complete tear of subscapularis. The novel test was the bear-hug test. The standard test was the Napoleon test. The modified tests were the belly-press test and the lift-off test. The sensitivity estimates ranged from 67% (95% CI: 9% to 99%) for the modified lift-off test to 100% for the novel bear-hug test, the standard Napoleon test and the modified belly-press test (95% CI 29 to 100% in each instance). The specificity estimates ranged from 80% (95%

CI 68% to 89%) for the novel bear-hug test to 98% (95% CI 91% to 100%) for the modified lift-off test.

One study evaluated a modified internal rotation lag sign for FTT of subscapularis, with a sensitivity of 47% (95% CI 21% to 73%) and specificity of 94% (95% CI 79% to 99%).

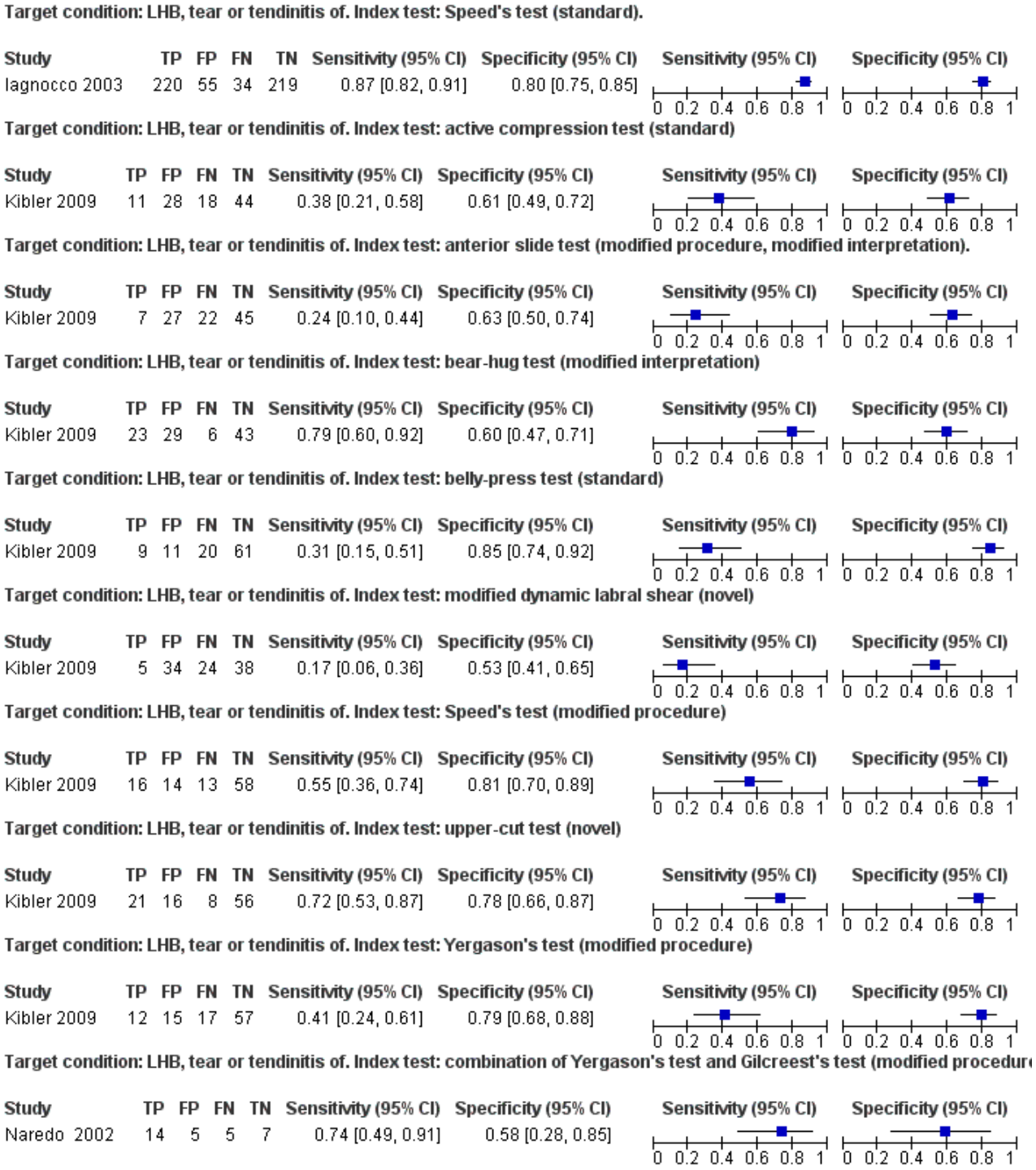
One study evaluated four novel, standard or modified tests for a partial tear of subscapularis. The novel test was the bear-hug test. The standard test was the Napoleon test. The modified tests were the belly-press test and the lift-off test. The sensitivity estimates ranged from 7% (95% CI 0% to 34%) for the modified lift-off test to 53% (95% CI 28% to 77%) for the novel bear-hug test, and the specificity estimates ranged from 92% (95% CI 80% to 98%) for the novel bear-hug test to 100% (95% CI 92% to 100%) for the modified lift-off test.

One study evaluated an incompletely defined combination comprising the lift-off test and resisted medial rotation from neutral rotation for subscapularis tendinitis, with a sensitivity of 50% (95% CI 12% to 88%) and a specificity of 88% (95% CI 69% to 97%).

LHB tendinopathy or tears (three studies)

The sensitivity and specificity estimates from each study for the tests of LHB tendinopathy or tears are shown in forest plots in Figure 14.

Figure 14. LHB tear or tendinitis



Three studies evaluated 10 novel, standard, modified or combination tests for LHB tears or tendinosis without attempting to differentiate between these diseases. There were no instances of a test being performed and interpreted similarly in two or more studies. The novel tests were modified dynamic labral shear and the upper-cut test. The standard tests were Speed's test, the active compression test and the belly-press test. The modified tests were the anterior slide test, the bear-hug test, Speed's test and Yergason's test. The combination test comprised Yergason's test and Gilcreest's test but was incompletely defined. The sensitivity estimates ranged from 17% (95% CI 6% to 36%) for the novel modified dynamic labral shear to 87% (95% CI 82% to 91%) for the

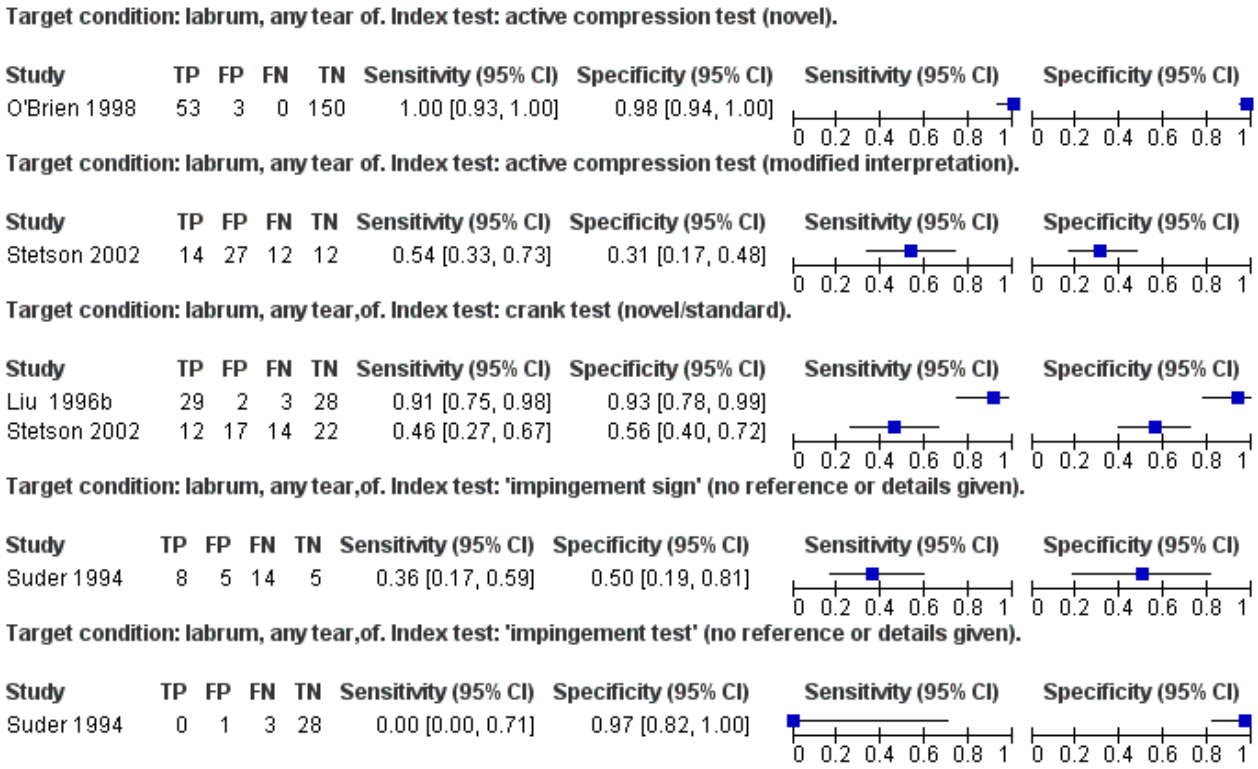
standard Speed's test. The specificity estimates ranged from 53% (95% CI 41% to 65%) for the novel modified dynamic labral shear to 85% (95% CI 74% to 92%) for the standard belly-press test.

Labral lesions (11 studies)

Non-specific labral lesions (four studies)

The sensitivity and specificity estimates from each study for the tests of non-specific labral lesions are shown in forest plots in Figure 15.

Figure 15. Glenoid labral lesion: non-specific labral lesion



Four studies evaluated five novel/standard or modified tests for labral lesions without attempting to discriminate between these lesions. The novel/standard tests were the active compression test and the crank test. The novel/standard crank test was performed and interpreted similarly in both studies, but the results were heterogeneous. The modified test was the active compression test. Two tests, an 'impingement sign' and an 'impingement test' were insufficiently defined to be classified. The sensitivity estimates ranged from 0% (95% CI 0% to 71%) for the undefined 'impingement test' to 100% (95% CI 93 to 100%) for the novel active compression test. The specificity estimates ranged from 31% (95% CI 17% to 48%) for the modified active compression test to 98% (95% CI 94% to 100%) for the novel active compression test.

Non-specific SLAP lesions (three studies)

The sensitivity and specificity estimates from each study for the tests of non-specific SLAP lesions are shown in forest plots in Figure 16.

Figure 16. Glenoid labral lesions: non-specific SLAP lesion

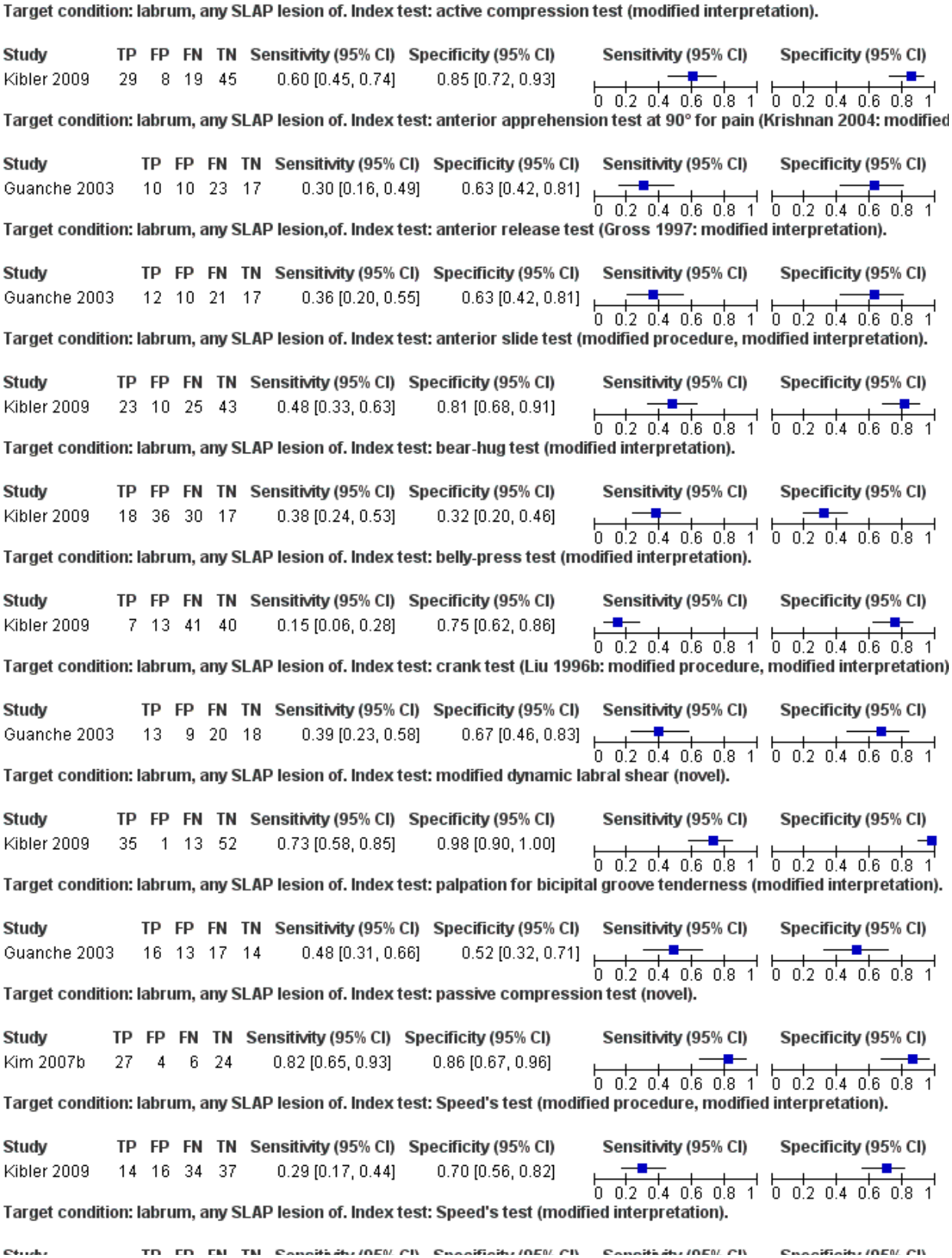
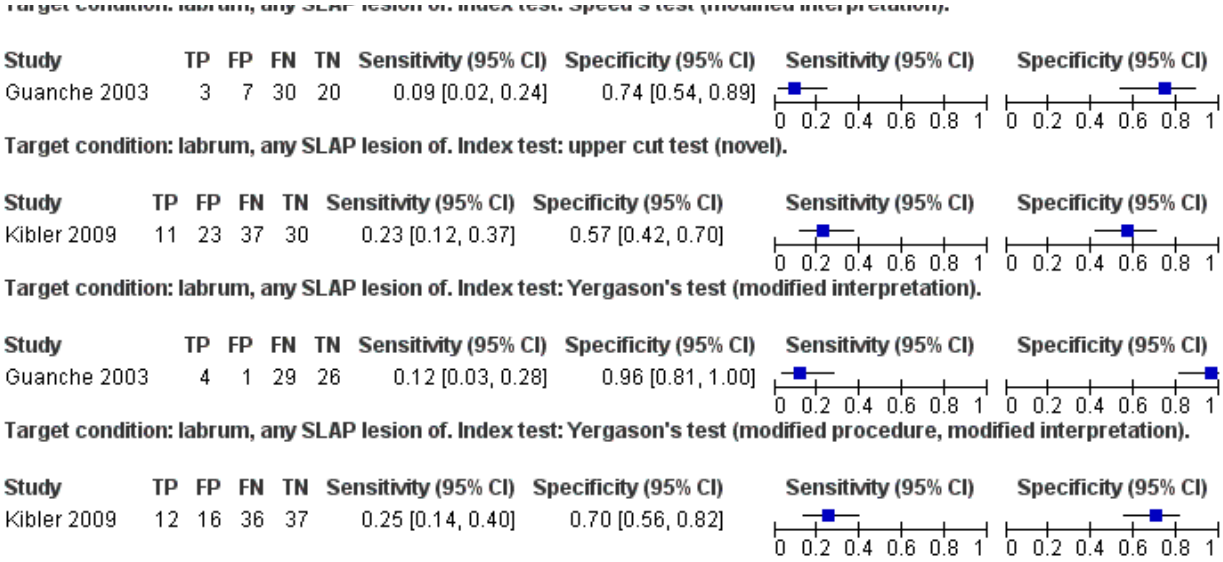


Figure 16. (Continued)

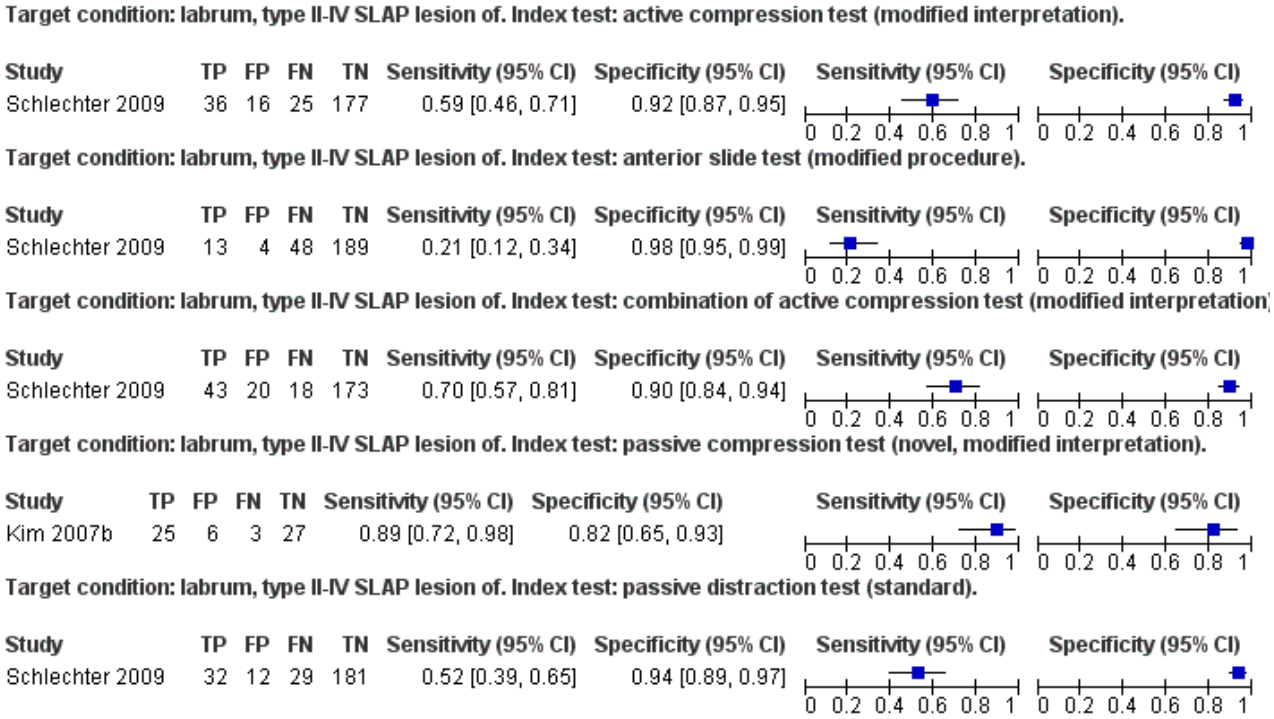


Three studies evaluated 15 novel or modified tests. There were no instances of a test being performed and interpreted similarly in two or more studies. The novel tests were modified dynamic labral shear, the passive compression test and the upper cut test. The modified tests were the active compression test, the anterior apprehension test at 90°, the anterior release test, the anterior slide test, the bear-hug test, the belly-press test, the crank test and palpation for bicipital groove tenderness (one variant each) and Speed's test and Yergason's test (two variants each). The sensitivity estimates ranged from 9% (95% CI 2% to 24%) for a modified Speed's test to 82% (95% CI 65% to 93%) for the novel passive compression test. The specificity estimates ranged from 32% (95% CI 20% to 46%) for the modified bear-hug test to 98% (95% CI 90% to 100%) for the novel modified dynamic labral shear.

Type II-IV SLAP lesions (two studies)

The sensitivity and specificity estimates from each study for the tests of type II-IV SLAP lesions are shown in forest plots in Figure 17.

Figure 17. Glenoid labral lesions: type II-IV SLAP lesion



Two studies evaluated five novel, standard, modified or combination tests for type II-IV SLAP lesions without attempting to differentiate between these types. None of the tests was performed and interpreted similarly in both studies. The novel test was the passive compression test. The standard test was the passive distraction test. The modified tests were the anterior slide test and the active compression test (one variant each). The combination test comprised the active compression test or the passive distraction test. The sensitivity estimates ranged from 21% (95% CI 12% to 34%) for the modified anterior slide test to 89% (95% CI 72% to 98%) for the novel passive compression test. The specificity estimates ranged from 82% (95% CI 65% to 93%) for the novel passive compression test to 98% (95% CI 95% to 99%) for the modified anterior slide test.

Type II SLAP lesions (three studies)

The sensitivity and specificity estimates from each study for the tests of type II SLAP lesions are shown in forest plots in Figure 18.

Figure 18. Glenoid labral lesions: type II SLAP lesion

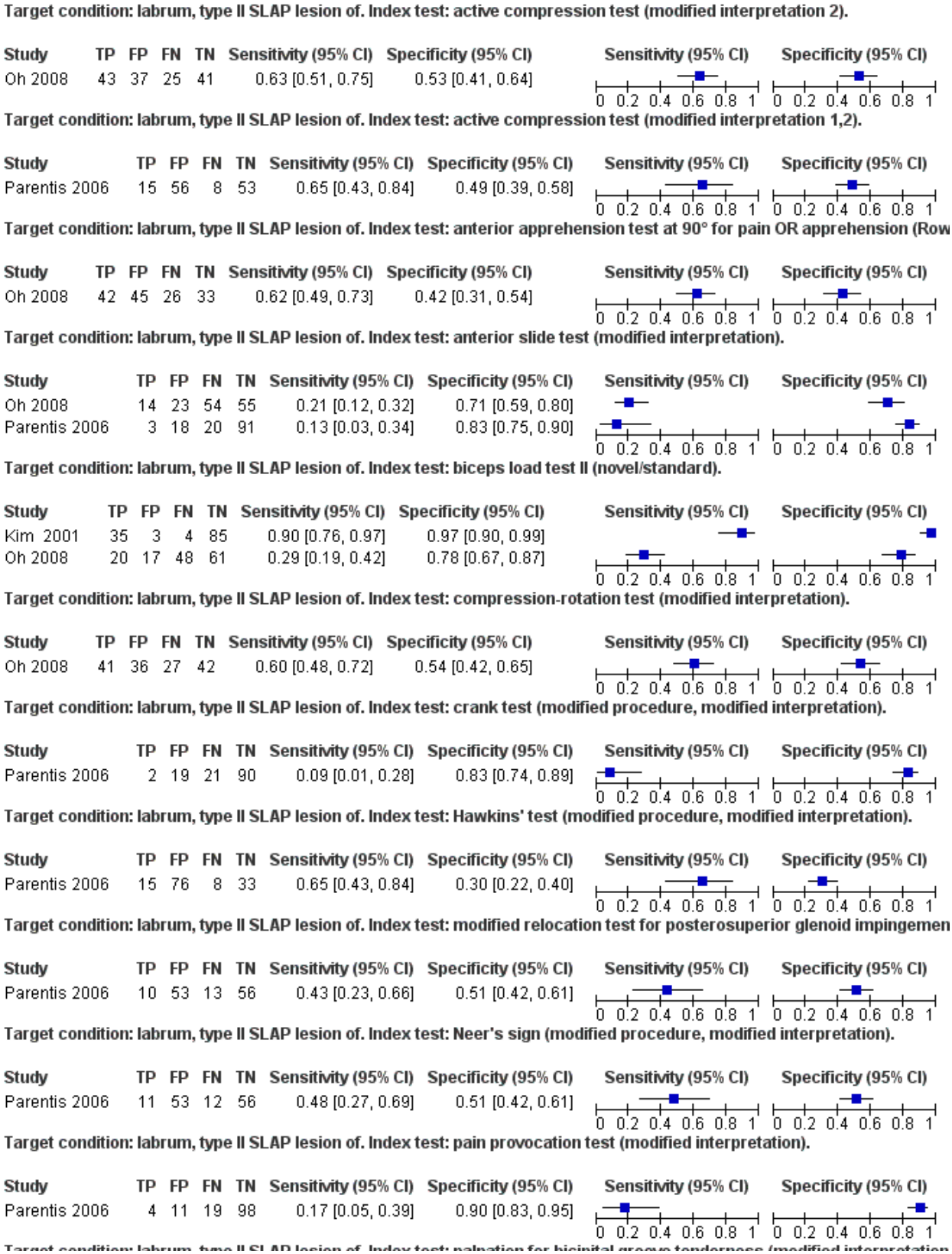
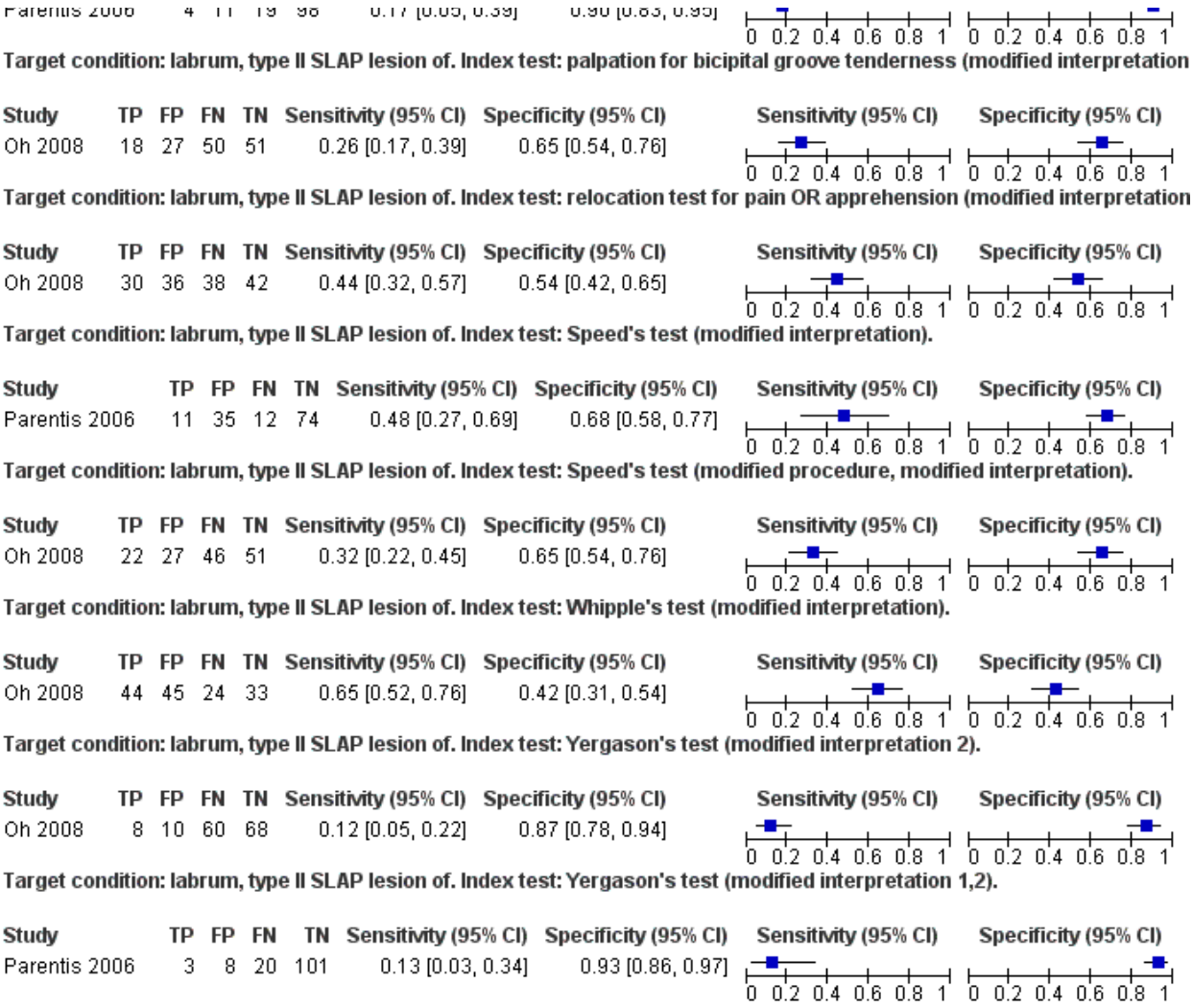


Figure 18. (Continued)



Three studies evaluated 18 novel/standard or modified tests for type II SLAP lesions. The novel/standard test was the biceps load II test. The modified tests were the active compression test, Speed's test and Yergason's (two variants each) and the anterior slide test, the compression-rotation test, the crank test, Hawkins' test, the modified relocation test for postero-superior glenoid impingement, Neer's sign, the pain provocation test, the relocation test for pain or apprehension and Whipple's test (one variant each). The sensitivity estimates ranged from 9% (95% CI 1% to 28%) for the modified crank test to 90% (95% CI 76% to 97%) for the novel/standard biceps load II test. The specificity estimates ranged from 30% (95% CI 22% to 40%) for the modified Hawkins' test to 97% (95% CI 90% to 99%) for the novel/standard biceps load II test.

Two tests were performed and interpreted similarly in two studies. These were the novel/standard biceps load II test by Kim 2001 and Oh 2008 and the modified anterior slide test by Oh 2008 and Parentis 2006.

Tests for multiple target conditions: undifferentiated (four studies)

The sensitivity and specificity estimates from each study for the tests of undifferentiated multiple target conditions are shown in forest plots in Figure 19.

Figure 19. Non-specific

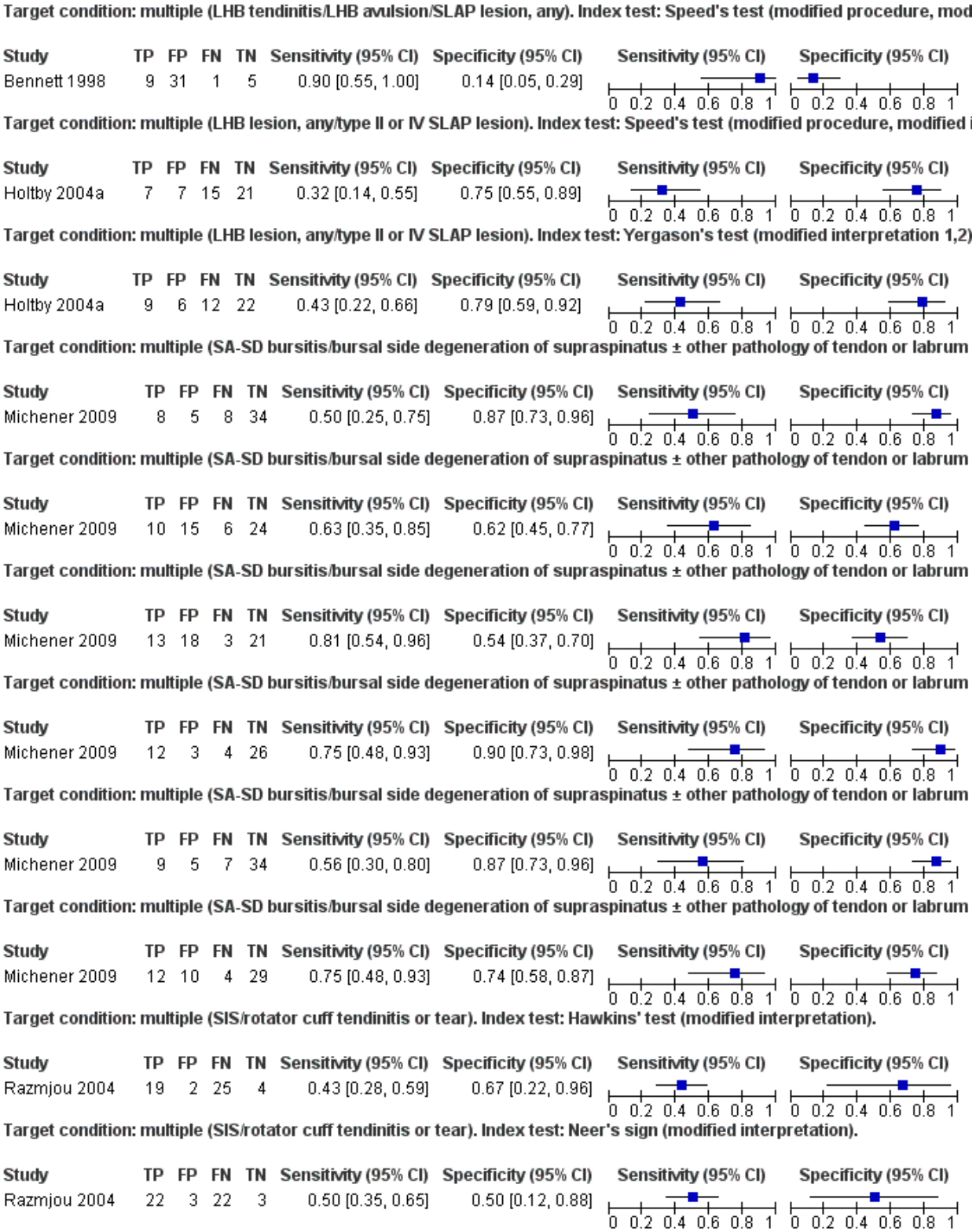


Figure 19. (Continued)

One study evaluated a modified Speed's test for LHB tendinitis or avulsion or any SLAP lesion without attempting to differentiate between these, with a sensitivity of 90% (95% CI 55% to 100%) and a specificity of 14% (95% CI 5% to 29%).

One study evaluated a modified Speed's test and a modified Yergason's test for any LHB lesion, a type II or a type IV SLAP lesion without attempting to distinguish between these diseases, with a sensitivity of 32% (95% CI 14% to 55%) for the modified Speed's test and 43% (95% CI 22% to 66%) for the modified Yergason's test; and specificity estimates of 75% (95% CI 55% to 89%) and 79% (95% CI 59% to 92%) respectively.

One study evaluated six standard, modified or combination tests for SA-SD bursitis or bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability, without attempting to distinguish between these diseases. The standard tests were Hawkins' test and the painful arc test. The modified tests were the empty can test, Neer's test and resisted lateral rotation from neutral rotation. The combination test comprised all the foregoing, of which three were required to be positive. The sensitivity estimates ranged from 50% (95% CI 25% to 75%) for the modified empty can test to 81% (95% CI: 54% to 96%) for the modified Neer's test. The specificity estimates ranged from 54% (95% CI 37% to 70%) for the modified Neer's test to 90% (95% CI 73% to 98%) for the standard painful arc test.

One study evaluated two modified tests, Hawkins' test and Neer's sign, for SIS or rotator cuff tendinitis or tear without attempting to discriminate between these diseases, with a sensitivity of 43% (95% CI 28 to 59%) for the modified Hawkins' test and 50% (95% CI 35 to 65%) for the modified Neer's sign; and specificity estimates of 67% (95% CI 22% to 96%) and 50% (95% CI 12% to 88%) respectively.

Between-tester reliability

Kappa coefficients were reported as 0.815 for the biceps load II test (Kim 2001); 0.39 (95% CI 0.12 to 0.65) for the empty can test (Michener 2009); respectively 0.47 (0.22 to 0.72) and 0.29 (0.180 to 0.398) for Hawkins' test (Michener 2009; Razmjou 2004); respectively 0.40 (0.13 to 0.67) and 0.506 (0.366 to 0.645) for Neer's test (Michener 2009; Razmjou 2004); 0.45 (0.18 to 0.72) for the painful arc test (Michener 2009); 0.771 for the passive compression test (Kim 2007b); 0.67 (0.40 to 0.94) for resisted lateral rotation from neutral rotation (Michener 2009).

DISCUSSION

We set out to identify and review studies evaluating the diagnostic accuracy of defined physical tests, whether applied singly or in combination, for shoulder impingements (subacromial or internal) or local lesions of bursa, rotator cuff or labrum that may accompany impingement, in people whose symptoms and/or history suggest any of these disorders. Our particular focus was primary care (while not excluding secondary or tertiary care, especially in the hospital outpatient setting).

Summary of main results

See Summary of findings 1. The 33 included studies, of 4002 shoulders in 3852 patients, incorporated numerous standard, modified, or combination index tests and 14 novel index tests. In consequence they embodied 170 target condition/index test combinations, with only six instances of any index test being performed and interpreted similarly in two studies. However, only two studies of a modified empty can test for FTT of the rotator cuff (Holtby 2004b; Kim 2006) and a modified anterior slide test for type II SLAP lesions (Kim 2001; Oh 2008) were clinically homogenous.

Variations in index tests' provenance, procedure and interpretation

The provenance given for index tests was diverse. Primary sources for at least one index test were cited in 20 studies (Barth 2006; Calis 2000; Castoldi 2009; Guanche 2003; Hertel 1996; Holtby 2004a; Itoi 1999; Itoi 2006; Kibler 2009; MacDonald 2000; Michener 2009; Miller 2008b; Naredo 2002; Oh 2008; Parentis 2006; Razmjou 2004; Schlechter 2009; Speer 1994; Stetson 2002; Wolf 2001); secondary/tertiary sources in eight (Bennett 1998; Calis 2000; Guanche 2003; Iagnocco 2003; Kibler 2009; Michener 2009; Naredo 2002; Oh 2008); and none in six (Frost 1999; Kim 2006; Norwood 1989; Oh 2008; Parentis 2006; Suder 1994). In two studies (studies reporting novel tests were not included in this) neither a reference nor a description of the index tests was offered (Norwood 1989; Suder 1994). We identified some misattributions, most commonly of the empty can test to Jobe 1982 (Holtby 2004b; Hertel 1996; Itoi 1999; Itoi 2006; Michener 2009). Neer's sign was misattributed to Neer 1972a, a report which gave no clear account of this index test, by Razmjou 2004. Speed's test was incorrectly referenced by Oh 2008. In one instance, a figure depicting Hawkins' test was apparently misinterpreted as a novel test, Yocum's test (Naredo 2002).

Some studies cited index tests' primary sources as well as providing a description, often revealing substantive, apparently unintentional inconsistencies in procedures, criteria for positive results, or both. This applied to the active compression test (Guanche 2003), the anterior release test (Guanche 2003), the anterior slide test (Kibler 2009; Schlechter 2009), the belly press test (Barth 2006), the crank test (Guanche 2003), the drop sign (Miller 2008b), Hawkins' test (Parentis 2006), Neer's sign (MacDonald 2000; Michener 2009; Parentis 2006), the palm up test (Naredo 2002), Yergason's test (Guanche 2003; Kibler 2009). There were also three instances in which the originators of index tests described the method of performance and interpretation differently across reports: the anterior slide test (Kibler 1995a; Kibler 2009), the crank test (Liu 1996a; Liu 1996b) and the lift-off test (Gerber 1991; Gerber 1996). In two of these instances, (Liu 1996a and Liu 1996b; Gerber 1991 and Gerber 1996), the differing descriptions apparently related to the same patient samples, revealing internal inconsistencies.

Whether intentional or unintentional, variations in index tests' procedure or interpretation were prevalent, such that, as observed above, there were only six instances of any index test being performed and interpreted (in terms of criteria for, and implications

of, a positive result) similarly in two studies; and no instances of three studies or more using any one test similarly.

As previously stated, 14 of the tests identified - a surprisingly large proportion - were novel, and we were particularly interested to explore their provenance. The justifications given included synthesis of empirical evidence from other studies (the internal rotation resistance strength test), application of biomechanical principles (the biceps load II test, the drop sign, the external rotation lag sign, the internal rotation lag sign, the passive compression test), or mechanisms of pain-provocation described by patients (the active compression test and the upper-cut test). Some tests appear to have been developed on a trial and error basis: active abduction, which was developed retrospectively from routinely collected data; the modified dynamic labral shear, which was adapted from an existing test; and the passive distraction test, the postulated mechanism of which was confirmed arthroscopically. For two tests (the crank test, the Gum-Turn test), no clear justification was reported.

Arithmetical discrepancies

Arithmetical discrepancies in reported statistics, over and above those attributable to rounding error, were prevalent. In three studies, substantial discrepancies in the 2 x 2 tables, as back-calculated from reported sensitivity and/or specificity, warranted exclusion (Ebinger 2008; Fodor 2009; Litaker 2000). Seven studies presented multiple but smaller errors of this type, multiple errors associated with other summary statistics, or both (Bennett 1998; Hertel 1996; Itoi 2006; Kim 2006; MacDonald 2000; Miller 2008b; Parentis 2006).

We adopted a policy of excluding studies in which, allowing for the possibility of an isolated typographical/transcription error, back-calculation of 2 x 2 tables from the reported sensitivity and specificity demonstrated greater-than 10% discrepancies in any cell (which was not attributable to unit-of-analysis issues). We excluded three studies (Ebinger 2008; Fodor 2009; Litaker 2000) on this basis.

Many of the remainder presented with errors which were either too small to warrant study exclusion or which related to summary statistics other than sensitivity or specificity. We divided these errors into two categories: isolated discrepancies within study reports which (extending the benefit of doubt) we attributed to typographical errors, or confirmed as such by communication with the authors; and multiple discrepancies within a report, which we attributed to miscalculation. Seven studies (Bennett 1998; Hertel 1996; Itoi 2006; Kim 2006; MacDonald 2000; Miller 2008b; Parentis 2006) fell within the former category and six within the latter (Calis 2000; Gumina 2008; Kibler 2009; Naredo 2002; Oh 2008; Schlechter 2009). We subcategorised each type of error according to its absolute magnitude in reported percentage terms: *see* Table 5.

Tests' potential to inform diagnoses

Sensitivity and specificity are test properties, and not directly applicable at the interface between clinician and patient. Useful statistics in this context are the positive and negative likelihood ratio (LR+, LR-). The LR+ may be defined as true positive rate/false positive rate = sensitivity/(1-specificity), and the LR- as false negative rate/true negative rate = (1-sensitivity)/specificity. These statistics facilitate a Bayesian approach, which is intuitive to clinicians (Gill 2005), enabling estimation of the likelihood of a target condition post-test when the pre-test probability of the condition is known, by means of a nomogram (Jaeschke 1994). We had intended to tabulate LR+ and LR- data to optimise the clinical utility of our review. However, in the light of our findings, we decided that this step would overplay the evidence. For the most part, this evidence derives from small, methodologically compromised, single studies; often conducted by tests' originators, with negative implications for reproducibility. Over and above these considerations is the fact that few of the results, and none from methodologically and arithmetically robust studies, are directly applicable to primary care.

Between-tester agreement

Few studies addressed this aspect, although it is fundamental to the validity of clinical tests. Agreement is best evaluated using the kappa coefficient, since this takes account of the fact that agreements may occur by chance. The coefficient ranges from 0 to 1, and interpretation has been recommended as follows by Altman 1991: less than 0.20 = poor; 0.21 to 0.40 = fair; 0.41 to 0.60 = moderate; 0.61 to 0.80 = good; 0.81 to 1 = very good. By these criteria, and based on point estimates, very good between-rater agreement was achieved for only one test, the biceps load II test (Kim 2001). Good agreement was obtained for the passive compression test (Kim 2007b) and resisted lateral rotation from neutral rotation (Michener 2009). Agreement for the painful arc test was moderate (Michener 2009), while that for Neer's test was fair to moderate (Michener 2009; Razmjou 2004). For the empty can test (Michener 2009) and Hawkins' test (Michener 2009; Razmjou 2004), agreement was only fair.

Calis 2000 reported interobserver reliability for their battery of index tests of 'above 98%', but gave no breakdown and presumably did not account for chance agreement in their calculations. Miller 2008b stated that ‘to ensure test quality, the clinical tests were practiced on five separate occasions with an orthopaedic surgeon with a special interest in shoulders on a separate subgroup of subjects’, but no statistical analysis was presented. In Naredo 2002, two rheumatologists performed independent assessments, then established clinical diagnoses by consensus. Finally, although evaluation of reproducibility was mentioned in the abstract of Schlechter 2009, this aspect was not addressed in the report.

Comparison with other systematic reviews

We identified 12 other systematic reviews and one non-systematic review which overlapped with aspects of ours. Two addressed multiple shoulder pathologies. These were Dinnes 2003, which covered impingement syndrome and rotator cuff tears (FTT, PTT or any) but excluded labral disorders and Hegedus 2008a, which had no exclusions. One systematic review (Alqunaee 2012) was nominally specific to subacromial impingement syndrome, but encompassed rotator cuff pathology tears (i.e. stage II and III impingement, according to the criteria of Neer 1977). Three (Beaudreuil 2009a; Hughes 2008a; Longo 2011) addressed rotator cuff disease. One, Munro 2009a, addressed labral disease. Disproportionately, considering the relatively low prevalence of this condition, six reviews (Dessaur 2008a; Calvert 2009a; Jones 2007a; Luime 2004b; Meserve 2009a; Walton 2008a) focused on SLAP lesions. Table 11 summarises these reviews and compares their main conclusions to our own.

Our own review differs from the remainder, except Dinnes 2003, in terms of its emphasis on relatively unselected populations such as might be encountered in primary care. We took this approach because it is in primary care that people with these problems are screened and, for most diagnoses, managed. We decided not exclude secondary or tertiary care, based in part upon the (correct) pragmatic assumption that few studies would use reference standards directly applicable to primary care. Extrapolation from the secondary and tertiary settings to primary care should be undertaken with caution, however. In primary care, patients will have travelled less far down the screening pathway, disease is likely to be less severe, and the expertise of clinicians conducting and interpreting the physical tests may be less. These aspects would tend to reduce diagnostic test accuracy. On the other hand, our results do have applicability to relatively unselected populations (in terms of occupation and sporting activity) in secondary and tertiary care.

Our review also differs in terms of its scope. We considered that shoulder impingements and those painful conditions that may be related to impingements present a coherent class of pathologies which would resonate with clinicians. Although we recognise that this class of pathologies may overlap with others (laxities and instabilities, capsular and acromio-clavicular conditions), we omitted these in order to maintain focus and optimise the utility of an already large review. We believe that we have achieved an appropriate compromise.

Strengths and weaknesses of the review

Strengths of the review

This review has a number of strengths, principal among which are the following.

Search strategy

Our search strategy was comprehensive.

Definition of physical tests

As an integral part of the review, we have described the performance and interpretation of physical tests, by reference, wherever possible, to the primary sources (Table 1). We have also been careful to evince the detail of index tests from the included studies.

Comparisons revealed that modifications in the procedures and/or interpretation of tests, sometimes intentional and sometimes not, are highly prevalent in the literature. Unclear is the extent to which a test's performance may be changed in terms of the starting position, the plane and range of movement, inclusion or exclusion of passive, active and resisted components, and the forcefulness of application, before it must be regarded as a different test. Our alertness to such procedural modifications, and modifications of interpretation (the criteria for a positive response, the implications of a positive response, or both) has enabled us to avoid pooling data which on superficial inspection seem suitable, but which are actually clinically heterogeneous; and has brought to light numerous internal inconsistencies.

Back-calculations

We double-checked the summary statistics presented in the included studies for each target condition/index test combination, back-calculating 2 x 2 tables. That we observed such a high frequency of arithmetical errors, some so serious as to warrant study exclusion (Ebinger 2008; Fodor 2009; Litaker 2000) and others sufficiently serious to cast doubt on the safety of results (Calis 2000; Gumina 2008; Kibler 2009; Naredo 2002; Oh 2008; Schlechter 2009), emphasises the inadvisability of uncritically accepting reported values.

Non-blinding of reviewers

Perhaps counter-intuitively, we consider non-blinding of reviewers to study authors was a strength of this review, because it facilitated identification of inconsistencies across multiple publications.

Weaknesses of the review

As predicted in our protocol, non-English literature was not included because of resource limitations, although we are aware that this may have led to selection bias. However, at minimum, very good quality and complete technical translations are required and even then some key subtleties (such as in population and test performance) may be missed. In practice, due to obscure presentation of data in much of the primary literature, extraction even in the English-language frequently presented a considerable challenge. This emphasizes the unfeasibility of extracting data with similar stringency from literature in other languages

While our search strategy was comprehensive, specialist feedback at editorial review highlighted that a greater use of subject headings in the database searches for the conditions under investigation and terms found in diagnostic test accuracy filters terms would have been desirable. Specialist feedback also suggested other databases and pointed to various grey literature sources such as MEDION (database of systematic reviews of diagnostic studies). We cannot say how many studies may have been missed through these potential deficiencies in our search strategy but reflect that our search through the reference lists of other reviews did not reveal any relevant omissions.

We recognise the possible inclusion of a large number of studies with an unacceptable delay between the index and reference tests, which we defined as an interval exceeding either the average duration of symptoms or one month (whichever was the shorter), as a potential source of bias. Specifically, there is a possibility of misclassification due to spontaneous recovery or progression of disease. Of the 33 trials included, 22 provided no details of the interval between the index and reference test. Only seven trials would have met the acceptable delay criterion, had it been a condition for inclusion. These were Calis 2000 (no delay assumed), Guanche 2003 (immediate), Iagnocco 2003 (a few days' delay), Miller 2008b (no delay), Naredo 2002 (≤ 1 week), Oh 2008 (mean delay 1 day), and Speer 1994 (index and reference tests were on the same day). The other four trials providing data would not have met the inclusion criterion. These were Holtby 2004a, Holtby 2004b (mean delay 23 weeks); Michener 2009 (mean 2.6 months); and Razmjou 2004 (mean 23 weeks). Given that the majority of studies did not report on the delay between the index and reference tests, our decision to include these and other studies with longer than desirable delays between index and reference tests was pragmatic, and necessitated by the evidence.

Applicability of findings to the review question

In the light of limitations of the primary evidence, the practical diagnostic advice that may be offered to the primary care clinician faced with an impingement-related disorder can only be very tentative. A first step may be to screen for those conditions, such as glenoid labral lesions and large rotator cuff tears, which are most likely to warrant surgical opinion, so that a timely referral made be made; even if the immediate plan, perhaps while the patient is on a waiting list, is to implement a trial of conservative treatment. Identifying the best tests for screening purposes, and thus ruling out these diseases, involves consideration of the reported sensitivity, but also the precision of the point estimate and the methodological quality of the reporting study (or studies). Taking these factors into account, the most promising screening test for glenoid labral lesions was the passive compression test (Kim 2007b) with a sensitivity of 89% (95% CI 72% to 98%). But the test was novel and the evaluation was by its originators, and firm conclusions as to its usefulness must await independent verification. For full thickness tears of the rotator cuff, the posterosuperior rotator cuff in general, supraspinatus or infraspinatus there were no strong candidate tests. Thus the decision whether to refer may rest upon patients' response to conservative intervention. For any tear of subscapularis, the sensitivity of the internal rotation lag sign was very high (97% (95% CI 82% to 100%)), although the evaluation was by the test's originators (Hertel 1996) and, as with the passive compression test for labral lesions, independent verification would be highly desirable.

The degree to which specific localization of lesions (and thus tests with high specificity) is necessary depends upon the therapeutic approach. For example, a diagnosis of impingement is sufficient to implement a programme of exercises which aims to centre the humeral head in the glenoid. On the other hand, administration of localized massage such as deep transverse friction (Cyriax 1984) would call for a tendon-specific diagnosis. In fact, there were no strong candidates for diagnosing impingement *per se*, and it may be that greater emphasis should be placed on making this diagnosis by excluding other main causes of shoulder pain such as frozen shoulder. Regarding specific rotator cuff tendons, resisted lateral rotation from neutral rotation, a conventional test for any disease of infraspinatus, had not only very high specificity (95% (95% CI 92% to 96%)) but also high sensitivity (94% (95% CI 87% to 98%)); likewise resisted medial rotation from neutral rotation, with a specificity of 99% (95% CI 97% to 100%)) and sensitivity of 96% (95% CI 78% to 100%) (Iagnocco 2003). There was no test which appeared to be equivalently useful for ruling in any disease of supraspinatus, but the empty can test appeared very useful for ruling it out (Iagnocco 2003; Naredo 2002) with sensitivities of 94% (95% CI 91% to 97%)) and 96% (95%CI 79% to100%) for an undefined test and a standard test, respectively.

Assessments of diagnostic test accuracy are only meaningful for tests which are replicable in different target users' hands. Thus between-tester reliability is a critical consideration. Few of the included studies (Calis 2000; Kim 2001; Michener 2009; Razmjou 2004) addressed this aspect, however. It is noteworthy that the sole test on which there was 'very good' agreement was novel, and reported by its originators, whom one would expect to be highly proficient in its application and interpretation. Likewise, of the two tests for which there was 'good' agreement, one was a novel test, and reported by its originators.

AUTHORS' CONCLUSIONS

Implications for practice

There is insufficient evidence upon which to base selection of physical tests for shoulder impingements, and local lesions of bursa, tendon or labrum that may accompany impingement, in primary care. Our discussion has given an indication of an approach that may be tentatively adopted; but it must be emphasized that this is very provisional.

Implications for research

Diagnostic test accuracy research should be approached with the same rigour as randomised controlled trials. Weak, often retrospective, designs, coupled with poor reporting, are prevalent; and it is cause for concern that not one of the 14 included studies published between 2004 and the present make explicit reference to STARD, the Standards for Reporting Studies of Diagnostic Accuracy, which were simultaneously published in seven journals in 2003 (Bossuyt 2003) and have been published in eight journals since. Editors should make compliance with these standards obligatory for diagnostic test accuracy research.

Another critical issue is the non-standard way in which tests are performed and interpreted across studies since (despite the large number of studies in the field) this hinders synthesis of the evidence and/or clinical applicability. Where possible, trialists should revisit the primary source in order to ensure that their perception of the test complies with the original description, because the descriptions in many secondary or tertiary sources, both 'how to do it' literature and reports of diagnostic test accuracy studies, are idiosyncratic.

Especially useful for primary care would be studies independently verifying the screening capabilities of the passive compression test for any labral lesion and the internal rotation lag sign for any tear of subscapularis, and to identify simple, accurate screening tests for full thickness tears of the rotator cuff, the posterosuperior rotator cuff in general, supraspinatus and infraspinatus. A test specific to subacromial impingement syndrome is still elusive, as is a test specific to any disease of the supraspinatus; and studies verifying the diagnostic accuracy of resisted lateral and medial rotation for any disease of infraspinatus and subscapularis are required.

Finally, investigators should consider conducting formal reliability studies within the context of, or alongside, diagnostic test accuracy studies: this important aspect is often overlooked.

ACKNOWLEDGEMENTS

We are grateful to the following for helpful feedback about the review: Lotty Hooft, James Robertson and Cathie Sherrington. We thank Lindsey Elstub and Joanne Elliott for their editorial help and support during the preparation of the review. We are grateful to Amy Kavanagh for obtaining copies of dozens of study reports and for some data entry of reference citations and trial characteristics.

We thank Riekie de Vet for her helpful comments on the first drafts of the review protocol and to her and Danielle van der Windt for sharing their protocol with us. We thank Riekie de Vet and Lesley Gillespie for their support during the editorial review process. We thank the following for helpful feedback at editorial review: Riekie

de Vet, Bill Gillespie, Lesley Gillespie, Petra Macaskill, Rob Scholten and Janet Wale.

REFERENCES

References to studies included in this review

Barth 2006 *{published data only}*

Barth JR, Burkhart SS, de Beer JF. The bear-hug test: a new and sensitive test for diagnosing a subscapularis tear. *Arthroscopy* 2006;**22**(10):1076-84.

Bennett 1998 *{published data only}*

Bennett WF. Specificity of the Speed's test: arthroscopic technique for evaluating the biceps tendon at the level of the bicipital groove. *Arthroscopy* 1998;**14**(8):789-96.

Calis 2000 *{published data only}*

Calis M, Akgun K, Birtane M, Karacan I, Calis H, Tuzun F. Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Annals of the Rheumatic Diseases* 2000;**59**(1):44-7.

Castoldi 2009 *{published data only}*

Castoldi F, Blonna D, Hertel R. External rotation lag sign revisited: accuracy for diagnosis of full thickness supraspinatus tear. *Journal of Shoulder & Elbow Surgery* 2009;**18**(4):529-34.

Frost 1999 *{published data only}*

Frost P, Andersen JH, Lundorf E. Is supraspinatus pathology as defined by magnetic resonance imaging associated with clinical sign of shoulder impingement?. *Journal of Shoulder & Elbow Surgery* 1999;**8**(6):565-8.

Guanche 2003 *{published data only}*

Guanche CA, Jones DC. Clinical testing for tears of the glenoid labrum. *Arthroscopy* 2003;**19**(5):517-23.

Gumina 2008 *{published data only}*

Gumina S, Bertino A, Di Giorgio G, Postacchini F. A new test of resistance in the diagnosis of postero-superior rotator cuff tears. *Chirurgia Degli Organi di Movimento* 2008;**91**(2):85-6.

Hertel 1996 *{published data only}*

Hertel R, Ballmer FT, Lombert SM, Gerber C. Lag signs in the diagnosis of rotator cuff rupture. *Journal of Shoulder and Elbow Surgery* 1996;**5**(4):307-13.

Holtby 2004a *{published data only}*

Holtby R, Razmjou H. Accuracy of the Speed's and Yergason's tests in detecting biceps pathology and SLAP lesions: comparison with arthroscopic findings. *Arthroscopy* 2004;**20**(3):231-6.

Holtby 2004b *{published data only}*

Holtby R, Razmjou H. Validity of the supraspinatus test as a single clinical test in diagnosing patients with rotator cuff pathology. *Journal of Orthopaedic & Sports Physical Therapy* 2004;**34**(4):194-200.

Iagnocco 2003 *{published data only}*

Iagnocco A, Coari G, Leone A, Valesini G. Sonographic study of painful shoulder. *Clinical & Experimental Rheumatology* 2003;**21**(1):355-8.

Itoi 1999 *{published data only}*

Itoi E, Kido T, Sano A, Urayama M, Sato K. Which is more useful, the "full can test" or the "empty can test," in detecting the torn supraspinatus tendon?. *American Journal of Sports Medicine* 1999;**27**(1):65-8.

Itoi 2006 *{published data only}*

Itoi E, Minagawa H, Yamamoto N, Seki N, Abe H. Are pain location and physical examinations useful in locating a tear site of the rotator cuff?. *American Journal of Sports Medicine* 2006;**34**(2):256-64.

Kibler 2009 *{published data only}*

Kibler WB, Sciasca A, Hester P, Dome d, Jacobs C. Clinical utility of traditional and new tests in the diagnosis of biceps tendon injuries and superior labrum anterior and posterior lesions in the shoulder. *American Journal of Sports Medicine* 2009;**37**(9):1840-7.

Kim 2001 *{published data only}*

Kim S-H, Ha K-I, Ahn J-H, Kim S-H, Choi H-J. Biceps Load Test II: A clinical test for SLAP lesions of the shoulder. *Arthroscopy* 2001;**17**(2):160-4.

Kim 2006 *{published data only}*

Kim E, Jeong HJ, Lee KW, Song JS. Interpreting positive signs of the supraspinatus test in screening for torn rotator cuff. *Acta Medica Okayama* 2006;**60**(4):223-8.

Kim 2007b *{published data only}*

Kim Y-S, Kim J-M, Ha K-Y, Choy S, Joo M-W, Chung Y-G. The passive compression test: A new clinical test for superior labral tears of the shoulder. *American Journal of Sports Medicine* 2007;**35**(9):1489-94.

Liu 1996b *{published data only}*

Liu SH, Henry MH, Nuccion S. A prospective evaluation of a new physical examination in predicting glenoid labral tears. *American Journal of Sports Medicine* 1996;**24**(6):721-5.

MacDonald 2000 *{published data only}*

MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement tests. *Journal of Shoulder and Elbow Surgery* 2000;**9**(4):299-301.

Michener 2009 *{published and unpublished data}*

Michener LA, Walsworth MK, Doukas WC, Murphy KP. Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. *Archives of Physical Medicine & Rehabilitation* 2009;**90**(11):1898-903.

Miller 2008b *{published data only}*

Elliott C. The validity of the lag signs in the diagnosis of rotator cuff dysfunction in patients with shoulder pain (protocol). www.controlled-trials.com 2006.

* Miller CA, Forrester GA, Lewis JS. The validity of the lag signs in diagnosing full-thickness tears of the rotator cuff:

a preliminary investigation. *Archives of Physical Medicine & Rehabilitation* 2008;**89**(6):1162-8.

Naredo 2002 *{published data only}*

Naredo E, Aguado P, De Miguel E, Uson J, Mayordomo L, Gijon-banos J, et al. Painful shoulder: comparison of physical examination and ultrasonographic findings. *Annals of the Rheumatic Diseases* 2002;**61**(2):132-6.

Norwood 1989 *{published data only}*

Norwood LA, Barrack R, Jacobson KE. Clinical presentation of complete tears of the rotator cuff. *Journal of Bone & Joint Surgery - American Volume* 1989;**71**(4):499-505.

O'Brien 1998 *{published data only}*

O'Brien SJ, Pagnani MJ, Fealy S, McGrory JE, Wilson JB. The active compression test: a new and effective test for diagnosing labral tears and acromioclavicular joint abnormality. *American Journal of Sports Medicine* 1998;**26**(5):610-3.

Oh 2008 *{published data only}*

Oh JH, Kim JY, Kim WS, Gong HS, Lee JH. The evaluation of various physical examinations for the diagnosis of type II superior labrum anterior and posterior lesion. *American Journal of Sports Medicine* 2008;**36**(2):353-9.

Parentis 2006 *{published data only}*

* Parentis MA, Glousman RE, Mohr KS, Yocum LA. An evaluation of the provocative tests for superior labral anterior posterior lesions. *American Journal of Sports Medicine* 2006;**34**(2):265-8.

Parentis MA, Mohr KJ, Elattrache NS. Disorders of the superior labrum: review and treatment guidelines. *Clinical Orthopaedics & Related Research* 2002;**400**:77-87.

Razmjou 2004 *{published data only}*

Razmjou H, Holtby R, Myhr T. Pain provocative shoulder tests: reliability and validity of the impingement tests. *Physiotherapy Canada* 2004;**56**(4):229-36.

Schlechter 2009 *{published data only}*

Schlechter JA, Summa S, Rubin BD. The passive distraction test: a new diagnostic aid for clinically significant superior labral pathology. *Arthroscopy* 2009;**25**(12):1374-9.

Speer 1994 *{published data only}*

Speer KP, Hannafin JA, Altchek DW, Warren RF. An evaluation of the shoulder relocation test. *American Journal of Sports Medicine* 1994;**22**(2):177-83.

Stetson 2002 *{published data only}*

Stetson WB, Templin K. The crank test, the O'Brien test, and routine magnetic resonance imaging scans in the diagnosis of labral tears. *American Journal of Sports Medicine* 2002;**30**(6):806-9.

Suder 1994 *{published data only}*

Suder PA, Hougaard K, Frich LH, Rasmussen OS, Lundorf E. Intraarticular findings in the chronically painful shoulder. A study of 32 posttraumatic cases. *Acta Orthopaedica Scandinavica* 1994;**65**(3):339-43.

Wolf 2001 *{published data only}*

Wolf EM, Agrawal V. Transdeltoid palpation (the rent test) in the diagnosis of rotator cuff tears. *Journal of Shoulder and Elbow Surgery* 2001;**10**(5):470-3.

Zaslav 2001 *{published data only}*

Zaslav KR. Internal rotation resistance strength test: a new diagnostic test to differentiate intra-articular pathology from outlet (Neer) impingement syndrome in the shoulder. *Journal of Shoulder and Elbow Surgery* 2001;**10**(1):23-7.

References to studies excluded from this review

Adolfsson 1991 *{published data only}*

Adolfsson L, Lysholm J. Arthroscopy for the diagnosis of shoulder pain. *International Orthopaedics* 1991;**15**(4):275-8.

Altchek 1990 *{published data only}*

Altchek DW, Warren RF, Wickiewicz TL, Skyhar MJ, Ortiz G, Schwartz E. Arthroscopic acromioplasty. Technique and results. [see comment]. *Journal of Bone & Joint Surgery - American Volume* 1990;**72**(8):1198-207.

Ansara 2006 *{published data only}*

Ansara S, El-Kawy S, Geeranavar S, Youssef B, Omar M. Diagnosis of rotator cuff tears: a comparison between clinical examination, MRI and operative findings. Our experience in a district general hospital. [abstract]. *Journal of Bone & Joint Surgery - British Volume* 2006;**88**(Suppl 2):277.

Ardic 2006 *{published data only}*

Aktas I, Akgun K. Re: Shoulder impingement syndrome: Relationships between clinical, functional, and radiologic findings. *American Journal of Physical Medicine and Rehabilitation* 2007;**86**(12):1035-6.

* Ardic F, Kahraman Y, Kacar M, Kahraman MC, Findikoglu G, Yorgancioglu ZR. Shoulder impingement syndrome: relationships between clinical, functional, and radiologic findings. *American Journal of Physical Medicine and Rehabilitation* 2006;**85**(1):53-60.

Beaudreuil 2009 *{published data only}*

Beaudreuil J. Reply to the letter by Silva about the review entitled "Contribution of clinical tests to the diagnosis of rotator cuff disease: A systematic literature review". *Joint Bone Spine: revue du rhumatisme* 2009;**76**(5):577-8.

* Beaudreuil J, Nizard R, Thomas T, Peyre M, Liotard JP, Boileau P, et al. Contribution of clinical tests to the diagnosis of rotator cuff disease: a systematic literature review. *Joint, Bone, Spine: Revue du Rhumatisme* 2009;**76**(1):15-9.

Silva L, Andreu JL. Comments on review by Beaudreuil et al "Contribution of clinical tests to the diagnosis of rotator cuff disease: a systematic review". *Joint, Bone, Spine: revue du rhumatisme* 2009;**76**(5):577; author reply 577-8.

Bedi 2008 *{published data only}*

Bedi A, Allen AA. Superior labral lesions anterior to posterior-evaluation and arthroscopic management. *Clinics in Sports Medicine* 2008;**27**(4):607-30.

Berbig 1999 *{published data only}*

Berbig R, Weishaupt D, Prim J, Shahin O. Primary anterior shoulder dislocation and rotator cuff tears. *Journal of Shoulder and Elbow Surgery* 1999;**8**(3):220-5.

Berg 1997 *{published data only}*

Berg EE, Ciullo JV. The SLAP lesion: a cause of failure after distal clavicle resection. *Arthroscopy* 1997;**13**(1):85-9.

Berg 1998 *{published data only}*

Berg EE, Ciullo JV. A clinical test for superior glenoid labral or 'SLAP' lesions. *Clinical Journal of Sport Medicine* 1998;**8**(2):121-3.

Bigliani 1997 *{published data only}*

Bigliani LU, Levine WN. Letter to the editor. *Journal of Bone & Joint Surgery - American Volume* 1998;**80**(12):1851-2.

* Bigliani LU, Levine WN. Subacromial impingement syndrome. *Journal of Bone & Joint Surgery - American Volume* 1997;**79**(12):1854-68.

Cohen J. Letter to the editor. *Journal of Bone & Joint Surgery - American Volume* 1998;**80**(12):1851-2.

Sarkisian GC. Letter to the editor. *Journal of Bone & Joint Surgery - American Volume* 1998;**80**(12):1851.

Birtane 2001 *{published data only}*

Birtane M, Calis M, Akgun K. The diagnostic value of magnetic resonance imaging in subacromial impingement syndrome. *Yonsei Medical Journal* 2001;**42**(4):418-24.

Blair 1996 *{published data only}*

Blair B, Rokito AS, Cuomo F, Jarolem K, Zuckerman JD. Efficacy of injections of corticosteroids for subacromial impingement syndrome. *Journal of Bone & Joint Surgery - American Volume* 1996;**78**(11):1685-9.

Blevins 1996a *{published data only}*

Blevins FT, Warren RF, Cavo C, Altchek DW. Arthroscopic assisted rotator cuff repair: results using a mini-open deltoid splitting approach. *Arthroscopy* 1996;**12**(1):50-9.

Blevins 1996b *{published data only}*

Blevins FT, Hayes WM, Warren RF. Rotator cuff injury in contact athletes. *American Journal of Sports Medicine* 1996;**24**(3):263-7.

Brasseur 2004 *{published data only}*

Brasseur JL, Lucidarme O, Tardieu M, Tordeur M, Montalvan B, Parier J, et al. Ultrasonographic rotator-cuff changes in veteran tennis players: the effect of hand dominance and comparison with clinical findings. *European Radiology* 2004;**14**(5):857-64.

Bron 2007 *{published data only}*

Bron C, Franssen J, Wensing M, Oostendorp RAB. Interrater reliability of palpation of myofascial trigger points in three shoulder muscles. *Journal of Manual and Manipulative Therapy* 2007;**15**(4):203-15.

Brownlow 2009 *{published data only}*

Brownlow HC, Smith C, Corner T, Neen D, Pennington R. Pain and stiffness in partial-thickness rotator cuff tears. *American Journal of Orthopedics* 2009;**38**(7):338-40.

Brox 2003 *{published data only}*

Brox JI. Regional musculoskeletal conditions: shoulder pain. *Best Practice & Research Clinical Rheumatology* 2003;**17**(1):33-56.

Bryant 2002 *{published data only}*

Bryant L, Shnier R, Bryant C, Murrell GA. A comparison of clinical estimation, ultrasonography, magnetic resonance imaging, and arthroscopy in determining the size of rotator cuff tears. *Journal of Shoulder & Elbow Surgery* 2002;**11**(3):219-24.

Buchberger 1999 *{published data only}*

Buchberger D J. Introduction of a new physical examination procedure for the differentiation of acromioclavicular joint lesions and subacromial impingement. *Journal of Manipulative and Physiological Therapeutics* 1999;**22**(5):316-21.

Burbank 2007 *{published data only}*

Burbank KM, Lemos MJ, Bell G, Lemos DW. Incidence of os acromiale in patients with shoulder pain. *American Journal of Orthopedics* 2007;**36**(3):153-5.

Burkhart 2000 *{published data only}*

Burkhart SS, Morgan CD, Kibler WB. Shoulder injuries in overhead athletes. The "dead arm" revisited. [Review]. *Clinics in Sports Medicine* 2000;**19**(1):125-58.

Burkhart 2002 *{published data only}*

Burkhart SS, Tehrany AM. Arthroscopic subscapularis tendon repair: Technique and preliminary results. *Arthroscopy* 18;**5**:454-63.

Burkhart 2003 *{published data only}*

Burkhart SS, Morgan CD, Kibler WB. The disabled throwing shoulder: spectrum of pathology. Part II: evaluation and treatment of SLAP lesions in throwers. *Arthroscopy* 2003;**19**(5):531-9.

Burkhart 2008 *{published data only}*

Burkhart SS, Ochoa Jr. Subscapularis tendon tears: Diagnosis and treatment strategies. *Current Orthopaedic Practice* 2008;**19**(5):542-7.

Callanan 2001 *{published data only}*

Callanan M, Tzannes A, Hayes K, Paxinos A, Walton J, Murrell GA. Shoulder instability. Diagnosis and management. *Australian Family Physician* 2001;**30**(7):655-61.

Calvert 2009 *{published data only}*

Calvert E, Chambers GK, Regan W, Hawkins RH, Leith JM. Special physical examination tests for superior labrum anterior posterior shoulder tears are clinically limited and invalid: a diagnostic systematic review. *Journal of Clinical Epidemiology* 2009;**62**(5):558-63.

Chronopoulos 2004 *{published data only}*

Chronopoulos E, Kim TK, Park HB, Ashenbrenner D, McFarland EG. Diagnostic value of physical tests for isolated chronic acromioclavicular lesions. *American Journal of Sports Medicine* 2004;**32**(3):655-61.

Cools 2008 *{published data only}*

Cools AM, Cambier D, Witvrouw EE. Screening the athlete's shoulder for impingement symptoms: a clinical reasoning algorithm for early detection of shoulder pathology. *British Journal of Sports Medicine* 2008;**42**(8):628-35.

Corso 1995 *{published data only}*

Corso G. Impingement relief test: an adjunctive procedure to traditional assessment of shoulder impingement syndrome. *Journal of Orthopaedic & Sports Physical Therapy* 1995;**22**(5):183-92.

Crenshaw 1966 *{published data only}*

Crenshaw AH, Kilgore WE. Surgical treatment of bicipital tenosynovitis. *Journal of Bone & Joint Surgery - American Volume* 1966;**48**(8):1496-502.

Cullen 2007 *{published data only}*

Cullen DM, Breidahl WH, Janes GC. Diagnostic accuracy of shoulder ultrasound performed by a single operator. *Australasian Radiology* 2007;**51**(3):226-9.

D'Alessandro 2000 *{published data only}*

D'Alessandro DF, Fleischli JE, Connor PM. Superior labral lesions: Diagnosis and management. *Journal of Athletic Training* 2000;**35**(3):286-92.

Dessaur 2008 *{published data only}*

Dessaur WA, Magarey ME. Diagnostic accuracy of clinical tests for superior labral anterior posterior lesions: a systematic review. *Journal of Orthopaedic & Sports Physical Therapy* 2008;**38**(6):341-52.

Deutsch 1997 *{published data only}*

Deutsch A, Altcheck DW, Veltri DM, Potter HG, Warren RF. Traumatic tears of the subscapularis tendon. Clinical diagnosis, magnetic resonance imaging findings, and operative treatment. *American Journal of Sports Medicine* 1997;**25**(1):13-22.

Diehr 2006 *{published data only}*

Diehr S, Ison D, Jamieson B, Oh R. Clinical inquiries. What is the best way to diagnose a suspected rotator cuff tear?. *Journal of Family Practice* 2006;**55**(7):621-4.

Dinnes 2003a *{published data only}*

Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. *Health Technology Assessment* 2003;**7**(29):iii, 1-166.

Ebinger 2008 *{published data only}*

Ebinger N, Magosch P, Lichtenberg S, Habermeyer P. A new SLAP test: the supine flexion resistance test. *Arthroscopy* 2008;**24**(5):500-5.

El Dalati 2005 *{published data only}*

El Dalati G, Castellarin G, Martone E, Ricci M, Vecchini E, Caffarri S, et al. Standard sonography and arthrosonography in the study of rotator cuff tears. *Radiologia Medica* 2005;**110**(5-6):616-22.

Fodor 2009 *{published data only}*

Fodor D, Poanta L, Felea I, Rednic S, Bolosiu H. Shoulder impingement syndrome: correlations between clinical tests and ultrasonographic findings. *Ortopedia Traumatologia Rehabilitacja* 2009;**11**(2):120-6.

Fukuda 1996 *{published data only}*

Fukuda H, Hamada K, Nakajima T, Yamada N, Tomonaga A, Goto M. Partial-thickness tears of the rotator cuff. A clinicopathological review based on 66 surgically verified cases. *International Orthopaedics* 1996;**20**(4):257-65.

Gartsman 2000 *{published data only}*

Gartsman GM. (iii) Partial thickness rotator cuff tears - evaluation and treatment. *Current Orthopaedics* 2000;**14**(3):167-72.

Gerber 1985 *{published data only}*

Gerber C, Terrier F, Ganz R. The role of the coracoid process in the chronic impingement syndrome. *Journal of Bone & Joint Surgery - British Volume* 1985;**67**(5):703-8.

Gerber 1991 *{published data only}*

Gerber C, Krushell RJ. Isolated rupture of the tendon of the subscapularis muscle. Clinical features in 16 cases. *Journal of Bone & Joint Surgery - British Volume* 1991;**73**(3):389-94.

Gerber 1996 *{published data only}*

Gerber C, Hersche O, Farron A. Isolated rupture of the subscapularis tendon. Results of operative repair. *Journal of Bone & Joint Surgery - American Volume* 1996;**78**(7):1015-23.

Gerber 2000 *{published data only}*

Gerber C, Sebesta A. Impingement of the deep surface of the subscapularis tendon and the reflection pulley on the anterosuperior glenoid rim: a preliminary report. *Journal of Shoulder & Elbow Surgery* 2000;**9**(6):483-90.

Giombini 1997 *{published data only}*

Giombini A, Rossi F, Pettrone FA, Dragoni S. Posterosuperior glenoid rim impingement as a cause of shoulder pain in top level waterpolo players. *Journal of Sports Medicine & Physical Fitness* 1997;**37**(4):273-8.

Gschwend 1988 *{published data only}*

Gschwend N, Ivosovic-Radovanovic D, Patte D. Rotator cuff tear--relationship between clinical and anatomopathological

findings. *Archives of Orthopaedic and Trauma Surgery* 1988;**107**(1):7-15.

Hagemann 2004 *{published data only}*

Hagemann G, Rijke AM, Mars M. Shoulder pathoanatomy in marathon kayakers. *British Journal of Sports Medicine* 2004;**38**(4):413-7.

Halbrecht 1999 *{published data only}*

Halbrecht JT, Tirman P. Internal impingement of the shoulder: comparison of findings between the throwing and nonthrowing shoulders of college baseball players. *Arthroscopy* 1999;**15**(3):253-8.

Hammer 2003 *{published data only}*

Hammer W. An excellent test for a SLAP Lesion. *Dynamic Chiropractic* 2003;**21**(19):1-3.

Hamner 2000 *{published data only}*

* Hamner DL, Pink MM, Jobe FW. A modification of the relocation test: arthroscopic findings associated with a positive test. *Journal of Shoulder & Elbow Surgery* 2000;**9**(4):263-7.

Kolbel R. Modification of the relocation test: arthroscopic findings associated with a positive test [comment]. *Journal of Shoulder & Elbow Surgery* 2001;**10**(5):497-8.

Handelberg 1998 *{published data only}*

Handelberg F, Willems S, Shahabpour M, Huskin JP, Kuta J. SLAP lesions: a retrospective multicenter study. *Arthroscopy* 1009;**14**(8):856-62.

Hawkins 1980 *{published data only}*

Hawkins RJ, Kennedy JC. Impingement syndrome in athletes. *American Journal of Sports Medicine* 1980;**8**(3):151-8.

Hayes 2003 *{published data only}*

Hayes KW, Petersen CM. Reliability of classifications derived from Cyriax's resisted testing in subjects with painful shoulders and knees. *Journal of Orthopaedic & Sports Physical Therapy* 2003;**33**(5):235-46.

Hegedus 2008 *{published data only}*

Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT III, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine* 2008;**42**(2):80-92.

Heyworth 2009 *{published data only}*

Heyworth BE, Williams RJ III. Internal impingement of the shoulder. *American Journal of Sports Medicine* 2009;**37**(5):1024-37.

Hughes 2008 *{published data only}*

Hughes PC, Taylor NF, Green RA. Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. *Australian Journal of Physiotherapy* 2008;**54**(3):159-70.

Hurschler 2004 *{published data only}*

Hertel R. Lag signs [comment]. *Journal of Shoulder & Elbow Surgery* 2005;**14**(3):343-4.

* Hurschler C, Walker N, Windhagen H, Hellmers N, Plumhoff P. Evaluation of the lag sign tests for external rotator function of the shoulder. *Journal of Shoulder and Elbow Surgery* 2004;**13**(3):298-304.

Jee 2001 *{published data only}*

Jee WH, McCauley TR, Katz lD, Matheny JM, Ruwe PA, Daigneault JP. Superior labral anterior posterior (SLAP) lesions of the glenoid labrum: reliability and accuracy of MR arthrography for diagnosis. *Radiology* 2001;**218**(1):127-32.

Jobe 1997 *{published data only}*

Jobe CM. Superior glenoid impingement. *Orthopaedic Clinics of North America* 1997;**28**(2):137-43.

Jobe 2000 *{published data only}*

Jobe CM, Coen MJ, Screnar P. Evaluation of impingement syndromes in the overhead-throwing athlete. *Journal of Athletic Training* 2000;**35**(3):293-9.

Johansson 2009 *{published data only}*

Johansson K, Ivarson S. Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. *Manual Therapy* 2009;**14**(2):231-9.

Jones 2007 *{published data only}*

Jones GL, Galluch DB. Clinical assessment of superior glenoid labral lesions: a systematic review. *Clinical Orthopaedics & Related Research* 2007;**455**:45-51.

Keener 2009 *{published data only}*

Keener JD, Brophy RH. Superior labral tears of the shoulder: pathogenesis, evaluation, and treatment. *Journal of the American Academy of Orthopaedic Surgeons* 2009;**17**(10):627-37.

Kibler 1995 *{published data only}*

* Kibler WB. Specificity and sensitivity of the anterior slide test in throwing athletes with superior glenoid labral tears. *Arthroscopy* 1995;**11**(3):296-300.

Wright SA, Hawkins RJ. The anterior slide test for identifying superior gleniod labral tears. Clinical Journal of Sport Medicine 1996; Vol. 6, issue 1:64.

Kibler 2006a *{published data only}*

Kibler WB. Scapular involvement in impingement: signs and symptoms. *Instructional Course Lectures* 2006;**55**:35-43.

Kibler 2006b *{published data only}*

Kibler WB, Sciascia A, Dome D. Evaluation of apparent and absolute supraspinatus strength in patients with shoulder injury using the scapular retraction test. *American Journal of Sports Medicine* 2006;**34**(10):1643-7.

Kim 1999 *{published data only}*

Kim S-H, Ha K-I, Han K-Y. Biceps load test: a clinical test for superior labrum anterior and posterior lesions in shoulders with recurrent anterior dislocations. *American Journal of Sports Medicine* 1999;**27**(3):300-3.

Kim 2004b *{published data only}*

Kim S-H, Park J-H, Park J-S, Oh I. Painful jerk test: a predictor of success in nonoperative treatment of posteroinferior instability of the shoulder. *American Journal of Sports Medicine* 2004;**32**(8):1849-55.

Kim 2005 *{published data only}*

Kim SH, Park JS, Jeong WK, Shin SK. The Kim test: a novel test for posteroinferior labral lesion of the shoulder--a comparison to the jerk test. *American Journal of Sports Medicine* 2005;**33**(8):1188-92.

Kim 2007a *{published data only}*

Kim HA, Kim SH, Seo YI. Ultrasonographic findings of the shoulder in patients with rheumatoid arthritis and comparison with physical examination. *Journal of Korean Medical Science* 2007;**22**(4):660-6.

Kim 2009 *{published data only}*

Kim HM, Teefey SA, Zelig A, Galatz LM, Keener JD, Yamaguchi K. Shoulder strength in asymptomatic individuals with intact compared with torn rotator cuffs. *Journal of Bone & Joint Surgery - American Volume* 2009;**91**(2):289-96.

Kirkley 2002 *{published data only}*

Kirkley A, Litchfield RB, Jackowski DM, Lo IK. The use of the impingement test as a predictor of outcome following subacromial decompression for rotator cuff tendinosis. *Arthroscopy* 2002;**18**(1):8-15.

Koester 2005 *{published data only}*

Koester MC, George MS, Kuhn JE. Shoulder impingement syndrome. *American Journal of Medicine* 2005;**118**(5):452-5.

Lafosse 2007 *{published data only}*

Lafosse L, Reiland Y, Baier GP, Toussaint B, Jost B. Anterior and posterior instability of the long head of the biceps tendon in rotator cuff tears: a new classification based on arthroscopic observations. *Arthroscopy* 2007;**23**(1):73-80.

Le Huec 1996 *{published data only}*

Le Huec JC, Schaeverbeke T, Moinard M, Kind M, Diard F, Dehais J, et al. Traumatic tear of the rotator interval. *Journal of Shoulder & Elbow Surgery* 1996;**5**(1):41-6.

Leroux 1995 *{published data only}*

Leroux JL, Thomas E, Bonnel F, Blotman F. Diagnostic value of clinical tests for shoulder impingement syndrome. *Revue Du Rhumatisme [English Edition]* 1995;**62**(6):423-8.

Lewis 2005 *{published data only}*

Lewis JS, Wright C, Green A. Subacromial impingement syndrome: the effect of changing posture on shoulder range of movement. *Journal of Orthopaedic & Sports Physical Therapy* 2005;**35**(2):72-87.

Lewis 2007 *{published data only}*

Lewis JS, Valentine RE. The pectoralis minor length test: A study of the intra-rater reliability and diagnostic accuracy in subjects with and without shoulder symptoms. *BMC Musculoskeletal Disorders* 2007;**8**(64).

Lewis 2009 *{published data only}*

Lewis JS. Rotator cuff tendinopathy/subacromial impingement syndrome: Is it time for a new method of assessment?. *British Journal of Sports Medicine* 2009;**43**(4):259-64.

Litaker 2000 *{published data only}*

* Litaker D, Pioro M, El Bilbeisi H, Brems J. Returning to the bedside: using the history and physical examination to identify rotator cuff tears. *Journal of the American Geriatric Society* 2000;**48**(12):1633-7.

Riddle D. A set of three clinical tests can detect the presence or absence of rotator cuff tears. *Australian Journal of Physiotherapy* 2001;**47**(4):297-8.

Liu 1996a *{published data only}*

* Liu SH, Henry MH, Nuccion S, Shapiro MS, Dorey F. Diagnosis of glenoid labral tears. A comparison between magnetic resonance imaging and clinical examinations. *American Journal of Sports Medicine* 1996;**24**(2):149-54.

Seeger LL, Yao L, Gold RH. Diagnosis of glenoid labral tears: a comparison between magnetic resonance imaging and clinical examinations [comment]. *American Journal of Sports Medicine* 1997;**25**(1):141-4.

Lo 2004 *{published data only}*

Lo IKY, Nonweiler B, Woolfrey M, Litchfiled R, Kirkley A. An evaluation of the apprehension, relocation, and surprise tests for anterior shoulder instability. *American Journal of Sports Medicine* 2004;**32**(2):301-7.

Luime 2004 *{published data only}*

Luime JJ, Verhagen AP, Miedema HS, Kuiper JI, Burdorf A, Verhaar JAN, et al. Does this patient have an instability of the shoulder or a labrum lesion?. *JAMA* 2004;**292**(16):1989-99.

Lyons 1992 *{published data only}*

Lyons AR, Tomlinson JE. Clinical diagnosis of tears of the rotator cuff. *Journal of Bone & Joint Surgery - British Volume* 1992;**74**(3):414-5.

Lyons 2005 *{published data only}*

Lyons RP, Green A. Subscapularis tendon tears. *Journal of the American Academy of Orthopaedic Surgeons* 2005;**13**(5):353-63.

Maffet 1995 *{published data only}*

Maffet MW, Gartsman GM, Moseley B. Superior labrum-biceps tendon complex lesions of the shoulder. *American Journal of Sports Medicine* 1995;**23**(1):93-8.

Malhi 2005 *{published data only}*

Malhi AM, Khan R. Correlation between clinical diagnosis and arthroscopic findings of the shoulder. *Postgraduate Medical Journal* 2005;**81**(960):657-9.

Maman 2009 *{published data only}*

Maman E, Harris C, White L, Tomlinson G, Shashank M, Boynton E. Outcome of nonoperative treatment of symptomatic rotator cuff tears monitored by magnetic resonance

imaging. *Journal of Bone & Joint Surgery - American Volume* 2009;**91**(8):1898-906.

Matava 2005 *{published data only}*

Matava MJ, Purcell DB, Rudzki JR. Partial-thickness rotator cuff tears. *American Journal of Sports Medicine* 2005;**33**(9):1405-17.

McCabe 2005 *{published data only}*

McCabe RA, Nicholas SJ, Montgomery KD, Finneran JJ, McHugh MP. The effect of rotator cuff tear size on shoulder strength and range of motion. *Journal of Orthopaedic & Sports Physical Therapy* 2005;**35**(3):130-5.

McFarland 1999 *{published data only}*

McFarland EG, Hsu C-Y, Neira C, O'Neil O. Internal impingement of the shoulder: A clinical and arthroscopic analysis. *Journal of Shoulder & Elbow Surgery* 1999;**8**(5):458-60.

McFarland 2001 *{published data only}*

McFarland EG, Neira CA, Gutierrez MI, Cosgarea AJ, Magee M. Clinical significance of the arthroscopic drive-through sign in shoulder surgery. *Arthroscopy* 2001;**17**(1):38-43.

McFarland 2008 *{published data only}*

McFarland EG, Tanaka MJ, Papp DF. Examination of the shoulder in the overhead and throwing athlete. *Clinics in Sports Medicine* 2008;**27**(4):553-78.

McFarland 2009 *{published data only}*

McFarland EG, Tanaka MJ, Garzon-Muvdi J, Jia X, Petersen SA. Clinical and imaging assessment for superior labrum anterior and posterior lesions. [Review] [51 refs]. *Current Sports Medicine Reports* 2009;**8**(5):234-9.

Meister 2004 *{published data only}*

Meister K, Buckley B, Batts J. The posterior impingement sign: diagnosis of rotator cuff and posterior labral tears secondary to internal impingement in overhand athletes. *American Journal of Orthopedics* 2004;**33**(8):412-5.

Meserve 2009 *{published data only}*

Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utility for assessing superior labral anterior posterior lesions. *American Journal of Sports Medicine* 2009;**37**(11):2252-8.

Miller 2008a *{published data only}*

Miller D, Frost A, Hall A, Barton C, Bhoora I, Kathuria V. A 'one-stop clinic' for the diagnosis and management of rotator cuff pathology: Getting the right diagnosis first time. *International Journal of Clinical Practice* 2008;**62**(5):750-3.

Mimori 1999 *{published data only}*

Mimori K, Muneta T, Nakagawa T, Shinomiya K. A new pain provocation test for superior labral tears of the shoulder. *American Journal of Sports Medicine* 1999;**27**(2):137-42.

Mirkovic 2005 *{published data only}*

Mirkovic M, Green R, Taylor N, Perrott M. Accuracy of clinical tests to diagnose superior labral anterior and posterior (SLAP) lesions. *Physical Therapy Reviews* 2005;**10**(1):5-14.

Mitchell 2005 *{published data only}*

Mitchell C, Adebajo A, Hay E, Carr A. Shoulder pain: diagnosis and management in primary care. *BMJ* 2005;**331**(7525):1124-8.

Moosikasuwan 2005 *{published data only}*

Moosikasuwan JB, Miller TT, Burke BJ. Rotator cuff tears: clinical, radiographic, and US findings. *Radiographics* 2005;**25**(6):1591-607.

Morgan 1998 *{published data only}*

Morgan CD, Burkhart SS, Palmeri M, Gillespie M. Type II SLAP lesions: three subtypes and their relationships to superior instability and rotator cuff tears. *Arthroscopy* 1998;**14**(6):553-65.

Morrison 1997 *{published data only}*

Morrison DS, Frogameni AD, Woodworth P. Non-operative treatment of shoulder impingement syndrome. *Journal of Bone and Joint Surgery - American Volume* 1997;**79**(5):732-7.

Morrissey 2005 *{published data only}*

Morrissey M. Shoulder diagnosis validation by motion analysis and laxity testing (protocol). National Research Register 2005 Issue 1 2005.

Munro 2009 *{published data only}*

Munro W, Healy R. The validity and accuracy of clinical tests used to detect labral pathology of the shoulder--a systematic review. *Manual Therapy* 2009;**14**(2):119-30.

Murrell 2001 *{published data only}*

* Murrell GA, Walton JR. Diagnosis of rotator cuff tears. *Lancet* 2001;**357**:769-70.

Riddle D. Individual tests from the history and physical examination are inaccurate in diagnosing rotator cuff tears. *Australian Journal of Physiotherapy* 2001;**47**(4):297.

Myers 2005 *{published data only}*

Myers TH, Zemanovic JR, Andrews JR. The resisted supination external rotation test: a new test for the diagnosis of superior labral anterior posterior lesions. *American Journal of Sports Medicine* 2005;**33**(9):1315-20.

Neer 1972 *{published data only}*

Neer CS. Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. *Journal of Bone & Joint Surgery - American Volume* 1972;**54**(1):41-50.

Neer 1977 *{published data only}*

Neer CS, Welsh RP. The shoulder in sports. *Orthopedic Clinics of North America* 1977;**8**(3):583-91.

Neer 1983a *{published data only}*

Neer CS 2nd. Impingement lesions. *Clinical Orthopaedics & Related Research* 1983;**(173)**:70-7.

Neri 2009 *{published data only}*

Neri BR, Chan KW, Kwon YW. Management of massive and irreparable rotator cuff tears. *Journal of Shoulder & Elbow Surgery* 2009;**18**(5):808-18.

Nomden 2009 *{published data only}*

Nomden JG, Slagers AJ, Bergman GJ, Winters JC, Kropmans TJ, Dijkstra PU. Interobserver reliability of physical examination of shoulder girdle. *Manual Therapy* 2009;**14**(2):152-9.

Norregaard 2002 *{published data only}*

Norregaard J, Krogsgaard MR, Lorenzen T, Jensen EM. Diagnosing patients with longstanding shoulder joint pain. *Annals of the Rheumatic Diseases* 2002;**61**(7):646-9.

O'Connor 2005 *{published data only}*

O'Connor PJ, Rankine J, Gibbon WW, Richardson A, Winter F, Miller JH. Interobserver variation in sonography of the painful shoulder. *Journal of Clinical Ultrasound* 2005;**33**(2):53-6.

Odom 2001 *{published data only}*

Odom CJ, Taylor AB, Hurd CE, Denegar CR. Measurement of scapular asymmetry and assessment of shoulder dysfunction using the Lateral Scapular Slide Test: a reliability and validity study. *Physical Therapy* 2001;**81**(2):799-809.

Oh 2007 *{published data only}*

Oh LS, Wolf BR, Hall MP, Levy BA, Marx RG. Indications for rotator cuff repair: a systematic review. *Clinical Orthopaedics & Related Research* 2007;**455**:52-63.

Olmsted 2003 *{published data only}*

Olmsted LC, Denegar C. Physical examination of the shoulder: Considerations of sensitivity and specificity. *Athletic Therapy Today* 2003;**8**(5):25-31.

Osbahr 2006 *{published data only}*

Osbahr DC, Murrell GA. The rotator cuff functional index. *American Journal of Sports Medicine* 2006;**34**(6):956-60.

Ostor 2005 *{published data only}*

Ostor AJ, Richards CA, Prevost AT, Speed CA, Hazleman BL. Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology* 2005;**44**(6):800-5.

Pandya 2008 *{published data only}*

Pandya NK, Colton A, Webner D, Sennett B, Huffman GR. Physical examination and magnetic resonance imaging in the diagnosis of superior labrum anterior-posterior lesions of the shoulder: a sensitivity analysis. *Arthroscopy* 2008;**24**(3):311-7.

Perez-Palomares 2009 *{published data only}*

Perez-Palomares S, Olivan-Blazquez B, Arnal-Burro AM, Mayoral-Del Moral O, Gaspar-Calvo E, de-la-Torre-Beldarrain ML, et al. Contributions of myofascial pain in diagnosis and treatment of shoulder pain. A randomized control trial. *BMC Musculoskeletal Disorders* 2009;**10**(92).

Piasecki 2008 *{published data only}*

Piasecki DP, Nicholson GP. Tears of the subscapularis tendon in athletes--diagnosis and repair techniques. *Clinics in Sports Medicine* 2008;**27**(4):731-45.

Pisan 2000 *{published data only}*

Pisan M, Gerber C. Clinical examination of the rotator cuff. *Current Orthopaedics* 2000;**14**(3):155-60.

Polimeni 2003 *{published data only}*

Polimeni V, Panuccio A, Furfari P, Crupi D, BarrecaG, Forgione C, et al. Preliminary study on the efficacy of various rehabilitation therapies for shoulder pain. *Europa Medicophysica* 2003;**39**(1):59-63.

Polsky 2006 *{published data only}*

Polsky MB, Stein J. Evaluation and management of superior labral anterior posterior tears in the overhead athlete. *Current Opinion in Orthopaedics* 2006;**17**(2):160-3.

Pugh 2009 *{published data only}*

Pugh S, Callaghan M, Hogg K. Towards evidence based emergency medicine: best BETs from the Manchester Royal Infirmary. BET 1: which is the best clinical test for diagnosing a full thickness rotator cuff tear?.. *Emergency Medicine Journal* 2009;**26**(12):881.

Rao 2003 *{published data only}*

Rao AG, Kim TK, Chronopoulos E, McFarland EG. Anatomical variants in the anterosuperior aspect of the glenoid labrum: a statistical analysis of seventy-three cases. *Journal of Bone and Joint Surgery - American Volume* 2003;**85**(4):635-59.

Razmjou 2006 *{published data only}*

Razmjou H, Haines T, Holtby R. Diagnostic and therapeutic decision-making: Exploring the role of pretest probability in patients with rotator cuff pathology. *Physiotherapy Canada* 2006;**58**(13):196-204.

Read 1998 *{published data only}*

Read JW, Perko M. Shoulder ultrasound: diagnostic accuracy for impingement syndrome, rotator cuff tear, and biceps tendon pathology. *Journal of Shoulder and Elbow Surgery* 1998;**7**(3):264-71.

Rhee 2005a *{published data only}*

Rhee YG, Lee DH, Lim CT. Unstable isolated SLAP lesion: Clinical presentation and outcome of arthroscopic fixation. *Arthroscopy - Journal of Arthroscopic and Related Surgery* 2005;**21**(9):1099.

Rowan 2007 *{published data only}*

Rowan KR, Andrews G, Spielmann A, Leith J, Forster BB. MR shoulder arthrography in patients younger than 40 years of age: frequency of rotator cuff tear versus labroligamentous pathology.[Erratum appears in Australas Radiol. 2007;51(4):402 Note: Forster, B B [added]]. *Australasian Radiology* 2007;**51**(3):257-9.

Ryu 2002 *{published data only}*

Ryu RK, Dunbar WH5th, Kuhn JE, McFarland EG, Chronopoulos E, Kim TK. Comprehensive evaluation and treatment of the shoulder in the throwing athlete. *Arthroscopy* 2002;**18**(9 Suppl 2):70-89.

Sandenbergh 2006 *{published data only}*

Sandenbergh R, Marais A. Reliability of Diagnostic Tests in Rotator Cuff Muscle Pathology. *South African Journal of Physiotherapy* 2006;**62**(3):6-12.

Savoie 2001 *{published data only}*

Savoie FH, Field LD, Atchinson S. Anterior superior instability with rotator cuff tearing: SLAC lesion. *Orthopaedic Clinics of North America* 2001;**32**(2):457-61.

Scheibel 2005 *{published data only}*

Scheibel M, Magosch P, Pritsch M, Lichtenberg S, Habermeyer P. The belly-off sign: A new clinical diagnostic sign for subscapularis lesions. *Arthroscopy* 2005;**21**(10):1229-35.

Schellingerhout 2008 *{published data only}*

Schellingerhout JM, Verhagen AP, Thomas S, Koes BW. Lack of uniformity in diagnostic labeling of shoulder pain: time for a different approach. *Manual Therapy* 2008;**13**(6):478-83.

Sileo 2006 *{published data only}*

Sileo M, Panchal A, Ruotolo CJ, Teitelbaum H, Lindenfeld TN, Wells L. A statistical correlation of shoulder physical examination maneuvers and arthroscopic findings [abstract]. Annual Meeting of the American Academy of Orthopaedic Surgeons; 2006 March 24-26: Chicago (IL). AAOS On-Line Service, 2006.

Silva 2008 *{published data only}*

Silva L, Andreu JL, Munoz P, Pastrana M, Millan I, Sanz J, et al. Accuracy of physical examination in subacromial impingement syndrome. *Rheumatology* 2008;**47**(5):679-83.

Skedros 2007 *{published data only}*

Skedros JG, Pitts TC. Temporal variations in a modified Neer impingement test can confound clinical interpretation. *Clinical Orthopaedics & Related Research* 2007;**460**:130-6.

Smith 2000 *{published data only}*

Smith KL, Harryman DT 2nd, Antoniou J, Campbell B, Sidles JA, Matsen FA 3rd. A prospective, multipractice study of shoulder function and health status in patients with documented rotator cuff tears. *Journal of Shoulder & Elbow Surgery* 2000;**9**(5):395-402.

Snyder 1990 *{published data only}*

Snyder S J, Karzel R P, Del Pizzo W, Ferkel R D, Friedman M J. SLAP lesions of the shoulder. *Arthroscopy* 1990;**6**(4):274-9.

Soncini 2000 *{published data only}*

Soncini G, Costantino C. Shoulder pain in baseball players. *Journal of Sports Traumatology & Related Research* 2000;**22**(2):45-57.

Sorensen 2007 *{published data only}*

Sorensen AK, Bak K, Krarup AL, Thune CH, Nygaard M, Jorgensen U, et al. Acute rotator cuff tear: do we miss the early diagnosis? A prospective study showing a high incidence of rotator cuff tears after shoulder trauma. *Journal of Shoulder & Elbow Surgery* 2007;**16**(2):174-80.

Sorohan 2009 *{published data only}*

Sorohan A, Mc Creesh K. Assessment of subacromial impingement syndrome in the primary care setting: development of an evidence based clinical framework. *Physiotherapy Ireland* 2009;**30**(1):16-22.

Tate 2008 *{published data only}*

Tate AR, McClure P, Kareha S, Irwin D. Effect of the scapula reposition text on shoulder impingement symptoms and elevation strength in overhead athletes. *Journal of Orthopaedic & Sports Physical Therapy* 2008;**38**(1):4-11.

Tennent 2003a *{published data only}*

Tennent TD, Beach WR, Meyers JF. A review of the special tests associated with shoulder examination: part I: the rotator cuff tests. *American Journal of Sports Medicine* 2003;**31**(1):154-60.

Tennent 2003b *{published data only}*

Tennent TD, Beach WR, Meyers JF. A review of the special tests associated with shoulder examination. Part II: laxity, instability, and superior labral anterior and posterior (SLAP) lesions. *American Journal of Sports Medicine* 2003;**31**(2):301-7.

Trantalis 2008 *{published data only}*

Trantalis JN, Lo IKY. Superior labral anterior-posterior (SLAP) tears: Recent advances and outcomes. *Current Orthopaedic Practice* 2008;**19**(5):530-4.

Tyler 2000 *{published data only}*

Tyler TF, Nicholas SJ, Roy T, Gleim GW. Quantification of posterior capsule tightness and motion loss in patients with shoulder impingement. *American Journal of Sports Medicine* 2000;**28**(5):668-73.

Vanderbeck 2007 *{published data only}*

Vanderbeck J, Fenlin J. Shoulder: what the orthopaedic doctor needs to know. *Seminars in Musculoskeletal Radiology* 2007;**11**(1):57-65.

Walch 1998 *{published data only}*

Walch G, Boulahia A, Calderone S, Robinson AHN. The 'dropping' and 'hornblower's' signs in evaluation of rotator-cuff tears. *Journal of Bone & Joint Surgery - British Volume* 1998;**80**(4):624-8.

Walsworth 2008 *{published data only}*

Walsworth MK, Doukas WC, Murphy KP, Mielcarek BJ, Michener LA. Reliability and diagnostic accuracy of history and physical examination for diagnosing glenoid labral tears. *American Journal of Sports Medicine* 2008;**36**(1):162-8.

Walton 2004 *{published data only}*

Walton J, Mahajan S, Paxinos A, Marshall J, Bryant C, Shnier R, et al. Diagnostic values of tests for acromioclavicular joint pain. *Journal of Bone & Joint Surgery - American Volume* 2004;**86**(4):807-12.

Walton 2008 *{published data only}*

Kemp K, Sheps DM. Clinical tests to identify SLAP lesions: A meta-analysis. Clinical Journal of Sport Medicine 2009; Vol. 19, issue 4:339-40.

* Walton DM, Sadi J. Identifying SLAP lesions: a meta-analysis of clinical tests and exercise in clinical reasoning. *Physical Therapy in Sport* 2008;**9**(4):167-76.

Wang 2000 *{published data only}*

Wang JC, Horner G, Brown ED, Shapiro MS. The relationship between acromial morphology and conservative treatment of patients with impingement syndrome. *Orthopedics* 2000;**23**(6):557-9.

Watson 1989 *{published data only}*

Watson M. Rotator cuff function in the impingement syndrome. *Journal of Bone & Joint Surgery - British Volume* 1989;**71**(3):361-6.

Wilk 2005 *{published data only}*

Wilk KE, Reinold MM, Dugas JR, Arrigo CA, Moser MW, Andrews JR. Current concepts in the recognition and treatment of superior labral (SLAP) lesions. *Journal of Orthopaedic & Sports Physical Therapy* 2005;**35**(5):273-91.

Wnorowski 1997 *{published data only}*

Wnorowski DC, Levinsohn EM, Chamberlain BC, McAndrew DL. Magnetic resonance imaging assessment of the rotator cuff: is it really accurate?. *Arthroscopy* 1997;**13**(6):710-9.

Wolff 2006 *{published data only}*

Wolff AB, Sethi P, Sutton KM, Covey AS, Magit DP, Medvecky M. Partial-thickness rotator cuff tears. *Journal of the American Academy of Orthopaedic Surgeons* 2006;**14**(13):715-25.

Yang 2006 *{published data only}*

Yang JL, Lin JJ. Reliability of function-related tests in patients with shoulder pathologies. *Journal of Orthopaedic & Sports Physical Therapy* 2006;**36**(8):572-6.

References to studies awaiting assessment

Gill 2007 *{published data only}*

Gill HS, El Rassi G, Bahk MS, Castillo RC, McFarland EG. Physical examination for partial tears of the biceps tendon. *American Journal of Sports Medicine* 2007;**35**(8):1334-40.

Jia 2008 *{published data only}*

Jia X, Ji JH, Petersen SA, Keefer J, McFarland EG. Clinical evaluation of the shoulder shrug sign. *Clinical Orthopaedics & Related Research* 2008;**466**(11):2813-9.

Jia 2009 *{published data only}*

Jia X, Petersen SA, Khosravi AH, Almareddi V, Pannirselvam V, McFarland EG. Examination of the shoulder: The past, the present, and the future. *Journal of Bone and Joint Surgery - American Volume* 2009;**91**(Suppl 6):10-8.

Kelly 2010 *{published data only}*

Kelly SM, Brittle N, Allen GM. The value of physical tests for subacromial impingement syndrome: a study of diagnostic accuracy. *Clinical Rehabilitation* 2010;**24**(2):149-58.

Kim 2003a *{published data only}*

Kim TK, Queale WS, Cosgarea AJ, McFarland EG. Clinical features of the different types of SLAP lesions: an analysis of one hundred and thirty-nine cases. *Journal of Bone & Joint Surgery - American Volume* 2003;**85**(1):66-71.

Kim 2003b *{published data only}*

Kim TK, Rauh PB, McFarland EG. Partial tears of the subscapularis tendon found during arthroscopic procedures on the shoulder: a statistical analysis of sixty cases. *American Journal of Sports Medicine* 2003;**31**(5):744-50.

Kim 2004a *{published data only}*

Kim TK, McFarland EG. Internal impingement of the shoulder in flexion. *Clinical Orthopaedics & Related Research* 2004;**(421)**:112-9.

McFarland 2002 *{published data only}*

McFarland EG, Kim TK, Savino RM. Clinical assessment of three common tests for superior labral anterior-posterior lesions. *American Journal of Sports Medicine* 2002;**30**(6):810-5.

Nanda 2008 *{published data only}*

Nanda R, Gupta S, Kanapathipillai P, Liow RYL, Rangan A. An assessment of the inter examiner reliability of clinical tests for subacromial impingement and rotator cuff integrity. *European Journal of Orthopaedic Surgery and Traumatology* 2008;**18**(7):495-500.

Park 2005 *{published data only}*

Park HB, Yokota A, Gill HS, Rasi GE, McFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. *Journal of Bone & Joint Surgery - American Volume* 2005;**87**(7):1446-55.

Additional references

Alqunaee 2012

Alqunaee M, Galvin R, Fahey T. Diagnostic accuracy of clinical tests for subacromial impingement syndrome: a systematic review and meta-analysis. *Archives of physical medicine and rehabilitation* 2012;**93**(2):229-36.

Altman 1991

Altman DG. Practical Statistics for Medical Research. London: Chapman and Hall, 1991.

Barth 2006a

Barth JR, Burkhart SS, de Beer JF. The bear-hug test: a new and sensitive test for diagnosing a subscapularis tear. *Arthroscopy* 2006;**22**(10):1076-84.

Beaudreuil 2009a

Beaudreuil J, Nizard R, Thomas T, Peyre M, Liotard JP, Boileau P, et al. Contribution of clinical tests to the diagnosis of rotator cuff disease: a systematic literature review. *Joint, Bone, Spine: Revue du Rhumatisme* 2009;**76**(1):15-9.

Berg 1998a

Berg EE, Ciullo JV. A clinical test for superior glenoid labral or 'SLAP' lesions. *Clinical Journal of Sport Medicine* 1998;**8**(2):121-3.

Blumenthal 2003

Blumenthal S, Nadig M, Gerber C, Borgeat A. Severe airway obstruction during arthroscopic shoulder surgery. *Anesthesiology* 2003;**99**(6):1455-6.

Boardman 1999

Boardman ND 3rd, Cofield RH. Neurologic complications of shoulder surgery. *Clinical Orthopaedics and Related Research* 1999;**(368)**:44-53.

Borgeat 2001

Borgeat A, Ekatodramis G, Kalberer F, Benz C. Acute and nonacute complications associated with interscalene block and shoulder surgery: a prospective study. *Anesthesiology* 2001;**95**(4):875-80.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou Paul P, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ (Clinical research ed.)* 2003;**326**(7379):41-4.

Broadhurst 2004

Broadhurst NA, Gialamas A, McElroy HJ, Beilby JJ. How do Australian GPs manage shoulder dysfunction?. *Australian Family Physician* 2004;**33**(10):861-2.

Calvert 2009a

Calvert E, Chambers GK, Regan W, Hawkins RH, Leith JM. Special physical examination tests for superior labrum anterior posterior shoulder tears are clinically limited and invalid: a diagnostic systematic review. *Journal of Clinical Epidemiology* 2009;**62**(5):558-63.

Codman 1934

Codman E A. Rupture of the supraspinatus tendon and other lesions in or about the subacromial bursa. Boston: Privately published, 1934.

Crenshaw 1966

Crenshaw AH, Kilgore WE. Surgical treatment of bicipital tenosynovitis. *Journal of Bone & Joint Surgery - American Volume* 1966;**48**(8):1496-502.

Cyriax 1982

Cyriax JH. Textbook of orthopaedic medicine: Diagnosis of soft tissue lesions. 8th Edition. Vol. **One**, London: Balliere Tindall, 1982.

Cyriax 1984

Cyriax J. Textbook of orthopaedic medicine volume two: Treatment by manipulation, massage and injection. London: Balliere Tindall, 1984.

de Vet 2005

De Vet HC, Riphagen I, Pai M, Aertgeerts B, Pewnser D. Chapter 5: Searching for studies. Cochrane Diagnostic Reviewers' Handbook. Version 0.3 [updated July 2005]. The Cochrane Collaboration, 2005.

Dessaur 2008a

Dessaur WA, Magarey ME. Diagnostic accuracy of clinical tests for superior labral anterior posterior lesions: a systematic review. *Journal of Orthopaedic & Sports Physical Therapy* 2008;**38**(6):341-52.

Dinnes 2003

Dinnes J, Loveman E, McIntyre L, Waugh N. The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review. *Health Technology Assessment* 2003;**7**(29):iii, 1-166.

Fowler 2010

Fowler EM, Horsley IG, Rolf CG. Clinical and arthroscopic findings in recreationally active patients. *Sports medicine, arthroscopy, rehabilitation, therapy & technology: SMARTT* 2010;**2**:2-2.

Fukuda 2003

Fukuda H. The management of partial-thickness tears of the rotator cuff. *Journal of Bone and Joint Surgery - British Volume* 2003;**85**(1):3-11.

Gerber 1991a

Gerber C, Krushell RJ. Isolated rupture of the tendon of the subscapularis muscle. Clinical features in 16 cases. *Journal of Bone and Joint Surgery - British Volume* 1991;**73**(3):389-94.

Gerber 1996

Gerber C, Hersche O, Farron A. Isolated rupture of the subscapularis tendon. *Journal of Bone and Joint Surgery - American Volume* 1996;**78**(7):1015-23.

Gilcreest 1936

Gilcreest EL. Dislocation and elongation of the long head of the biceps brachii: an analysis of six cases. *Annals of Surgery* 1936;**104**(1):118-38.

Gilecreest 1939

Gilecreest EL, Albi P. Unusual lesions of muscles and tendons of the shoulder girdle and upper arm. *Surgery, Gynecology & Obstetrics* 1939;**68**:903-17.

Gill 2005

Gill C J, Sabin L, Schmid C H. Why clinicians are natural bayesians. *BMJ (Clinical research ed.)* 2005;**330**(7499):1080-3.

Green 2003

Green S, Buchbinder R, Hetrick SE. Physiotherapy interventions for shoulder pain. *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD004258]

Gross 1997

Gross ML, Distefano MC. Anterior release test: A new test for occult shoulder instability. *Clinical Orthopaedics & Related Research* 1997;**1**(339):105-8.

Gumina 2008a

Gumina S, Bertino A, Di Giorgio G, Postacchini F. A new test of resistance in the diagnosis of postero-superior rotator cuff tears. *Chirurgia Degli Organi di Movimento* 2008;**91**(2):85-6.

Hamner 2000

Hamner DL, Pink MM, Jobe FW. A modification of the relocation test: arthroscopic findings associated with a positive test. *Journal of Shoulder and Elbow Surgery* 2000;**9**(4):263-7.

Hanchard 2004

Hanchard NCA, Cummins J, Hanson C. Evidence-based clinical guidelines for the diagnosis, assessment and physiotherapy management of shoulder impingement syndrome. London: Chartered Society of Physiotherapy London, 2004.

Hanchard 2005

Hanchard NC, Howe TE, Gilbert MM. Diagnosis of shoulder pain by history and selective tissue tension: agreement between assessors. *Journal of Orthopaedic and Sports Physical Therapy* 2005;**35**(3):147-53.

Hawkins 1980

Hawkins RJ, Kennedy JC. Impingement syndrome in athletes. *American Journal of Sports Medicine* 1980;**8**(3):151-8.

Hegedus 2008a

Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT III, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *British Journal of Sports Medicine* 2008;**42**(2):80-92.

Hertel 1996a

Hertel R, Ballmer FT, Lombert SM, Gerber C. Lag signs in the diagnosis of rotator cuff rupture. *Journal of Shoulder and Elbow Surgery* 1996;**5**(4):307-13.

Hughes 2008a

Hughes PC, Taylor NF, Green RA. Most clinical tests cannot accurately diagnose rotator cuff pathology: a systematic review. *Australian Journal of Physiotherapy* 2008;**54**(3):159-70.

Jaeschke 1994

Jaeschke R, Guyatt GH. 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**:703-7.

Jobe 1982

Jobe F W, Moynes D R. Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. *The American journal of sports medicine* 1982;**10**(6):336-9.

Jobe 1983

Jobe F W, Jobe C M. Painful athletic injuries of the shoulder. *Clinical Orthopaedics & Related Research* 1983;**(173)**:117-24.

Jobe 1989

Jobe FW, Kvitne RS, Giangarra CE. Shoulder pain in the overhand or throwing athlete. The relationship of anterior instability and rotator cuff impingement. *Orthopaedic review* 1989;**18**(9):963-75.

Jobe 1995

Jobe C M. Posterior superior glenoid impingement: expanded spectrum. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association* 1995;**11**(5):530-6.

Jobe 1996

Jobe CM. Superior glenoid impingement. Current concepts. *Clinical Orthopaedics and Related Research* 1996;**(330)**:98-107.

Jones 2007a

Jones GL, Galluch DB. Clinical assessment of superior glenoid labral lesions: a systematic review. *Clinical Orthopaedics & Related Research* 2007;**455**:45-51.

Kelly 1996

Kelly BT, Kadrmas WR, Speer KP. The manual muscle examination for rotator cuff strength. An electromyographic investigation. *American Journal of Sports Medicine* 1996;**24**(5):581-8.

Kessel 1977

Kessel L, Watson M. The painful arc syndrome: clinical classification as a guide to management. *Journal of Bone and Joint Surgery - British Volume* 1977;**59**(2):166-72.

Kibler 1995a

Kibler WB. Specificity and sensitivity of the anterior slide test in throwing athletes with superior glenoid labral tears. *Arthroscopy* 1995;**11**(3):296-300.

Kim 2001

Kim SH, Ha KI, Ahn JH, Kim SH, Choi HJ. Biceps load test II: A clinical test for SLAP lesions of the shoulder. *Arthroscopy* 2001;**17**(2):160-4.

Krishnan 2004

Krishnan SG, Hawkins RJ, Bokor DJ. Clinical evaluation of shoulder problems. In: Rockwood Jr CA, Matsen FA, Wirth MA, Lippitt SB editor(s). The Shoulder. 3. Philadelphia: Saunders, 2004:145-85.

Liu 1996b

Liu SH, Henry MH, Nuccion SL. A prospective evaluation of a new physical examination in predicting glenoid labral tears. *American Journal of Sports Medicine* 1996;**24**(6):721-5.

Liu 1996c

Liu SH, Henry MH, Nuccion S, Shapiro MS, Dorey F. Diagnosis of glenoid labral tears. A comparison between magnetic resonance imaging and clinical examinations. *American Journal of Sports Medicine* 1996;**24**(2):149-54.

Lo 2003

Lo IKY, Burkhart SS. The comma sign: An arthroscopic guide to the torn subscapularis tendon. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association* 2003;**19**(3):334-7.

Longo 2011

Longo UG, Berton A, Ahrens PM, Maffulli N, Denaro V. Clinical tests for the diagnosis of rotator cuff disease. *Sports Medicine and Arthroscopy Review* 2011;**19**(3):266-78.

Luime 2004a

Luime JJ, Koes BW, Hendriksen IJ, Burdorf A, Verhagen AP, Miedema HS, et al. Prevalence and incidence of shoulder pain in the general population; a systematic review. *Scandinavian Journal of Rheumatology* 2004;**33**(2):73-81.

Luime 2004b

Luime JJ, Verhagen AP, Miedema HS, Kuiper JI, Burdorf A, Verhaar JA, et al. Does this patient have an instability of the shoulder or a labrum lesion?. *JAMA* 2004;**292**(16):1989-99.

MacDonald 2000a

MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. *Journal of Shoulder and Elbow Surgery* 2000;**9**(4):299-301.

Meister 2004

Meister K, Buckley B, Batts J. The posterior impingement sign: diagnosis of rotator cuff and posterior labral tears secondary to internal impingement in overhand athletes. *American Journal of Orthopedics* 2004;**33**(8):412-5.

Meserve 2009a

Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utility for assessing superior labral anterior posterior lesions. *American Journal of Sports Medicine* 2009;**37**(11):2252-8.

Milgrom 1995

Milgrom C, Schaffler M, Gilbert S, van Holsbeeck M. Rotator-cuff changes in asymptomatic adults. The effect of age, hand dominance and gender. *Journal of Bone and Joint Surgery - British Volume* 1995;**77**(2):296-8.

Mimori 1999a

Mimori K, Muneta T, Nakagawa T, Shinomiya K. A new pain provocation test for superior labral tears of the shoulder. *American Journal of Sports Medicine* 1999;**27**(2):137-42.

Mohtadi 2004

Mohtadi NG, Jager FL, Sasyniuk TM, Hollinshead RM, Fick GH. The reliability between surgeons comparing arthroscopic and video evaluation of patients with shoulder impingement syndrome. *Arthroscopy* 2004;**20**(10):1055-62.

Morgan 1998a

Morgan CD, Burkhart SS, Palmeri M, Gillespie M. Type II SLAP lesions: three subtypes and their relationships to superior instability and rotator cuff tears. *Arthroscopy* 1998;**14**(6):553-65.

Morrison 1997a

Morrison DS, Frogameni AD, Woodworth P. Non-operative treatment of subacromial impingement syndrome. *Journal of Bone and Joint Surgery - American Volume* 1997;**79**(5):732-7.

Munro 2009a

Munro W, Healy R. The validity and accuracy of clinical tests used to detect labral pathology of the shoulder--a systematic review. *Manual Therapy* 2009;**14**(2):119-30.

Naredo 2002

Naredo E, Aguado P, De Miguel E, Uson J, Mayordomo L, Gijon-banos J, et al. Painful shoulder: comparison of physical examination and ultrasonographic findings. *Annals of the Rheumatic Diseases* 2002;**61**(2):132-6.

Neer 1972a

Neer CS. Anterior acromioplasty for the chronic impingement syndrome in the shoulder: a preliminary report. *Journal of Bone & Joint Surgery - American Volume* 1972;**54**(1):41-50.

Neer 1977

Neer CS 2nd, Welsh RP. The shoulder in sports. *Orthopedic Clinics of North America* 1977;**8**(3):583-91.

Neer 1983

Neer CS 2nd. Impingement lesions. *Clinical Orthopaedics and Related Research* 1983;**(173)**:70-7.

O'Brien 1998a

O'Brien SJ, Pagnani MJ, Fealy S, McGlynn SR, Wilson JB. The active compression test: a new and effective test for diagnosing labral tears and acromioclavicular joint abnormality. *American Journal of Sports Medicine* 1998;**26**(5):610-3.

Patte 1987

Patte D, Gerber C. Pathologie du defile sous acromial et coraco-humeral du jeune. Pathologie de l'appareil locomoteur liee au sport. Paris: Pfizer, 1987.

Rowe 1981

Rowe CR, Zarins B. Recurrent transient subluxation of the shoulder. *Journal of Bone and Joint Surgery - American volume* 1981;**63**(6):863.

Rubin 2002

Rubin Benjamin D, Ben Kibler W. Fundamental principles of shoulder rehabilitation: Conservative to postoperative management. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2002;**18**(9):29-39.

Savoie 2001

Savoie FH, Field LD, Atchinson S. Anterior superior instability with rotator cuff tearing: SLAC lesion. *Orthopaedic Clinics of North America* 2001;**32**(2):457-61.

Schwamborn 1999

Schwamborn T, Imhoff AB. Diagnostik und klassifikation der rotatorenmanchettenlasionen. In: Imhoff AB, Konig U editor(s). Schulterinstabilitat-Rotatorenmanschette. Darmstadt: Steinkopff Verlag, 1999:193-5.

Sheon 1987

Sheon RP, et al. Soft Tissue Rheumatic Pain. Philadelphia: Lea & Febinger, 1987.

Sher 1995

Sher JS, Uribe JW, Posada A, Murphy BJ, Zlatkin MB. Abnormal findings on magnetic resonance images of asymptomatic shoulders. *Journal of Bone and Joint Surgery - American Volume* 1995;**77**(1):10-5.

Snyder 1990a

Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ. SLAP lesions of the shoulder. *Arthroscopy* 1990;**6**(4):274-9.

Tokish 2003

Tokish JM, Decker MJ, Ellis HB, Torry MR, Hawkins RJ. The belly-press test for the physical examination of the subscapularis muscle: electromyographic validation and comparison to the lift-off test. *Journal Of Shoulder And Elbow Surgery / American Shoulder And Elbow Surgeons* 2003;**12**(5):427-30.

Uchiyama 2010

Uchiyama Y, Hamada K, Khruekarnchana P, Handa A, Nakajima T, Shimpuku E, et al. Surgical treatment of confirmed intratendinous rotator cuff tears: retrospective analysis after an average of eight years of follow-up. *Journal of Shoulder & Elbow Surgery* 2010;**19**(6):837-46.

Van der Heijden 1997

Van der Heijden GJ, Van der Windt DA, De Winter AF. Physiotherapy for patients with soft tissue shoulder disorders: a systematic review of randomised clinical trials. *BMJ* 1997;**315**(7099):25-30.

van Moppes 1995

van Moppes F I, Veldkamp O, Roorda J. Role of shoulder ultrasonography in the evaluation of the painful shoulder. *European Journal Of Radiology* 1995;**19**(2):142-6.

Walton 2008a

Walton DM, Sadi J. Identifying SLAP lesions: a meta-analysis of clinical tests and exercise in clinical reasoning. *Physical Therapy in Sport* 2008;**9**(4):167-76.

Whiting 2003

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003;**3**:25.

Whiting 2005a

Whiting P, Mallett S, Bachmann L, Egger M, Pewsner D. Chapter 7: Collecting data. Cochrane Diagnostic Reviewers' Handbook. Version 0.3 [updated July 2005]. The Cochrane Collaboration, 2005.

Whiting 2005b

Whiting P, Mallett S, Bachmann L, Egger M, Pewsner D. Appendix 7.1: Collection of 2x2 table data. Cochrane Diagnostic Reviewers' Handbook. Version 0.3 [updated July 2005]. The Cochrane Collaboration, 2005.

Wofford 2005

Wofford JL, Mansfield RJ, Watkins RS. Patient characteristics and clinical management of patients with shoulder pain in U.S. primary care settings: secondary data analysis of the National Ambulatory Medical Care Survey. *BMC Musculoskeletal Disorders* 2005;**6**:4.

Yergason 1931

Yergason RM. Supination sign. *Journal of Bone and Joint Surgery - American Volume* 1931;**13**(1):160.

Yocum 1983

Yocum Lewis A. Assessing the shoulder: history, physical examination, differential diagnosis, and special tests used. *Clinics in Sports Medicine* 1983;**2**(2):281-9.

Zaslav 2001

Zaslav KR. Internal rotation resistance strength test: a new diagnostic test to differentiate intra-articular pathology from outlet (Neer) impingement syndrome in the shoulder. *Journal of Shoulder and Elbow Surgery* 2001;**10**(1):23-7.

Zlatkin 1989

Zlatkin M B, Iannotti J P, Roberts M C, Esterhai J L, Dalinka M K, Kressel H Y, et al. Rotator cuff tears: diagnostic performance of MR imaging. *Radiology* 1989;**172**(1):223-9.

References to other published versions of this review

Hanchard 2008

Hanchard NCA, Handoll HHG. Physical tests for shoulder impingements and local lesions of bursa, tendon or labrum that may accompany impingement. *Cochrane Database of Systematic Reviews* 2008, Issue 4. [DOI: 10.1002/14651858.CD007427]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Barth 2006

Clinical features and settings	***Inclusion criteria*** [1] Arthroscopy ***Exclusion criteria*** [1] Previous surgery on shoulder, [2] shoulder scheduled for capsular release and lysis of adhesions ***Duration of symptoms*** Not reported

Barth 2006 *(Continued)*

	Previous treatments Not reported (although previous surgery on the shoulder was an exclusion) ***Care setting*** Secondary or tertiary care
Participants	USA (01-03/2004) 68 shoulders in 68 patients (72% male), mean age 45 years (SD 15, range 16-76)
Study design	Prospective, consecutive, cross sectional study
Target condition and reference standard(s)	Subscapularis tears, [1] any, [2] full, [3] partial ***Arthroscopy*** *Procedure* The patient was in the lateral decubitus position, with 5-10 lbs of balanced suspension with the arm in 20-30° abduction and 20° flexion. A senior author (SSB) undertook a complete arthroscopic examination of the shoulder joint and subacromial space through a standard posterior portal. Subscapularis was evaluated with a 30° and a 70° arthroscope with an arthroscopy pump maintaining pressure at 60mmHg. The subscapularis insertion and its footprint were readily visualised by abduction and internal rotation (reference given to Lo 2003). *Interpretation* An area of 'bare footprint' indicated a partial or complete subscapularis lesion. The size of this lesion was assessed by measuring the superior-to-inferior length of bare footprint. This was then converted to a percentage of the complete insertion length by dividing it by 2.5cm (the mean superior-to-inferior length of subscapularis footprint). ***Interval between index and reference test*** Not reported ***Tester(s)*** An orthopaedic surgeon specializing in arthroscopic surgery and reconstructive procedures of the shoulder ***Prevalence*** 29% (any subscapularis tears), 26% (partial thickness subscapularis tears in subgroup with complete tears excluded: see Note 1) 4.4% (complete subscapularis tears)
Index and comparator tests	***Lift-off test (modified interpretation 1)*** Referenced to the primary source (Gerber 1991a) and described *Procedure* As in the primary source *Interpretation* As in the primary source except that, in addition to inability to lift the hand off the back, the ability to do so only by extending the shoulder or elbow was also considered positive. ***Belly-press test (modified procedure)*** Referenced to the primary source (Gerber 1996) and described *Procedure* As in the primary source except that, reportedly, the test was performed with 'the arm at the side'. The test was originally depicted in slight shoulder abduction (Gerber 1996). An addition in this study was use of a digital tensiometer for bilateral comparisons of force. Resistance was applied perpendicular to the distal forearm via a padded sling. *Interpretation* As in the primary source, the test was considered positive if the patient utilised elbow or shoulder extension. In addition, force was measured and compared to the other shoulder. ***Napoleon test (standard)*** Referenced to Schwamborn 1999, the German language primary source, and Burkhart 2002, who refined the test's interpretation. The test was also described. *Procedure* The patient was required to adopt a Napoleonic pose, palm on abdomen. *Interpretation* A negative (normal) result was where the patient could press against the abdomen without flexing the wrist. A positive result was defined as inability to push against the abdomen without flexing the wrist to 90°. An intermediate result, which has been correlated with partial tears of subscapularis Burkhart 2002, was also recorded as positive. ***Bear-hug test (novel)***

Barth 2006 *(Continued)*

	This was a novel test. *Procedure* The patient placed the palm of the affected limb on the opposite shoulder, fingers extended. He or she was then required to sustain the position while the tester applied an external rotation force perpendicular to the forearm. An electronic digital tensiometer was used to measure force via a padded sling applied just proximal to the wrist. *Interpretation* The test was considered positive if the patient was unable to position the arm, or demonstrated more than a 20% strength deficit compared to the other side. ***Tester(s)*** No information given. However the testers' expertise at the bear hug test, which they originated, would be expected to be high.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** This sub-grouping presupposes that full thickness tears have been identified and excluded. However, a. this subgroup may be more representative of the primary care population than the study group as a whole; and b. the increasing availability of diagnostic ultrasound raises the possibility that FTT may indeed be excluded, even in the primary care setting. ***2 x 2 tables and summary statistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	Only a broad outline was given.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	All patients were accounted for as having undergone the reference test.
Differential verification avoided? All tests	Yes	All patients underwent the same reference test.
Incorporation avoided? All tests	Yes	

Barth 2006 *(Continued)*

Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Unclear	The study was prospective and it is unclear whether there was blinding, but the index test preceded the reference standard.
Reference standard results blinded? All tests	No	There does not appear to have been blinding.
Relevant clinical information? All tests	Unclear	The index tests were conducted on patients already listed for arthroscopy.
Uninterpretable results reported? All tests	Yes	The study was prospective, recruitment was consecutive and (with the exception of the lift-off test), results were reported for all initially included patients.
Withdrawals explained? All tests	Yes	Except in relation to the lift-off test

Bennett 1998

Clinical features and settings	***Inclusion criteria*** [1] Shoulder pain, [2] arthroscopy (implied) ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	USA (1/10/1994-28/2/1995). 46 shoulders in 45 patients (67% male), mean age 56 years (SD not reported, range 16-80)
Study design	Prospective, cross-sectional study. The statement, "In light of the results of this series, patients not entered into the study who responded to conservative treatment may have had positive Speed test findings" implies non-consecutiveness in terms of surgery.
Target condition and reference standard(s)	[1] Intra-articular LHB tendon/labral injury or macroscopic bicipital groove and/or LHB tendon inflammation at the level of the groove ***Arthroscopy*** *Procedure* The beach-chair position was used. By flexing the elbow (and elevating the shoulder), tension was released on the LHB tendon, thus allowing the LHB tendon at the level of the bicipital groove to be directly visualized in the intra-articular portion of the glenohumeral joint (Figure given). Visual examination of the tendon at the level of the bicipital groove for inflammation and/or fraying was facilitated by placing a neuroprobe into and through the anterior portal and placing it just under the superior glenohumeral ligament and then over the LHB tendon. The application of a caudal force to the tendon with a neuroprobe levers the nonarticular portion of the LHB tendon an additional 3 to 4 cm into the glenohumeral joint (Figure given). The normal intra-articular portion of the LHB tendon was also evaluated for inflammation, fraying, avulsion, or SLAP lesions. *Interpretation* Any macroscopic inflammation and/or tearing of the LHB tendon at any level, evidence of any type of SLAP lesion (*these are undefined*), or a complete avulsion of the LHB tendon were considered positive results. ***Interval between index and reference test*** Not reported

Bennett 1998 *(Continued)*

	Tester(s) No information given ***Prevalence*** 22% (LHB tendon/labral injury)
Index and comparator tests	***Speed's test (modified procedure, modified interpretation 1)*** Referenced to a secondary source (van Moppes 1995) and described: *Procedure* A downward pressure was applied to the upper extremity with the shoulder elevated 90°, full supination of the forearm, and the elbow extended. (This interpretation of the test as an isometric technique appears to differ to that of Crenshaw 1966, who cited a personal communication with the test's originator. They wrote, 'it is performed by having the patient flex his shoulder (elevate it anteriorly) against resistance.') *Interpretation* 'Pain experienced in the anterior shoulder of the patient when pressure is applied' and 'pain in the proximal portion of the shoulder during the application of a downward force applied to the arm' were variously stated to indicate a positive Speed's test result. (According to Crenshaw 1966, 'the test is positive when pain is localized to the bicipital groove'.) ***Tester(s)*** No information given
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** There is an error in table 3. The contents of the lower two cells are transposed. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** Yes ***If applicable, could 2 x 2 tables be back-calculated?*** N/A ***Were the reported summary statistics confirmed as accurate?*** There was a borderline discrepancy in specificity, but the other summary statistics were confirmed as accurate.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	Very unspecific description: patients with shoulder pain and listed for arthroscopy
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	No	The test was applied to some patients who were treated conservatively.
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	

Bennett 1998 *(Continued)*

Sufficient description of index tests? All tests	Yes	Though poorly reported
Sufficient description of reference test? All tests	Yes	Generally very clear though SLAP lesions were undefined
Index test results blinded? All tests	Unclear	There is no statement of blinding, but the study was prospective.
Reference standard results blinded? All tests	No	There does not appear to have been blinding.
Relevant clinical information? All tests	Yes	Probably
Uninterpretable results reported? All tests	Unclear	Insufficient information - probably not consecutive recruitment (see comments in Study design above)
Withdrawals explained? All tests	Unclear	Insufficient information (see comments in Study design above)

Calis 2000

Clinical features and settings	***Inclusion criteria*** [1] Shoulder pain, [2] aged 18-70, [3] standard radiographic assessment ***Exclusion criteria*** [1] Inflammatory or systemic diseases, [2] acute traumatic conditions, [3] postoperative conditions, [4] neck or elbow disorders ***Duration of symptoms*** Not reported, but specified to be equivalent for the groups with and without impingement ***Previous treatments*** Not reported ***Care setting*** Mixed: primary, secondary and tertiary
Participants	Turkey (period not reported) 125 shoulders in 120 patients (40% male), mean age 52 years (SD not reported, range not reported)
Study design	Prospective, cross-sectional study: unclear whether recruitment was consecutive
Target condition and reference standard(s)	[1] Subacromial impingement syndrome, [2] complete disruption of supraspinatus tendon ***Reassessment following blind subacromial local anaesthetic injection*** *Procedure* 1% lidocaine was injected under the acromion by an anterior approach 'by experienced hands'. Using the ACJ and the anterior edge of the acromion as landmarks, the aim was to place the needle tip directly below the anterior edge of the acromion. Gentle longitudinal traction through the arm was used to facilitate needle entry. *Interpretation* Marked relief of pain and almost total improvement in passive or active ROM after 30 minutes was interpreted as positive, provided that there were no calcific lesions on adjuvant X-rays. ***MRI*** *Procedure* Not reported *Interpretation* Used Zlatkin stages (reference given to Zlatkin 1989), with stage 1 defined as increased signal intensity in the tendon without any thinning, irregularity or discontinuity, stage 2 as increased

Calis 2000 *(Continued)*

	signal intensity with irregularity and thinning in the tendon, and stage 3 as complete disruption of the supraspinatus tendon. ***Interval between index and reference test*** Not reported, but the X-ray and MRI appear to have preceded the index tests, and the injection test was presumably immediately afterwards. ***Tester(s)*** Subacromial injection test was performed by 'experienced hands'; MRI by a radiologist with experience on the skeletal system, especially shoulder imaging ***Prevalence*** 70% (subacromial impingement syndrome); 17% (complete disruption of supraspinatus tendon)
Index and comparator tests	Tests were referenced to primary and/or secondary sources and described. All were conducted/interpreted as standard. ***Neer's sign, Hawkins' test, passive horizontal adduction*** ***painful arc test, drop arm test, Yergason's test*** ***Speed's test (all standard)*** ***Tester(s)*** Two physicians with four and eight years' experience in shoulder management
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. There were borderline discrepancies between most reported sensitivities, specificities and other summary statistics and those derived from back-calculated 2 x 2 tables.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Unclear	Proportion of referred to self-referred not known
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Unclear	Blind (unguided) injection reference test. MRI is acceptable for full thickness tears in low-prevalence samples.
Acceptable delay between tests? All tests	Unclear	Likely, but no information given
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	

Calis 2000 *(Continued)*

Incorporation avoided? All tests	Yes	For the physical tests the same procedures were used, but with or without local anaesthesia. 'Yes' for MRI.
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	Injection 'Yes', MRI 'No'
Index test results blinded? All tests	Unclear	It is unclear whether there was blinding, but the study was prospective and the index tests preceded the injection aspect of the reference standard. (MRI appears to have preceded the index tests, however.)
Reference standard results blinded? All tests	No	
Relevant clinical information? All tests	Unclear	No information
Uninterpretable results reported? All tests	Unclear	Unclear
Withdrawals explained? All tests	Unclear	Unclear

Castoldi 2009

Clinical features and settings	***Inclusion criteria*** [1] Painful shoulder, [2] having had arthroscopy or open surgery ***Exclusion criteria*** [1] Acute traumatic conditions, [2] acute postoperative conditions, [3] fractures, [4] frozen shoulder with a great deficit of range of motion, [5] neurologic disorders (plexus injury), [6] notes unavailable ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Italy or Switzerland? (01/2004-06/2006) 395 shoulders in 390 patients (57% male), mean age 50 years (SD 16, range 16-89)
Study design	Retrospective, consecutive, cross sectional study
Target condition and reference standard(s)	Supraspinatus tears: [1] FTT, full-width; [2] PTT, isolated ***Arthroscopy or open surgery*** *Procedure* No information *Interpretation* "The size of any tear was estimated depending on the number of involved tendons. The topographic description of the tear was performed by dividing the sagittal plane of the rotator cuff in 7 zones. Partial nontransmural tears of the supraspinatus were allocated to a separate group." ***Interval between index and reference test*** Not reported

Castoldi 2009 *(Continued)*

	Tester(s) No information given ***Prevalence*** 29% (isolated partial thickness supraspinatus tears), 17% (full width, full thickness supraspinatus tears)
Index and comparator tests	***External rotation lag sign (standard)*** Referenced to the primary source (Hertel 1996a) and described *Procedure* As in the primary source *Interpretation* As in the primary source. Emphasis was placed on the need to make bilateral comparisons in order to identify small lags. ***Tester(s)*** A 'single skilled examiner'. A co-author was the originator of the test.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Table 1 reports 199 cuff tears plus 2 cuff tears with instability, making a total of 201 cuff tears (70 left, 131 right) in 395 shoulders. On this basis, 395 - 201 = 194 cuffs were intact. However in Table 2 the number of intact cuffs is reported as 157, and this is the value used in the diagnostic test accuracy calculations. No explanation is given for this discrepancy of 37 shoulders. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** Yes. For isolated partial thickness tears of supraspinatus, only shoulders with that specific lesion (n = 65) or a normal cuff (n = 157) were given in the 2 x 2 table (total = 222). For isolated full width, full thickness tears of supraspinatus, only shoulders with that specific lesion (n = 32) or a normal cuff (n = 157) were included in the 2 x 2 table (total = 189). In relation to the latter target condition, this restrictive approach to inclusion limits the applicability of the results to the clinical setting, where intermediate cases would be encountered. ***If applicable, could 2 x 2 tables be back-calculated?*** N/A ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The study was retrospective. Also the reference test (arthroscopy or open surgery) was invasive.
Selection criteria described? All tests	Unclear	The inclusion criteria were limited.
Acceptable reference standard? All tests	Yes	Probably, though no procedural details were given
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided?	Yes	The reference tests used were probably equivalent.

Castoldi 2009 *(Continued)*

All tests		
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	No information was given on the procedure for the reference test.
Index test results blinded? All tests	No	The study was retrospective.
Reference standard results blinded? All tests	No	Very unlikely, since the study was retrospective
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	The study was retrospective. Also see below.
Withdrawals explained? All tests	No	The study was retrospective and there are unexplained missing cases from the analysis

Frost 1999

Clinical features and settings	***Inclusion criteria*** Based on a questionnaire screening workers for shoulder symptoms; [1] cases, comprising individuals who had suffered ≥ months' shoulder pain in the past year, and [2] controls, who had experienced no long lasting shoulder symptoms. Controls were age matched to cases. ***Exclusion criteria*** [1] Claustrophobia, [2] size too large to fit in the MRI ***Duration of symptoms*** 3 months in the past year for the cases ***Previous treatments*** Not reported ***Care setting*** The individuals recruited were not necessarily undergoing care
Participants	Denmark (period not reported). From among 167 individuals who met the criteria for 'patients' and 110 who met the criteria for controls, 73 shoulders in 73 participants (81% male; mean age and range not reported) ultimately underwent the index and reference tests.
Study design	Prospective, case control study with non-consecutive recruitment
Target condition and reference standard(s)	Supraspinatus: [1] FTT, degeneration or tendinitis; [2] FTT or degeneration; [3] FTT ***MRI*** *Procedure* This utilised 2 identical 1.5 Tesla Gyroscans and surface shoulder coils. The supine patient's arm was held in neutral. The scans were done in 3 sequences: a scout in the coronal plane was used to place the co-ordinates for the transverse scans of the area from the undersurface of the acromion to the lower glenoid margin. Coronal scans were performed in the plane parallel to the supraspinatus tendon. Scans were done with TR/msec/TE msec 1800/30 proton and 1800/90 T2-weighted sequences, slice thickness 4 mm, spacing 0.4 mm, matrix 266 X 256, field view 200 mm, and number of excitations 1.

Frost 1999 *(Continued)*

	Interpretation used criteria adapted from Zlatkin 1989, collapsing the possible diagnoses relating to supraspinatus pathologies into 4 categories. ***Interval between index and reference test*** Not reported ***Tester*** A radiologist whose level of expertise was not reported, but who was blinded to participants' age and clinical status. ***Prevalence*** 53% (any supraspinatus lesions), 9.6% (full thickness tears)
Index and comparator tests	***Hawkins' test (modified procedure, modified interpretation 2)*** *Procedure* With the arm in 90° of scaption [versus the standard starting position for Hawkins test, which is 90° of flexion], the shoulder was passively medially rotated. *Interpretation* The test was considered positive if there was provocation or aggravation of pain at the anterolateral aspect of the shoulder. ***Tester(s)*** 2 examiners from the Department of Occupational Medicine: no further information given
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The primary purpose of this study was not diagnostic, but to compare MRI appearances of supraspinatus in symptomatic subjects aged < and ≥ 50 years, respectively, who had impingement against that in a control group of asymptomatic subjects of corresponding age without impingement. The results showed no difference in the prevalence of supraspinatus pathology between the impingement sign positive/ impingement sign negative groups (though full thickness tears were reported to be more common in the former) but a strong association with age. This emphasises the questionable suitability of MRI (and possibly, by corollary, ultrasound, arthroscopy and open surgery) as a reference standard in impingement related diagnostic studies. ***Note 2*** From a diagnostic perspective, this study is weakened by its case control design. The small numbers within the subgroups 'tendonitis', 'degeneration' and 'full thickness tear' precluded their individual analysis. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No summary diagnostic test accuracy statistics were reported.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	Case control design. Community - not a care setting as such.
Selection criteria described? All tests	No	Very unspecific description: 3 months' shoulder pain in the past year
Acceptable reference standard? All tests	Yes	MRI: acceptable for full thickness tears
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	No	A non-random selection of participants underwent a reference test

Frost 1999 *(Continued)*

Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Unclear	Probably, given the unambiguous nature of the index test
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	No	There was blinding to the results of the questionnaires
Uninterpretable results reported? All tests	No	Recruitment was not consecutive
Withdrawals explained? All tests	No	

Guanche 2003

Clinical features and settings	***Inclusion criteria*** [1] Patients undergoing shoulder arthroscopy for a variety of activity related complaints, [2] complete physical examination immediately prior to surgery, [3] full range of motion attainable under anaesthesia and (4) failure to respond to previous conservative treatment. ***Exclusion criteria*** Previous surgery on affected shoulder ***Duration of symptoms*** Not reported ***Previous treatments*** Non-operative care. The protocol for subacromial impingement was subacromial steroid injections (maximum of three) followed by physical therapy; that for suspected ACJ arthritis was steroid injection; and that for presumptive labral pathology was a course of progressive physical therapy and activity modification. ***Care setting*** Secondary or tertiary care
Participants	USA (period not reported). 60 shoulders in 59 patients (82% male); mean age 38 years (SD not reported, range 15-76).
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	[1] SLAP lesion (any) ***Arthroscopy*** *Procedure* The procedure described is internally inconsistent. Thus, with the patient in the beach-chair position, 'in all cases the glenohumeral joint was first evaluated and the pathology addressed as ap-

Guanche 2003 *(Continued)*

propriate. Subsequently the subacromial space was evaluated... and appropriate treatments were performed. Finally lesions of the glenohumeral joint were addressed.'

Interpretation Classification of SLAP lesions was according to Snyder 1990a. Labral lesions inferior to the 10 o'clock to 2 o'clock segment defined as SLAP were recorded as 'labral tears'.
Interval between index and reference test The index tests were performed in the holding area of the operating theatre, immediately before administration of anaesthetic or narcotic agent.
Tester(s) No information given
Prevalence 55% (Types I-IV SLAP lesions)

Index and comparator tests

Active compression test (modified procedure, modified interpretation 2)

Procedure Referenced to primary source (O'Brien 1998a), but the illustration and caption indicated that the second part of the test was conducted with the shoulder in neutral rotation rather than full external rotation, and this differs from the test as originally described.

Interpretation As in the primary source, deep pain in the shoulder which was less in 'external' than internal rotation was considered positive. Questioning as to the site of pain was overlooked in 27 patients, reducing the data available for this test.

Anterior apprehension test at 90° for pain (modified interpretation 2)

Referenced to a chapter by Hawkins and Bokor in Rockwood and Matsen (1990) and faithfully described. (Rockwood and Matsen (1990) was unavailable to the review authors, but the description was compatible with that given in the third edition of that book (Krishnan 2004).)

Procedure The supine patient's arm was positioned in 90° abduction and full external rotation. (As distinct from the version of this test described by Jobe 1989, no anterior pressure was applied to the humeral head.)

Interpretation Pain was considered a positive result.

Palpation for bicipital groove tenderness (modified interpretation 2)

Procedure and Interpretation Tenderness was sought by palpation.

Crank test (modified procedure, modified interpretation 2)

Referenced to one of the primary sources (Liu 1996b) and described

Procedure Performed with the patient lying supine and the shoulder in full abduction. This differs from both of the definitions given by the originators of the test (Liu 1996c; Liu 1996b) - *see* further in this table.

Interpretation Standard
Anterior release test (modified interpretation 2)

The authors reportedly evaluated the relocation test, which they referenced to Jobe et al (1990); but the manoeuvre illustrated is the anterior release test (Gross 1997).

Speed's test (modified interpretation 1,2)

Referenced to a tertiary source (Bennett 1998) which itself described the test incompletely, and illustrated.

Procedure The procedure appears to have been consistent with that in the primary source, but neither report specified the arc of movement through which the test was performed.

Interpretation As distinct from the primary source, pain was accepted as a positive result regardless of whether it was localized to the bicipital groove.

Yergason's test (modified interpretation 1,2)

Secondarily referenced to a chapter by Hawkins and Bokor in Rockwood and Matsen (1990) and illustrated.

Guanche 2003 *(Continued)*

	Procedure As in the primary source (Yergason 1931) with the addition of 'simultaneous [palpation] of the biceps tendon'. Rockwood and Matsen (1990) was unavailable to the review authors, but in the third edition of that book there is no mention of palpation as part of the test (Krishnan 2004). *Interpretation* As in the primary source (Yergason 1931) with the additional requirement that 'unusual' pain in the biceps tendon be felt in order to signify a positive result. This qualification is not mentioned by Krishnan 2004. ***Tester(s)*** All tests were performed by the first author, an orthopaedic surgeon.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The references to 'tears' in the title and Table 3 of this article are misleading, because Type I SLAP lesions (which are not tears) are pooled with other types of SLAP lesion (which are tears) in the data analyses. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes, except for the active compression test (for which the numbers of disease positive and disease negative patients were not reported) ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	Unclear	The description was not completely clear.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Yes	The index tests were performed immediately pre-operatively.
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test?	Yes	Brief but referenced

Guanche 2003 *(Continued)*

All tests		
Index test results blinded? All tests	Unclear	There is no statement of blinding, but the study was prospective.
Reference standard results blinded? All tests	No	There does not appear to have been blinding.
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	

Gumina 2008

Clinical features and settings	***Inclusion criteria*** Diagnoses relating to shoulder pain and weakness (impingement syndrome and postero-superior cuff tears). It is unclear whether either or both were required for inclusion. ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Italy (period not reported). 120 shoulders in 120 patients (gender distribution not reported); mean age 64 years (SD not reported, range 46-79).
Study design	Unclear whether prospective. Cross sectional study with consecutive inclusion.
Target condition and reference standard(s)	Subacromial impingement and full thickness postero-superior rotator cuff tears. ***Arthroscopy*** *Procedure* No information given *Interpretation* No information given ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 38% (subacromial impingement), 62% (rotator cuff tears)
Index and comparator tests	***Gum-Turn test*** A novel test. *Procedure* Starting in the empty can test position, the patient traced a 20 cm wide spiral drawn on the wall, from centre to periphery and back 10 times, resting for one minute, then repeating the procedure. *Interpretation* The test was positive if weakness or pain prevented completion. For positive results, the number of turns completed was recorded, but it is unclear how these data were used. Results were compared with the contralateral arm but, again, it is unclear how these data were used. ***Tester(s)*** Presumably expert, as the authors were the originators of the test
Follow-up	***Adverse events*** None mentioned

Gumina 2008 *(Continued)*

Notes

2 x 2 tables and summarystatistics

Were 2 x 2 tables reported? No

If applicable, could 2 x 2 tables be back-calculated? Yes

Were the reported summary statistics confirmed as accurate? No. Sensitivity was reported accurately throughout, but there were borderline discrepancies between some reported specificities and those derived from back-calculated 2 x 2 tables.

For numerous other summary statistics there were substantial discrepancies.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The selection criteria were unclearly described and no exclusion criteria were reported.
Acceptable reference standard? All tests	Yes	In principle, though no procedural or interpretative details were given
Acceptable delay between tests? All tests	Unclear	No information given
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Although it is unclear how the number of turns completed or comparison with the other arm were used, these issues do not appear critical to interpretation of the test.
Sufficient description of reference test? All tests	No	No information given
Index test results blinded? All tests	No	There is no statement of blinding and the study may not have been prospective.
Reference standard results blinded? All tests	No	There does not appear to have been blinding.
Relevant clinical information?	Yes	Most probably

Gumina 2008 *(Continued)*
All tests

Uninterpretable results reported? All tests	Unclear	The study may not have been prospective.
Withdrawals explained? All tests	No	The study may not have been prospective.

Hertel 1996

Clinical features and settings	***Inclusion criteria*** [1] Unilateral subacromial impingement (Neer's stage I-3), [2] subsequent open or arthroscopic rotator cuff exploration ***Exclusion criteria*** Impaired passive range of glenohumeral motion ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Unclear but likely to be tertiary care
Participants	Berne, Switzerland (03/1992-12/1993). '100 consecutive patients' (74% male) with unilateral subacromial impingement syndrome (stages 1-3); but see further information in 'Index and comparator tests'. Median age 51 years (range 16-79)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Rotator cuff ruptures in 2, probably overlapping, groups: [1] ruptures affecting the supraspinatus and infraspinatus tendons and [2] ruptures affecting the subscapularis tendon. ***Open surgery or arthroscopy*** *Procedure* No information given *Interpretation* Criterion for a positive result was "rotator cuff rupture". No further details were given. ***Interval between index and reference test*** Unclear ***Tester(s)*** No details given ***Prevalence*** 63% (lesions of any type involving the postero-superior rotator cuff), 29% (lesions of any type involving subscapularis)
Index and comparator tests	The ***empty can test*** and two novel tests, the ***external rotation lag sign (ERLS)*** (to assess the function of supraspinatus and infraspinatus) and the ***drop sign*** (to assess the posterosuperior rotator cuff), were evaluated in a subgroup (n = 87) found not to have subscapularis lesions in isolation. The ***lift-off sign*** (to assess the function of subscapularis) and a novel test, the ***internal rotation lag sign (IRLS)*** (to assess for cranial or large tears of subscapularis) were evaluated in a subgroup (n = 53) found not to have supraspinatus and/or infraspinatus lesions in isolation. *Procedure and interpretation* The reference erroneously given for the empty can test is Jobe 1982, in which an isotonic manoeuvre was described for strengthening and testing strength. The manoeuvre described by Hertel 1996 is Jobe's isometric test for supraspinatus integrity, the primary source for which is Jobe 1983. In the context of Hertel 1996 it appears that weakness was taken as a positive result, which is compatible with the primary source. The lift-off sign is referenced to its primary source (Gerber 1991a) but not described further. ***Tester(s)*** No details given though, apart from the empty can test, all of the tests evaluated had been originated by one of the study's authors
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The empty can test, ERLS, and drop sign were not reported for isolated subscapularis lesions (n = 13); the lift-off test and IRLS were not reported for lesions of supraspinatus +/- infraspinatus (n = 47). ***Note 2*** There is an error in Table 1. The accuracy of the ERLS should read 67/87, not 37/87 as reported; and the accuracy of the drop sign should read 37/87, not 68/87 as reported.

Hertel 1996 *(Continued)*

2 x 2 tables and summarystatistics

Were 2 x 2 tables reported? No

If applicable, could 2 x 2 tables be back-calculated? Yes

Were the reported summary statistics confirmed as accurate? Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was probably tertiary and the reference test (open surgery or arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Yes	Open surgery or arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	All patients underwent surgery
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Probably, though there is no mention of how the lag was measured
Sufficient description of reference test? All tests	No	No details given
Index test results blinded? All tests	Unclear	There is no statement of blinding and the analysis is structured towards presenting results with knowledge of the reference test results
Reference standard results blinded? All tests	No	There does not appear to have been blinding
Relevant clinical information? All tests	Yes	Probably
Uninterpretable results reported?	Yes	

Hertel 1996 *(Continued)*

All tests	
Withdrawals explained? All tests	Yes

Holtby 2004a

Clinical features and settings	***Inclusion criteria*** [1] Persistent pain and functional disability > 6 months, not responsive to adequate conservative treatment, with a positive impingement test confirmed with local anaesthetic or clinical or investigative signs of rotator cuff tears, labral or LHB lesions, [2] symptoms referred to the ACJ > 6 months with radiographic changes in the joint [3] symptomatic shoulder instabilities, [4] informed consent. ***Exclusion criteria*** History of previous shoulder surgery or upper extremity fractures ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	Canada (period not reported) Patients complaining of shoulder pain, referred by family physicians or orthopaedic surgeons for assessment, and meeting the criteria for surgery. 50 shoulders in 50 patients (68% male); mean age 50 years (SD 14, range 24-79).
Study design	Prospective cross sectional study with consecutive recruitment
Target condition and reference standard(s)	LHB tendon pathology or SLAP lesions ***Arthroscopy*** *Procedure* Lateral position with arm suspended from a robot traction device with 12 lbs of weight. The arthroscope was inserted through the superior portal to inspect the articular surfaces, anterior and inferior labrum, LHB, and articular surface of the rotator cuff. *Interpretation* Criteria for a positive result were [1] LHB tendon fibrillation, adhesion in the bicipital groove, subluxation or dislocation, partial tear or complete rupture, [2] visualisation of type II or IV SLAP lesions. Types I and III SLAP lesions were not considered positive findings. ***Interval between index and reference test*** Mean approximately 23 weeks ***Tester(s)*** No details given ***Prevalence*** 42-43% (LHB pathologies or SLAP type II or IV lesions)
Index and comparator tests	***Speed's test*** Referenced to Gilcreest 1936, which we have been unable to obtain, and described. *Procedure* The examiner pressed downwards on the arm with the shoulder flexed to 90°, the forearm fully supinated and the elbow extended. (This interpretation of the test as an isometric technique appears to differ to that of Crenshaw 1966, who cited a personal communication with the test's originator, and wrote, "it is performed by having the patient flex his shoulder (elevate it anteriorly) against resistance.") *Interpretation* Pain in the anterior shoulder or glenohumeral joint was considered a positive result. (According to Crenshaw 1966, 'the test is positive when pain is localized to the bicipital groove'.) ***Yergason's test*** Referenced to primary source (Yergason 1931) and described. *Procedure* As in primary source *Interpretation* As in the primary source, pain in the bicipital groove area was considered a positive result. Pain in the glenohumeral joint was also considered positive. ***Tester(s)*** No information given

Holtby 2004a *(Continued)*

Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** MRI and X-rays taken beforehand were examined only after completion of the data collection form. ***Note 2*** Same study population as Holtby 2004b and Razmjou 2004. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** Yes ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	No	Mean interval from examination to surgery approximately 23 weeks
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	

Holtby 2004a *(Continued)*

Relevant clinical information? All tests	Yes	MRI and X-rays taken beforehand were examined only after completion of the data collection form.
Uninterpretable results reported? All tests	Unclear	See below
Withdrawals explained? All tests	No	[1] Some problems regarding handling of results for bicipital tendonitis; [2] 1 patient's data missing from the Yergason's test results.

Holtby 2004b

Clinical features and settings	***Inclusion criteria*** Undergoing surgery for at least one of: [1] persistent pain and functional disability for > 6 months, not responsive to adequate conservative intervention, and a positive impingement test, confirmed with local anaesthesia; [2] clinical signs of rotator cuff tears, labrum or LHB lesions; [3] symptoms referred to the ACJ lasting > 6 months, with visible radiographic changes in the joint, or [4] symptomatic shoulder instabilities ***Exclusion criteria*** History of previous shoulder surgery or upper extremity fractures ***Duration of symptoms*** 1 month to > 5 years ***Previous treatments*** 'Adequate conservative intervention' ***Care setting*** Tertiary care
Participants	Canada (period not reported) 50 shoulders in 50 patients (68% male); mean age 50 years (SD 14, range 24-79)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Tendinitis and partial thickness rotator cuff tears, full thickness rotator cuff tears, large or massive full thickness rotator cuff tears All patients were initially examined by ***arthroscopy*** , proceeding to ***open surgery*** if required for management of large or massive rotator cuff tears. *Procedure* The bursal and articular aspects of the rotator cuff were inspected for inflammation, fraying, or tears. *Interpretation* Tears were categorized as partial thickness or full thickness. Full details are given in an appendix. ***Interval between index and reference test*** Mean 23 weeks (range 5 weeks to 11 months) ***Tester(s)*** The orthopaedic surgeon ***Prevalence*** 34% (full thickness supraspinatus tears)
Index and comparator tests	***Jobe's empty can test*** *Procedure and interpretation* The reference erroneously given for the empty can test is Jobe 1982, in which an isotonic manoeuvre was described for strengthening and testing strength. The primary source for the empty can test of supraspinatus' integrity is Jobe 1983. The *Procedure* described is compatible with the primary source, but *Interpretation* is broadened to the 'rotator cuff'. ***Tester(s)*** Independent examinations were performed by two testers, an orthopaedic surgeon and a physical therapist (for assessment of reliability) but only the surgeon's findings were included in the analysis of test diagnostic performance.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Same study population as Holtby 2004a and Razmjou 2004 ***2 x 2 tables and summarystatistics***

Holtby 2004b *(Continued)*

Were 2 x 2 tables reported? Yes

If applicable, could 2 x 2 tables be back-calculated? N/A

Were the reported summary statistics confirmed as accurate? Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Yes	Athroscopy
Acceptable delay between tests? All tests	No	Mean interval from examination to surgery was approximately 23 weeks
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Yes	

Holtby 2004b *(Continued)*

Withdrawals explained? All tests	Yes

Iagnocco 2003

Clinical features and settings	***Inclusion criteria*** Patients referred to a rheumatology unit with shoulder pain ***Exclusion criteria*** None reported ***Duration of symptoms*** Mean 5.3 months (SD not reported, range 1-19 months) ***Previous treatments*** Analgesics (n=182), NSAIDs (n=122) ***Care setting*** Secondary care
Participants	Italy (period not reported) 528 shoulders in 425 patients (34% male, mean age 58 years, range 18-90 years) referred to rheumatology with painful shoulders (103 had bilateral involvement) ***Subgroups*** [1] 228 had clinically diagnosed peri-articular disorders, [2] 13 rheumatoid arthritis, [3] 10 seronegative spondarthritis [4] 14 osteoarthritis, [5] 15 previous trauma, [6] 5 chondrocalcinosis, [7] 6 SLE, [8] 3 systemic sclerosis, [9] 131 not yet clinically diagnosed
Study design	Prospective, cross sectional study with consecutive recruitment.
Target condition and reference standard(s)	[1] Pathological changes of LHB, supraspinatus, infraspinatus and subscapularis tendons and acromioclavicular joints, [2] calcifications. The SA-SD bursa and glenohumeral joint were also listed among structures evaluated ultrasonographically (*see* below), but were not included in calculations of diagnostic test accuracy. ***Ultrasound*** *Procedure* Examinations were repeated twice by each of 2 independent operators using 7.5 MHz linear transducers. *Interpretation* Criteria for a positive result were hyperechoic areas with possible posterior shadows (calcifications); thinning, hypo-echoicity of tendons (partial rupture); thickening, hyperechoicity of tendons (tendinitis); hyperechoic discontinuity of tendons, non-visualisation of tendons (rupture); empty bicipital groove with identification of LHB tendon in the surrounding area (subluxation); hypo-anechoic area between LHB tendon and its sheath (tenosynovitis); hypo-anechoic fluid in bursae with > 2mm thickening (bursitis); effusion in or bony irregularities of the ACJ; effusion in the glenohumeral joint on transaxillary scan. References were given (was the procedure referenced too?). ***Interval between index and reference test*** A 'few days' ***Tester(s)*** 2 rheumatologists experienced in sonography ***Prevalence*** 48% (changes in LHB tendon), 60% (changes in supraspinatus tendon), 17% (changes in infraspinatus tendon), 4.4% (changes in subscapularis tendon)
Index and comparator tests	A single general reference was given for all index tests (Dalton, 1998) ***Speed's test*** for the LHB tendon *Procedure* The examiner resisted flexion of the patient's shoulder, with the elbow extended and the forearm supinated. This is compatible with the primary source (Crenshaw 1966). *Interpretation* No additional information given ***Resisted abduction*** for supraspinatus *Procedure* This test was performed with the arm abducted to 90°, flexed to 30° and internally rotated (i.e. ***empty can test***). *Interpretation* No additional information given ***Resisted lateral rotation (presumably from neutral rotation)*** for infraspinatus, and ***resisted medial rotation (presumably from neutral rotation)*** for subscapularis, were named but not described.

Iagnocco 2003 *(Continued)*

	Tester(s) An experienced rheumatologist
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. For all tests evaluated there were borderline discrepancies between the reported sensitivities and those derived from back-calculated 2 x 2 tables.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary care.
Selection criteria described? All tests	No	
Acceptable reference standard? All tests	Unclear	Ultrasonography
Acceptable delay between tests? All tests	Yes	A 'few days'
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Unclear	The reference test was conducted independently.

Iagnocco 2003 *(Continued)*

Relevant clinical information? All tests	Yes	Probably
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	

Itoi 1999

Clinical features and settings	***Inclusion criteria*** Various shoulder symptoms ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Japan (05/1996-07/1997) 143 shoulders in 136 patients (77% male); mean age 43 years, (SD not reported, range 13-80)
Study design	Unclear whether prospective or retrospective. Cross sectional study with reportedly consecutive recruitment.
Target condition and reference standard(s)	Full thickness tears of supraspinatus **MRI** *Procedure* Using a 1.5-T MR imager, spin-echo T1 weighted images and fast spin echo T2 weighted images were obtained in the coronal and sagittal planes in 3 mm thick 1 mm gap slices with 10 to 15 cm fields of view and 256 x 128 matrix pixel size. A combination of dual phased array coils and a small filed of view achieved high resolution MRI scans. (Reference given) *Interpretation* High signal intensity occupying the full thickness layer of the rotator cuff tendon on T2 weighted image was diagnosed as a full thickness rotator cuff tear. The torn tendon was also determined by the location of high signal intensity on T2 weighted image in both the coronal and sagittal planes. (Reference given.) ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 25% (full thickness supraspinatus tears)
Index and comparator tests	***Full can test*** and ***empty can test*** *Procedure* Full can test referenced to its primary source (Kelly 1996) and described in keeping with this. The reference erroneously given for the empty can test is Jobe 1982, in which an isotonic manoeuvre was described for strengthening and testing strength. The primary source for the empty can test of supraspinatus' integrity is Jobe 1983. *Interpretation* Criteria for a positive result: [1] pain, [2] weakness (less than normal ability to withstand an applied force), [3] pain or weakness or both ***Tester(s)*** No information given
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes

Itoi 1999 *(Continued)*

Were the reported summary statistics confirmed as accurate? Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was a specialist shoulder clinic.
Selection criteria described? All tests	No	'Various shoulder symptoms'. No other inclusion or exclusion criteria were given.
Acceptable reference standard? All tests	Unclear	MRI appears acceptably accurate in samples with a low prevalence of full thickness rotator cuff tears, but it is not clear that this sample has a low prevalence.
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Primary references given
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	No	Not reported, and it is uncertain whether the study was prospective
Reference standard results blinded? All tests	No	No statement of blinding
Relevant clinical information? All tests	Yes	Probably
Uninterpretable results reported? All tests	Unclear	[1] Unclear whether study was prospective; [2] not explicitly stated that there were no withdrawals
Withdrawals explained? All tests	Unclear	See above

Itoi 2006

Clinical features and settings	***Inclusion criteria*** [1] Patients with shoulder pain retrospectively diagnosed as rotator cuff injuries. The definitive diagnosis was by arthroscopy. [2] Cuff tendinitis was diagnosed when all three of the following held true: painful arc; positive impingement test [Neer's sign? Neer's test?]; no detectable rotator cuff tears on arthroscopy. ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Japan (January 2000-December 2004) 160 shoulders in 149 patients (gender distribution not reported), mean age 53 years (SD not reported, range 16-86)
Study design	Retrospective, cross sectional study. Unclear whether inclusion consecutive.
Target condition and reference standard(s)	Supraspinatus tears (partial thickness or full thickness), infraspinatus tears (partial thickness or full thickness), subscapularis tears (partial thickness or full thickness). ***Arthroscopy*** *Procedure* No information given *Interpretation* Presence or absence of a rotator cuff tear was determined. By inference, the labrum, LHB, ACJ and glenohumeral joint were also examined. ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 81% (supraspinatus tears, partial thickness or full thickness), 50% (infraspinatus tears, partial thickness or full thickness), 18% (subscapularis tears, partial thickness or full thickness).
Index and comparator tests	***Full can test*** *Procedure* Referenced to the primary source (Kelly 1996) but not described *Interpretation* MMT grading from 0 to 5 (referenced), where 0 = no contraction; 1 = contraction without joint motion; 2 = ability to move the segment through its range of motion with decreased gravity [presumably with assistance from examiner]; 3 = ability to move the segment through its range of motion against gravity; 4 = more than 3 but less than 5; 5 = normal resistance to applied force. Grades were further subdivided by adding + or -. For the analysis, grades 0-3 were evaluated as criteria for a positive test. Pain was also noted. ***Empty can test*** *Procedure* The reference erroneously given for the empty can test is Jobe 1982, in which an isotonic manoeuvre was described for strengthening and testing strength. (The primary source for the empty can test of supraspinatus' integrity is Jobe 1983.) *Interpretation* As for full can test. This is compatible with the interpretation described in the primary source. ***External rotation strength*** *Procedure* Described as being performed with the elbow at the side and referenced to a textbook (Daniels & Worthington, 1980) *Interpretation* 'The strength ... was recorded.' For the analysis, grades 0-3 were evaluated as criteria for a positive test. It is unclear how grade 3 would have been determined unless the test was conducted in the side lying position. Pain was also noted.

Itoi 2006 *(Continued)*

	Strength to lift the hand off the back [lift-off test with force] *Procedure* Referenced to the primary source for the lift-off test (Gerber 1991a) but not the lift-off test with force (Kelly 1996) and not described. *Interpretation* 'The strength ... was recorded.' For the analysis, grades 0-3 were evaluated as criteria for a positive test. It is unclear how grade 3 would have been determined unless the test was conducted in the prone lying position. Pain was also noted. ***Tester(s)*** No information given
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** Yes, except the sensitivity of external rotation strength for pain to detect infraspinatus tears is a typographical error and should read 46%, not 54% (confirmed by communication with authors)

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The study was retrospective. Also the setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The exclusion criterion was applied retrospectively.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Mostly. some uncertainty was introduced by the application of strength grades to the tests.
Sufficient description of reference test? All tests	No	No technical information was given.

Itoi 2006 *(Continued)*

Index test results blinded? All tests	Yes	The study was retrospective, but results of physical tests were routinely documented on a standardised form at the time of collection.
Reference standard results blinded? All tests	No	The study was retrospective. It is unlikely that the arthroscopic results were interpreted without knowledge of the index tests.
Relevant clinical information? All tests	Yes	Most likely
Uninterpretable results reported? All tests	No	The study was retrospective.
Withdrawals explained? All tests	No	The study was retrospective.

Kibler 2009

Clinical features and settings	***Inclusion criteria*** [1] Shoulder pain, [2] consent to participate ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	USA (period not reported) 101 shoulders in 101 patients (58% male), mean age 49 years (SD 15, range 24-64)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	LHB and SLAP lesions ***Arthroscopy*** *Procedure* No information given *Interpretation* [1] Diagnoses of interest were: subscapularis tear (full or partial upper third), supraspinatus tear (full or partial articular sided), LHB injury (tear, instability or intrasubstance tendinopathy) and SLAP types I-IV. [2] For supraspinatus, subscapularis and LHB tendon injuries, criteria were based on observable anatomic disruption of the tendon. For SLAP lesions, the criteria included the 4-type classification of Snyder, the presence of a 'peel back' lesion and/or chondral damage on the superior glenoid rim. A degenerative appearance of the labrum was not included unless it met the other criteria (*See* Note 1). ***Interval between index and reference test*** Not reported ***Tester(s)*** One of 3 fellowship trained orthopaedic surgeons ***Prevalence*** 29% (LHB injury), 48% (labral injury)
Index and comparator tests	***TRADITIONAL TESTS*** ***Yergason's test*** Referenced to primary source (Yergason 1931) but *procedure* included a component of resisted external rotation at the shoulder which was not described by the test's originator.

Kibler 2009 *(Continued)*

Interpretation Pain over the bicipital groove was interpreted as a positive result, in accordance with the primary source. A supplementary positive finding, not described in the primary source, was subluxation of the LHB.

Speed's test

Referenced to a secondary source (Bennett 1998).

Procedure Performed from a starting position of 90°shoulder flexion.

Interpretation As standard

Belly press test

Referenced to a secondary source (Tokish 2003) but described as standard

Bear hug test

Referenced to the primary source (Barth 2006a)

Procedure As in the primary source

Interpretation A positive result was defined as inability to hold the hand on the shoulder. (In the primary source, but omitted here, is 'weakness of > 20% compared to the other side'.)

O'Brien's test [Active compression test]

Referenced to the primary source (O'Brien 1998a) and faithfully described

Anterior slide

Referenced to the primary source (Kibler 1995a)

Procedure Despite the first author of this paper having originated the test, it is described differently here, omitting the active component provided by the patient.

Interpretation A positive result was defined as pain or a painful click on the anterior or posterior joint line. Note the slightly different criteria in the original description: 'pain localized to the front of the shoulder under the examiner's hand, and/or a pop or click in the same area, was considered to be a positive test. This test is also positive if the athlete reports a subjective feeling that this testing maneuver reproduces the symptoms that occur during overhead activity' (Kibler 1995a).

NOVEL TESTS

Upper cut

Procedure The patient's shoulder was positioned in neutral, and with the elbow flexed to 90° he/she was asked to make a fist. The examiner, with his hand placed over the fist, applied isotonic resistance as the patient attempted to rapidly bring the hand up towards the chin, in the manner of a boxing upper-cut.

Interpretation Pain or a painful pop over the anterior portion of the involved shoulder during the resisted movement was interpreted as a positive result.

Modified dynamic labral shear [MDLS]

Procedure The patient was standing. The affected arm was flexed 90° at the elbow, elevated to above 120° of scaption, and externally rotated to tightness. It was then guided into maximal horizontal aBduction. The examiner then applied a shear load by maintaining external rotation and horizontal abduction while lowering the arm to 60° of scaption. Reportedly, this differs from the test described by O'Driscoll (no further citation information given) in that the arm was not placed into maximal horizontal aBduction until it was elevated above 120° (Reportedly, in pilot testing this modification was found to reduce the high number of false positive tests, with pain throughout the whole motion.)

*Interpretation R*eproduction of the pain and/or a painful click or catch along the posterior joint line between 120° and 90° scaption was interpreted as a positive result.

Kibler 2009 *(Continued)*

	Tester(s) One of 3 fellowship-trained orthopaedic surgeons who had practiced and agreed upon the performance and interpretation of the tests prior to commencement of the study. The authors were the originators of the upper cut test; and of the DLS in its modified form [MDLS].
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** It is unclear whether type I SLAP lesions were included or not. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. There were borderline discrepancies between numerous reported sensitivities and specificities and those derived from back-calculated 2 x 2 tables. There were numerous discrepancies, major and minor, in other summary statistics.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The information was very limited: shoulder pain. No exclusion criteria.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	No information given
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Yergason's and anterior slide tests differed from their primary sources in their execution; and Yergason's, the bear hug and the anterior slide test differed from their primary sources in their interpretation.
Sufficient description of reference test? All tests	No	No procedural information was given.
Index test results blinded? All tests	Yes	'At the time of examination, each physician recorded the results of each test by indicating whether they were positive or negative according to the established study criteria.'

Kibler 2009 *(Continued)*

Reference standard results blinded? All tests	No	'The operating physician, who was usually also the original evaluating physician, was aware of the preoperative clinical diagnosis but was not allowed to review the preoperative examination data sheet when completing the postoperative surgical data sheet.'
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	There do not appear to have been any withdrawals.

Kim 2001

Clinical features and settings	***Inclusion criteria*** Patients with shoulder pain ***Exclusion criteria*** [1] Stiff shoulder or [2] history of shoulder dislocation ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	South Korea (period not reported) 127 shoulders in 127 patients (70% male); mean age 31 years (SD not reported, range 15-52).
Study design	Prospective (most likely), cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Type II SLAP lesions **Arthroscopy** *Procedure* Not referenced or reported *Interpretation* Not referenced or reported ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 31% (Type II SLAP lesions)
Index and comparator tests	A novel test, the **biceps load test II** was evaluated. ***Tester(s)*** 2 independent testers. No other information given.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Interpretation of the biceps load test II was done independently of any other tests: clinical, radiographic and MRI. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description

Kim 2001 *(Continued)*

Representative spectrum? All tests	No	The care setting was tertiary and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	Inclusion and exclusion criteria very incompletely described
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	No information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	No information
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	No	Probably not. Testers were blinded to other data when they performed the test.
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Unclear	No information given, though the study was seemingly prospective and recruitment consecutive

Kim 2006

Clinical features and settings	***Inclusion criteria*** Patients with shoulder pain for > 3 months. MRI.

Kim 2006 *(Continued)*

	Exclusion criteria [1] 'Acute pain within 3 months', [2] bilateral shoulder pain, [3] previous surgery on the shoulder, [4] fracture, [5] inflammatory arthritis, [6] infection, [7] cervicogenic pain ***Duration of symptoms*** > 3 months (mean, SD and range not reported) ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Republic of Korea (Feb 2004–August 2005) 200 shoulders in 200 patients (42% male); mean age 60 years (SD not reported, range 37-83).
Study design	Prospective, consecutive cross sectional study
Target condition and reference standard(s)	Rotator cuff tears: partial thickness and full thickness ***MRI*** *Procedure* A 1.5-tesla magnetic-resonance scanner (General Electric, Milwaukee, WI, USA) was used. *Interpretation* A full thickness tear was diagnosed if a high signal intensity occupied the full thickness layer of the rotator cuff tendon on T2-weighted image in both the coronal and sagittal planes. A partial thickness tear was diagnosed when the fluid intensity signal within the tendons was in contact with only one of the surfaces. ***(Surgery)*** Was performed on a subgroup of 61 patients (see *Note 2*) *Procedure* no information given *Interpretation* No information given ***Interval between index and reference test*** Not reported ***Tester(s)*** MRI findings were interpreted by a musculoskeletal radiologic specialist. ***Prevalence*** 33% (complete tear in the rotator cuff), 69% (partial or complete tear in the rotator cuff)
Index and comparator tests	Neither test was referenced ***Empty can test*** *Procedure* 'The empty-can test was performed by subjects seated with arms abducted 90 degrees horizontally and rotated 45 degrees internally'. The test is usually performed in scaption, and this non-standard approach undermines clinical relevance. *Interpretation* Weakness was considered present if the patient could not resist a downward pressure applied by the examiner, or if strength was less than that of the contralateral side. Presence or absence of pain was also recorded. Thus there were 4 categories of positive response: 1. Pain, whether with or without weakness (P) 2. Weakness, whether with or without pain (W) 3. Just one sign of pain or weakness (P or W) 4. Pain and weakness combined (P and W) ***Full can test*** *Procedure* 'The full-can test was done with arms abducted 90 horizontally and rotated 45 degrees externally'. The test is usually performed in scaption, and this non-standard approach undermines clinical relevance. *Interpretation* As for the empty can test ***Tester***

Kim 2006 *(Continued)*

	No information given
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Subgrouping by age was: Patients aged 30-39: 2 Patients aged 40-49: 43 Patients aged 50-59: 72 Patients aged 60-69: 43 Patients aged 70-79: 43 Note that this totals 233, not the stated sample size of 200, and is inconsistent with the age range specified (whereby at least one patient was in his/her eighties) ***Note 2*** 'Sixty-one ... of the 200 shoulders were subsequently operated on, and their diagnoses were mostly re-confirmed.' More detail on agreement within the subgroup is given in the Discussion, but these incidental data only serve to emphasise that MRI, the study reference standard, was not perfectly accurate. ***Note 3*** The labels for the first two columns of data in Table 2 are transposed. The third and fourth columns are correctly labelled. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** There were no discrepancies between the reported sensitivities and specificities and those derived from back-calculated 2 x 2 tables. There were borderline discrepancies in some other summary statistics.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care.
Selection criteria described? All tests	No	The selection criteria were very incompletely described.
Acceptable reference standard? All tests	No	We considered MRI acceptable for full thickness tears in samples with a likely low prevalence. In this instance, MRI was the reference standard for both partial and full thickness tears and the prevalence was high. Moreover subsequent surgery on a subsample of the patients in this study (61/200) indicated imperfect agreement between MRI and surgery for partial and full thickness tears.
Acceptable delay between tests? All tests	Unclear	No information given

Kim 2006 *(Continued)*

Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	No	61 patients underwent surgery to verify their disease status. The report does not specify whether the indication for surgery was based on the MRI findings or the results of the index test. The diagnostic test accuracy analysis was based on the results of both MRI and surgery.
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	But, as described, the tests are not those in standard use.
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	No	There was no clear statement of blinding and it was also unclear whether the index test preceded the reference test ('The inclusion criterion was shoulder pain for more than 3 months which were undergone MRI').
Reference standard results blinded? All tests	Unclear	The reference test appears to have been interpreted independently.
Relevant clinical information? All tests	Unclear	Unclear
Uninterpretable results reported? All tests	Unclear	Although the study was prospective and recruitment was reportedly consecutive, there was a discrepancy between the age frequency distribution (n > 233) and the stated sample size (n = 200).
Withdrawals explained? All tests	No	See above

Kim 2007b

Clinical features and settings	***Inclusion criteria*** Painful shoulder joint ***Exclusion criteria*** [1] Fractures around the shoulder joint, [2] frozen shoulder ***Duration of symptoms*** Not reported ***Previous treatments*** All patients had undergone non-operative treatments for at least 2 months before surgery.This included various combinations of physical therapy, activity modification and steroid injection. Steroid injections were done at ≥ 6 week intervals up to a maximum of 3 injections. ***Care setting*** Secondary or tertiary care
Participants	Korea (April 2005 to February 2006) 61 shoulders in 61 patients (93% male); mean age 33 years (SD not reported, range 19-54).
Study design	Prospective, consecutive cross sectional study

Kim 2007b *(Continued)*

Target condition and reference standard(s)	SLAP lesions ***Arthroscopy*** 'The surgical procedures were performed with the patients in the lateral decubitus position. The operated arm was pulled with 10 lb (4.5 kg) of weight. In all cases, the glenohumeral joints were evaluated and definitive treatments performed on the lesions. We investigated the whole joint thoroughly with a scope via a posterior viewing portal and examined the stability of the superior labrum with a probe inserted through the superoanterior portal. All the SLAP lesions and the combined lesions were managed arthroscopically. Subsequently, the subacromial space was evaluated, and any coexisting pathological lesions were addressed.' The criterion for a positive results were instability of the glenoid labrum on probing and, implicitly, visible abnormalities of the shoulder joint, labrum or subacromial space. SLAP lesions were defined as disruption of the superior labrum between the 10.00 and 2.00 positions of the glenoid. These were classified into four types according to Snyder et al (1990). ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 46% (any SLAP lesion with or without other pathology)
Index and comparator tests	A novel test, the ***passive compression test*** , was evaluated. *Procedure* The patient was positioned in side lying with the affected shoulder uppermost, and asked to remain relaxed throughout. The tester stabilised the upper shoulder by holding the ACJ, controlling the patient's elbow with the other hand. With the patient's shoulder abducted 30°, and proximal pressure applied through the humerus, the examiner rotated the joint while moving it into extension. *Interpretation* Pain or a painful click elicited in the glenohumeral joint was interpreted as a positive result. ***Testers*** Two physicians who were co-originators of the test (Chung, 2011 (personal communication))
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The data used in the DTA calculations were derived from one tester who was randomly selected following data collection (Chung, 2011 (personal communication)). ***Note 2*** Evaluation of agreement between the 2 testers yielded a Kappa value of 0.771. This is in the range defined by Altman (1991) as 'good' agreement. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The selection criteria were very incompletely described.
Acceptable reference standard?	Yes	

Kim 2007b *(Continued)*

All tests		
Acceptable delay between tests? All tests	No	All patients had undergone non-operative treatment for at least 2 months before surgery.
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	No explicit statement of blinding was made, but this must have been in place for the inter-tester agreement aspect of the study.
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	No	The tests were conducted at initial evaluation before any other investigation (including history).
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	

Liu 1996b

Clinical features and settings	***Inclusion criteria*** Patients with activity related shoulder pain, worse on overhead motions, who had failed a 3-month trial of conservative management ***Exclusion criteria*** [1] History of traumatic dislocation or [2] clinical evidence of a rotator cuff tear (i.e. weak scaption, external rotation or lift-off) evident from the history or physical examination ***Duration of symptoms*** Not reported ***Previous treatments*** Most patients had undergone physical therapy elsewhere before attending the clinic. Episodes of care within the clinic commenced with conservative treatment, which involved NSAIDs, activity modification and physical therapy. Patients were selected after failure of pain relief despite 3 months' adherence to such treatment. ***Care setting*** Tertiary care
Participants	USA (April 1993 to August 1995)

Liu 1996b *(Continued)*

	62 shoulders in 62 patients (65% male); mean age 28 years (SD not reported, range 18-57) 85% had involvement of the dominant arm; 81% were recreational athletes; 56% had a history of trauma (without dislocation)
Study design	Reportedly prospective, cross sectional study. Unclear whether consecutive.
Target condition and reference standard(s)	Labral tears ***Arthroscopy*** *Procedure* In the lateral decubitus position, the glenohumeral joint, then the subacromial space, were inspected. *Interpretation* Cuff fraying, and Types II-IV SLAP and other labral tears were identified, with a distinction being made between labral tears and normal anatomical variants in the anterosuperior quadrant. ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given. ***Prevalence*** 52% (labral tears)
Index and comparator tests	A novel test, the ***crank test*** was evaluated. *Procedure* The patient was upright with the arm elevated to 160° in the scapular plane. While the glenohumeral joint was loaded along the axis of the humerus, humeral rotations were passively performed. However, 'The test should be repeated in the supine position, where the patient is usually more relaxed. Frequently a positive crank test in the upright position will also be positive in the supine position.' According to the Discussion, it may be necessary to repeat the manoeuvre several times. *Interpretation* The test was considered positive if there was [1] pain during its performance, usually on external rotation, with or without a click or [2] reproduction of the patient's symptoms, usually pain or catching. ***Tester(s)*** No information given, though the tester(s) may have been the originator(s) of the test
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** There is overlap between the period of this study and Liu 1996a, which also evaluated the crank test. if some patients were included in both this would involve an internal inconsistency, because the crank test technique and mode of interpretation differ between the two reports. (In Liu 1996a it was performed in 'maximal forward flexion' and a positive result was determined on the basis of clicking.) ***Note 2*** In addition, Liu 1996a was disputed in terms of patient numbers and actual primary diagnoses (as recorded at the time of MRI) by Seager (1997), from the Radiology Dept of the same institution; and the response (Liu 1997) reveals problems in methodology, especially patient flow, in that paper and potentially, by corollary, Liu 1996b. ***2 x 2 tables and summar ystatistics*** ***Were 2 x 2 tables reported?*** Yes ***If applicable, could 2 x 2 tables be back-calculated?*** N/A ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described?	Unclear	Exclusion criteria incomplete e.g. frozen shoulder not excluded

Liu 1996b *(Continued)*

All tests		
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	Partial description given
Index test results blinded? All tests	Yes	Unclear, but study was reportedly prospective
Reference standard results blinded? All tests	No	No statement of blinding
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Insufficient information
Withdrawals explained? All tests	Unclear	Insufficient information

MacDonald 2000

Clinical features and settings	***Inclusion criteria*** Patients with shoulder pain ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Canada (period not reported)

MacDonald 2000 *(Continued)*

	85 shoulders in 85 patients (73% male); mean age 40 years (SD not reported, range16-72)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Subacromial impingement: [1] SA-SD bursitis and [2] rotator cuff pathosis ***Arthroscopy*** *Procedure* Not reported *Interpretation* 'An appearance suggestive of subacromial tendonitis and bursitis' was recorded if all of three conditions were met. These were [1] erythema and bleeding of subacromial tissue, [2] swelling and difficult visualization, and [3] vascular engorgement of the bursal sides of the rotator cuff, causing bleeding from the surface of the cuff. ***Interval between index and reference test*** Not reported ***Tester*** A surgeon ***Prevalence*** 28% (bursitis), 28% (rotator cuff tendinosis)
Index and comparator tests	Tests were evaluated separately and in combination. ***Hawkins' test*** was referenced to its primary source (Hawkins 1980) and faithfully described. ***Neer's sign*** was referenced to a primary source (Neer 1983) but modified, being done in internal rotation, and the plane of elevation was not specified; but interpretation as in the primary source was implied. ***Tester*** The operating surgeon
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes, except in relation to Hawkins' test and Neer's sign as separately applied to the diagnosis of bursitis or rotator cuff pathosis. For these there were insufficient data for back-calculations. ***Were the reported summary statistics confirmed as accurate?*** Yes, except for the NPV of Hawkins and Neer's test (combined) for bursitis which was reported as 55.7% but back-calculated as 89.5%. Also the sensitivity of Neer's test for rotator cuff tendinosis was misreported in the abstract as 85% (versus 83.3% in the table).

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary care and the reference standard was more than minimally invasive.
Selection criteria described? All tests	No	Aside from being scheduled for surgery
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information

MacDonald 2000 *(Continued)*

Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Unclear	No statement of blinding but study reportedly prospective
Reference standard results blinded? All tests	No	No statement of blinding
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Insufficient information
Withdrawals explained? All tests	Unclear	Insufficient information

Michener 2009

Clinical features and settings	***Inclusion criteria*** [1] A primary complaint of shoulder pain, [2] shoulder pain present for > 1 week ***Exclusion criteria*** None reported ***Duration of symptoms*** Mean 33.8 months (SD 48.9 months, range 2-230 months) ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	USA (period not reported) 55 shoulders in 55 patients (86% male); mean age 41 years (SD 15.1, range18-83)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Subacromial impingement syndrome ***Arthroscopy*** *Procedure* No information given

Michener 2009 *(Continued)*

	Interpretation [1] Any combination of a visually enlarged or fibrotic bursa or degeneration of the bursal side of supraspinatus were considered criteria for a positive result. [2] Patients with additional pathology (e.g. partial thickness tears, full thickness tears, instability, labral tears and labral fraying) were not excluded. ***Interval between index and reference test*** Mean 2.6 months (SD 2.7 months, range 1 day-8 months) ***Tester*** 'An operative surgeon' ***Prevalence*** 29% (subacromial impingement syndrome)
Index and comparator tests	Tests were evaluated separately and, based on ROC analysis, 3 or more positive of the 5 tests was also evaluated. ***Neer's 'test' [Neer's sign]*** Referenced to a primary source (Neer 1983) but *procedure* described in a non-standard manner (the manoeuvre was conducted into flexion, not scaption). *Interpretation* as standard. ***Hawkins' test*** was referenced to its primary source (Hawkins & Kennedy, 1980) and faithfully described. ***Painful arc*** Referenced to Kessel 1977 with - as stated in that source - pain between 60° and 120° of active abduction being considered positive. ***Empty can test*** The reference erroneously given for the empty can test is Jobe 1982, in which an isotonic manoeuvre was described for strengthening and testing strength. (The primary source for the empty can test of supraspinatus' integrity is Jobe 1983.) But the *Procedure* and *Interpretation* described are compatible with the primary source. ***Resisted lateral rotation from neutral rotation*** Referenced to Park 2005 *Procedure* With the patient's elbow at his or her side and flexed to 90°, resistance to attempted external rotation was applied at the wrist. *Interpretation* Weakness was interpreted as a positive result. ***Tester*** Each patient underwent a standardised history and clinical examination by two clinicians; one surgeon, board certified in orthopaedic surgery, and one physical therapist, board certified in orthopaedics, with 17 and 18 years' experience in musculoskeletal examinations, respectively. The orthopaedic surgeon's results were used for the diagnostic test accuracy calculations (correspondence with author).
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Agreement (kappa) between the testers was: Hawkins' test 0.39 (0.12-0.65) Neer's test 0.40 (0.13-0.67) Painful arc 0.45 (0.18-0.72) Empty can 0.47 (0.22-0.72) Resisted lateral rotation from neutral rotation 0.67 (0.40-0.94) This level of agreement (according to Altman 1991 'fair' for Hawkins' and Neer's tests; 'moderate' for painful arc and the empty can test; and 'good' only for resisted lateral rotation from neutral rotation), despite training to standardise technique and interpretation, undermines the validity of the diagnostic test accuracy results.

Michener 2009 *(Continued)*

Note 2 A ROC curve is presented from which are derived the cut point (number of + or - tests) for ruling in or out 'subacromial impingement syndrome'.

2 x 2 tables and summarystatistics

Were 2 x 2 tables reported? No

If applicable, could 2 x 2 tables be back-calculated? Yes

Were the reported summary statistics confirmed as accurate? Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The selection criteria were very unclearly described ('patients with shoulder pain') and the exclusion criteria were undescribed.
Acceptable reference standard? All tests	Yes	
Acceptable delay between tests? All tests	No	Mean 2.6 months
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	No procedural information was given.
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	There was a clear statement of blinding.
Relevant clinical information? All tests	Yes	

Michener 2009 *(Continued)*

Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	There were none.

Miller 2008b

Clinical features and settings	***Inclusion criteria*** [1] Shoulder pain (c5-6 dermatome), [2] full passive range of movement of the affected shoulder, [3] age > 18 years ***Exclusion criteria*** Previous surgery to the affected upper limb ***Duration of symptoms*** [1] All: mean 37.5 months (SD not reported, range 4-120 months). [2] No full thickness tear: mean 29.8 months (SD 38.5 months, range not reported). [3] Full thickness tear: mean 48.2 months (SD 55.6 months, range not reported). ***Previous treatments*** Not reported ***Care setting*** Probably secondary
Participants	UK ('a period of approximately 6 months' but date range not reported) 46 shoulders in 37 patients (43% male); mean age 55.5 years (SD not reported, range 20-86)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Full thickness rotator cuff tears ***Ultrasound*** *Procedure* Using a Philips ATL HDI5000a and 5–12 MHz linear array transducer with compound imaging, each component of the rotator cuff was examined in a standardized manner to minimize any random errors. *Interpretation* A full thickness tear was diagnosed if the tendon was completely absent or if the tear traversed from the articular to the bursal aspect of the tendon. ***Interval between index and reference test*** Physical tests immediately preceded ultrasound ***Tester*** A single consultant radiologist who specialized in ultrasonography of the shoulder performed all ultrasound examinations. A retrospective audit compared the ultrasonographer's diagnosis of a rotator cuff tear with the diagnosis at arthroscopy performed by the consultant orthopedic surgeon specializing in shoulder conditions. These results demonstrated that the ultrasonographer's diagnosis was identical to the diagnosis made during arthroscopy on each occasion in 10 subjects. ***Prevalence*** 33% (full thickness tear of supraspinatus and infraspinatus combined), 9% (full thickness tear of subscapularis)
Index and comparator tests	***External rotation lag sign (ERLS)*** Referenced to the primary source (Hertel 1996a) and faithfully described except that *Interpretation* was simplified: 'Inability to maintain this position would suggest a full thickness tear of [the supraspinatus and infraspinatus] tendons'. ***Internal rotation lag sign (IRLS)*** As above ***Drop sign*** Referenced to the primary source (Hertel 1996a) and faithfully described except that the starting position was full external rotation (the primary source described the arm being positioned just short of full range)

Miller 2008b *(Continued)*

	Tester 'A specialist physiotherapist with 5 years of experience in outpatient musculoskeletal practice and a special interest in shoulder pathology. To ensure test quality, the clinical tests were practiced on 5 separate occasions with an orthopedic surgeon with a special interest in shoulders on a separate group of subjects (n=10).'
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** There was a borderline discrepancy between the reported sensitivity and that derived from the back-calculated 2 x 2 table. There were also borderline discrepancies between the reported NPV for the drop sign and the PPV for the IRLS and those derived from the respective back-calculated 2 x 2 tables. The reported calculations for false-positive and false-negative rates in respect of the ERLS and drop sign were incorrect, however.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Unclear	The care setting is uncertain
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Unclear	Ultrasonography
Acceptable delay between tests? All tests	Yes	The reference test immediately followed the index test.
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded?	Yes	

Miller 2008b *(Continued)*

All tests		
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	There appear to have been none.

Naredo 2002

Clinical features and settings	***Inclusion criteria*** Patients with 'first flare' of shoulder pain (although this seems incompatible with the reported range of durations of symptoms, below) clinically diagnosed as a 'peri-articular' disorder ***Exclusion criteria*** [1] Previous trauma, [2] chronic inflammatory arthritis ***Duration of symptoms*** 1-48 months ***Previous treatments*** Not reported ***Care setting*** Secondary care
Participants	Spain (period not reported) 31 shoulders in 31 patients (13% male); mean age 58 years (SD not reported, range 21-77)
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Periarticular shoulder lesions: subacromial impingement in real time, SA-SD bursitis, tendinitis, rotator cuff tears ***Ultrasonography*** *Procedure* Using a 7.5 MHz linear array transducer, transverse and longitudinal planes from the bicipital groove, rotator cuff and SA-SD bursa were scanned bilaterally, and a dynamic evaluation conducted for impingement syndrome. References are given. *Interpretation* Standard criteria were used, based on those widely described in the literature. These criteria are described and tabulated in the report, and references given. ***Interval between index and reference test*** ≤ 1 week ***Tester(s)*** A third rheumatologist, who was experienced in ultrasonographic examination ***Prevalence*** 94% (supraspinatus lesions), 58% (supraspinatus tendinitis), 52% (supraspinatus tears), 55% infraspinatus lesions, 23% (infraspinatus tendinitis), 35% (infraspinatus tears), 39% (subscapularis lesions), 19% (subscapularis tendinitis), 23% (subscapularis tears), 61% (LHB tendinitis), 45% (SA-SD bursitis), 65% (subacromial impingements)
Index and comparator tests	Evaluated in combination were ***Hawkins' test*** and ***Neer's sign*** , referenced to their primary sources (Hawkins 1980; Neer 1977) and faithfully described. '***Yocum's' test*** , mistakenly attributed to Yocum 1983. *Procedure* According to Naredo 2002 the test involves the patient placing the palm of the affected upper limb on his or her other shoulder and, while keeping the point of the affected shoulder down, raises that elbow. ***Empty can test*** was referenced to its primary source (Jobe 1983) and faithfully described. ***Patte's test*** was referenced to the primary English-language source Leroux 1995 and faithfully described. *Procedure* The examiner supports the patient's elbow in 90° scaption and asks the patient to rotate the arm laterally. *Interpretation* Strength is compared bilaterally. The ability to resist despite pain was interpreted as

Naredo 2002 *(Continued)*

tendinitis, while the inability to resist with gradual lowering of the arm or forearm was interpreted as tendon rupture.

Lift-off test Referenced to a primary source (Gerber 1991a) and described. *Procedure* The standard simplified form of the test, with the tester lifting the patient's hand clear of his or her lumbar spine and the patient trying to maintain this position. *Interpretation* The patient's inability to hold the hand off the lumbar spine was taken to denote complete rupture of subscapularis. As distinct from the primary source, no intermediate stage was considered. ***Resisted internal rotation*** Referenced to Gerber 1991a though no details of procedure or interpretation are given in that paper, and not described further.

Yergason's test Referenced to a secondary source (Sheon 1987) and described in keeping with the primary source (Yergason 1931). There was also a supplementary procedure for subluxation of the LHB tendon. The patient was asked to perform a combined movement of elbow flexion and shoulder medial rotation against the examiner's resistance. A positive response was defined as reproduction of the patient's abnormal sensation.

Palm up test Referenced to its primary source (Gilcreest 1936) and described. *Procedure* The patient was asked to elevate the arm anteriorly against resistance, elbow extended and palm upward. *Interpretation* Positive if the patient feels pain along the course of the LHB. Omitted were 2 elements described in the primary source: [1] lowering through the coronal plane, between 110° and 90° of which a snap might be audible accompanied by a sharp pain in the shoulder and bicipital groove; and [2] palpation throughout for tendon subluxation.
Tester(s) 2 rheumatologists performed independent assessments, then established clinical diagnoses by consensus.

Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** No data were reported for the diagnostic accuracy of individual index tests. An overall result from a physical examination, which in addition to the above tests included evaluation of active and passive range and inspection for the Popeye sign, were presented. ***Note 2*** 'Surgical results were not available [as a potential reference standard] for most of the patients'. It is unclear at what stage surgery was performed on any of these patients. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. There were numerous discrepancies between reported summary statistics and those derived from the back-calculated 2 x 2 tables. Most were borderline, but some were substantial and one of the latter involved sensitivity (for a subscapularis tear, reported as 50% *versus* the back-calculated value of 42.86%).

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	Setting was secondary care
Selection criteria described? All tests	Unclear	'First flare' of shoulder pain seems unlikely at 48 months.
Acceptable reference standard? All tests	Unclear	Ultrasonography. There is some mention of surgery in some patients: unclear.

Naredo 2002 *(Continued)*

Acceptable delay between tests? All tests	Yes	Within 1 week
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	Prospective and consecutive; 2 blinded rheumatologists
Reference standard results blinded? All tests	Yes	Rheumatologist experienced in technique and with no knowledge of physical examination findings
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Consensus, but no information given on individuals' test results and differences between them
Withdrawals explained? All tests	Unclear	Insufficient information

Norwood 1989

Clinical features and settings	***Inclusion criteria*** Operative repair or reconstruction for a complete rotator cuff tear ***Exclusion criteria*** Rotator cuff tear associated with fracture or dislocation ***Duration of symptoms*** Most of Group 1 (*see* Participants, below) had had symptoms for a mean of 30 months before operative treatment. ***Previous treatments*** 75% of Group 1 had received previous 'extensive treatment' for chronic symptoms; only 33% of Group 2 had had previous symptoms sufficiently severe for them to seek treatment. ***Care setting*** Tertiary care
Participants	USA (1976-1984) 103 shoulders in 103 patients (75% male); mean age 54.1 years (SD not reported, range 18-73). These were subgrouped into individuals with single tendon tears (n = 28; Group 1) and those with multiple tendon tears (n = 75; Group 2). There was an additional control group, data for which were not presented in a usable format.

Norwood 1989 *(Continued)*

Study design	Retrospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Distinguishing single tendon tears from multiple tendon tears ***Operation (presumably open surgery)*** *Procedure* No details were given. *Interpretation* Criteria for a positive result were [1] tears restricted to the tendinous insertion above the spine of scapula (defined as single tendon tears) and [2] tears extending posteriorly below this level (involving infraspinatus with or without teres major) or anteriorly (involving subscapularius) (defined as multiple tendon tears). [3] Additionally noted were calcific deposits, ruptures of the LHB, and tears of the glenoid labrum. ***Interval between index and reference test*** Not reported ***Testers*** No information given ***Prevalence*** 73% (multiple tendon tears)
Index and comparator tests	***Active abduction*** *Procedure* No details were given. *Interpretation* The criterion for a positive result was inability to abduct to ≥ 90°. This criterion may have been defined retrospectively.
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No summary diagnostic accuracy statistics were reported, but these were calculated as sensitivity = 84.29, specificity = 76.92, PPV = 90.77 and NPV = 64.52.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (surgery) was more than minimally invasive. Also the sample was selected (all had cuff tears).
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Yes	Surgery
Acceptable delay between tests? All tests	No	Most of group 1 had had symptoms for a mean of 30 months before operative treatment.
Partial verification avoided? All tests	Yes	There was an non-operated control group which was disregarded for the purposes of this review.
Differential verification avoided? All tests	Yes	

Norwood 1989 *(Continued)*

Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Probably
Sufficient description of reference test? All tests	No	No procedural details given
Index test results blinded? All tests	No	No statement of blinding and study was retrospective
Reference standard results blinded? All tests	No	There does not appear to have been blinding.
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Study was retrospective
Withdrawals explained? All tests	No	Study was retrospective and withdrawals were unexplained

O'Brien 1998

Clinical features and settings	***Inclusion criteria*** Shoulder pain ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Unclear: possibly elements of primary, secondary and tertiary care
Participants	USA (period not reported) Of the patients who underwent the active compression test, 206 were categorized as testing either [1] positive for labral tears or [2] negative for both labral tears and ACJ abnormality. These patients were included in the calculations of diagnostic test accuracy for labral tears. A further subgroup [3] of 62 patients, who tested positive for ACJ abnormality, might presumably have been included as 'test negative' in the calculation of diagnostic test accuracy for labral tears, but were not. A group of 62 patients believed to have ACJ disorders on the basis of the index test were also excluded from the diagnostic test accuracy calculations for SLAP lesions. (Fifty-five of these were verified by different combinations of X-ray, MRI and surgery - 32 by radiography alone.) No breakdown of patients by gender or age was given.
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Labral tears Various combinations of ***radiography, MRI*** and ***surgery*** . No *procedural* or *interpretative* details were given. ***Interval between index and reference test*** Not reported ***Tester(s)*** No information given ***Prevalence*** 26% (any labral tears)

O'Brien 1998 *(Continued)*

Index and comparator tests	A novel test, the ***active compression test*** , was evaluated. ***Tester(s)*** No information given, though the tester may have been the originator of the test
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The report is contradictory. Thus, "... we performed an active compression test on 268 consecutive patients with shoulder pain who had had no prior diagnostic evaluation", but the "examinations were performed without the examiner knowing the radiographic examination of the shoulder". ***Note 2*** The statement that 50 patients "who had not had complete or properly performed examinations to confirm the accuracy of the test were not included in the study" is a matter for concern. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** The reported results and our own were discrepant because we excluded the control group (with knee pain) from our calculations. Thus specificity was recalculated as 150/203 = 98.0%, and NPV as 153/206 = 74.3%.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary and one of the reference tests used (surgery) was more than minimally invasive
Selection criteria described? All tests	No	'Shoulder pain'
Acceptable reference standard? All tests	Unclear	Varied according to condition
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	No	Not all patients received a reference test
Differential verification avoided? All tests	No	Patients underwent different reference tests depending on the index test result
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	No information given

O'Brien 1998 *(Continued)*

Index test results blinded? All tests	Yes	Probably
Reference standard results blinded? All tests	No	There is no statement of blinding
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	No	There were 50 incomplete and improperly conducted tests
Withdrawals explained? All tests	Unclear	Insufficient information

Oh 2008

Clinical features and settings	***Inclusion criteria*** Patients who had undergone shoulder arthroscopy ***Exclusion criteria*** [1] Adhesive capsulitis, [2] septic shoulder, [3] fractured greater tuberosity ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	Korea (January 2004-July 2005) 146 shoulders in 146 patients 68 with SLAP lesions (73.5% male), mean age 45 years (SD not reported, n = 28 < 40 years, n = 40 > 40 years); 78 without SLAP lesions (56.4% male), mean age 44 years (SD not reported, n = 33 < 40 years, n = 45 > 40 years).
Study design	Retrospective, case-control type accuracy study with non-consecutive recruitment
Target condition and reference standard(s)	Type II SLAP lesions ***Arthroscopy*** *Procedure* 'Patients were in the lateral decubitus position with the affected arm in an arm holder under 10 lb of traction. The superior labrum complex was palpated with a probe to determine the type of SLAP lesion. *Interpretation* SLAP lesions were graded according to Snyder 1990a. When the "peel back" phenomenon was observed in the abduction and external rotation position, the superior labrum was elevated more than 5 mm with a cartilaginous crack, and a haemorrhagic spot or inflammatory granulation tissue beneath the detached superior labrum were observed, the lesion was diagnosed as a type II SLAP lesion. ***Interval between index and reference test*** Mean 1 day ***Tester(s)*** The senior author, an orthopaedic surgeon ***Prevalence*** 47% (type II SLAP lesions)
Index and comparator tests	***Speed's test*** Incorrectly referenced. Performed isometrically at 90° of shoulder flexion with pain into the biceps region being interpreted as a positive result. ***Yergason's test*** No reference given, but performed and interpreted as standard

Oh 2008 *(Continued)*

Anterior apprehension test Referenced to Rowe 1981, whose description of the test differs from that of Jobe 1989 in terms of technique (an anterior force is applied to the humerus) and interpretation, whereby both pain and instability are considered positive. (Jobe 1989 considered a positive response for anterior subluxation 'pain but not apprehension' and subsequent reports by Jobe 1995; Jobe 1996 explicitly link pain 'rather than apprehension' on this test to posterosuperior glenoid impingement.)

Relocation test Secondarily referenced to Kvitne & Jobe (1993) and described and interpreted as standard

Compression-rotation test (Snyder 1990a), ***Anterior slide test*** (Kibler 1995a) and ***biceps load II test*** (Kim 2001 were referenced to their primary sources, as indicated, and faithfully described.

The ***active compression test*** was referenced to its primary source (O'Brien 1998a) but not clearly described: there was no clear indication of the direction in which the resistance was applied.

The ***Whipple test*** was referenced to (Savoie 2001) and described.

Bicipital groove tenderness was sought by gentle palpation over the bicipital groove with the shoulder abducted to about 10°. A report of pain was considered a positive finding.

Tester(s) The first author: an orthopaedic surgeon

Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. There were some borderline discrepancies between reported sensitivities, specificities and other summary statistics and those derived from the back-calculated 2 x 2 tables. There were also numerous instances where the reported PPV or NPV differed substantially from the back-calculated value.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The study was case-control and therefore not prospective. The reference test (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The inclusion criteria were very limited.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Yes	Surgery was performed on the day after the physical assessment.
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	

Oh 2008 *(Continued)*

Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	No	The study was retrospective.
Reference standard results blinded? All tests	No	
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	The study was not prospective.
Withdrawals explained? All tests	No	The study was not prospective.

Parentis 2006

Clinical features and settings	***Inclusion criteria*** Patients undergoing shoulder arthroscopy as the initial step in their surgical procedure ***Exclusion criteria*** Adhesive capsulitis ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary care
Participants	USA (October 1999 to April 2000). 132 shoulders in 132 patients (74% male); mean age 42 years (SD not reported, range 15-71).
Study design	Prospective, cross sectional study with consecutive recruitment
Target condition and reference standard(s)	Type II SLAP lesions ***Arthroscopy*** *Procedure* and *interpretation* were referenced to Snyder et al (1990) ***Interval between index and reference test*** Not reported ***Tester(s)*** No details given ***Prevalence*** 17% (Type II SLAP lesions)
Index and comparator tests	The ***crank test*** (in 160° flexion), The ***modified relocation test for posterosuperior glenoid impingement*** (mislabelled ***Jobe's relocation test)*** and the ***pain provocation test*** were referenced to their primary sources (Liu 1996b; Hamner 2000; Mimori 1999a) and faithfully described. The remaining index tests were also referenced to their primary sources but modified, or possibly so. In the ***active compression test*** (O'Brien 1998a), the patient was not asked to distinguish between pain inside or on top of the shoulder, and the test was interpreted as positive if the patient's pain was reproduced. In the ***anterior slide test*** the tester's hand was placed 'anteriorly over the shoulder with the fin-

Parentis 2006 *(Continued)*

	gers at the acromion', and it is unclear from the wording whether this position tallies with that in the primary source (Kibler 1995a). ***Hawkins' test*** (Hawkins 1980) was modified by the addition of a shoulder adduction component. ***Neer's sign*** (Neer 1983) was modified both in terms of its plane of movement - 'passive forward flexion of the shoulder' - and its apparent omission of shoulder girdle stabilization. It was interpreted as positive if the patient's pain was reproduced. ***Speed's test*** (referenced to Gilecreest 1939: not obtained by us) and ***Yergason's test*** (unreferenced) were modified in terms of their interpretation, such that reproduction of patients' pain (as opposed to provocation of pain localized to the bicipital groove) was considered positive. ***Tester(s)*** Three orthopaedic surgeons who had completed 3 months of their 1-year sports medicine fellowship at the outset of data collection. For uniformity, each was instructed in the techniques of the test procedures.
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The 2002 and 2006 articles report the same study, respectively as a brief report embedded in a review and a detailed report. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No. There were no discrepancies in sensitivity or specificity, but one borderline and one substantial discrepancy in NPVs derived from the back-calculated 2 x 2 tables.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	Very broad: patients undergoing shoulder arthroscopy. Only patients with frozen shoulder were excluded.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Physical examination was conducted 'pre-operatively'
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests?	Yes	Mostly yes. Standard tests, most referenced and adequately described.

Parentis 2006 *(Continued)*

All tests		
Sufficient description of reference test? All tests	Yes	Not directly, but a reference was given
Index test results blinded? All tests	Yes	Data were recorded on a standardized form that was included with the patient's preoperative paperwork.
Reference standard results blinded? All tests	No	The preoperative data sheets appear to have been available at the time of surgery (*see* below).
Relevant clinical information? All tests	Yes	Most probably
Uninterpretable results reported? All tests	Yes	Apparently so
Withdrawals explained? All tests	Yes	As above

Razmjou 2004

Clinical features and settings	***Inclusion criteria*** Patients with shoulder problems referred to either of two centres and meeting the criteria for surgery: [1] persistent pain and functional disability > 6 months, unresponsive to conservative treatment and with a positive impingement test confirmed with local anaesthesia; [2] clinical signs of rotator cuff tears, labral tears (Bankart, superior labrum) or LHB lesions; [3] symptoms referred to the ACJ > 6 months, with visible radiographic changes in the joint; or [4] symptomatic shoulder instability. Written informed consent. ***Exclusion criteria*** Previous shoulder surgery or upper extremity fractures ***Duration of symptoms*** 10 < 6 months; 30 = 6 months-5 years; 10 > 5 years ***Previous treatments*** Conservative ***Care setting*** Tertiary care
Participants	Canada (period not reported) 50 (68% male), median age 50 years (range 24-79)
Study design	Prospective, cross-sectional study with consecutive recruitment.
Target condition and reference standard(s)	Impingement syndrome ***Open or arthroscopic surgery*** *Procedure* No details given *Interpretation* Criteria for a positive result included: [1] rotator cuff tendinitis (inflammation or swelling); [2] subacromial impingement secondary to existing anterior acromial spur, arthritis in the ACJ, and/or coracoacromial ligament impingement; [3] partial thickness tears of supraspinatus, whether on the bursal or articular surface; [4] full thickness tears of supraspinatus, with or without involvement of other cuff tendons. ***Tester(s)*** One experienced surgeon
Index and comparator tests	***Neer's 'test' [sic]: Neer's sign is probably intended*** Referenced to Neer 1972a which does not clearly describe Neer's test, stating only 'pain at the anterior edge of the acromion on forced elevation'. From the description given in Razmjou 2004, Neer's sign was intended, and this was performed and interpreted as standard.

Razmjou 2004 *(Continued)*

	Hawkins' test Referenced to the primary source (Hawkins 1980) and faithfully described. ***Interval between index and reference test*** Mean 23 weeks (range 5 weeks-11 months) ***Tester(s)*** Experienced physiotherapist and surgeon ***Prevalence*** 88% (impingement syndrome)
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Intertester agreement was also evaluated. For Neer's test kappa was 0.506 (CI 0.366-0.645), denoting 'moderate' agreement (Altman, 1991). For Hawkins' test kappa was 0.29 (CI 0.180-0.398), denoting 'fair' agreement (Altman, 1991). ***Note 2*** Of 150 patients in the study, 50 met the criteria for surgery. ***Note 3*** Thirty-two patients had weakness. Symptoms were improving in 9 patients and worsening in 20. ***Note 4*** The study population was the same as in Holtby 2004a; Holtby 2004b. ***Note 5*** Long term results were reported but disregarded for the purposes of this review. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** Yes ***If applicable, could 2 x 2 tables be back-calculated?*** N/A ***Were the reported summary statistics confirmed as accurate?*** There were borderline discrepancies in specificity and PPV for Hawkins' test.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (open or arthroscopic surgery) was more than minimally invasive.
Selection criteria described? All tests	Yes	
Acceptable reference standard? All tests	Yes	
Acceptable delay between tests? All tests	No	23 weeks
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests?	Yes	

Razmjou 2004 *(Continued)*

All tests		
Sufficient description of reference test? All tests	No	No procedural details given
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	'The examining surgeon had no access to the initial assessment form or results of the specific tests at the time of completing the surgical data collection forms.' The same surgeon conducted both index and reference tests, but the interval between the two was sufficient to ensure blinding.
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Unclear
Withdrawals explained? All tests	Yes	None

Schlechter 2009

Clinical features and settings	***Inclusion criteria*** Patients who had undergone arthroscopy for shoulder pain/dysfunction ***Exclusion criteria*** [1] Inability to elevate the arm to 150° or inability to do so comfortably, [2] previous arthroscopy on the same shoulder ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported (except that none had previously undergone arthroscopy) ***Care setting*** Secondary or tertiary care
Participants	USA (May 2001 to November 2003) 254 shoulders in 246 patients (65% male), mean age 44 years (SD not reported, range 13-84)
Study design	Retrospective, cross-sectional study with reportedly consecutive inclusion.
Target condition and reference standard(s)	SLAP lesions types II, III or IV or LHB anchor instability (considered an incomplete type II SLAP lesion) ***Arthroscopic surgery*** *Procedure* No details given *Interpretation* 'A positive arthroscopic finding for a SLAP lesion was recorded if the glenoid labrum showed clinically significant type II, III or IV changes by use of the original classification of Snyder et al [reference given] or if it was believed that the LHB anchor was unstable and the pathology matched the history, clinical presentation and symptoms. LHB anchor instability was diagnosed by applying traction to the LHB tendon with a nerve hook and observing either 'fish mouthing' at the labrum-glenoid interface or significant splitting of the superior labrum. This arthroscopic finding may be considered an incomplete type II SLAP tear.' ***Tester(s)*** The senior author, an orthopaedic surgeon
Index and comparator tests	The tests were conducted as part of a comprehensive assessment which included history. ***Active compression test*** Referenced to its primary source (O'Brien 1998a) and faithfully described

Schlechter 2009 *(Continued)*

Anterior slide test

Though referenced to its primary source (Kibler 1995a), the test was not described as standard, the patient-applied resistance being absent. Interpretation was standard, however.

Passive distraction test

This test was referenced to Rubin (2003), one of the reports authors, and described thus: the shoulder was elevated to 150° in the coronal plane with the elbow extended, the forearm supinated, and the arm stabilised to prevent humeral rotation. If this position was reasonably comfortable, the forearm was pronated. Pain reported deep inside the shoulder joint, either anteriorly or posteriorly, was interpreted as positive.

Interval between index and reference test Not reported
Tester(s) The senior author, an orthopaedic surgeon
Prevalence 24% (type II, III or IV SLAP lesion)

Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** Although evaluation of reproducibility was mentioned in the abstract, this aspect was not addressed in the report. ***Note 2*** For the passive distraction test and active compression test in combination, 16/234 (6.30%) of the results were indeterminate (situations where there must have been one test positive and the other negative). This left just 238 shoulders to report in the 2 x 2 table. ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** There were no discrepancies in sensitivity or specificity derived from the back-calculated 2 x 2 tables, but there were borderline discrepancies in other summary statistics.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference standard (arthroscopic surgery) was more than minimally invasive.
Selection criteria described? All tests	No	The selection criteria were very limited.
Acceptable reference standard? All tests	Yes	Athroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided?	Yes	

Schlechter 2009 *(Continued)*

All tests		
Incorporation avoided? All tests	No	At operation the pathology was 'matched [to] the history, clinical presentation and symptoms.'
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	No	There was no clear statement of blinding, and the study was retrospective.
Reference standard results blinded? All tests	No	
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	The study was not prospective. In relation to the passive distraction test and active compression test interpreted in combination, indeterminate results could be deduced.
Withdrawals explained? All tests	No	The study was not prospective.

Speer 1994

Clinical features and settings	***Inclusion criteria*** Patients undergoing shoulder surgery ***Exclusion criteria*** Multiple diagnoses e.g. multidirectional instability or co-existing anterior instability and treatable or observable rotator cuff lesions. Test not performed if patient had received intravenous sedation. ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	USA (period not reported) 100 shoulders in 100 patients (gender and age distribution not reported)
Study design	Prospective, cross-sectional study. Unclear whether consecutive recruitment.
Target condition and reference standard(s)	Rotator cuff disease ***Open or arthroscopic surgery*** *Procedure* No details given *Interpretation* No details given ***Tester(s)*** Not stated, but the procedures took place at a Hospital for Special Surgery
Index and comparator tests	***Relocation test for pain (modified)*** As in the primary source (Jobe 1989) but lacking an anteriorly applied force

Speer 1994 *(Continued)*

Procedure The patient lay supine with the lateral half of their scapula over the edge of the couch. The shoulder was positioned in 90° abduction and 90° of external rotation and pain assessed ('yes' or 'no'). Patients who could not attain this position were excluded. A posterior force was then applied to the humerus in order to determine whether there was a reduction in pain.

Interpretation Criteria for a positive result: a report of pain, diminished on the posteriorly applied force. Criteria for a negative result: no pain in the 90°/90° position or no reduction of pain with the posteriorly applied force.

Relocation test for pain

As in the primary source (Jobe 1989)

Procedure In the 90°/90° position an anterior force was applied to the humerus, and pain assessed ('yes' or 'no'). Patients who could not attain this position were excluded. A posterior force was then applied to the humerus in order to determine whether there was a reduction in pain.

Interpretation Criteria for a positive result: a report of pain, diminished on the posteriorly applied force. Criteria for a negative result: no pain in the 90°/90°*ANT position or no reduction of pain with the posteriorly applied force.

Interval between index and reference test Same day

Tester(s) Two of the authors, using a standardised protocol.

Prevalence 34% (rotator cuff disease)

Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No summary diagnostic test accuracy statistics were reported for rotator cuff disease. These were calculated as [1] 90°/90°: sensitivity 82.35%, specificity 46.97%; [2] 90°/90°*RELOC: sensitivity 44.12%, specificity 66.67%; [3] 90°/90°:ANT: sensitivity 88.24%, specificity 33.33%; [4] 90°/90°:ANT*RELOC: sensitivity 55.88%, specificity 46.97%.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (open or arthroscopic surgery) was more than minimally invasive.
Selection criteria described? All tests	No	Very unclear: 'patients who underwent shoulder surgery'
Acceptable reference standard? All tests	Yes	
Acceptable delay between tests? All tests	Yes	Same day
Partial verification avoided? All tests	Yes	
Differential verification avoided?	Yes	

Speer 1994 *(Continued)*

All tests		
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	No information
Index test results blinded? All tests	Unclear	No statement of blinding but study was prospective and the test response (painful, yes or no) required no interpretation by the tester.
Reference standard results blinded? All tests	No	There does not appear to have been blinding
Relevant clinical information? All tests	Yes	Assuming the tester was the surgeon about to undertake the operation
Uninterpretable results reported? All tests	Unclear	Insufficient information
Withdrawals explained? All tests	Unclear	Not enumerated

Stetson 2002

Clinical features and settings	***Inclusion criteria*** 'Initial' symptoms of shoulder pain in patients who (by implication) had not responded to conservative therapy ***Exclusion criteria*** None reported ***Duration of symptoms*** Mean 12 months ***Previous treatments*** A minimum of 3 months' non-operative treatment including rest, physical therapy, non-steroidal anti-inflammatory medication and a subacromial steroid injection when indicated for a localized bursitis or partial cuff tear ***Care setting*** Secondary or tertiary care
Participants	USA (period not reported) 65 (69% male); mean age 50 years (SD not reported, range 18-75)
Study design	Prospective, cross-sectional study. Unclear whether inclusion consecutive.
Target condition and reference standard(s)	Labral tears ***Arthroscopy*** *Procedure* A complete 15-point glenohumeral examination was performed. Special attention was directed to the glenoid labrum and LHB anchor. All results were recorded. *Interpretation* Slight fraying of the superior labrum consistent with a Type I SLAP lesion was evaluated a normal variant, as was fraying of the anterior and posterior labrum. All other superior labral tears were evaluated pathologic and classified as II-IV or complex (combinations of II-IV). Anterior and posterior labral tears were also recorded. Snyder's classification of SLAP lesions is included and referenced.

Stetson 2002 *(Continued)*

	Interval between index and reference test Not reported ***Tester(s)*** No information given ***Prevalence*** 40% (labral tears)
Index and comparator tests	The ***active compression test*** and the ***crank test (at 160°)*** , both referenced to their primary sources (O'Brien 1998a; Liu 1996b) and faithfully described except that, for the active compression test, not only pain but also clicking was interpreted as positive. ***Tester(s)*** One surgeon: the 'senior' author
Follow-up	***Adverse events*** None mentioned
Notes	***Note 1*** The authors note the difference in mean age of their sample and that of Liu 1996b (45.9 years versus 28 years); also that they did not exclude patients with rotator cuff tears (71% had either a full or partial thickness rotator cuff tear) whereas Liu 1996b did exclude most of this group (only 10 patients had as partial thickness rotator cuff tear). ***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** Cell contents were given for 2 x 2 tables, although these were not presented in tabular format. ***If applicable, could 2 x 2 tables be back-calculated?*** N/A ***Were the reported summary statistics confirmed as accurate?*** Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference standard (arthroscopy) was more than minimally invasive
Selection criteria described? All tests	No	'Initial' symptoms of shoulder pain
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	

Stetson 2002 *(Continued)*

Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Unclear	Unclear whether there was blinding, but study was prospective
Reference standard results blinded? All tests	No	There does not appear to have been blinding
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Unclear	Recruitment may not have been consecutive
Withdrawals explained? All tests	Unclear	As above

Suder 1994

Clinical features and settings	***Inclusion criteria*** [1] Chronic shoulder pain, [2] single-shoulder trauma and [3] > 6 months' non-operative treatment ***Exclusion criteria*** [1] Previous dislocation, [2] previous shoulder trauma, or [3] radiographic signs of degenerative shoulder lesions ***Duration of symptoms*** Mean 24 months (> 6 months) ***Previous treatments*** > 6 months' non-operative treatment, the nature of which was not specified ***Care setting*** Secondary or tertiary care
Participants	Denmark (duration 18 months) 31 (91% male); mean age 32 years (SD not reported, range 17-55)
Study design	Prospective, cross-sectional study with consecutive recruitment
Target condition and reference standard(s)	Labral tears, partial thickness rotator cuff tears (with differentiation between bursal side and joint side tears), full thickness rotator cuff tears. In each case these were, ultrasonography, MRI and arthroscopy: only the latter is considered here. ***Arthroscopy*** Referenced to Andrews et al (1984); Hurley and Anderson (1990); Adolfsson 1991) and described. *Procedure* The patient was in the lateral decubitus position with a traction device mounted on the arm placed at approximately 70° abduction. The standard approach using a 5 mm, 30° or 70° arthroscope was used, and bursography was performed in all patients. *Interpretation* 'Lesions of the glenoid labrum were described as either total detachment of the labrum from the glenoid rim and capsule or partial, with rupture of the labrum (flap tear). degenerative lesions of the labrum were also described. The lesions of the rotator cuff were described as either partial with a joint side or a bursal side tear or total, with complete rupture of the rotator cuff.' ***Interval between index and reference test*** Not reported ***Tester(s)*** 'Each examination was carried out by a different person and no information regarding clinical and diagnostic findings by the other methods was provided before each examination.'

Suder 1994 *(Continued)*

	Prevalence 19% (partial thickness, joint side rotator cuff tears), 9% (full thickness rotator cuff tears), 69% (labral tears)
Index and comparator tests	The ***impingement sign*** and ***impingement test***. Neither were referenced or described, but probably intended were Neer's sign and Neer's test. ***Tester(s)*** 'Each examination was carried out by a different person and no information regarding clinical and diagnostic findings by the other methods was provided before each examination.'
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes ***Were the reported summary statistics confirmed as accurate?*** No summary diagnostic test accuracy statistics were reported. For the impingement sign these were calculated from individual patient data as [1] sensitivity 100% and specificity 65.52 for complete rotator cuff tear, [2] sensitivity 66.67% and specificity 65.38% for partial rotator cuff tear, [3] sensitivity 77.78% and specificity 73.91% for any rotator cuff tear, [4] sensitivity 36.36% and specificity 50% for labral tear. For the impingement test these were calculated from individual patient data as [1] sensitivity 0.00% and specificity 96.55% for complete rotator cuff tear, [2] sensitivity 0.00% and specificity 96.15% for partial rotator cuff tear, [3] sensitivity 0.00% and specificity 95.65% for any rotator cuff tear, [4] sensitivity 0.00% and specificity 90.00% for labral tear.

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The information was very limited.
Acceptable reference standard? All tests	Yes	Arthroscopy. Special radiographic, MRI and ultrasonographic assessments were also conducted, but these were evaluated against the reference standard of arthroscopy.
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	No	No information

Suder 1994 *(Continued)*

Sufficient description of reference test? All tests	Yes	
Index test results blinded? All tests	Yes	
Reference standard results blinded? All tests	Yes	
Relevant clinical information? All tests	Unclear	Unclear
Uninterpretable results reported? All tests	Yes	
Withdrawals explained? All tests	Yes	There do not appear to have been any withdrawals.

Wolf 2001

Clinical features and settings	***Inclusion criteria*** Patients undergoing arthroscopy for diagnoses relating to shoulder pain and weakness e.g. impingement and rotator cuff tear. (It is unclear whether inclusion required both pain and weakness to be present in each case.) ***Exclusion criteria*** None reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Unclear: secondary or tertiary care
Participants	USA (August 1999 to September 2000) 119 (62% male); mean age 51 years (SD not reported, range 29-86)
Study design	Unclear whether prospective. Cross-sectional study in which induction was consecutive.
Target condition and reference standard(s)	Full thickness rotator cuff tears ***Arthroscopy*** No details given ***Interval between index and reference test*** Not reported ***Tester(s)*** One surgeon. No other information given. ***Prevalence*** 42% (full thickness rotator cuff tears)
Index and comparator tests	***Transdeltoid palpation (rent test)*** Referenced to a primary source (Codman 1934) and faithfully described ***Tester(s)*** One surgeon - the reportedly very experienced first author
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** No ***If applicable, could 2 x 2 tables be back-calculated?*** Yes

Wolf 2001 *(Continued)*

Were the reported summary statistics confirmed as accurate? Yes

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	It is unclear whether eligibility required shoulder pain and weakness in combination, or whether one or the other was acceptable
Acceptable reference standard? All tests	Yes	Arthroscopy: hence in general terms, yes, though no procedural or interpretative details were given
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	Probably. The test is described and original references given. The test appears very subjective, however.
Sufficient description of reference test? All tests	No	No information
Index test results blinded? All tests	No	No statement of blinding and unclear whether prospective
Reference standard results blinded? All tests	No	There is no clear statement of blinding
Relevant clinical information? All tests	Yes	Presumably
Uninterpretable results reported? All tests	Unclear	Unclear whether study was prospective
Withdrawals explained? All tests	No	Unclear whether study was prospective

Zaslav 2001

Clinical features and settings	***Inclusion criteria*** [1] Positive 'Neer overhead impingement sign' and by implication, [2] failure to respond to conservative treatment ***Exclusion criteria*** Primary diagnosis of avascular necrosis. No other exclusion criteria reported. ***Duration of symptoms*** Mean 11 months (range 2-48 months) ***Previous treatments*** Average 16 weeks of conservative treatment (range 2-25 months), comprising a single corticosteroid injection preceding a programme of physical therapy (progressive resistance exercises for the rotator cuff and scapular stabilizers plus, where appropriate, proprioceptive neuromuscular facilitation and sport specific exercises), with oral non-steroidal anti-inflammatory medication being used for pain relief. Two patients had previously had open rotator cuff repairs elsewhere. ***Care setting*** Secondary or tertiary care
Participants	USA (period not reported) 110 (59% male); mean age 44 years (SD not reported, range 17-76)
Study design	Prospective, cross-sectional study with consecutive recruitment.
Target condition and reference standard(s)	Impingement (subacromial and internal) ***Arthroscopy*** *Procedure* No details given *Interpretation* [1] Outlet impingement. Thickened and inflamed subacromial bursa, erosions on the coraco-acromial ligament and the undersurface of the acromion, bursal side partial or full thickness (*sic*) tearing of the rotator cuff; [2] Internal impingement. Anterior glenoid erosions, labral tears, middle glenohumeral ligament tears, undersurface rotator cuff partial tears, posterosuperior labral lesions, SLAP lesions. Also ***Examination under anaesthetic*** *Procedure* The (intra-operative) excursion of the humeral head on abduction/external rotation testing or inferior sulcus testing in the beach chair position was noted. *Interpretation* Classification was as grade I, II or III according to Warren's scale (Warren, 1983). Grade II or III subluxation in a shoulder with a pristine subacromial space was taken as positive for internal impingement. ***Interval between index and reference test*** Not reported ***Tester(s)*** One surgeon. No other information given. ***Prevalence*** 76% (subacromial impingement); 24% (internal impingement)
Index and comparator tests	A novel test, the ***internal rotation resistance strength test (IRRST)*** ***Tester(s)*** Either the single surgeon who performed the arthroscopies or an assistant. The surgeon may/may not have been the originator of the test: this was not specified.
Follow-up	***Adverse events*** None mentioned
Notes	***2 x 2 tables and summarystatistics*** ***Were 2 x 2 tables reported?*** For patients of all ages yes; for patients aged ≤ 50 years no. ***If applicable, could 2 x 2 tables be back-calculated?*** For patients of all ages N/A; for patients aged ≤ 50 years yes (approximately). ***Were the reported summary statistics confirmed as accurate?*** For patients of all ages there were borderline discrepancies between the reported sensitivity and PPV and those derived from the back-calculated 2 x 2 tables. For patients aged ≤ 50 years only accuracy was reported. This value was reproducible by two different 2 x 2 table configurations: TP 23 and TN 47 (sensitivity 95.83%, specificity 90.38%), or TP 22 and TN 48 (sensitivity 91.67%, specificity 92.31%).

Zaslav 2001 *(Continued)*

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	The setting was secondary or tertiary care and the reference standard (arthroscopy) was more than minimally invasive.
Selection criteria described? All tests	No	The information was very limited.
Acceptable reference standard? All tests	Yes	Arthroscopy
Acceptable delay between tests? All tests	Unclear	Insufficient information
Partial verification avoided? All tests	Yes	
Differential verification avoided? All tests	Yes	
Incorporation avoided? All tests	Yes	
Sufficient description of index tests? All tests	Yes	
Sufficient description of reference test? All tests	No	Interpretative information is given, but no detail on procedure of the test
Index test results blinded? All tests	Unclear	Unclear, but the study was prospective
Reference standard results blinded? All tests	No	No statement of blinding
Relevant clinical information? All tests	Yes	
Uninterpretable results reported? All tests	Yes	Probably, though not explicitly stated
Withdrawals explained? All tests	Yes	n = 5 due to [1] negative impingement test (n = 4); [2] avascular necrosis (n = 1).

LHB = long head of biceps

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Adolfsson 1991	It is unclear which tendinitis tests were used.
Altchek 1990	Not a diagnostic test accuracy study.
Ansara 2006	Abstract only. No full-text report available.
Ardic 2006	No information is provided on how the Neer's and Hawkins-Kennedy tests were concertedly interpreted. In relation to Speed's test for any LHB pathology, the test is insufficiently defined to be informative, and 20% of the sample (56% of the disease negative portion) is unreported without explanation.
Beaudreuil 2009	Not a diagnostic test accuracy study. A systematic review.
Bedi 2008	Not a diagnostic test accuracy study.
Berbig 1999	Relates to rotator cuffs following traumatic dislocation: a highly selected population, and not impingement.
Berg 1997	Not a diagnostic test accuracy study.
Berg 1998	A highly selected population: 100% prevalence of SLAP lesions.
Bigliani 1997	Not a diagnostic test accuracy study.
Birtane 2001	Not a diagnostic test accuracy study of physical tests.
Blair 1996	Not a diagnostic test accuracy study.
Blevins 1996a	Not a diagnostic test accuracy study.
Blevins 1996b	Not a diagnostic test accuracy study.
Brasseur 2004	A highly selected population: veteran tennis players.
Bron 2007	Not a diagnostic test accuracy study.
Brownlow 2009	Not a diagnostic test accuracy study.
Brox 2003	Not a diagnostic test accuracy study.
Bryant 2002	Not a diagnostic test accuracy study.
Buchberger 1999	Not a diagnostic test accuracy study.
Burbank 2007	Not a diagnostic test accuracy study and not related to impingement.
Burkhart 2000	A highly selected population: 100% prevalence of SLAP lesions.
Burkhart 2002	A highly selected population: 100% prevalence of rotator cuff tears.

Study	Reason for exclusion
Burkhart 2003	Not a diagnostic test accuracy study.
Burkhart 2008	Not a diagnostic test accuracy study.
Callanan 2001	Not a diagnostic test accuracy study.
Calvert 2009	Not a diagnostic test accuracy study. A systematic review.
Chronopoulos 2004	Not a diagnostic tests accuracy study for impingement.
Cools 2008	Not a diagnostic test accuracy study.
Corso 1995	Not a diagnostic test accuracy study.
Crenshaw 1966	Not a diagnostic test accuracy study.
Cullen 2007	Not a diagnostic test accuracy study of physical tests.
D'Alessandro 2000	Not a diagnostic test accuracy study.
Dessaur 2008	Not a diagnostic test accuracy study. A systematic review.
Deutsch 1997	Not a diagnostic test accuracy study.
Diehr 2006	Not a diagnostic test accuracy study.
Dinnes 2003a	Not a diagnostic test accuracy study.
Ebinger 2008	For one of three tests evaluated there were substantial discrepancies in the back-calculated 2 x 2 table (exceeding 10% in one or more cells), and this was not attributable to unit-of-analysis issues. For all three tests there were incorrectly reported summary statistics.
El Dalati 2005	Not a diagnostic test accuracy study of physical tests.
Fodor 2009	For all of the tests evaluated there were substantial discrepancies in the back-calculated 2 x 2 tables (exceeding 10% in one or more cells), and this was not attributable to unit-of-analysis issues. For all tests there were incorrectly reported summary statistics.
Fukuda 1996	A highly selected population: 100% prevalence of partial thickness rotator cuff tears.
Gartsman 2000	Not a diagnostic test accuracy study.
Gerber 1985	Not a diagnostic test accuracy study.
Gerber 1996	Not a diagnostic test accuracy study.
Gerber 2000	Not a diagnostic test accuracy study.
Gerber 1991	Unsatisfactory reference test.
Giombini 1997	Not a diagnostic test accuracy study.
Gschwend 1988	No disease-negative group: no specificity data.
Hagemann 2004	Not a diagnostic test accuracy study.

Study	Reason for exclusion
Halbrecht 1999	Not a diagnostic test accuracy study.
Hammer 2003	Not a diagnostic test accuracy study.
Hamner 2000	A highly selected population: overhead throwing athletes.
Handelberg 1998	Not a diagnostic test accuracy study.
Hawkins 1980	Not a diagnostic test accuracy study.
Hayes 2003	Not a diagnostic test accuracy study.
Hegedus 2008	Not a diagnostic test accuracy study. A systematic review.
Heyworth 2009	Not a diagnostic test accuracy study.
Hughes 2008	Not a diagnostic test accuracy study. A systematic review.
Hurschler 2004	Not a diagnostic test accuracy study.
Jee 2001	Not a diagnostic test accuracy study of physical tests.
Jobe 1997	Not a diagnostic test accuracy study.
Jobe 2000	Not a diagnostic test accuracy study.
Johansson 2009	Not a diagnostic test accuracy study.
Jones 2007	Not a diagnostic test accuracy study. A systematic review.
Keener 2009	Not a diagnostic test accuracy study.
Kibler 1995	A highly selected population: all athletes.
Kibler 2006a	Not a diagnostic test accuracy study.
Kibler 2006b	Not a diagnostic test accuracy study.
Kim 1999	A highly selected population: all following anterior dislocation.
Kim 2004b	Not impingement.
Kim 2005	Concerns postero-inferior labral lesions: not impingement.
Kim 2007a	Concerns patients with rheumatoid disease.
Kim 2009	Not a diagnostic test accuracy study.
Kirkley 2002	Not a diagnostic test accuracy study.
Koester 2005	Not a diagnostic test accuracy study.
Lafosse 2007	The focus of this study was LHB instability but also LHB tendon lesions. However, since all patients were undergoing arthroscopic cuff repairs, the diagnostic performance results depend on fore-knowledge of a cuff tear. This militates against their applicability to primary care.

Study	Reason for exclusion
Le Huec 1996	Not a diagnostic test accuracy study.
Leroux 1995	All patients had surgery for outlet impingement, so only sensitivity data are calculable for identification of impingement. The study could have been included for localisation data but for irreconcilable and very large discrepancies in the statistical analyses.
Lewis 2005	Not a diagnostic test accuracy study.
Lewis 2007	Not a diagnostic test accuracy study for impingement.
Lewis 2009	Not a diagnostic test accuracy study.
Litaker 2000	There were substantial arithmetical discrepancies in back-calculated 2 x 2 tables (discrepancies exceeding 10% in any cell). There were also incorrectly reported summary statistics, with back-calculated T+ values very discrepant from those reported (impingement sign = 426 versus 310 reported; empty can test = 244 versus 226; weakness of external rotation = 291 versus 284; painful arc = 426 versus 299). This did not appear attributable to unit of analysis error.
Liu 1996a	A highly selected population: rotator cuff tendinitis and tears were excluded from this retrospective study, removing the main cause of diagnostic uncertainty.
Lo 2004	Unsatisfactory reference test.
Luime 2004	Not a diagnostic test accuracy study. A systematic review.
Lyons 1992	Determination of tear size in a population with known rotator cuff tears.
Lyons 2005	Not a diagnostic test accuracy study.
Maffet 1995	Not a diagnostic test accuracy study.
Malhi 2005	A (retrospective) diagnostic test accuracy study, but the nature of testing is insufficiently defined to be informative.
Maman 2009	Not a diagnostic test accuracy study.
Matava 2005	Not a diagnostic test accuracy study.
McCabe 2005	Special equipment (a hand held dynamometer) was used.
McFarland 1999	Not a diagnostic test accuracy study.
McFarland 2001	Not a diagnostic test accuracy study.
McFarland 2008	Not a diagnostic test accuracy study.
McFarland 2009	Not a diagnostic test accuracy study.
Meister 2004	A highly selected population: all overhead athletes.
Meserve 2009	Not a diagnostic test accuracy study. A systematic review.
Miller 2008a	The nature of testing is insufficiently defined to be informative.
Mimori 1999	A highly selected population: all throwing injuries.

Study	Reason for exclusion
Mirkovic 2005	Not a diagnostic test accuracy study.
Mitchell 2005	Not a diagnostic test accuracy study.
Moosikasuwan 2005	Not a diagnostic test accuracy study.
Morgan 1998	A highly selected population: 100% prevalence of SLAP lesions.
Morrison 1997	Not a diagnostic test accuracy study.
Morrissey 2005	No published report was available as of 3 May 2012 and the contact person was no longer at the email address given.
Munro 2009	Not a diagnostic test accuracy study. A systematic review.
Murrell 2001	Information on how to perform the tests is no longer available on the journal's web site. No reply was received to a request to authors for further information.
Myers 2005	A highly selected population: all athletes.
Neer 1977	Not a diagnostic test accuracy study.
Neer 1972	Not a diagnostic test accuracy study.
Neer 1983a	Not a diagnostic test accuracy study.
Neri 2009	Not a diagnostic test accuracy study.
Nomden 2009	Not a diagnostic test accuracy study.
Norregaard 2002	Inadequately defined index text.
O'Connor 2005	Not a diagnostic test accuracy study of physical tests.
Odom 2001	Not a diagnostic test accuracy study for impingement.
Oh 2007	Not a diagnostic test accuracy study. A systematic review on indications for surgery.
Olmsted 2003	Not a diagnostic test accuracy study.
Osbahr 2006	Special equipment (a hand-held dynamometer) was used.
Ostor 2005	Not a diagnostic test accuracy study.
Pandya 2008	All patients had SLAP lesions: hence no specificity data.
Perez-Palomares 2009	Not a diagnostic test accuracy study.
Piasecki 2008	Not a diagnostic test accuracy study.
Pisan 2000	Not a diagnostic test accuracy study.
Polimeni 2003	No summary measures of test performance were given, and 2 x 2 tables were neither reported nor calculable.

Study	Reason for exclusion
Polsky 2006	Not a diagnostic test accuracy study.
Pugh 2009	Not a diagnostic test accuracy study. A "short-cut review".
Rao 2003	Not a diagnostic test accuracy study.
Razmjou 2006	Not a diagnostic test accuracy study.
Read 1998	Excluded patients if physical tests negative, hence no specificity data.
Rhee 2005a	All patients were disease-positive (unstable SLAP lesion), hence no specificity data.
Rowan 2007	No test-specific data.
Ryu 2002	Not a diagnostic test accuracy study.
Sandenbergh 2006	2 x 2 tables are neither presented not calculable. There are key inconsistencies in the description of the sample.
Savoie 2001	Not a diagnostic test accuracy study.
Scheibel 2005	Inadequate 'control': contralateral shoulder.
Schellingerhout 2008	Not a diagnostic test accuracy study. A systematic review concerning the diagnostic labelling of shoulder pain.
Sileo 2006	Abstract only. No full-text report available.
Silva 2008	Unacceptability of MRI as reference standard for impingement.
Skedros 2007	Not a diagnostic test accuracy study.
Smith 2000	Not a diagnostic test accuracy study.
Snyder 1990	Not a diagnostic test accuracy study.
Soncini 2000	Not a diagnostic test accuracy study.
Sorensen 2007	Presentation of data does not allow calculation of sensitivity and specificity.
Sorohan 2009	Not a diagnostic test accuracy study.
Tate 2008	Not a diagnostic test accuracy study.
Tennent 2003a	Not a diagnostic test accuracy study.
Tennent 2003b	Not a diagnostic test accuracy study.
Trantalis 2008	Not a diagnostic test accuracy study.
Tyler 2000	Not a diagnostic test accuracy study.
Vanderbeck 2007	Not a diagnostic test accuracy study.
Walch 1998	Not a diagnostic test accuracy study.

Study	Reason for exclusion
Walsworth 2008	Highly selected population: 73% on active military duty.
Walton 2004	Not a diagnostic test accuracy study for impingement.
Walton 2008	Not a diagnostic test accuracy study. A systematic review.
Wang 2000	Not a diagnostic test accuracy study. An evaluation of acromial type as a predictor of the outcome of conservative treatment for impingement.
Watson 1989	All patients had subacromial impingement: no disease-negative group, hence no specificity data.
Wilk 2005	Not a diagnostic test accuracy study.
Wnorowski 1997	No description of tests used for clinical diagnosis.
Wolff 2006	Not a diagnostic test accuracy study.
Yang 2006	Not a diagnostic test accuracy study. A reliability study which does not concern tests of impingement or impingement-related conditions.

LHB = long head of biceps

Characteristics of studies awaiting classification *[ordered by study ID]*

Gill 2007

Clinical features and settings	***Inclusion criteria*** Arthroscopy ***Exclusion criteria*** [1] Fracture, [2] previous surgery on the same shoulder, [3] isolated open procedures for any reason, [4] SLAP lesion type II, III, or IV, [5] subluxation, dislocation or complete rupture of the LHB tendon. ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Probably secondary or tertiary
Participants	Probably USA (1994-2004) 850 shoulders (?) in 847 patients, approximately 54% male; mean age 44 years (SD 17.2) in no tear group, 59 years (SD 11.8) in partial tear group
Study design	Probably retrospective, cross sectional study with reportedly consecutive inclusion
Target condition and reference standard(s)	Partial tears of the LHB tendon ***Arthroscopy*** ***Prevalence*** 5% (partial tears of the LHB tendon)
Index and comparator tests	***Palpation for bicipital groove tenderness, Speed's test, lift-off test, belly-press test, Neer's sign, Hawkins' test, crank test, active compression test, anterior slide test***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] For two of the tests evaluated (active compression test, belly press test), the primary references in the literature postdate the start of the study. [2] Substantial discrepan-

Gill 2007 *(Continued)*

	cy between the description of the anterior slide test in the report and that in the cited source (compromised internal validity). [3] Very long (10-year), probably retrospective data collection period presents a substantial threat to the consistency of test application/interpretation. [4] Evidence of inconsistency in test application/interpretation across reports, where the reports describe the same group of patients (cf Hawkins' test v. Park 2005, and compression-rotation test v. McFarland 2002). [5] Description of patient database inconsistent with other reports.

Jia 2008

Clinical features and settings	***Inclusion criteria*** Shoulder surgery and the presence of shrug test data ***Exclusion criteria*** Not reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Probably secondary or tertiary
Participants	Probably USA (1994-2006) 982 shoulders in 982 patients, 55.2% male; mean age 57 years (SD 14.6) in shrug pos. group, 40 years (SD 17.2) in shrug neg. group
Study design	Retrospective, cross sectional study with reportedly consecutive inclusion
Target condition and reference standard(s)	Rotator cuff disease ***Surgery*** ***Prevalence*** 8% (tendinosis), 9% (partial rotator cuff tear), 27% (full thickness rotator cuff tear), 5% (massive rotator cuff tear), 3% (SLAP lesions)
Index and comparator tests	***Shrug sign***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Very long (12-year), retrospective data collection period presents a substantial threat to the consistency of test application/interpretation. [2] Description of patient database inconsistent with other reports.

Jia 2009

Clinical features and settings	***Inclusion criteria*** Shoulder surgery ***Exclusion criteria*** Not reported ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Probably secondary or tertiary
Participants	Probably USA (1995-2008) 1913 patients (denominators reportedly vary by test, but are not specified). Gender and age distribution not reported.

Jia 2009 *(Continued)*

Study design	Retrospective, cross sectional study. Unclear whether inclusion consecutive.
Target condition and reference standard(s)	Rotator cuff disease, SLAP lesions and LHB tendon tears ***Surgery*** (arthroscopic or open) ***Prevalence*** Insufficient data
Index and comparator tests	***Drop arm sign, shoulder shrug sign, Neer's sign, Hawkins' test, Speed's test, active compression test, anterior slide test, lift-off test, painful arc sign, Whipple test, external rotation lag sign***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Very long (13-year), retrospective data collection period presents a substantial threat to the consistency of test application/interpretation. [2] Overlap with McFarland 2002, Kim 2003a, Gill 2007 and Jia 2008, and possible multiple counting. [3] Denominators unspecified for index test subgroups. [4] Description of patient database inconsistent with other reports.

Kelly 2010

Clinical features and settings	***Inclusion criteria*** [1] Referred for diagnostic ultrasound for shoulder pain, [2] aged 20-70 years, [3] able to follow instructions ***Exclusion criteria*** [1] Trauma to shoulder, [2] neurological pain or weakness originating from the cervical spine, [3] inflammatory joint disease ***Duration of symptoms*** > 4 months ***Previous treatments*** Not reported ***Care setting*** Probably secondary or tertiary
Participants	UK (June-September 2006) 34 shoulders in 34 patients, 58.82% male; median age 57 years (interquartile range 44-63)
Study design	Prospective cross sectional study with consecutive inclusion.
Target condition and reference standard(s)	Subacromial impingement syndrome: full thickness rotator cuff tear, partial thickness rotator cuff tear and subacromial bursitis. ***Ultrasonography*** ***Prevalence*** 32% (full thickness rotator cuff tear), 21% (partial thickness rotator cuff tear), 35% (subacromial bursitis only)
Index and comparator tests	***Neer's sign, Hawkins' test, painful arc test, empty can test, full can test, resisted abduction, resisted external rotation***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** Insufficient data. Return to authors for 2 x 2 tables.

Kim 2003a

Clinical features and settings	***Inclusion criteria*** Various diagnoses warranting arthroscopy

Kim 2003a *(Continued)*

	Exclusion criteria Revision arthroscopy ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	USA (1992-2000). 555 shoulders in 544 patients (57% male); mean age 44 (12-86).
Study design	Probably retrospective, cross-sectional study with reportedly consecutive inclusion
Target condition and reference standard(s)	SLAP lesions Types I-IV (Snyder classification). Primary diagnoses were based on a combination of ***history , physical examination*** and ***arthroscopy*** . Physical examination included the index tests and evaluations of passive ranges of motion. ***Prevalence*** 28% (Types I-IV SLAP lesions), 7-8% (Types II-IV SLAP lesions)
Index and comparator tests	***Active compression test, anterior slide test,compression-rotation test,Hawkins' test,Neer's 'sign' or 'test'*** (the latter terms were used interchangeably, and in this context probably both referred to Neer's sign), ***relocation test, apprehension test*** , ***painful arc test, Speed's test***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Large, unexplained losses in patient flow. [2] For one of the tests evaluated (active compression test) the primary reference in the literature postdates the start of the study. [3] Overlap with McFarland 2002 and Jia 2009 and possible multiple counting.

Kim 2003b

Clinical features and settings	***Inclusion criteria*** No tear or a partial width tear of subscapularis identified at arthroscopy ***Exclusion criteria*** [1] Non-arthroscopic procedures. [2] Revision procedures. [3] Complete rupture of subscapularis identified at surgery. ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	USA (1994-2000) 314 shoulders in 314 patients (56% male); mean age 44 (13-82)
Study design	Probably retrospective, effectively cross-sectional study, with reportedly consecutive inclusion.
Target condition and reference standard(s)	Partial width tears of subscapularis, whether partial or full thickness ***Arthroscopy*** ***Prevalence*** 19% (partial width tears of subscapularis, whether partial or full thickness)
Index and comparator tests	***Apprehension test,drop arm test,Hawkins' test, lift-off test, Neer's sign, painful arc test, Jobe's relocation test, Speed's test***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Large, unexplained losses in patient flow. [2] Description of patient database inconsistent with other reports

Kim 2004a

Clinical features and settings	***Inclusion criteria*** Diagnostic shoulder arthroscopy ***Exclusion criteria*** [1] Previous surgery of shoulder. [2] Arthroscopic observation of internal impingement in flexion unreliable due to severe synovial hypertrophy or labral fraying. ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	Country unclear: Korea or USA (1995-2000) 376 shoulders in 376 patients (58% male); mean age 42 years, SD 17 (13-86)
Study design	Probably retrospective, cross-sectional study. Unclear whether consecutive.
Target condition and reference standard(s)	Internal impingement in flexion ***Arthroscopy*** ***Prevalence*** 74% (internal impingement in flexion)
Index and comparator tests	***Active compression***,***anterior slide***, ***apprehension*** *(probably for pain rather than apprehension),* ***compression-rotation***,***drop arm test, Hawkins' test, relocation test***,***Neer's sign***, ***painful arc*** *test,* ***Speed's test***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Large, unexplained losses in patient flow. [2] For one of the tests evaluated (active compression test) the primary reference in the literature postdates the start of the study. [3] Description of database inconsistent with other reports.

McFarland 2002

Clinical features and settings	***Inclusion criteria*** Shoulder arthroscopy ***Exclusion criteria*** [1] No arthroscopy for clinical reasons (fracture, patient undergoing arthroplasty or isolated open surgical procedures). [2] Previous surgery on the same shoulder. ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Tertiary care
Participants	USA (1994-2000) 426 shoulders in 426 patients (59% male). Age distribution not reported.
Study design	Retrospective, reportedly consecutive, cross-sectional study.
Target condition and reference standard(s)	Types II-IV SLAP lesions ***Arthroscopy*** ***Prevalence*** 10% (types II-IV SLAP lesions)
Index and comparator tests	***Active compression test, anterior slide test, compression-rotation test***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** [1] Large, unexplained losses in patient flow. [2] For one of the tests evaluated (active compression test) the primary reference in the literature postdates the start of the study. [3] Substantial discrepancy between the description of the anterior slide and active compression tests and that in the cited sources (compromised internal validity). [4] Evidence of inconsistency in test application/interpretation across reports, where the reports described the same group of patients (cf description of compression-rotation test in Gill 2007). [5] Overlap with Kim 2003a and Jia 2009, and possible multiple counting. [6] Description of database inconsistent with other reports.

Nanda 2008

Clinical features and settings	***Inclusion criteria*** [1] Patients with weakness and pain with overhead activities, [2] pain localised to the anterolateral aspects of the shoulder with radiation towards the lateral aspect of the upper arm ***Exclusion criteria*** [1] Gross restriction of external rotation, [2] symptomatic instability, [3] acute traumatic conditions, [4] associated neck or elbow disorders ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	UK (period not reported, but duration approximately 8 months) 50 shoulders in 50 patients (gender and age distribution not reported)
Study design	Prospective, consecutive, cross-sectional study.
Target condition and reference standard(s)	Impingement syndrome, rotator cuff tears, LHB lesions ***Arthroscopy*** ***Prevalence*** 90% (impingement), 36% (full thickness tears of supraspinatus), 24% (partial thickness tears of supraspinatus), 40% (attrition of LHB), 18% (tear in infraspinatus), 6% (tear in subscapularis)
Index and comparator tests	***Painful arc test, Hawkins' test, Neer's sign, Neer's test, drop arm test, resisted abduction, resisted external rotation from neutral, empty can test, Gerber's test with force, Yergason's test, Speed's test.***
Follow-up	***Adverse events*** None mentioned
Notes	***Issues and concerns*** Insufficient data. Return to authors for 2 x 2 tables.

Park 2005

Clinical features and settings	***Inclusion criteria*** Shoulder arthroscopy ***Exclusion criteria*** Previous shoulder surgery, impingement complicated by ACJ arthritis, incomplete physical examination due to limited motion or extreme pain, impingement complicated by SLAP lesions, impingement complicated by instability ***Duration of symptoms*** Not reported ***Previous treatments*** Not reported ***Care setting*** Secondary or tertiary care
Participants	USA (August 1992-June 2003) 552 patients (gender and age distribution not reported)
Study design	Retrospective, cross-sectional study. Unclear whether consecutive.
Target condition and reference standard(s)	Subacromial impingement syndrome: bursitis alone, partial thickness rotator cuff tears, full thickness rotator cuff tears ***Arthroscopy*** ***Prevalence*** 20% (bursitis alone), 20% (partial thickness rotator cuff tear), 60% (full thickness rotator cuff tear)
Index and comparator tests	***Neer's sign, Hawkins' test, painful arc test, empty can test, Speed's test, scarf test, drop arm test, resisted external rotation from neutral, external rotation lag sign***
Follow-up	***Adverse events*** None mentioned

Park 2005 *(Continued)*

Notes	***Issues and concerns*** [1] Very long (11-year), retrospective data collection period presents a substantial threat to the consistency of test application/interpretation. [2] Evidence of inconsistency in test application/interpretation across reports (where the reports describe the same group of patients). [3] Description of patient database inconsistent with other reports.

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Target condition: SIS. Index test: combination of ALL 7 tests +ve (*see* table 7).	1	125
2 Target condition: SIS. Index test: combination of Hawkins' test AND Neer's sign (modified procedure) +ve.	1	85
3 Target condition: SIS. Index test: combination of Hawkins' test OR Neer's sign (modified procedure) +ve.	1	85
4 Target condition: SIS. Index test: drop arm test (modified interpretation).	1	125
5 Target condition: SIS. Index test: Gum-turn test (novel)	1	120
6 Target condition: SIS. Index test:. Hawkins' test (standard).	2	210
7 Target condition: SIS. Index test: Neer's sign (standard).	1	125
8 Target condition: SIS. Index test: Neer's sign (modified procedure).	1	85
9 Target condition: SIS. Index test: painful arc test (standard).	1	125
10 Target condition: SIS. Index test: passive horizontal adduction (modified interpretation).	1	125
11 Target condition: SIS. Index test: Speed's test (modified interpretation).	1	125
12 Target condition: SIS. Index test: Yergason's test (modified interpretation).	1	125
13 Target condition: SIS (SA-SD bursitis). Index test: combination of Hawkins' test, Neer's sign, 'Yocum's (impingement) test' (overall criterion for +ve result not given).	1	31
14 Target condition: SIS *versus* internal impingement, differentiation. Index test: internal rotation resistance strength test (novel).	1	110
15 Target condition: rotator cuff, any disease of. Index test: relocation test for pain (Jobe 1989: standard).	1	100
16 Target condition: rotator cuff, any disease of. Index test: relocation test for pain (Jobe 1989: modified procedure).	1	100

Test	No. of studies	No. of participants
17 Target condition: rotator cuff, FTT or PTT of. Index test: combination of Hawkins' test (modified interpretation) OR Neer's sign (modified procedure, modified interpretation) +ve.	1	85
18 Target condition: rotator cuff, FTT or PTT of. Index test: combination of Hawkins' test (modified interpretation) AND Neer's sign (modified procedure,modified interpretation) +ve.	1	85
19 Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain ± weakness (modified interpretation).	1	200
20 Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain OR weakness (ONE ONLY) (modified interpretation).	1	200
21 Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain AND weakness (BOTH) (modified interpretation).	1	200
22 Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for weakness ± pain (modified interpretation).	1	200
23 Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain ± weakness (modified interpretation).	1	202
24 Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain OR weakness (ONE ONLY) (modified interpretation).	1	200
25 Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain AND weakness (BOTH) (modified interpretation).	1	200
26 Target condition: rotator cuff, FTT or PTT of. Index test: full can test for weakness ± pain (standard).	1	200
27 Target condition: rotator cuff, FTT or PTT of. Index test: Hawkins' test (modified interpretation).	1	85
28 Target condition: rotator cuff, FTT or PTT of. Index test: 'Impingement sign' (no reference or details given).	1	32
29 Target condition: rotator cuff, FTT or PTT of. Index test: 'Impingement test' (no reference or details given).	1	32
30 Target condition: rotator cuff, FTT or PTT of. Index test: Neer's sign (modified procedure, modified interpretation).	1	85
31 Target condition: rotator cuff, PTT or tendinitis of. Index test: empty can test for pain WITHOUT weakness (modified interpretation).	1	50
32 Target condition: rotator cuff, FTT of. Index test: empty can test for pain ± weakness (modified interpretation).	1	200
33 Target condition: rotator cuff, FTT of. Index test: empty can test for pain OR weakness (ONE ONLY) (modified interpretation).	1	200
34 Target condition: rotator cuff, FTT of. Index test: empty can test for pain AND weakness (BOTH) (modified interpretation).	1	200

Test	No. of studies	No. of participants
35 Target condition: rotator cuff, FTT of. Index test: empty can test for weakness ± pain (modified interpretation).	2	250
36 Target condition: rotator cuff, FTT of. Index test: full can test for pain ± weakness (modified interpretation).	1	200
37 Target condition: rotator cuff, FTT of. Index test: full can test for pain OR weakness (ONE ONLY) (modified interpretation).	1	200
38 Target condition: rotator cuff, FTT of. Index test: full can test for pain AND weakness (BOTH) (modified interpretation).	1	200
39 Target condition: rotator cuff, FTT of. Index test: full can test for weakness ± pain (modified interpretation).	1	200
40 Target condition: rotator cuff, FTT of. Index test: 'impingement sign' (no reference or details given).	1	32
41 Target condition: rotator cuff, FTT of. Index test: 'impingement test' (no reference or details given).	1	32
42 Target condition: rotator cuff, FTT of. Index test: rent test (standard).	1	109
43 Target condition: rotator cuff, FTT of, massive or large. Index test: empty can test for weakness ± pain (modified interpretation).	1	50
44 Target condition: rotator cuff, PTT of. Index test: 'Impingement sign' (no reference or details given).	1	32
45 Target condition: rotator cuff, PTT of. Index test: 'Impingement test' (no reference or details given).	1	32
46 Target condition: rotator cuff, postero-superior (supraspinatus AND infraspinatus), FTT of. Index test: Gum-turn test (novel).	1	120
47 Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: drop sign (novel).	1	87
48 Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: empty can test for weakness ± pain (modified interpretation).	1	87
49 Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: external rotation lag sign (novel).	1	87
50 Target condition: rotator cuff, postero-superior, FTT of. Index test: drop sign (modified interpretation).	1	46
51 Target condition: rotator cuff, postero-superior, FTT of. Index test: external rotation lag sign (modified interpretation).	1	46
52 Target condition: rotator cuff, postero-superior, FTT of. Index test: Gum-turn test (novel).	1	120
53 Target condition: rotator cuff, FTT, multiple- *versus* single-tendon. Index test: active abduction range (novel).	1	96

Test	No. of studies	No. of participants
54 Target condition: supraspinatus, any disease of, including calcification. Index test: empty can test (no reference or details given).	1	528
55 Target condition: supraspinatus, FTT, degeneration or tendinitis,of. Index test: Hawkins' test (modified procedure, modified interpretation).	1	73
56 Target condition: supraspinatus, FTT, PTT or tendinitis,of. Index test: empty can test for pain AND/OR weakness (standard).	1	26
57 Target condition: supraspinatus, FTT or degeneration of. Index test: Hawkins' test (modified procedure, modified interpretation).	1	73
58 Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for pain ± weakness (modified interpretation).	1	160
59 Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for weakness ± pain (standard).	2	191
60 Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for weakness (< grade 3) ± pain.(modified interpretation)	1	160
61 Target condition: supraspinatus, FTT or PTT of. Index test: full can test for pain ± weakness (modified interpretation).	1	160
62 Target condition: supraspinatus, FTT or PTT of. Index test: full can test for weakness (< grade 3) ± pain (modified interpretation).	1	160
63 Target condition: supraspinatus, FTT or PTT of. Index test: full can test for weakness ± pain (standard).	1	160
64 Target condition: supraspinatus, FTT of. Index test: drop arm test (standard).	1	125
65 Target condition: supraspinatus, FTT of. Index test: empty can test for pain ± weakness (modified interpretation).	1	143
66 Target condition: supraspinatus, FTT of. Index test: empty can test for pain AND/OR weakness (modified interpretation).	1	143
67 Target condition: supraspinatus, FTT of. Index test: empty can test for weakness ± pain (standard).	1	143
68 Target condition: supraspinatus, FTT of. Index test: full can test for pain ± weakness (modified interpretation).	1	143
69 Target condition: supraspinatus, FTT of. Index test: full can test for pain AND/OR weakness (modified interpretation).	1	143
70 Target condition: supraspinatus, FTT of. Index test: full can test for weakness ± pain (standard).	1	143
71 Target condition: supraspinatus, FTT of. Index test: Gum-turn test (novel).	1	120
72 Target condition: supraspinatus, FTT of. Index test: Hawkins' test (modified interpretation).	1	125

Test	No. of studies	No. of participants
73 Target condition: supraspinatus, FTT of. Index test: Hawkins' test (modified procedure, modified interpretation).	1	73
74 Target condition: supraspinatus, FTT of. Index test: Neer's sign (modified interpretation).	1	125
75 Target condition: supraspinatus, FTT of. Index test: painful arc test (modified interpretation).	1	125
76 Target condition: supraspinatus, FTT of. Index test: passive horizontal adduction (standard).	1	125
77 Target condition: supraspinatus, FTT of. Index test: Speed's test (modified interpretation).	1	125
78 Target condition: supraspinatus, FTT of. Index test: Yergason's test (modified interpretation).	1	125
79 Target condition: supraspinatus, FTT of, full-width. Index test: external rotation lag sign (standard).	1	189
80 Target condition: supraspinatus, isolated PTT of. Index test: external rotation lag sign (standard).	1	222
81 Target condition: supraspinatus, tendinitis of. Index test: empty can test for pain WITHOUT weakness (standard).	1	31
82 Target condition: infraspinatus, any disease of, including calcification. Index test: resisted lateral rotation from neutral rotation (no reference or details given).	1	528
83 Target condition: infraspinatus, FTT, PPT or tendinitis,of. Index test: Patte's test for pain AND/OR weakness (standard).	1	31
84 Target condition: infraspinatus, FTT or PTT of. Index test: Patte's test for weakness ± pain (standard).	1	31
85 Target condition: infraspinatus, FTT or PTT of. Index test: resisted lateral rotation from neutral rotation for weakness < grade 3 ± pain (modified interpretation).	1	160
86 Target condition: infraspinatus, tendinitis of. Index test: Patte's test for pain WITHOUT weakness (standard).	1	31
87 Target condition: subscapularis, any disease of, including calcification. Index test: resisted medial rotation from neutral rotation (no reference or details given).	1	528
88 Target condition: subscapularis, any tear or tendinitis of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given).	1	31
89 Target condition: subscapularis, any tear of. Index test: bear-hug test (novel)	1	68

Test	No. of studies	No. of participants
90 Target condition: subscapularis, any tear of. Index test: belly-press test (standard)	1	68
91 Target condition: subscapularis, any tear of. Index test: internal rotation lag sign (novel).	1	53
92 Target condition: subscapularis, any tear of. Index test: lift-off test (Gerber 1991: modified interpretation).	1	63
93 Target condition: subscapularis, any tear of. Index test: lift-off test (Gerber 1991: probably standard)	1	53
94 Target condition: subscapularis, any tear of. Index test: Napoleon test (Burkhart 2002: standard).	1	68
95 Target condition: subscapularis, any tear of. Index test:: lift-off test with force for weakness < grade 2 ± pain (modified procedure, modified interpretation).	1	159
96 Target condition: subscapularis, any tear of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given).	1	31
97 Target condition: subscapularis, complete tear of. Index test: bear-hug test (novel).	1	68
98 Target condition: subscapularis, complete tear of. Index test: belly-press test (modified procedure).	1	68
99 Target condition: subscapularis, complete tear of. Index test: lift-off test (Gerber 1991: modified interpretation).	1	63
100 Target condition: subscapularis, complete tear of. Index test: Napoleon test (Burkhart 2002: standard).	1	68
101 Target condition: subscapularis, FTT of. Index test: internal rotation lag sign (modified interpretation).	1	46
102 Target condition: subscapularis, partial tear of. Index test: bear-hug test (novel).	1	65
103 Target condition: subscapularis, partial tear of. Index test: belly-press test (modified procedure).	1	65
104 Target condition: subscapularis, partial tear of. Index test: lift-off test (Gerber 1991: modified interpretation).	1	60
105 Target condition: subscapularis, partial tear of. Index test: Napoleon test (Burkhart 2002: standard).	1	65
106 Target condition: subscapularis, tendinitis of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given).	1	31
107 Target condition: LHB, tear or tendinitis of. Index test: Speed's test (standard).	1	528

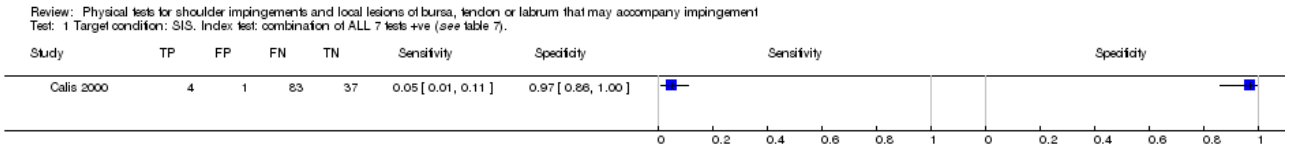
Test	No. of studies	No. of participants
108 Target condition: LHB, tear or tendinitis of. Index test: active compression test (standard)	1	101
109 Target condition: LHB, tear or tendinitis of. Index test: anterior slide test (modified procedure, modified interpretation).	1	101
110 Target condition: LHB, tear or tendinitis of. Index test: bear-hug test (modified interpretation)	1	101
111 Target condition: LHB, tear or tendinitis of. Index test: belly-press test (standard)	1	101
112 Target condition: LHB, tear or tendinitis of. Index test: modified dynamic labral shear (novel)	1	101
113 Target condition: LHB, tear or tendinitis of. Index test: Speed's test (modified procedure)	1	101
114 Target condition: LHB, tear or tendinitis of. Index test: upper-cut test (novel)	1	101
115 Target condition: LHB, tear or tendinitis of. Index test: Yergason's test (modified procedure)	1	101
116 Target condition: LHB, tear or tendinitis of. Index test: combination of Yergason's test and Gilcreest's test (modified procedure, modified interpretation).(Overall criterion for +ve result not given.)	1	31
117 Target condition: labrum, any tear of. Index test: active compression test (novel).	1	206
118 Target condition: labrum, any tear of. Index test: active compression test (modified interpretation).	1	65
119 Target condition: labrum, any tear,of. Index test: crank test (novel/standard).	2	127
120 Target condition: labrum, any tear,of. Index test: 'impingement sign' (no reference or details given).	1	32
121 Target condition: labrum, any tear,of. Index test: 'impingement test' (no reference or details given).	1	32
122 Target condition: labrum, any SLAP lesion of. Index test: active compression test (modified interpretation).	1	101
123 Target condition: labrum, any SLAP lesion of. Index test: anterior apprehension test at 90° for pain (Krishnan 2004: modified interpretation).	1	60
124 Target condition: labrum, any SLAP lesion,of. Index test: anterior release test (Gross 1997: modified interpretation).	1	60
125 Target condition: labrum, any SLAP lesion of. Index test: anterior slide test (modified procedure, modified interpretation).	1	101

Test	No. of studies	No. of participants
126 Target condition: labrum, any SLAP lesion of. Index test: bear-hug test (modified interpretation).	1	101
127 Target condition: labrum, any SLAP lesion of. Index test: belly-press test (modified interpretation).	1	101
128 Target condition: labrum, any SLAP lesion of. Index test: crank test (Liu 1996b: modified procedure, modified interpretation).	1	60
129 Target condition: labrum, any SLAP lesion of. Index test: modified dynamic labral shear (novel).	1	101
130 Target condition: labrum, any SLAP lesion of. Index test: palpation for bicipital groove tenderness (modified interpretation).	1	60
131 Target condition: labrum, any SLAP lesion of. Index test: passive compression test (novel).	1	61
132 Target condition: labrum, any SLAP lesion of. Index test: Speed's test (modified procedure, modified interpretation).	1	101
133 Target condition: labrum, any SLAP lesion of. Index test: Speed's test (modified interpretation).	1	60
134 Target condition: labrum, any SLAP lesion of. Index test: upper cut test (novel).	1	101
135 Target condition: labrum, any SLAP lesion of. Index test: Yergason's test (modified interpretation).	1	60
136 Target condition: labrum, any SLAP lesion of. Index test: Yergason's test (modified procedure, modified interpretation).	1	101
137 Target condition: labrum, type II-IV SLAP lesion of. Index test: active compression test (modified interpretation).	1	254
138 Target condition: labrum, type II-IV SLAP lesion of. Index test: anterior slide test (modified procedure).	1	254
139 Target condition: labrum, type II-IV SLAP lesion of. Index test: combination of active compression test (modified interpretation) OR passive distraction test (standard).	1	254
140 Target condition: labrum, type II-IV SLAP lesion of. Index test: passive compression test (novel, modified interpretation).	1	61
141 Target condition: labrum, type II-IV SLAP lesion of. Index test: passive distraction test (standard).	1	254
142 Target condition: labrum, type II SLAP lesion of. Index test: active compression test (modified interpretation 2).	1	146
143 Target condition: labrum, type II SLAP lesion of. Index test: active compression test (modified interpretation 1,2).	1	132

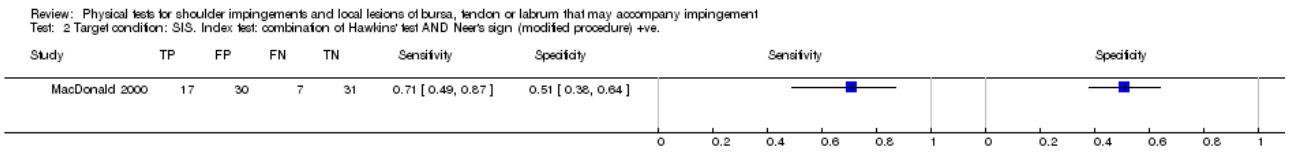
Test	No. of studies	No. of participants
144 Target condition: labrum, type II SLAP lesion of. Index test: anterior apprehension test at 90° for pain OR apprehension (Rowe 1981: modified interpretation).	1	146
145 Target condition: labrum, type II SLAP lesion of. Index test: anterior slide test (modified interpretation).	2	278
146 Target condition: labrum, type II SLAP lesion of. Index test: biceps load test II (novel/standard).	2	273
147 Target condition: labrum, type II SLAP lesion of. Index test: compression-rotation test (modified interpretation).	1	146
148 Target condition: labrum, type II SLAP lesion of. Index test: crank test (modified procedure, modified interpretation).	1	132
149 Target condition: labrum, type II SLAP lesion of. Index test: Hawkins' test (modified procedure, modified interpretation).	1	132
150 Target condition: labrum, type II SLAP lesion of. Index test: modified relocation test for posterosuperior glenoid impingement (modified interpretation).	1	132
151 Target condition: labrum, type II SLAP lesion of. Index test: Neer's sign (modified procedure, modified interpretation).	1	132
152 Target condition: labrum, type II SLAP lesion of. Index test: pain provocation test (modified interpretation).	1	132
153 Target condition: labrum, type II SLAP lesion of. Index test: palpation for bicipital groove tenderness (modified interpretation).	1	146
154 Target condition: labrum, type II SLAP lesion of. Index test: relocation test for pain OR apprehension (modified interpretation).	1	146
155 Target condition: labrum, type II SLAP lesion of. Index test: Speed's test (modified interpretation).	1	132
156 Target condition: labrum, type II SLAP lesion of. Index test: Speed's test (modified procedure, modified interpretation).	1	146
157 Target condition: labrum, type II SLAP lesion of. Index test: Whipple's test (modified interpretation).	1	146
158 Target condition: labrum, type II SLAP lesion of. Index test: Yergason's test (modified interpretation 2).	1	146
159 Target condition: labrum, type II SLAP lesion of. Index test: Yergason's test (modified interpretation 1,2).	1	132
160 Target condition: multiple (LHB tendinitis/LHB avulsion/SLAP lesion, any). Index test: Speed's test (modified procedure, modified interpretation 1).	1	46
161 Target condition: multiple (LHB lesion, any/type II or IV SLAP lesion). Index test: Speed's test (modified procedure, modified interpretation 1,2).	1	50

Test	No. of studies	No. of participants
162 Target condition: multiple (LHB lesion, any/type II or IV SLAP lesion). Index test: Yergason's test (modified interpretation 1,2).	1	49
163 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Empty can test (modified interpretation).	1	55
164 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Hawkins' test (standard).	1	55
165 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Neer's sign (modified procedure).	1	55
166 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: painful arc test (standard).	1	45
167 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability) Index test: resisted lateral rotation from neutral rotation for weakness ± pain (modified interpretation 1,2).	1	55
168 Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: combination of 3 or more tests +ve (*see* table 11).	1	55
169 Target condition: multiple (SIS/rotator cuff tendinitis or tear). Index test: Hawkins' test (modified interpretation).	1	50
170 Target condition: multiple (SIS/rotator cuff tendinitis or tear). Index test: Neer's sign (modified interpretation).	1	50

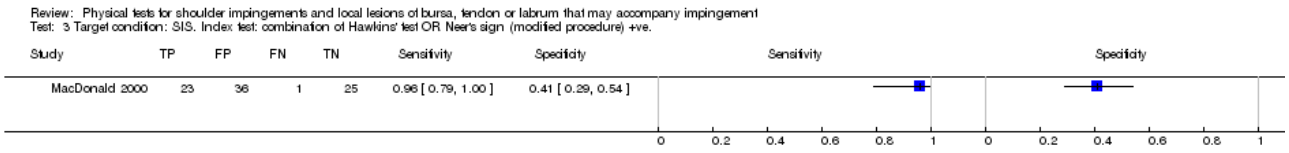
Test 1. Target condition: SIS. Index test: combination of ALL 7 tests +ve (*see* table 7)..



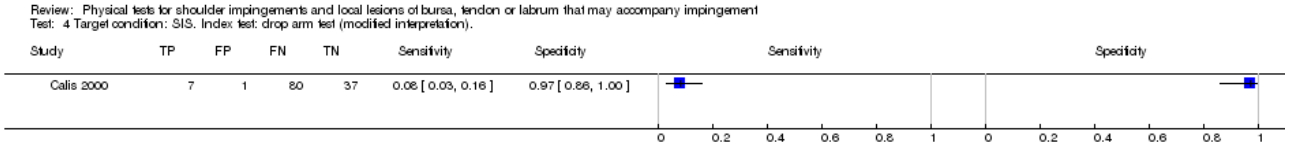
Test 2. Target condition: SIS. Index test: combination of Hawkins' test AND Neer's sign (modified procedure) +ve..



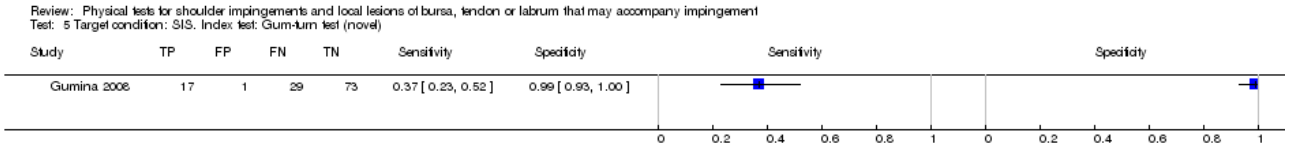
Test 3. Target condition: SIS. Index test: combination of Hawkins' test OR Neer's sign (modified procedure) +ve..



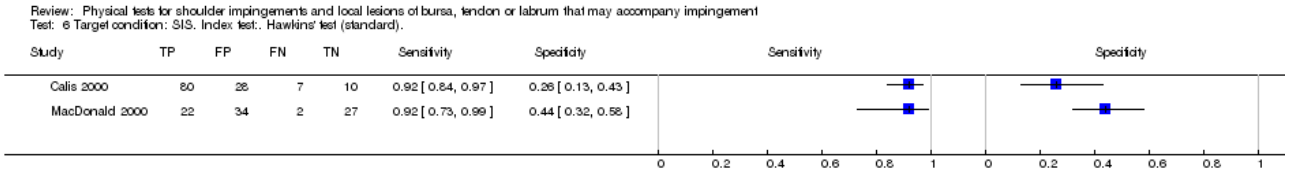
Test 4. Target condition: SIS. Index test: drop arm test (modified interpretation)..



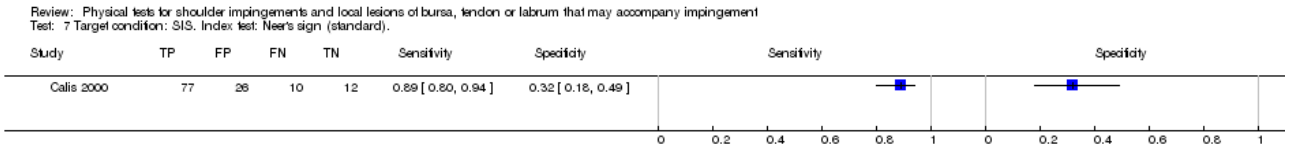
Test 5. Target condition: SIS. Index test: Gum-turn test (novel).



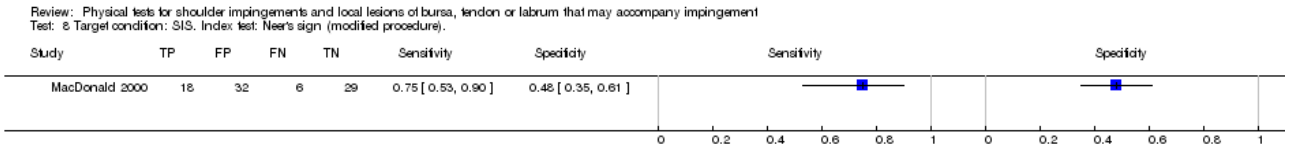
Test 6. Target condition: SIS. Index test:. Hawkins' test (standard)..



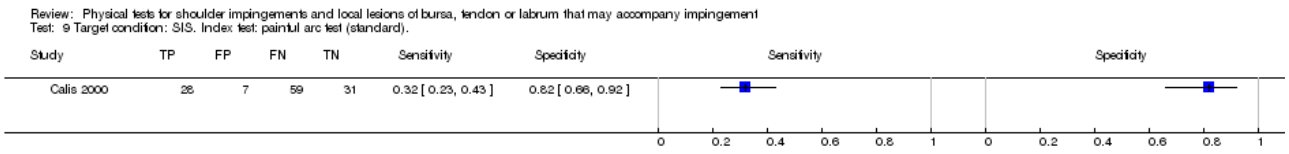
Test 7. Target condition: SIS. Index test: Neer's sign (standard)..



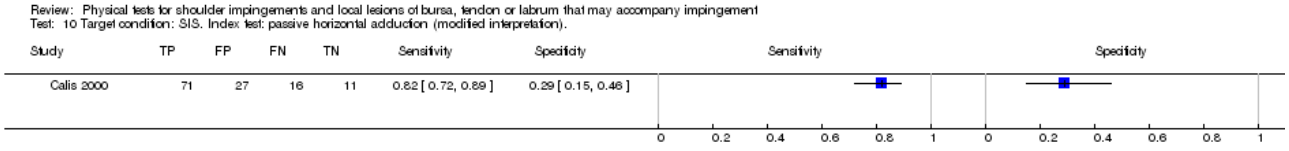
Test 8. Target condition: SIS. Index test: Neer's sign (modified procedure)..



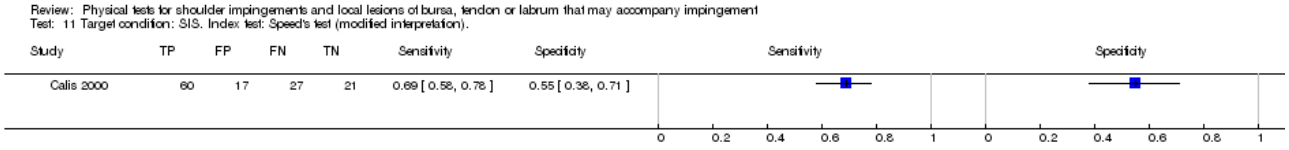
Test 9. Target condition: SIS. Index test: painful arc test (standard)..



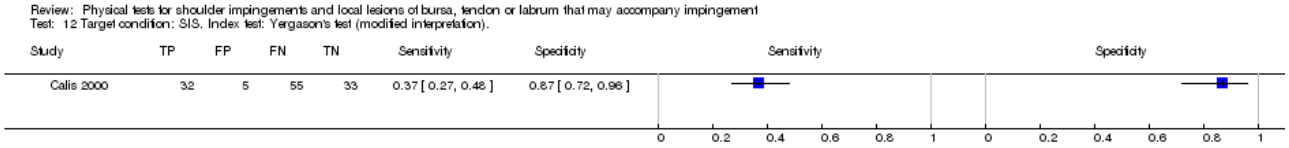
Test 10. Target condition: SIS. Index test: passive horizontal adduction (modified interpretation)..



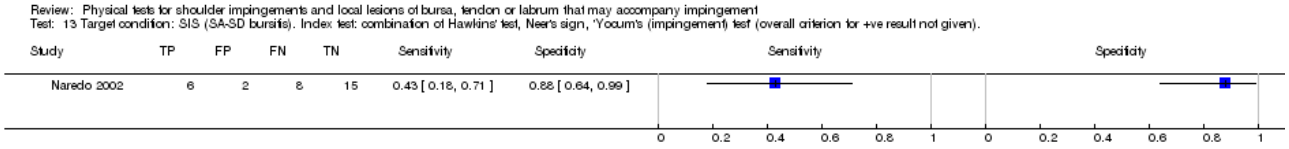
Test 11. Target condition: SIS. Index test: Speed's test (modified interpretation)..



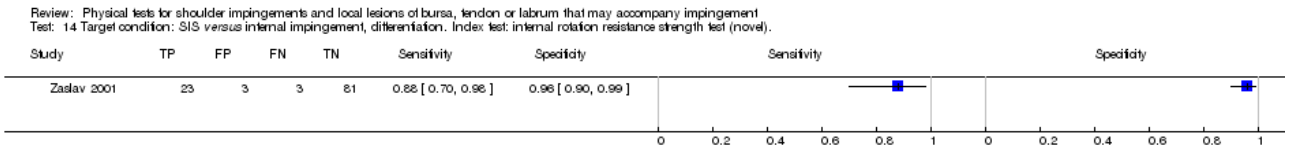
Test 12. Target condition: SIS. Index test: Yergason's test (modified interpretation)..



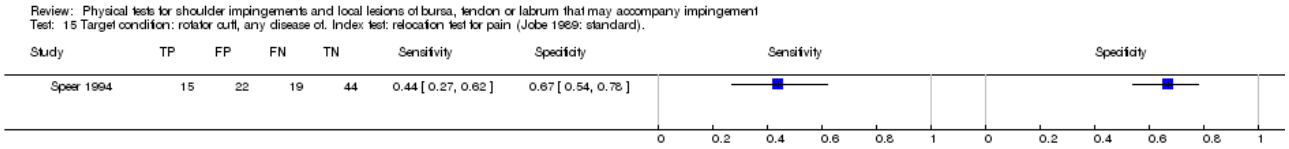
Test 13. Target condition: SIS (SA-SD bursitis). Index test: combination of Hawkins' test, Neer's sign, 'Yocum's (impingement) test' (overall criterion for +ve result not given)..



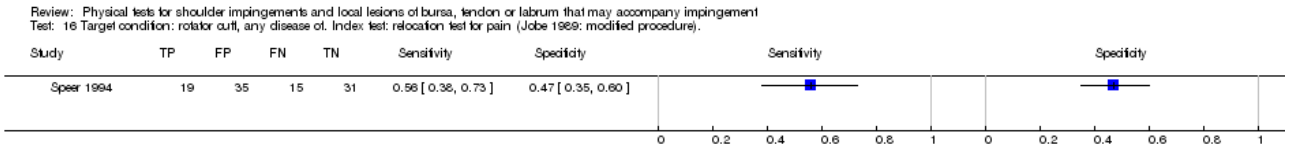
Test 14. Target condition: SIS *versus* internal impingement, differentiation. Index test: internal rotation resistance strength test (novel)..



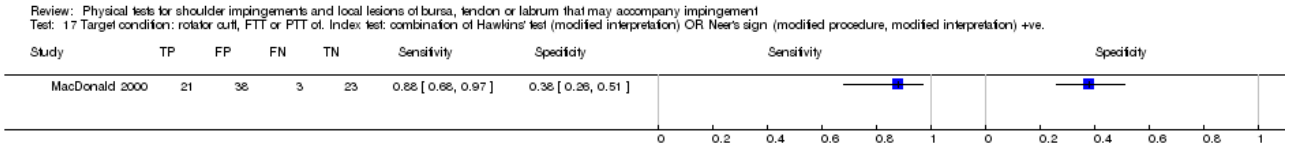
Test 15. Target condition: rotator cuff, any disease of. Index test: relocation test for pain (Jobe 1989: standard)..



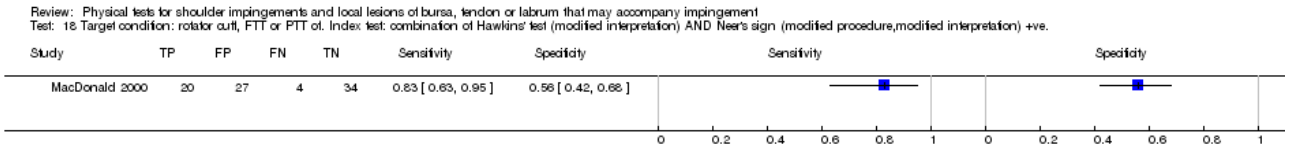
Test 16. Target condition: rotator cuff, any disease of. Index test: relocation test for pain (Jobe 1989: modified procedure)..



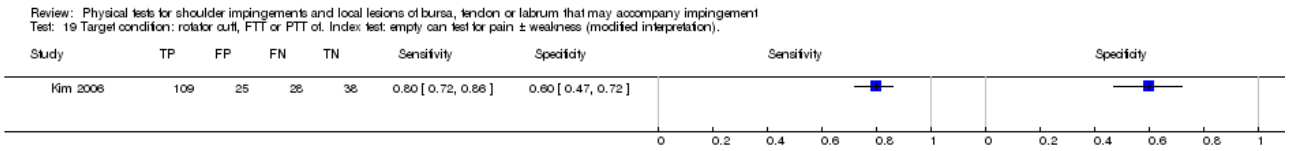
Test 17. Target condition: rotator cuff, FTT or PTT of. Index test: combination of Hawkins' test (modified interpretation) OR Neer's sign (modified procedure, modified interpretation) +ve..



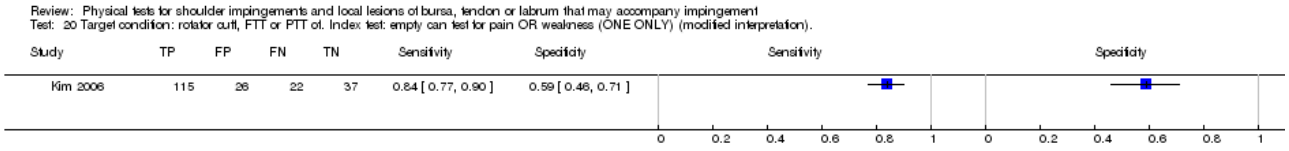
Test 18. Target condition: rotator cuff, FTT or PTT of. Index test: combination of Hawkins' test (modified interpretation) AND Neer's sign (modified procedure,modified interpretation) +ve..



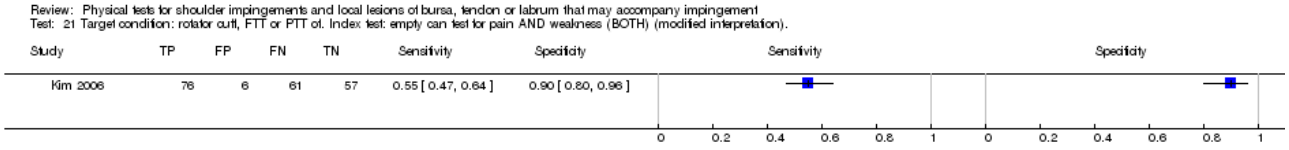
Test 19. Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain ± weakness (modified interpretation)..



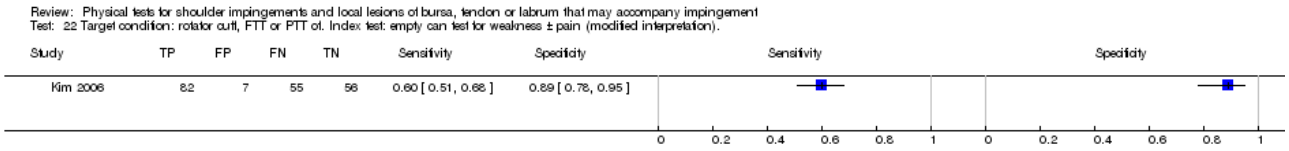
Test 20. Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain OR weakness (ONE ONLY) (modified interpretation)..



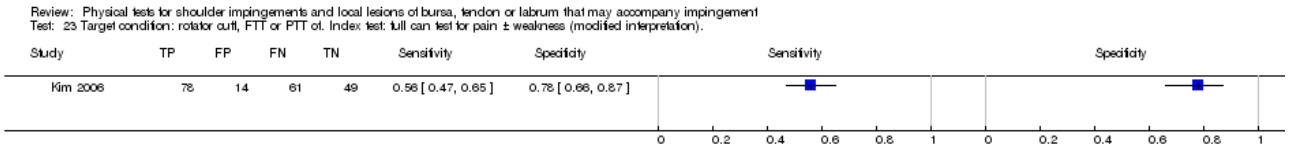
Test 21. Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for pain AND weakness (BOTH) (modified interpretation)..



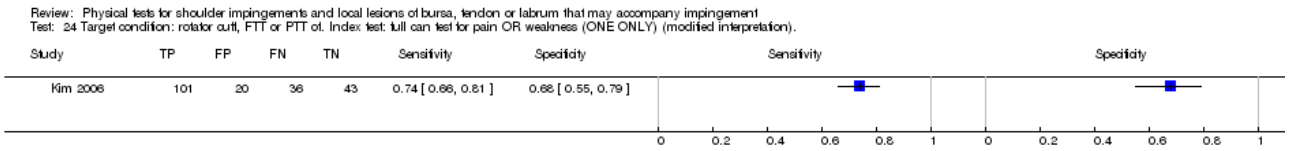
Test 22. Target condition: rotator cuff, FTT or PTT of. Index test: empty can test for weakness ± pain (modified interpretation)..



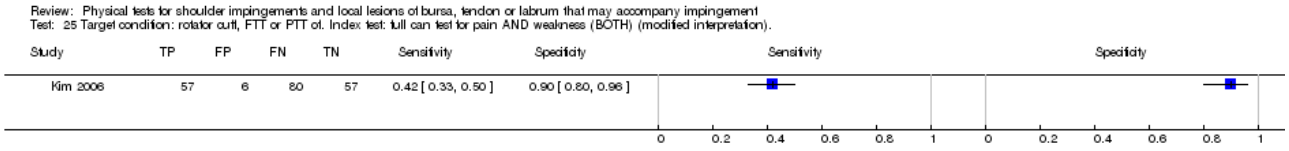
Test 23. Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain ± weakness (modified interpretation)..



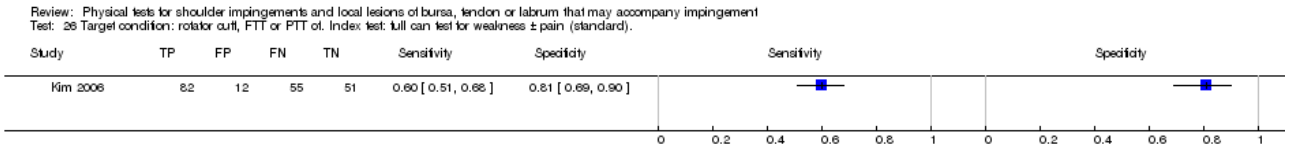
Test 24. Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain OR weakness (ONE ONLY) (modified interpretation)..



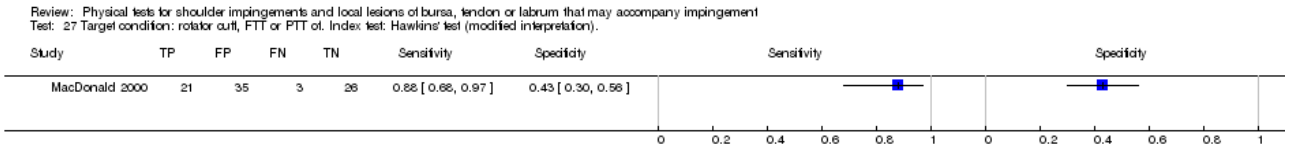
Test 25. Target condition: rotator cuff, FTT or PTT of. Index test: full can test for pain AND weakness (BOTH) (modified interpretation)..



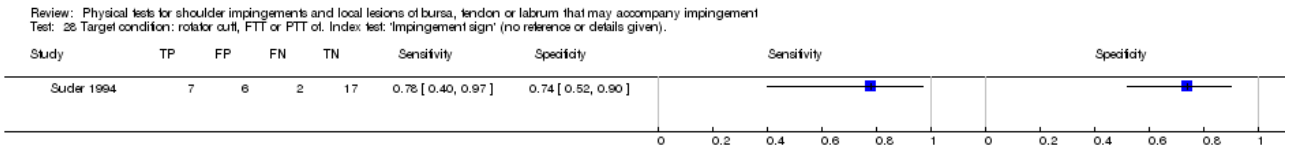
Test 26. Target condition: rotator cuff, FTT or PTT of. Index test: full can test for weakness ± pain (standard)..



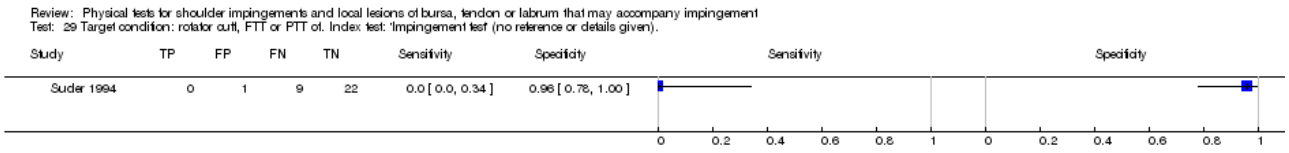
Test 27. Target condition: rotator cuff, FTT or PTT of. Index test: Hawkins' test (modified interpretation)..



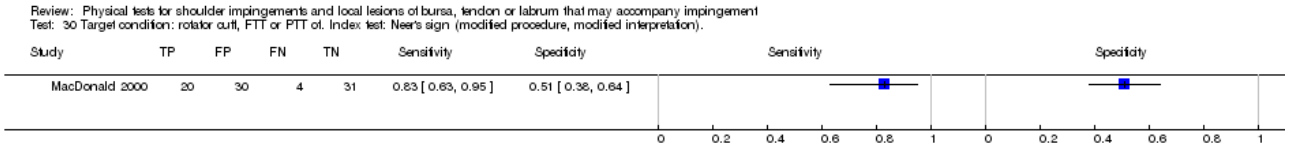
Test 28. Target condition: rotator cuff, FTT or PTT of. Index test: 'Impingement sign' (no reference or details given)..



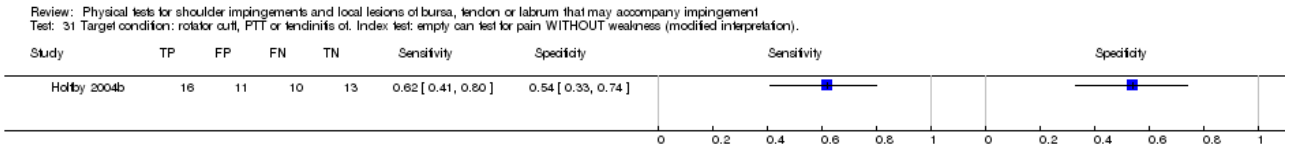
Test 29. Target condition: rotator cuff, FTT or PTT of. Index test: 'Impingement test' (no reference or details given)..



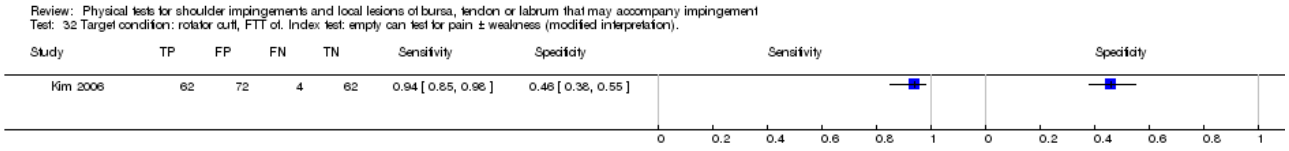
Test 30. Target condition: rotator cuff, FTT or PTT of. Index test: Neer's sign (modified procedure, modified interpretation)..



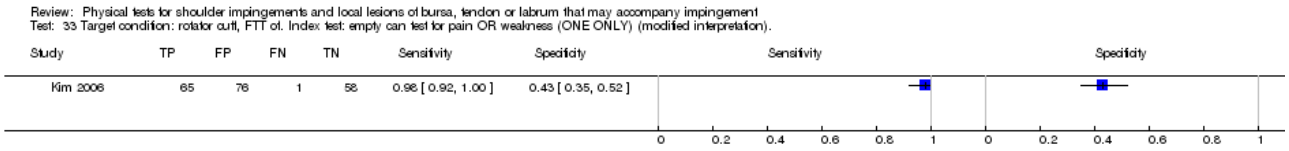
Test 31. Target condition: rotator cuff, PTT or tendinitis of. Index test: empty can test for pain WITHOUT weakness (modified interpretation)..



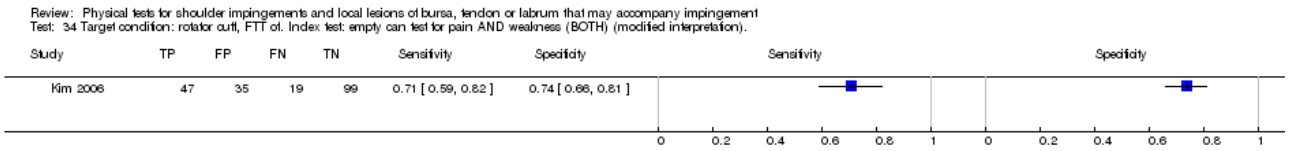
Test 32. Target condition: rotator cuff, FTT of. Index test: empty can test for pain ± weakness (modified interpretation)..



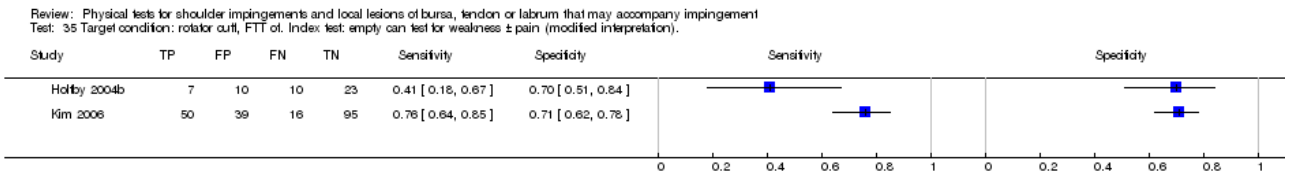
Test 33. Target condition: rotator cuff, FTT of. Index test: empty can test for pain OR weakness (ONE ONLY) (modified interpretation)..



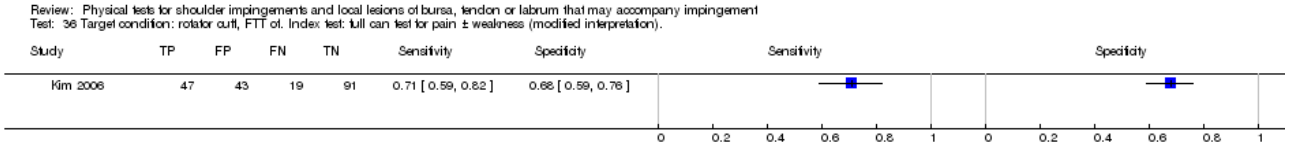
Test 34. Target condition: rotator cuff, FTT of. Index test: empty can test for pain AND weakness (BOTH) (modified interpretation)..



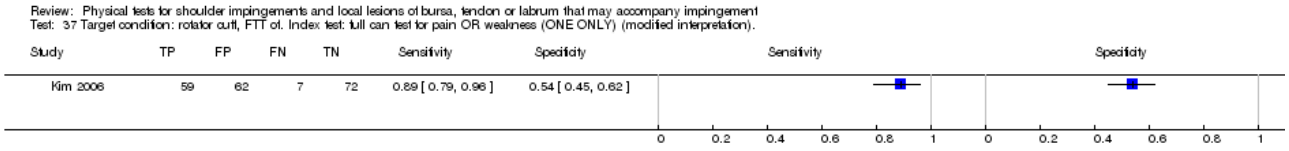
Test 35. Target condition: rotator cuff, FTT of. Index test: empty can test for weakness ± pain (modified interpretation)..



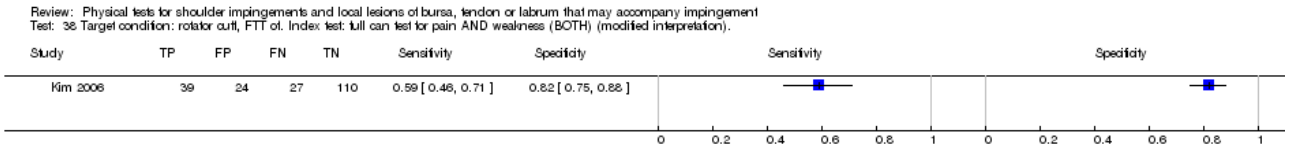
Test 36. Target condition: rotator cuff, FTT of. Index test: full can test for pain ± weakness (modified interpretation)..



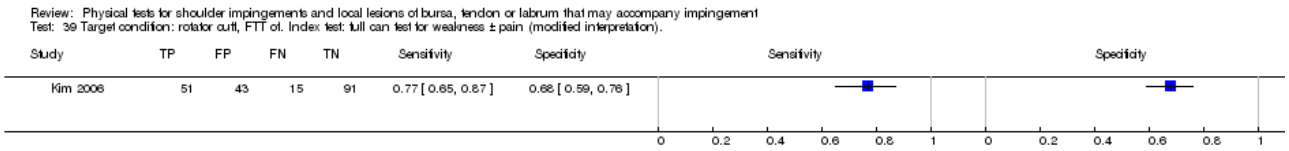
Test 37. Target condition: rotator cuff, FTT of. Index test: full can test for pain OR weakness (ONE ONLY) (modified interpretation)..



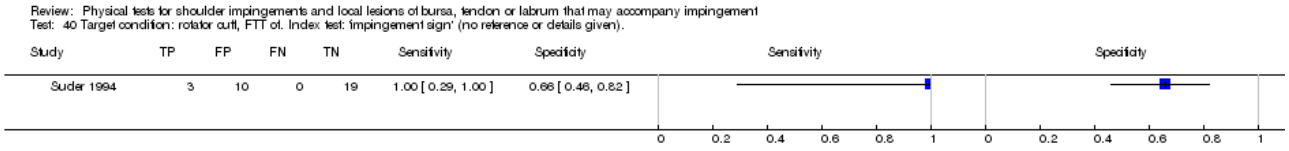
Test 38. Target condition: rotator cuff, FTT of. Index test: full can test for pain AND weakness (BOTH) (modified interpretation)..



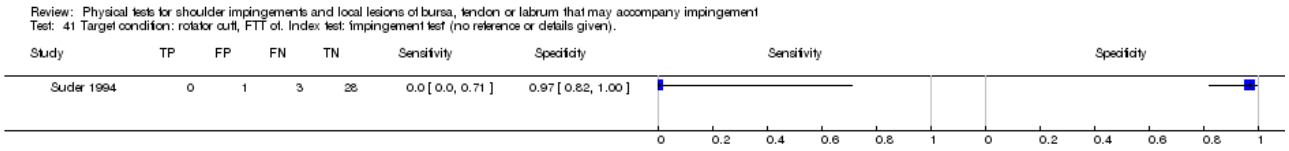
Test 39. Target condition: rotator cuff, FTT of. Index test: full can test for weakness ± pain (modified interpretation)..



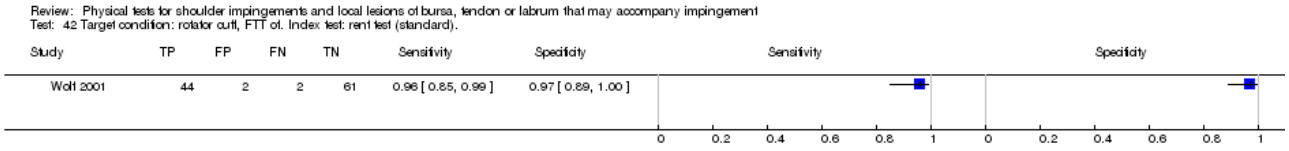
Test 40. Target condition: rotator cuff, FTT of. Index test: 'impingement sign' (no reference or details given)..



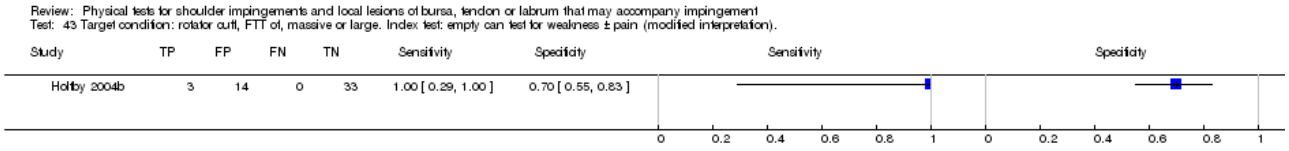
Test 41. Target condition: rotator cuff, FTT of. Index test: 'impingement test' (no reference or details given)..



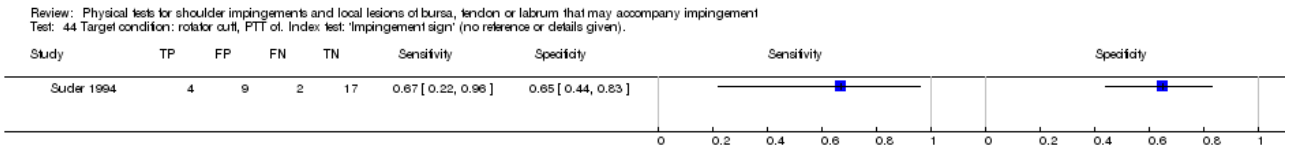
Test 42. Target condition: rotator cuff, FTT of. Index test: rent test (standard)..



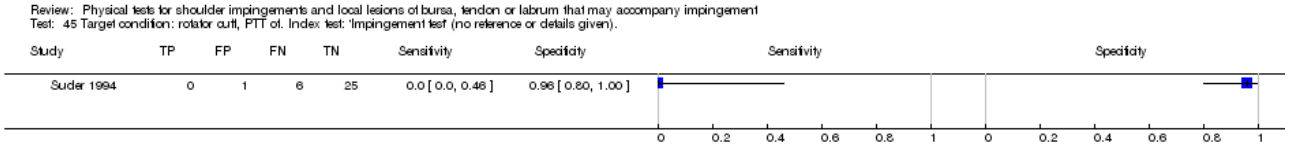
Test 43. Target condition: rotator cuff, FTT of, massive or large. Index test: empty can test for weakness ± pain (modified interpretation)..



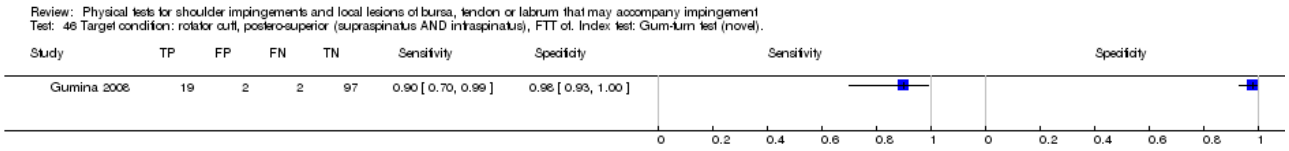
Test 44. Target condition: rotator cuff, PTT of. Index test: 'Impingement sign' (no reference or details given)..



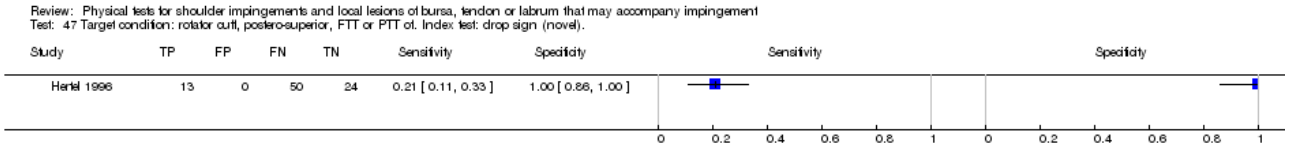
Test 45. Target condition: rotator cuff, PTT of. Index test: 'Impingement test' (no reference or details given)..



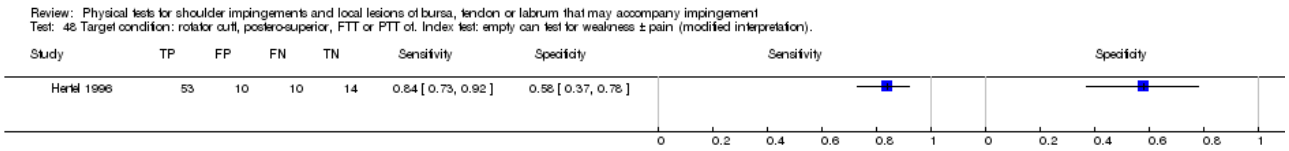
Test 46. Target condition: rotator cuff, postero-superior (supraspinatus AND infraspinatus), FTT of. Index test: Gum-turn test (novel)..



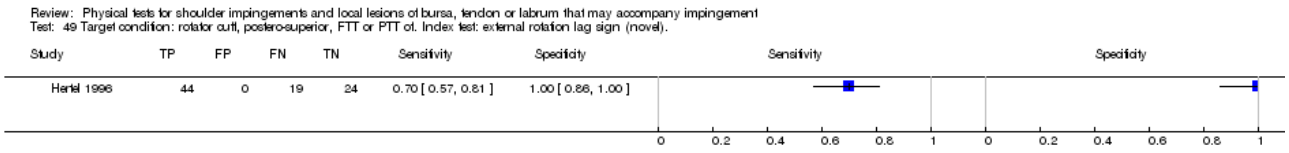
Test 47. Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: drop sign (novel)..



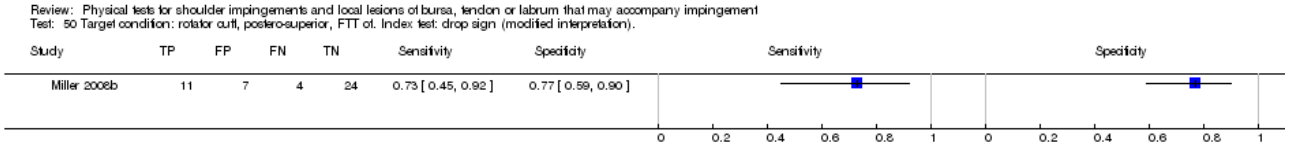
Test 48. Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: empty can test for weakness ± pain (modified interpretation)..



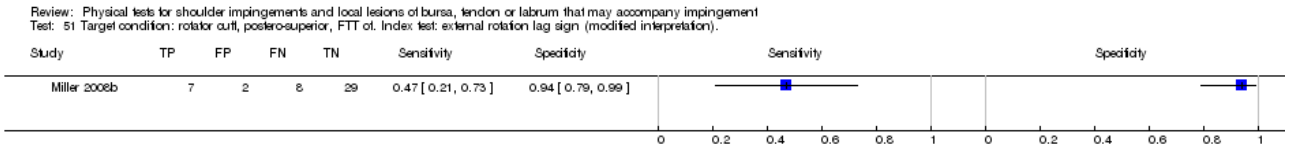
Test 49. Target condition: rotator cuff, postero-superior, FTT or PTT of. Index test: external rotation lag sign (novel)..



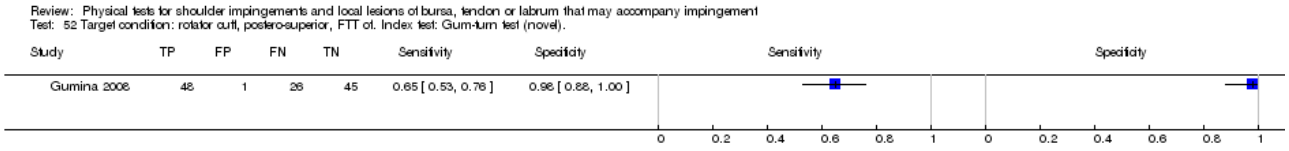
Test 50. Target condition: rotator cuff, postero-superior, FTT of. Index test: drop sign (modified interpretation)..



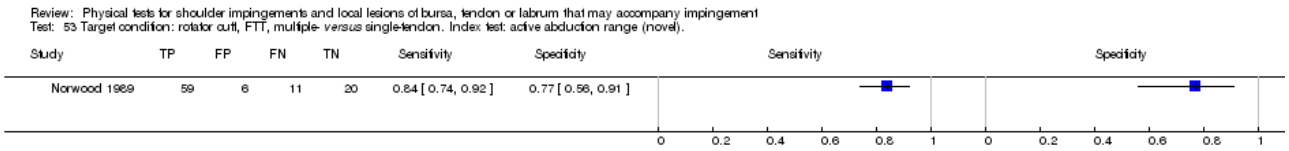
Test 51. Target condition: rotator cuff, postero-superior, FTT of. Index test: external rotation lag sign (modified interpretation)..



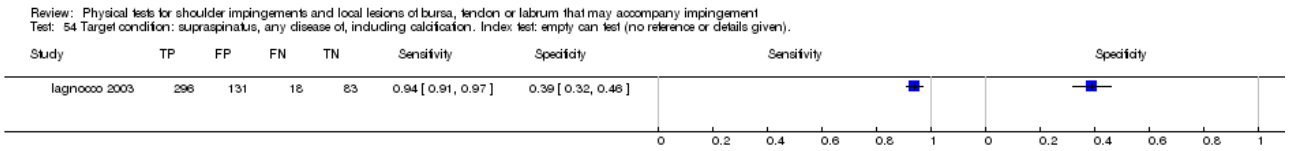
Test 52. Target condition: rotator cuff, postero-superior, FTT of. Index test: Gum-turn test (novel)..



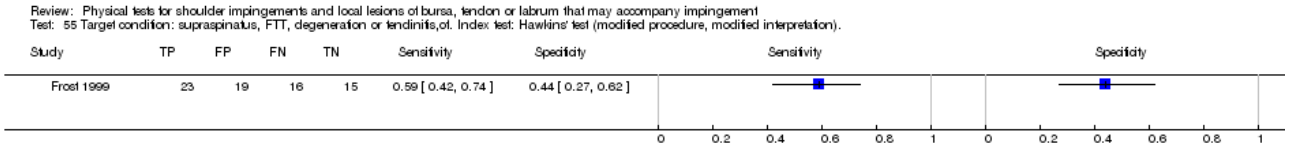
Test 53. Target condition: rotator cuff, FTT, multiple- *versus* single-tendon. Index test: active abduction range (novel)..



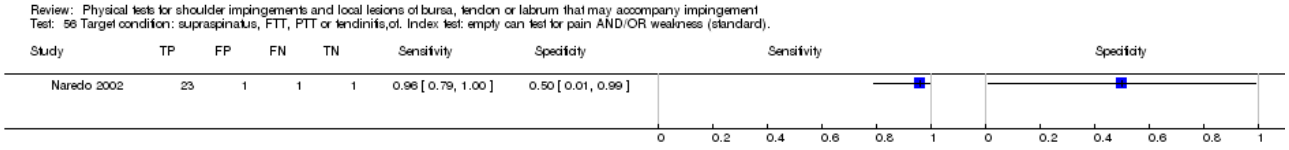
Test 54. Target condition: supraspinatus, any disease of, including calcification. Index test: empty can test (no reference or details given)..



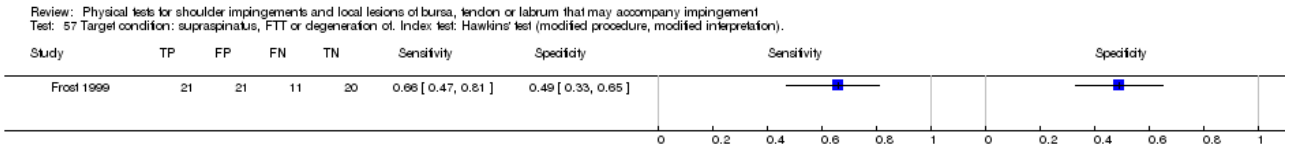
Test 55. Target condition: supraspinatus, FTT, degeneration or tendinitis,of. Index test: Hawkins' test (modified procedure, modified interpretation)..



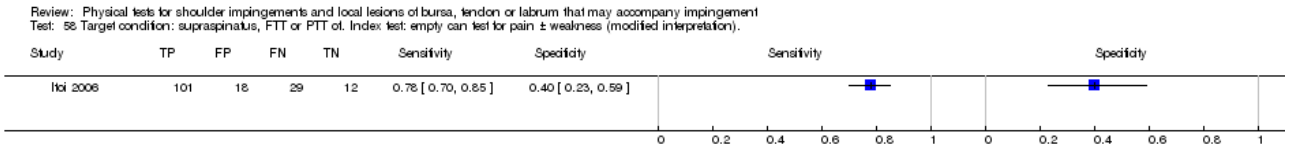
Test 56. Target condition: supraspinatus, FTT, PTT or tendinitis,of. Index test: empty can test for pain AND/OR weakness (standard)..



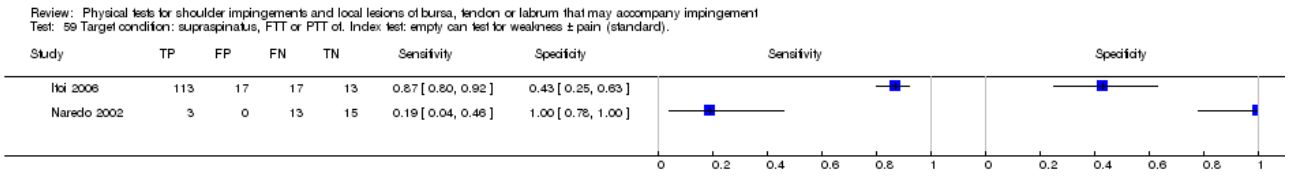
Test 57. Target condition: supraspinatus, FTT or degeneration of. Index test: Hawkins' test (modified procedure, modified interpretation)..



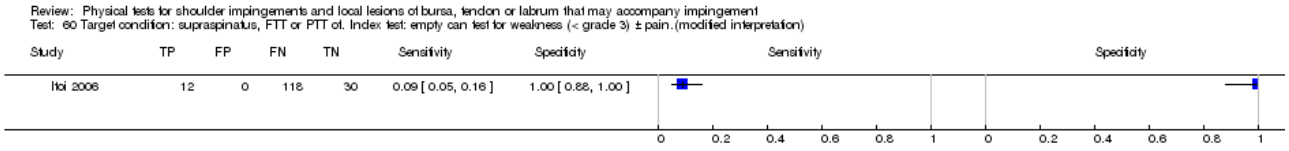
Test 58. Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for pain ± weakness (modified interpretation)..



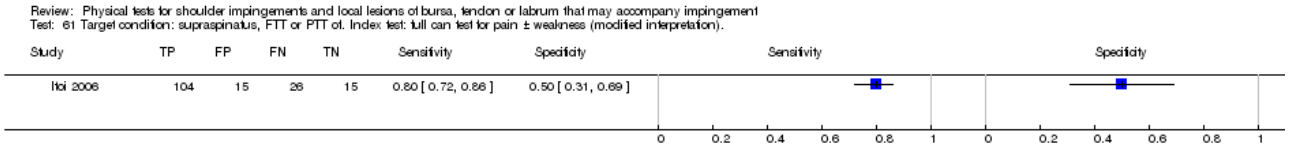
Test 59. Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for weakness ± pain (standard)..



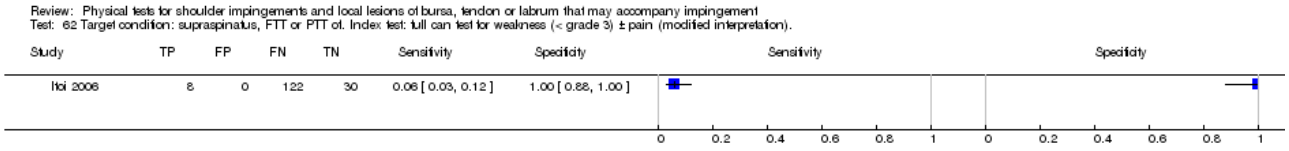
Test 60. Target condition: supraspinatus, FTT or PTT of. Index test: empty can test for weakness (< grade 3) ± pain.(modified interpretation).



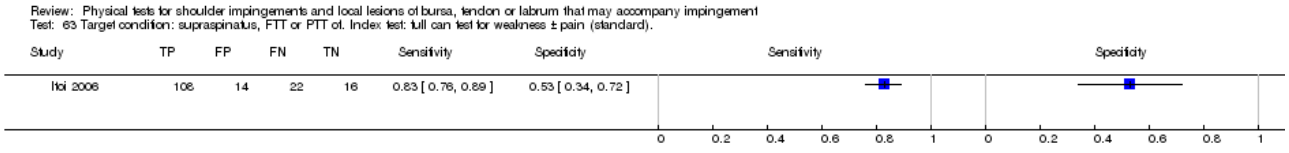
Test 61. Target condition: supraspinatus, FTT or PTT of. Index test: full can test for pain ± weakness (modified interpretation)..



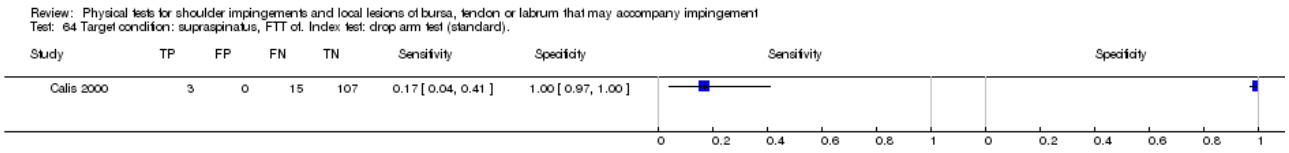
Test 62. Target condition: supraspinatus, FTT or PTT of. Index test: full can test for weakness (< grade 3) ± pain (modified interpretation)..



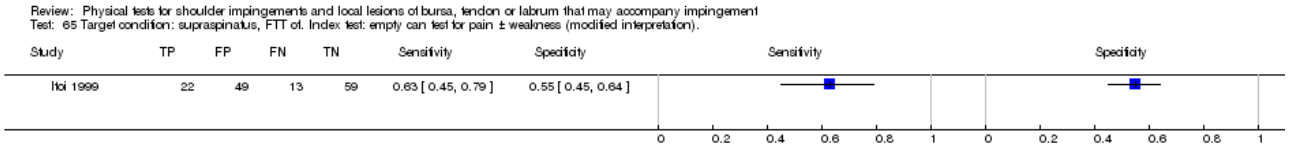
Test 63. Target condition: supraspinatus, FTT or PTT of. Index test: full can test for weakness ± pain (standard)..



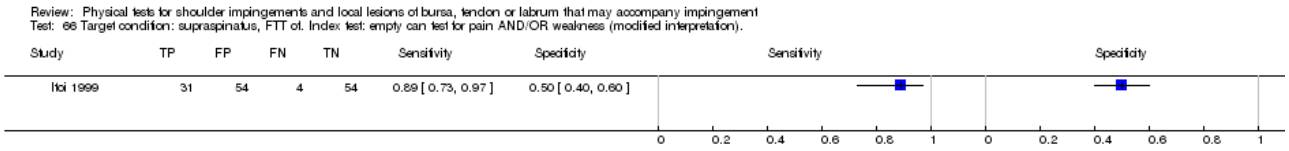
Test 64. Target condition: supraspinatus, FTT of. Index test: drop arm test (standard)..



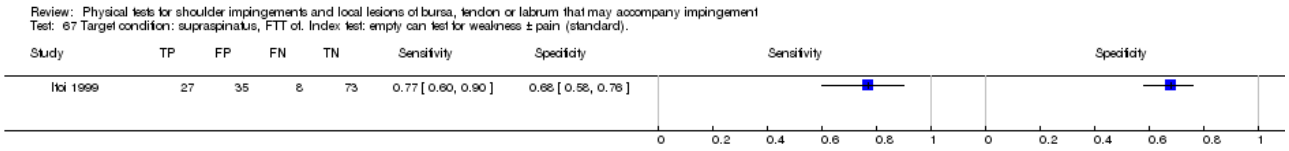
Test 65. Target condition: supraspinatus, FTT of. Index test: empty can test for pain ± weakness (modified interpretation)..



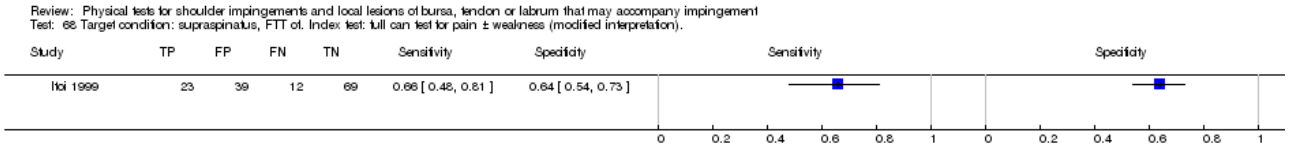
Test 66. Target condition: supraspinatus, FTT of. Index test: empty can test for pain AND/OR weakness (modified interpretation)..



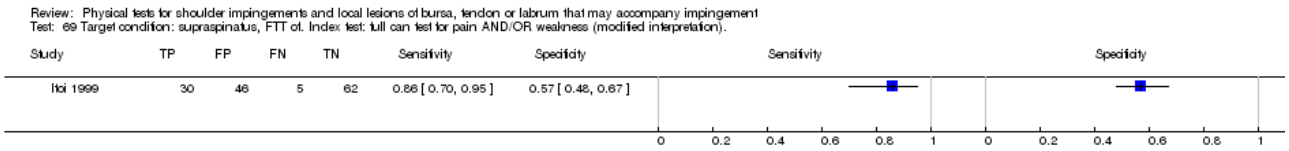
Test 67. Target condition: supraspinatus, FTT of. Index test: empty can test for weakness ± pain (standard)..



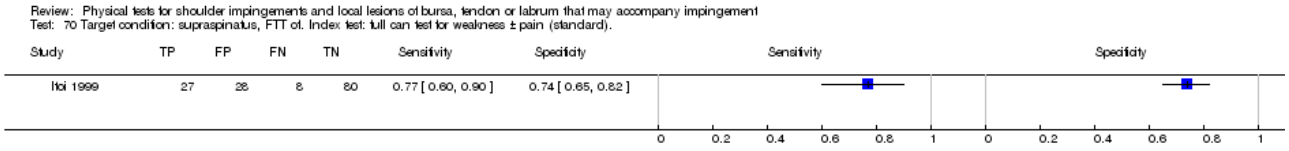
Test 68. Target condition: supraspinatus, FTT of. Index test: full can test for pain ± weakness (modified interpretation)..



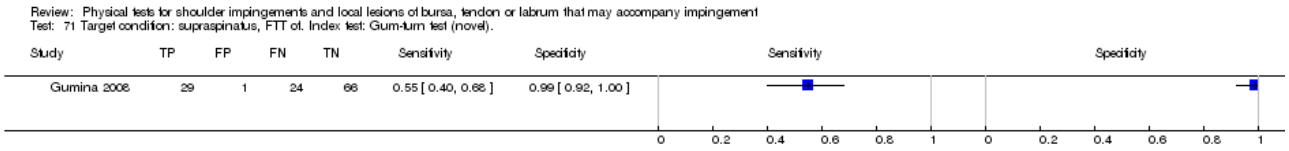
Test 69. Target condition: supraspinatus, FTT of. Index test: full can test for pain AND/OR weakness (modified interpretation)..



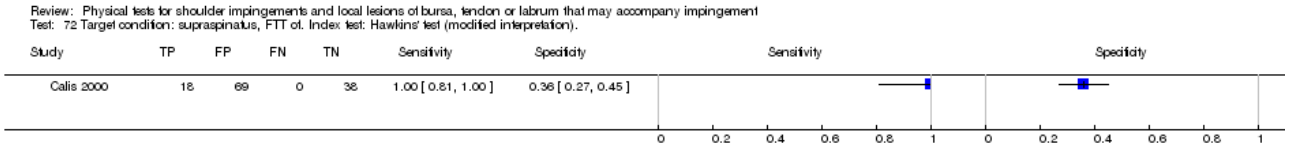
Test 70. Target condition: supraspinatus, FTT of. Index test: full can test for weakness ± pain (standard)..



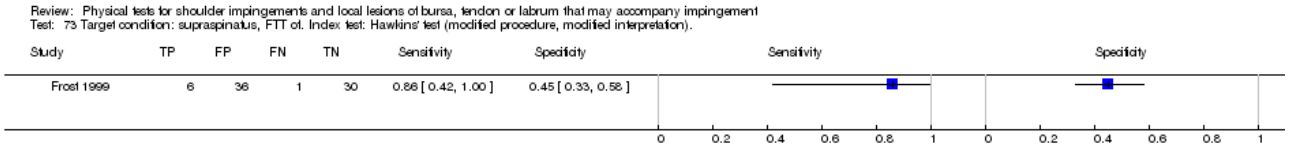
Test 71. Target condition: supraspinatus, FTT of. Index test: Gum-turn test (novel)..



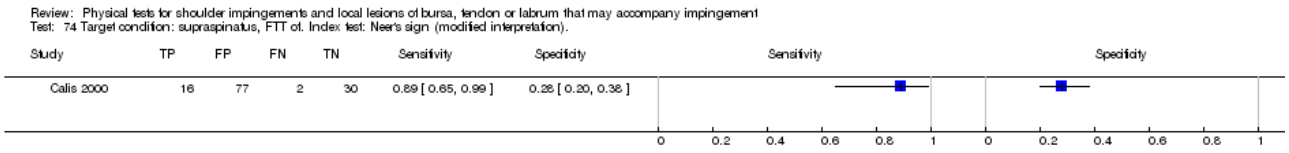
Test 72. Target condition: supraspinatus, FTT of. Index test: Hawkins' test (modified interpretation)..



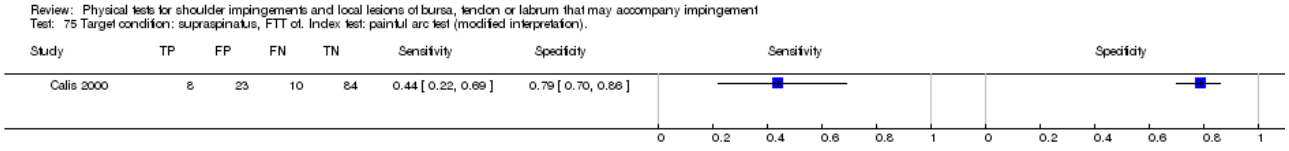
Test 73. Target condition: supraspinatus, FTT of. Index test: Hawkins' test (modified procedure, modified interpretation)..



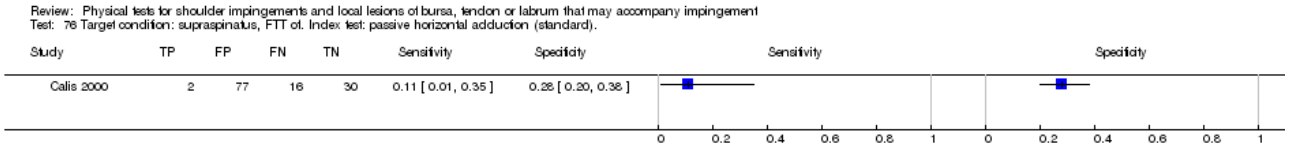
Test 74. Target condition: supraspinatus, FTT of. Index test: Neer's sign (modified interpretation)..



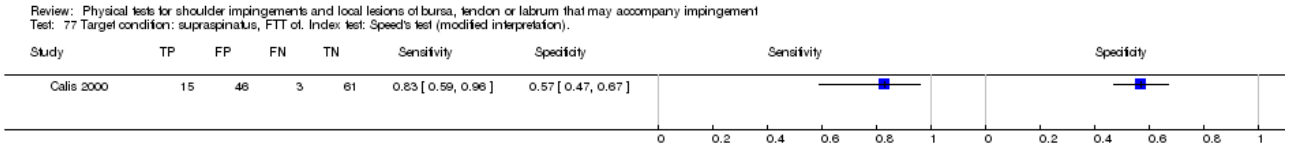
Test 75. Target condition: supraspinatus, FTT of. Index test: painful arc test (modified interpretation)..



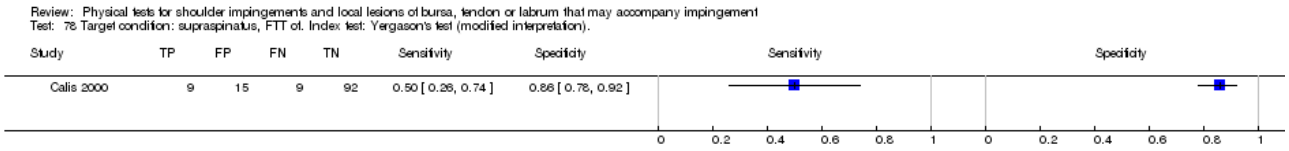
Test 76. Target condition: supraspinatus, FTT of. Index test: passive horizontal adduction (standard)..



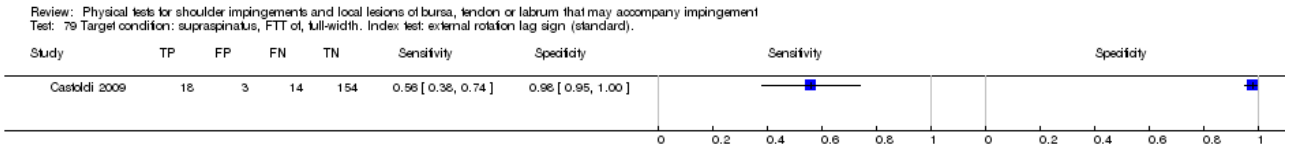
Test 77. Target condition: supraspinatus, FTT of. Index test: Speed's test (modified interpretation)..



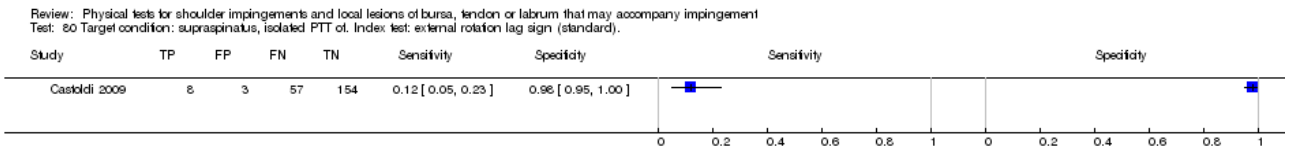
Test 78. Target condition: supraspinatus, FTT of. Index test: Yergason's test (modified interpretation)..



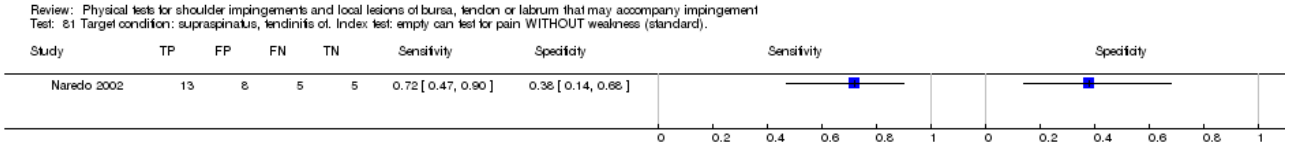
Test 79. Target condition: supraspinatus, FTT of, full-width. Index test: external rotation lag sign (standard)..



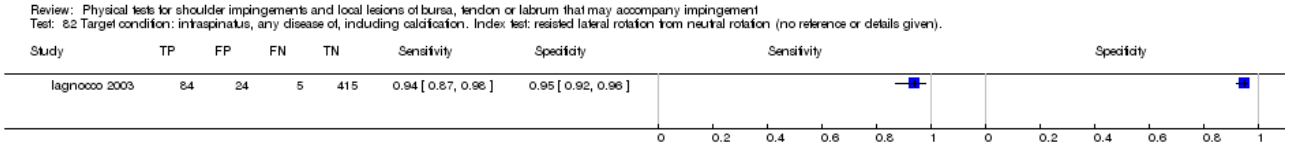
Test 80. Target condition: supraspinatus, isolated PTT of. Index test: external rotation lag sign (standard)..



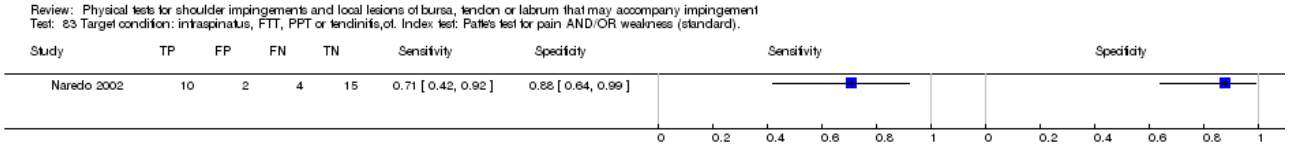
Test 81. Target condition: supraspinatus, tendinitis of. Index test: empty can test for pain WITHOUT weakness (standard)..



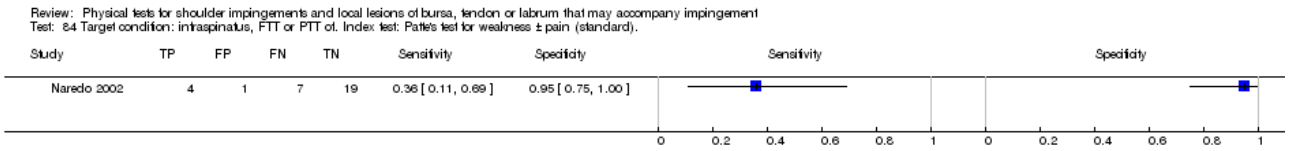
Test 82. Target condition: infraspinatus, any disease of, including calcification. Index test: resisted lateral rotation from neutral rotation (no reference or details given)..



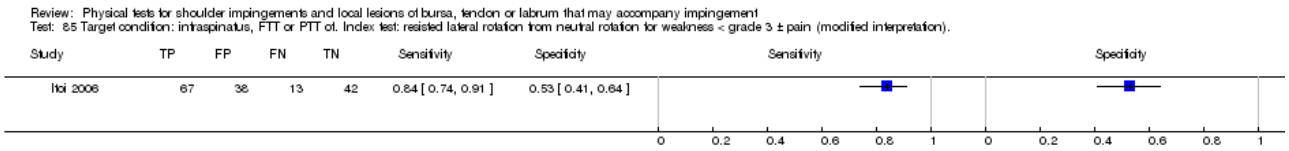
Test 83. Target condition: infraspinatus, FTT, PPT or tendinitis,of. Index test: Patte's test for pain AND/OR weakness (standard)..



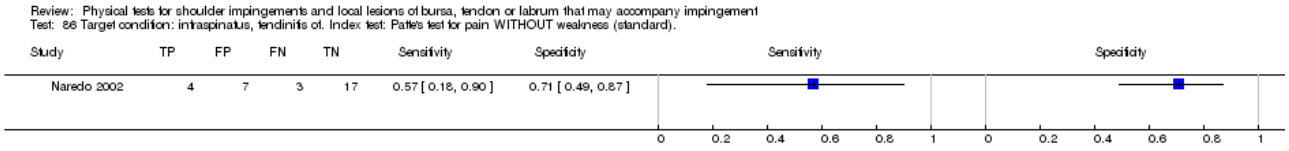
Test 84. Target condition: infraspinatus, FTT or PTT of. Index test: Patte's test for weakness ± pain (standard)..



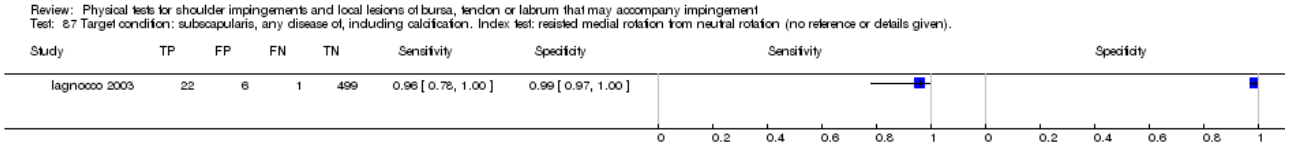
Test 85. Target condition: infraspinatus, FTT or PTT of. Index test: resisted lateral rotation from neutral rotation for weakness < grade 3 ± pain (modified interpretation)..



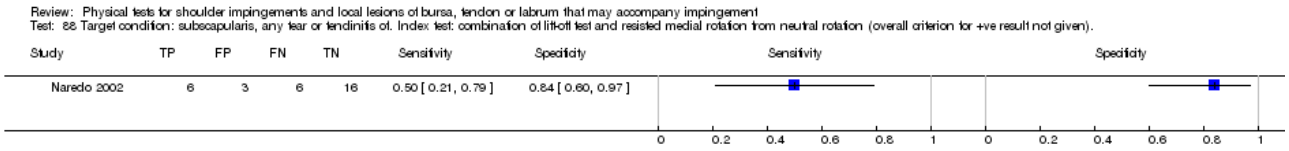
Test 86. Target condition: infraspinatus, tendinitis of. Index test: Patte's test for pain WITHOUT weakness (standard)..



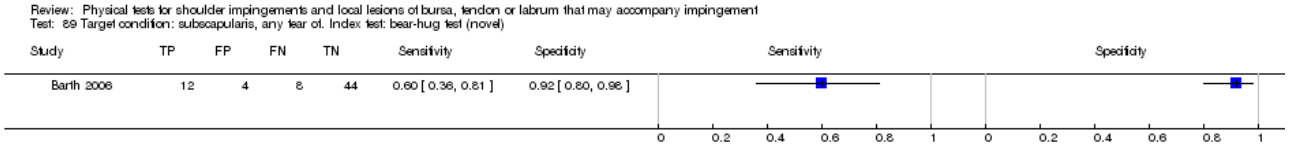
Test 87. Target condition: subscapularis, any disease of, including calcification. Index test: resisted medial rotation from neutral rotation (no reference or details given)..



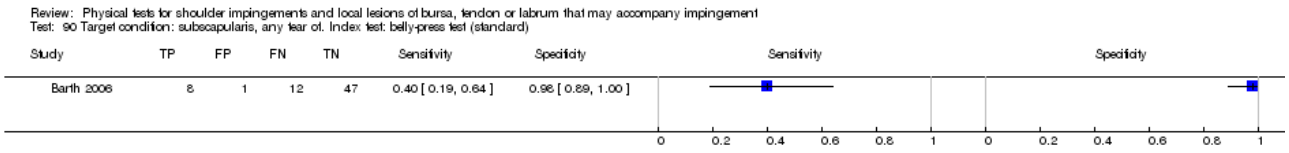
Test 88. Target condition: subscapularis, any tear or tendinitis of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given)..



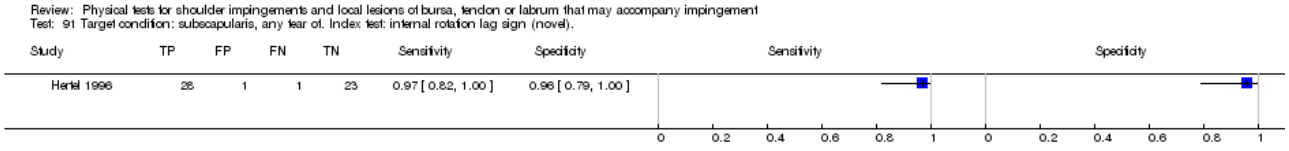
Test 89. Target condition: subscapularis, any tear of. Index test: bear-hug test (novel).



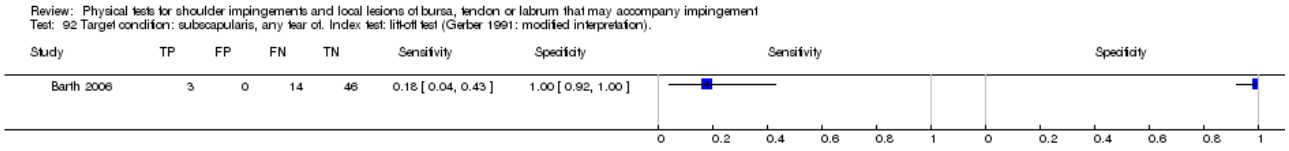
Test 90. Target condition: subscapularis, any tear of. Index test: belly-press test (standard).



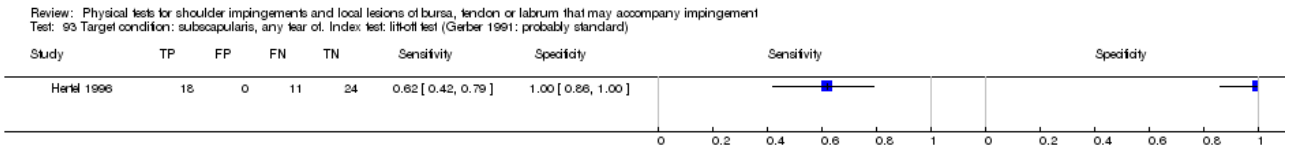
Test 91. Target condition: subscapularis, any tear of. Index test: internal rotation lag sign (novel)..



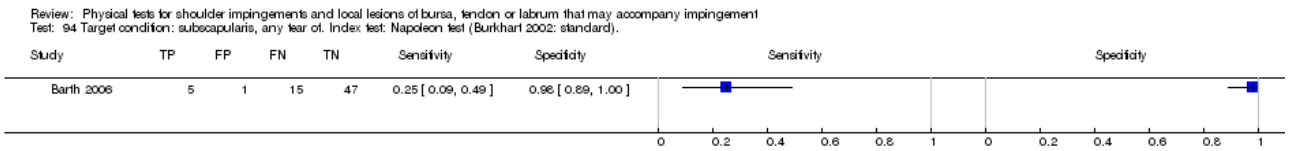
Test 92. Target condition: subscapularis, any tear of. Index test: lift-off test (Gerber 1991: modified interpretation)..



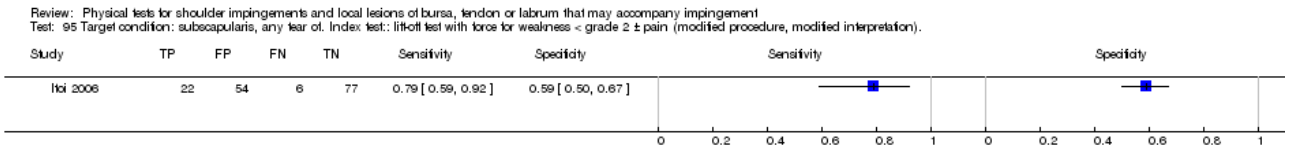
Test 93. Target condition: subscapularis, any tear of. Index test: lift-off test (Gerber 1991: probably standard).



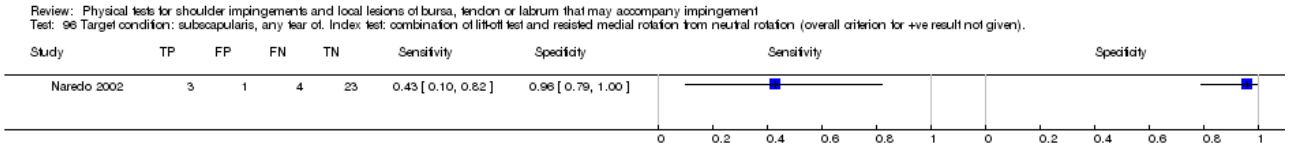
Test 94. Target condition: subscapularis, any tear of. Index test: Napoleon test (Burkhart 2002: standard)..



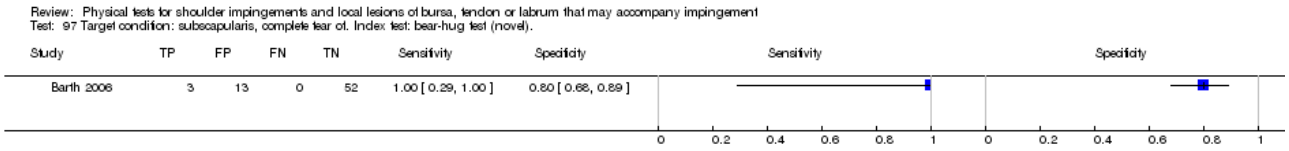
Test 95. Target condition: subscapularis, any tear of. Index test:: lift-off test with force for weakness < grade 2 ± pain (modified procedure, modified interpretation)..



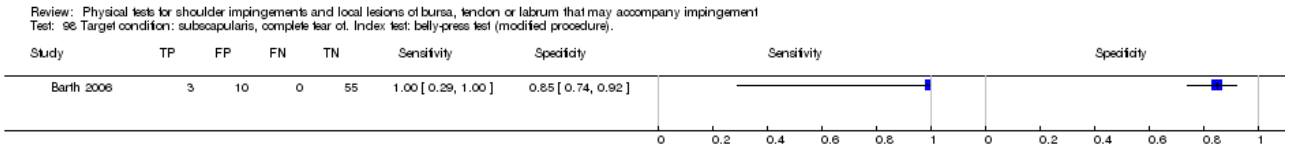
Test 96. Target condition: subscapularis, any tear of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given)..



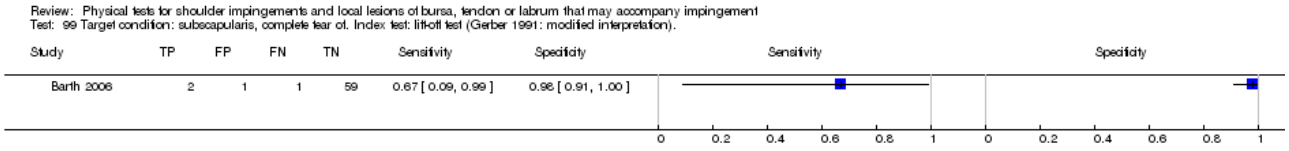
Test 97. Target condition: subscapularis, complete tear of. Index test: bear-hug test (novel)..



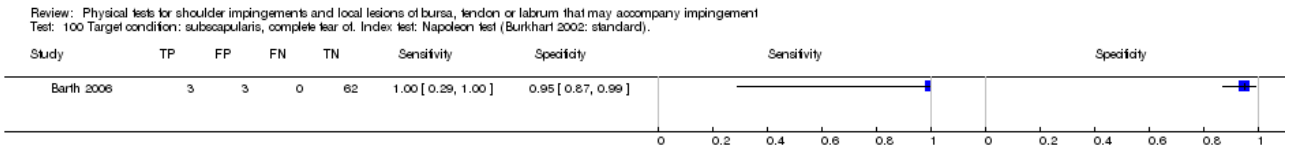
Test 98. Target condition: subscapularis, complete tear of. Index test: belly-press test (modified procedure)..



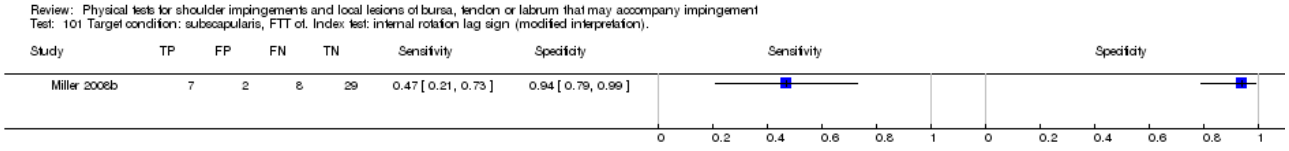
Test 99. Target condition: subscapularis, complete tear of. Index test: lift-off test (Gerber 1991: modified interpretation)..



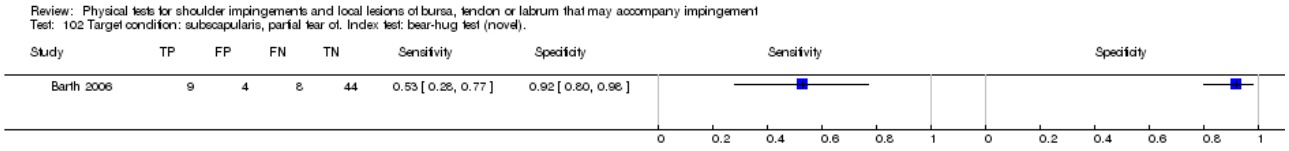
Test 100. Target condition: subscapularis, complete tear of. Index test: Napoleon test (Burkhart 2002: standard)..



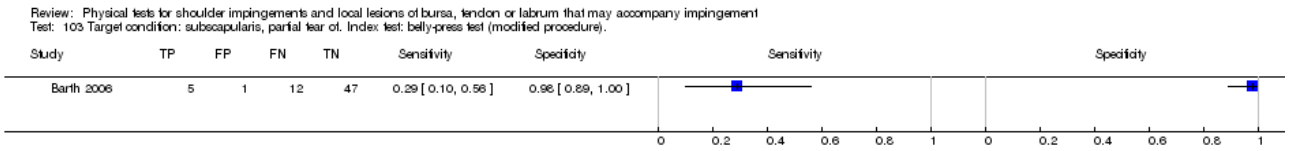
Test 101. Target condition: subscapularis, FTT of. Index test: internal rotation lag sign (modified interpretation)..



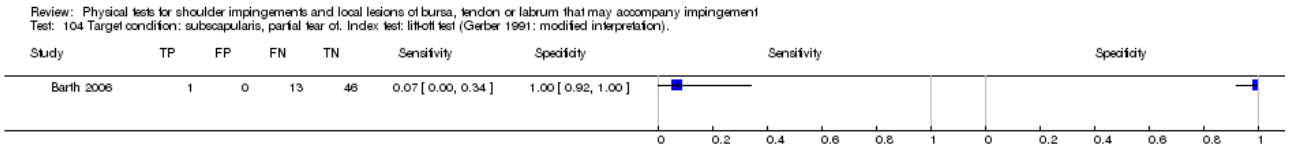
Test 102. Target condition: subscapularis, partial tear of. Index test: bear-hug test (novel)..



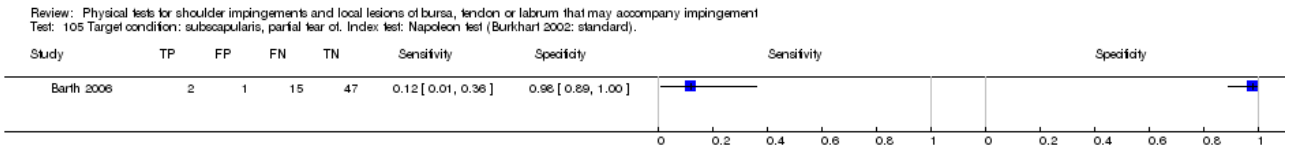
Test 103. Target condition: subscapularis, partial tear of. Index test: belly-press test (modified procedure)..



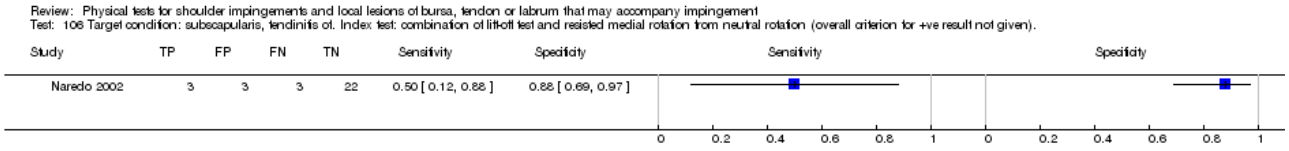
Test 104. Target condition: subscapularis, partial tear of. Index test: lift-off test (Gerber 1991: modified interpretation)..



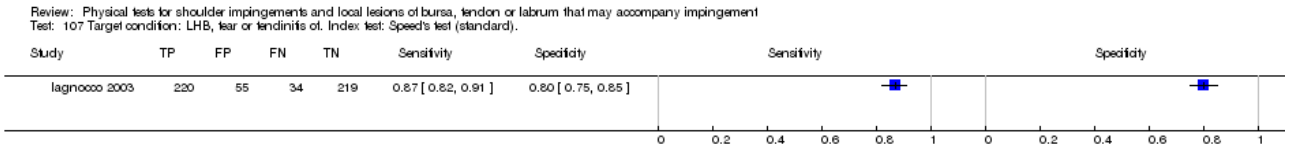
Test 105. Target condition: subscapularis, partial tear of. Index test: Napoleon test (Burkhart 2002: standard)..



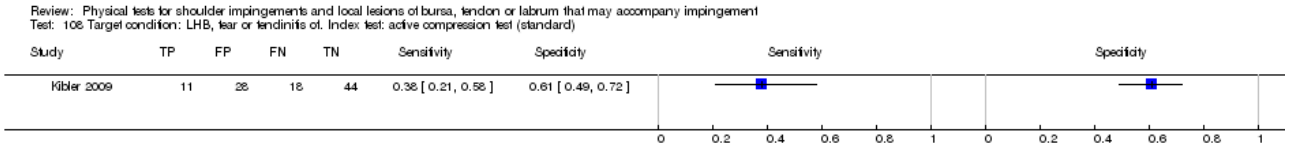
Test 106. Target condition: subscapularis, tendinitis of. Index test: combination of lift-off test and resisted medial rotation from neutral rotation (overall criterion for +ve result not given)..



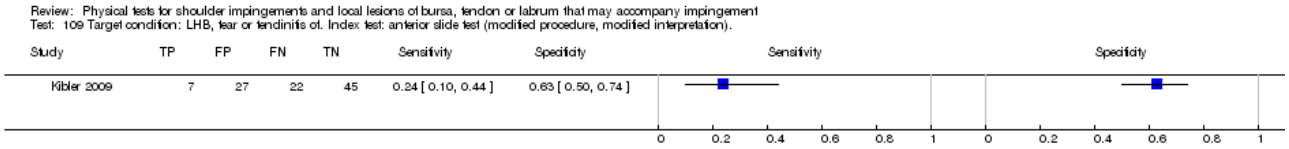
Test 107. Target condition: LHB, tear or tendinitis of. Index test: Speed's test (standard)..



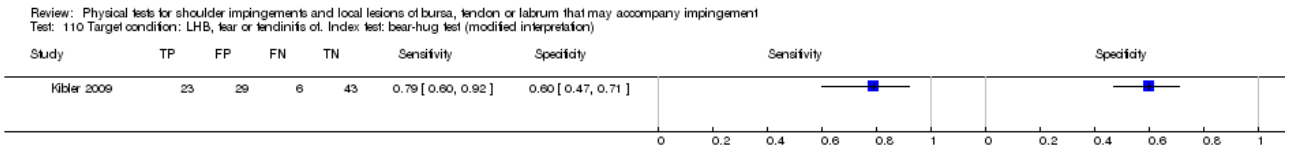
Test 108. Target condition: LHB, tear or tendinitis of. Index test: active compression test (standard).



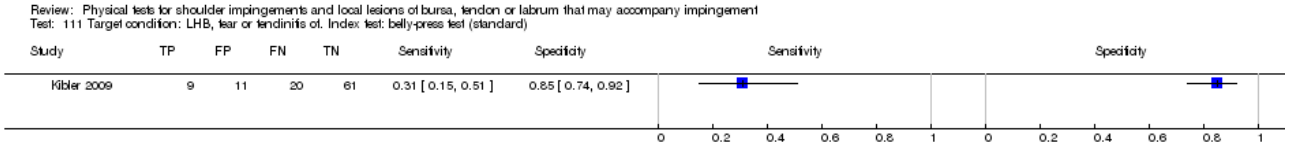
Test 109. Target condition: LHB, tear or tendinitis of. Index test: anterior slide test (modified procedure, modified interpretation)..



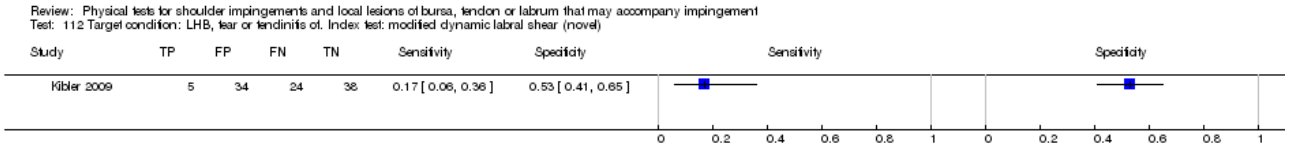
Test 110. Target condition: LHB, tear or tendinitis of. Index test: bear-hug test (modified interpretation).



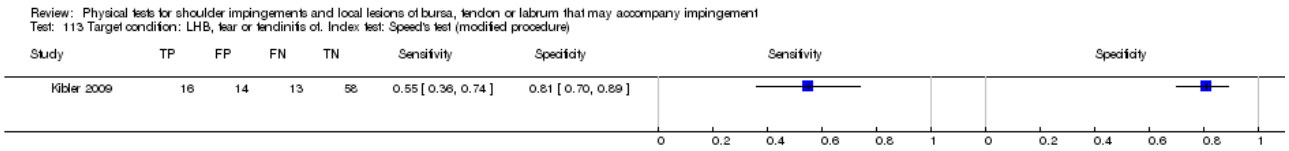
Test 111. Target condition: LHB, tear or tendinitis of. Index test: belly-press test (standard).



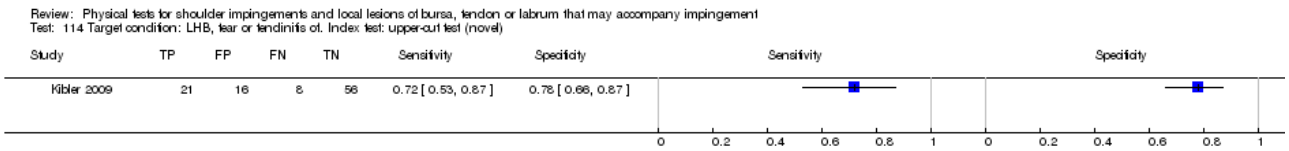
Test 112. Target condition: LHB, tear or tendinitis of. Index test: modified dynamic labral shear (novel).



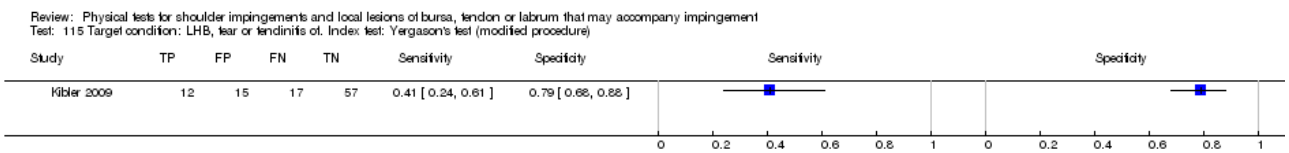
Test 113. Target condition: LHB, tear or tendinitis of. Index test: Speed's test (modified procedure).



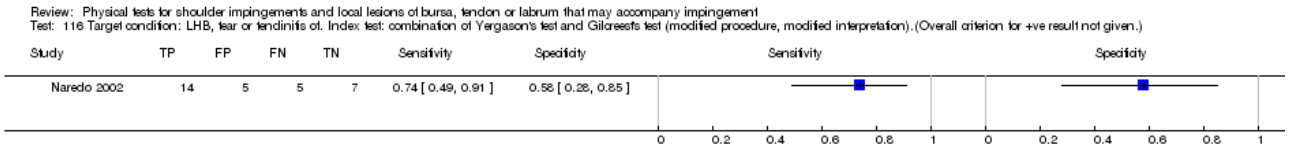
Test 114. Target condition: LHB, tear or tendinitis of. Index test: upper-cut test (novel).



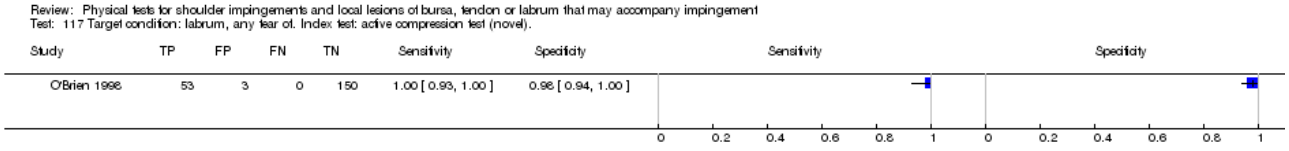
Test 115. Target condition: LHB, tear or tendinitis of. Index test: Yergason's test (modified procedure).



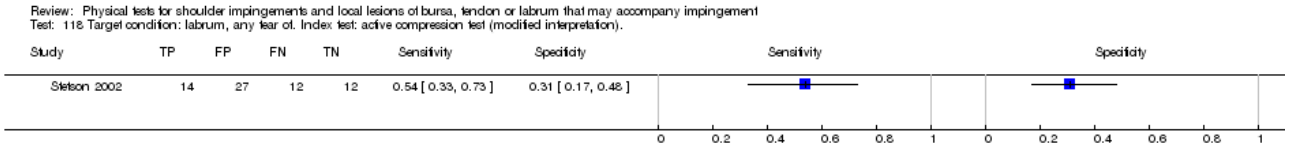
Test 116. Target condition: LHB, tear or tendinitis of. Index test: combination of Yergason's test and Gilcreest's test (modified procedure, modified interpretation).(Overall criterion for +ve result not given.).



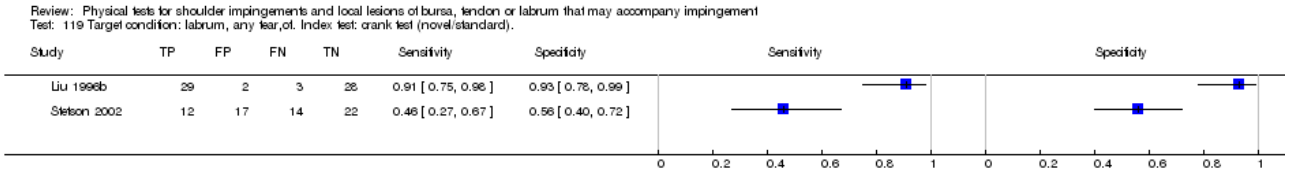
Test 117. Target condition: labrum, any tear of. Index test: active compression test (novel)..



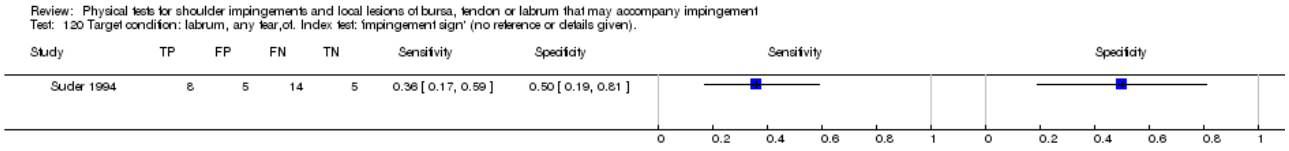
Test 118. Target condition: labrum, any tear of. Index test: active compression test (modified interpretation)..



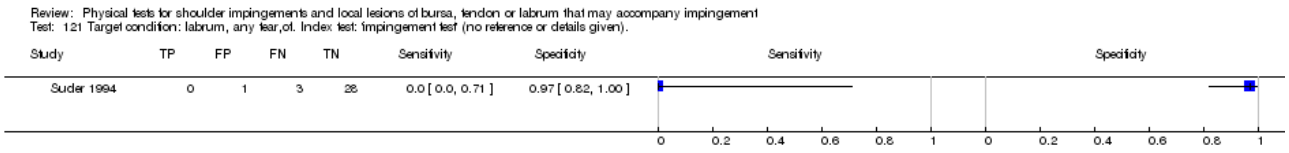
Test 119. Target condition: labrum, any tear,of. Index test: crank test (novel/standard)..



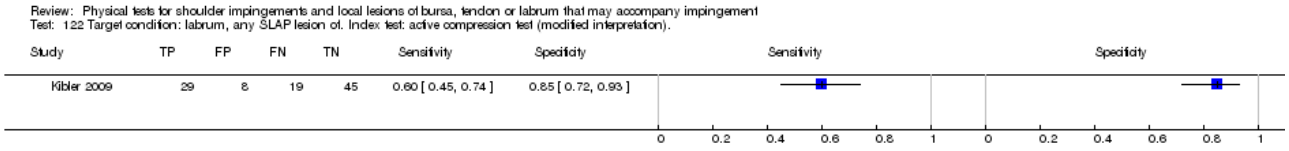
Test 120. Target condition: labrum, any tear,of. Index test: 'impingement sign' (no reference or details given)..



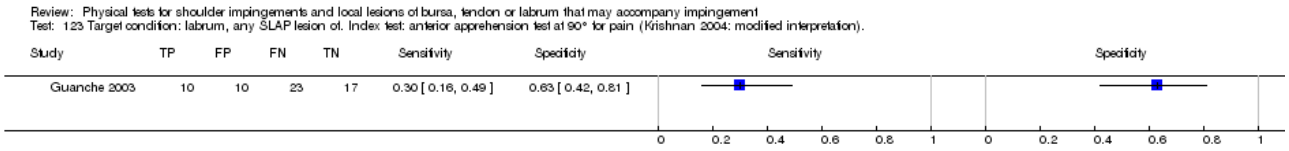
Test 121. Target condition: labrum, any tear,of. Index test: 'impingement test' (no reference or details given)..



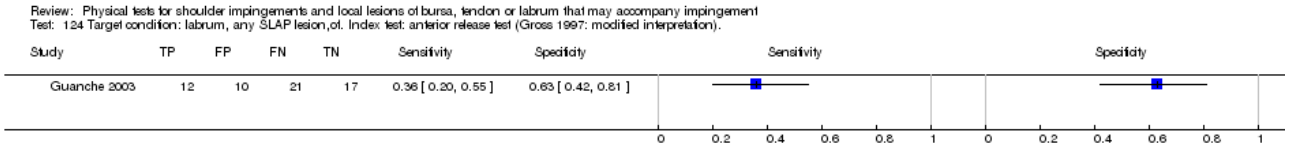
Test 122. Target condition: labrum, any SLAP lesion of. Index test: active compression test (modified interpretation)..



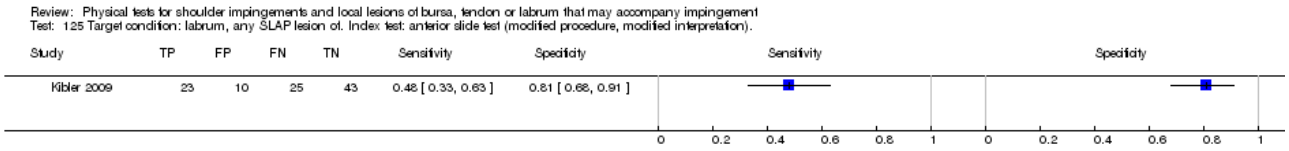
Test 123. Target condition: labrum, any SLAP lesion of. Index test: anterior apprehension test at 90° for pain (Krishnan 2004: modified interpretation)..



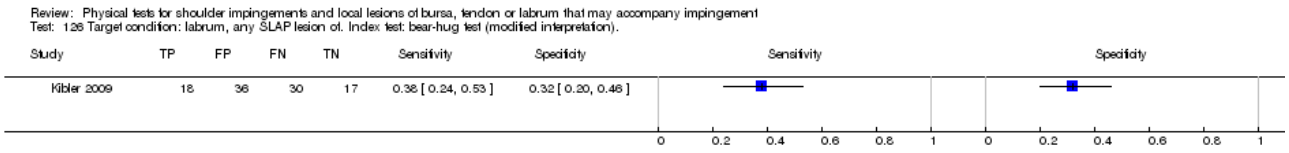
Test 124. Target condition: labrum, any SLAP lesion,of. Index test: anterior release test (Gross 1997: modified interpretation)..



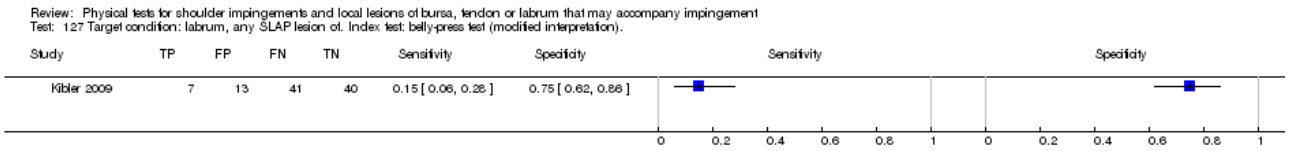
Test 125. Target condition: labrum, any SLAP lesion of. Index test: anterior slide test (modified procedure, modified interpretation)..



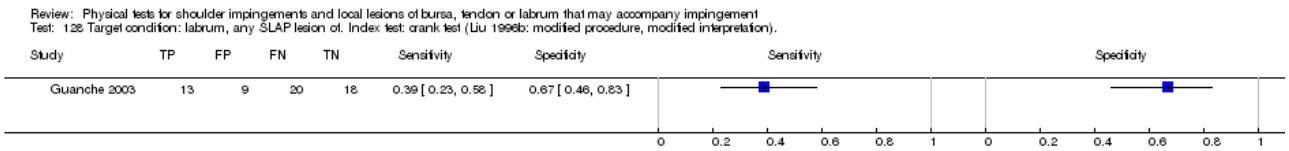
Test 126. Target condition: labrum, any SLAP lesion of. Index test: bear-hug test (modified interpretation)..



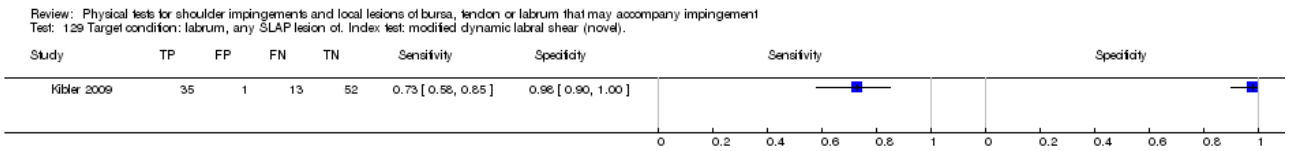
Test 127. Target condition: labrum, any SLAP lesion of. Index test: belly-press test (modified interpretation)..



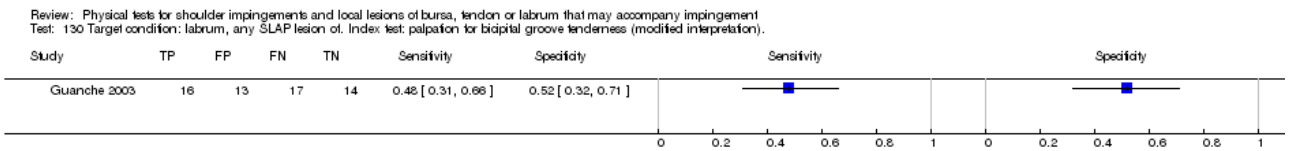
Test 128. Target condition: labrum, any SLAP lesion of. Index test: crank test (Liu 1996b: modified procedure, modified interpretation)..



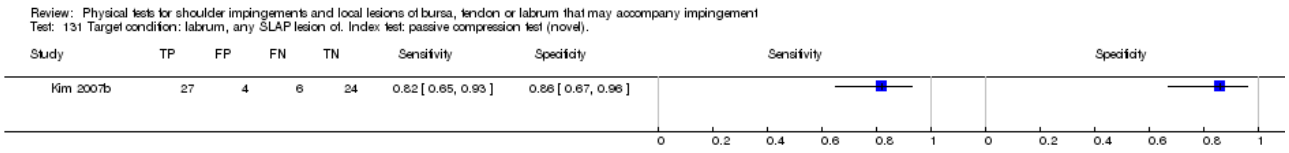
Test 129. Target condition: labrum, any SLAP lesion of. Index test: modified dynamic labral shear (novel)..



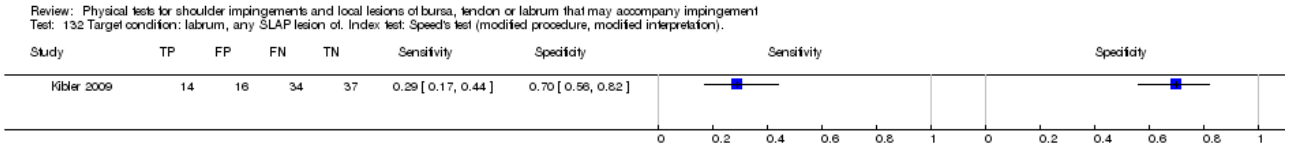
Test 130. Target condition: labrum, any SLAP lesion of. Index test: palpation for bicipital groove tenderness (modified interpretation)..



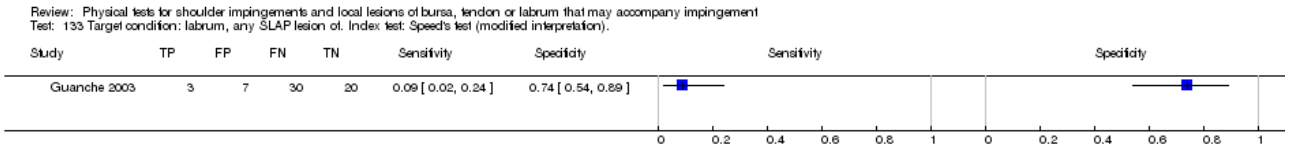
Test 131. Target condition: labrum, any SLAP lesion of. Index test: passive compression test (novel)..



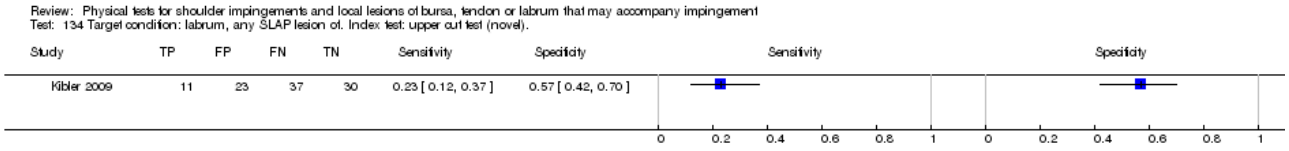
Test 132. Target condition: labrum, any SLAP lesion of. Index test: Speed's test (modified procedure, modified interpretation)..



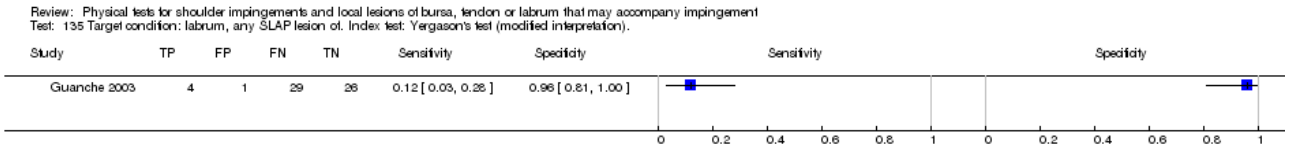
Test 133. Target condition: labrum, any SLAP lesion of. Index test: Speed's test (modified interpretation)..



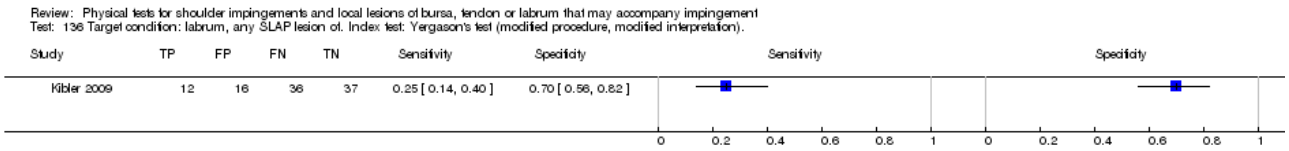
Test 134. Target condition: labrum, any SLAP lesion of. Index test: upper cut test (novel)..



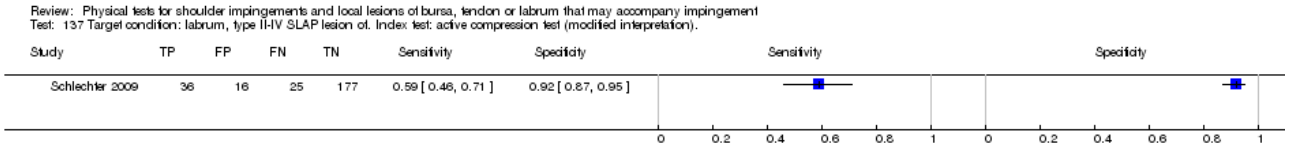
Test 135. Target condition: labrum, any SLAP lesion of. Index test: Yergason's test (modified interpretation)..



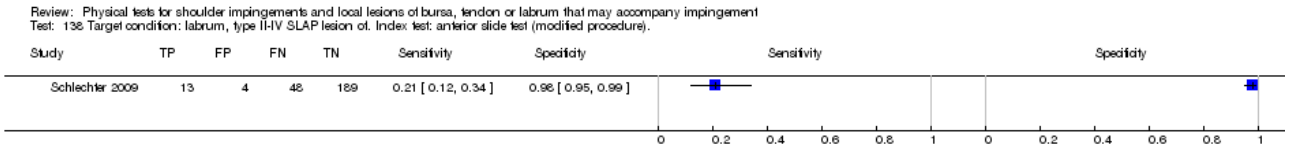
Test 136. Target condition: labrum, any SLAP lesion of. Index test: Yergason's test (modified procedure, modified interpretation)..



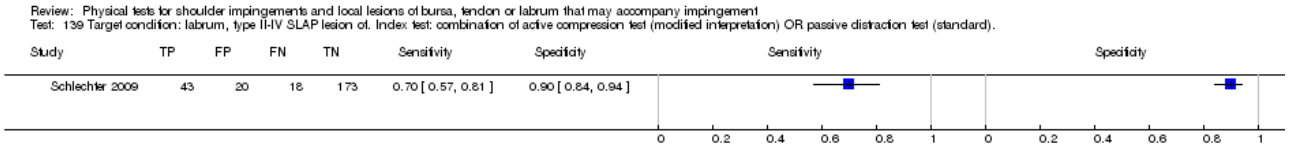
Test 137. Target condition: labrum, type II-IV SLAP lesion of. Index test: active compression test (modified interpretation)..



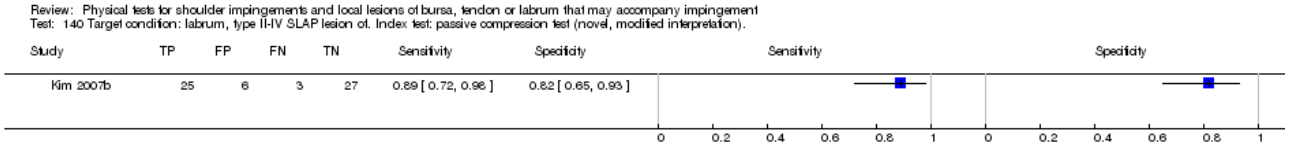
Test 138. Target condition: labrum, type II-IV SLAP lesion of. Index test: anterior slide test (modified procedure)..



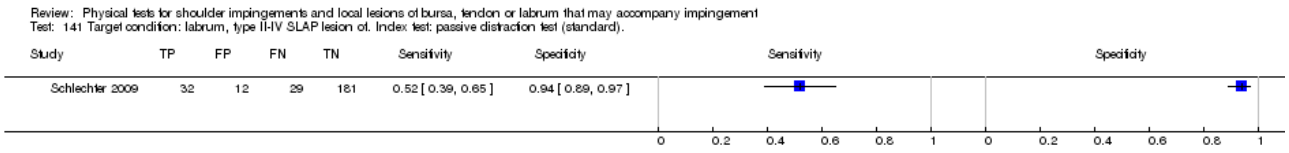
Test 139. Target condition: labrum, type II-IV SLAP lesion of. Index test: combination of active compression test (modified interpretation) OR passive distraction test (standard)..



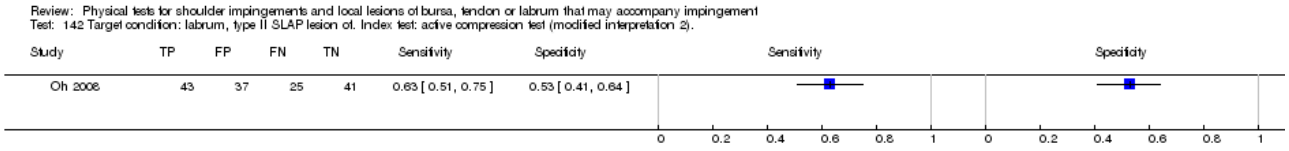
Test 140. Target condition: labrum, type II-IV SLAP lesion of. Index test: passive compression test (novel, modified interpretation)..



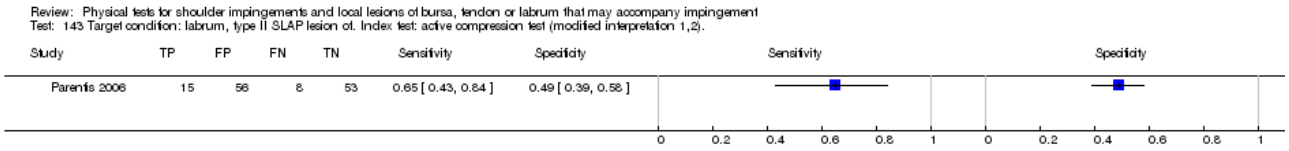
Test 141. Target condition: labrum, type II-IV SLAP lesion of. Index test: passive distraction test (standard)..



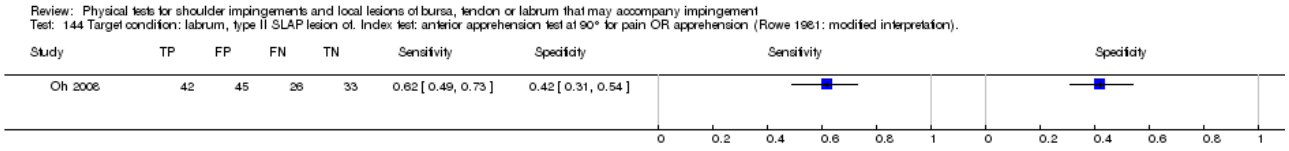
Test 142. Target condition: labrum, type II SLAP lesion of. Index test: active compression test (modified interpretation 2)..



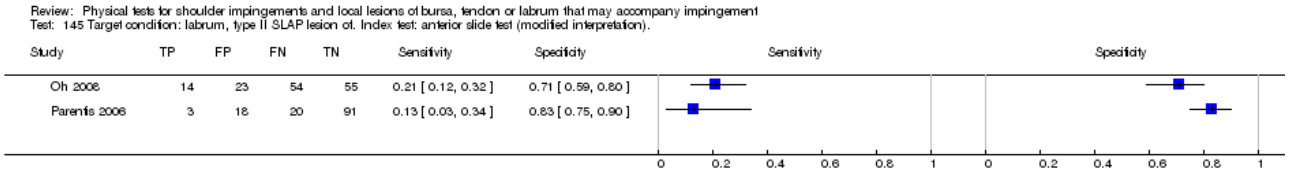
Test 143. Target condition: labrum, type II SLAP lesion of. Index test: active compression test (modified interpretation 1,2)..



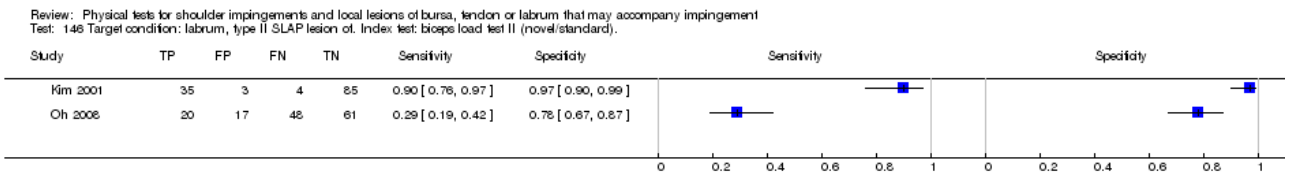
Test 144. Target condition: labrum, type II SLAP lesion of. Index test: anterior apprehension test at 90° for pain OR apprehension (Rowe 1981: modified interpretation)..



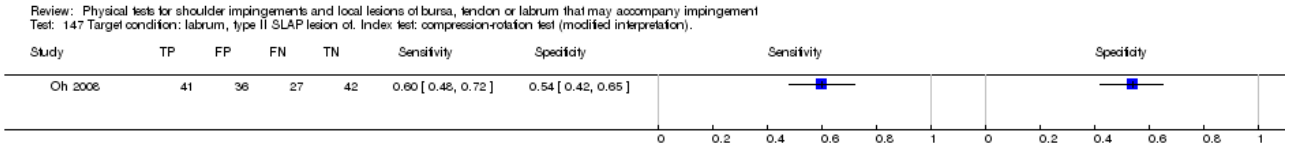
Test 145. Target condition: labrum, type II SLAP lesion of. Index test: anterior slide test (modified interpretation)..



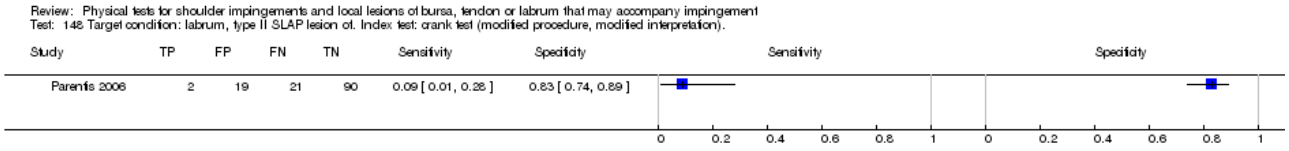
Test 146. Target condition: labrum, type II SLAP lesion of. Index test: biceps load test II (novel/standard)..



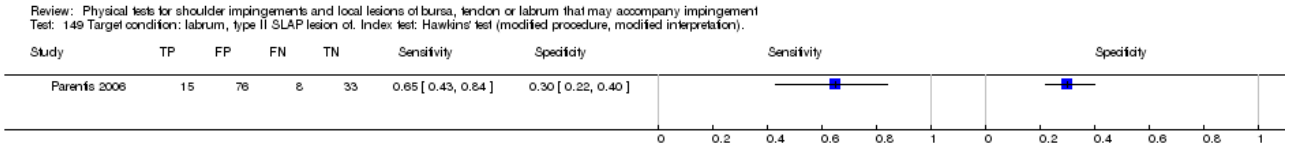
Test 147. Target condition: labrum, type II SLAP lesion of. Index test: compression-rotation test (modified interpretation)..



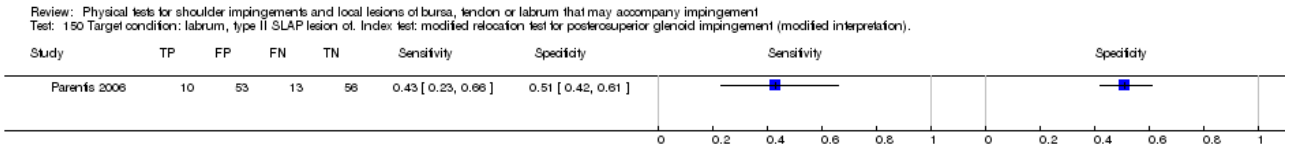
Test 148. Target condition: labrum, type II SLAP lesion of. Index test: crank test (modified procedure, modified interpretation)..



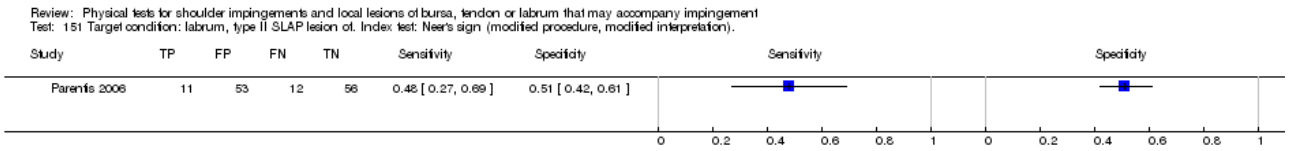
Test 149. Target condition: labrum, type II SLAP lesion of. Index test: Hawkins' test (modified procedure, modified interpretation)..



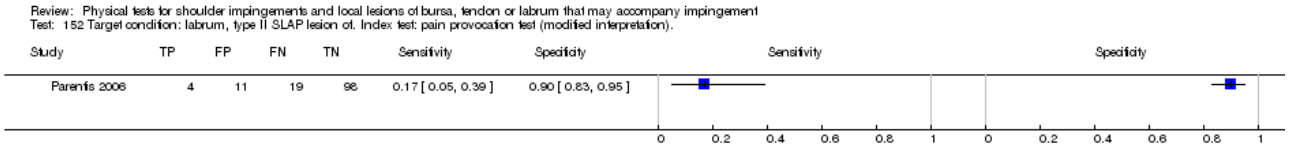
Test 150. Target condition: labrum, type II SLAP lesion of. Index test: modified relocation test for posterosuperior glenoid impingement (modified interpretation)..



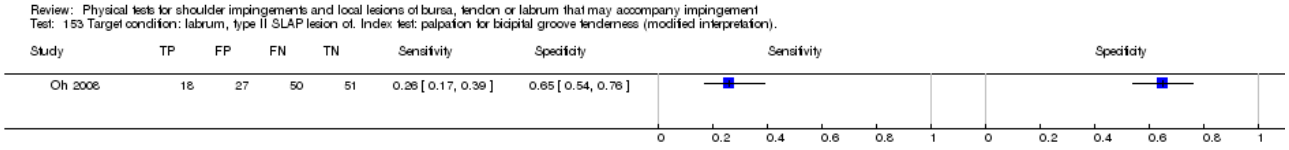
Test 151. Target condition: labrum, type II SLAP lesion of. Index test: Neer's sign (modified procedure, modified interpretation)..



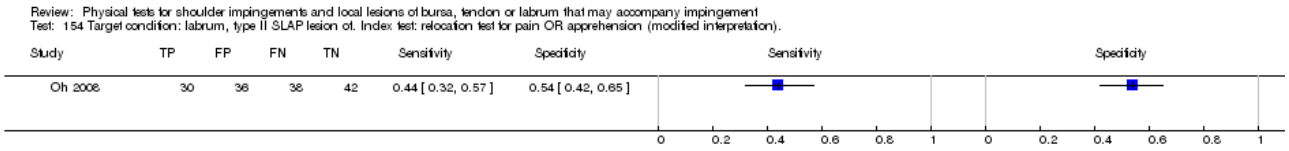
Test 152. Target condition: labrum, type II SLAP lesion of. Index test: pain provocation test (modified interpretation)..



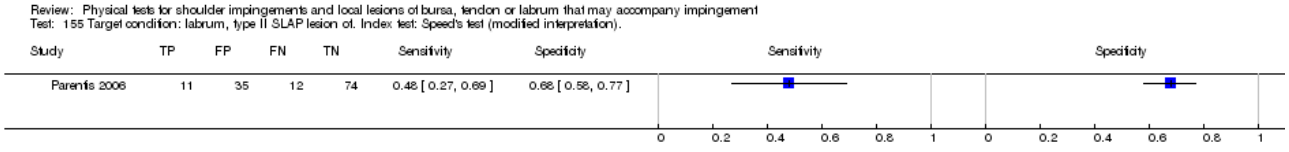
Test 153. Target condition: labrum, type II SLAP lesion of. Index test: palpation for bicipital groove tenderness (modified interpretation)..



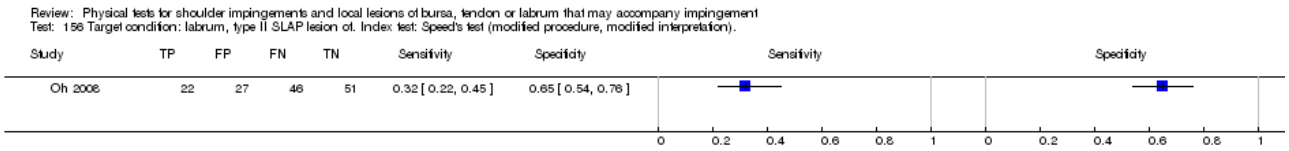
Test 154. Target condition: labrum, type II SLAP lesion of. Index test: relocation test for pain OR apprehension (modified interpretation)..



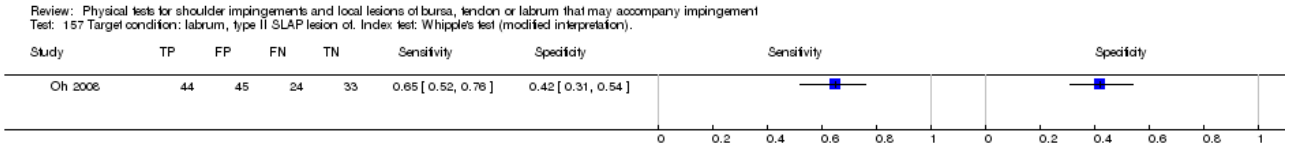
Test 155. Target condition: labrum, type II SLAP lesion of. Index test: Speed's test (modified interpretation)..



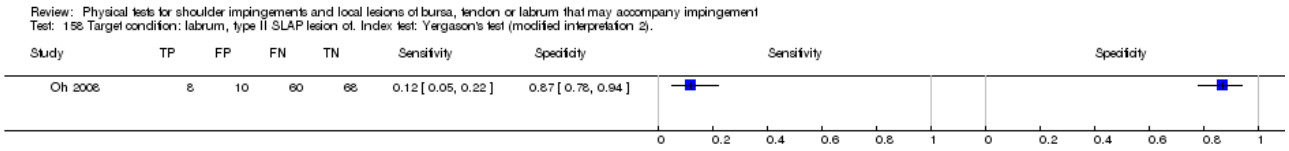
Test 156. Target condition: labrum, type II SLAP lesion of. Index test: Speed's test (modified procedure, modified interpretation)..



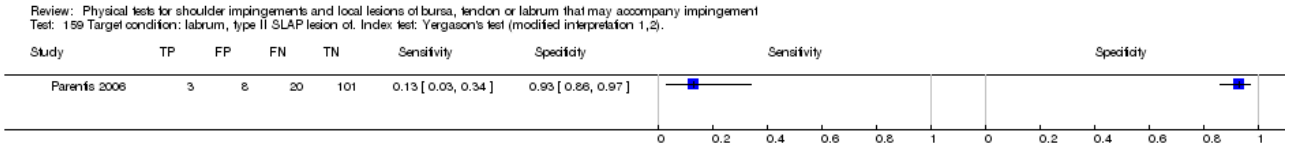
Test 157. Target condition: labrum, type II SLAP lesion of. Index test: Whipple's test (modified interpretation)..



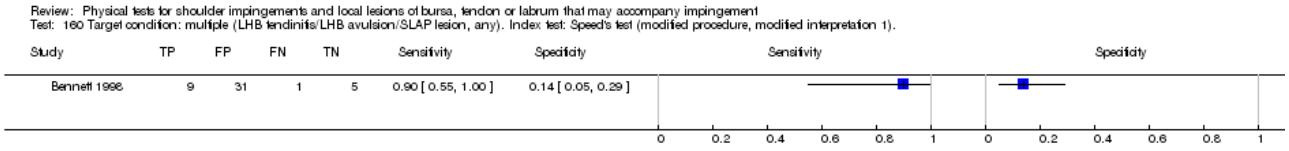
Test 158. Target condition: labrum, type II SLAP lesion of. Index test: Yergason's test (modified interpretation 2)..



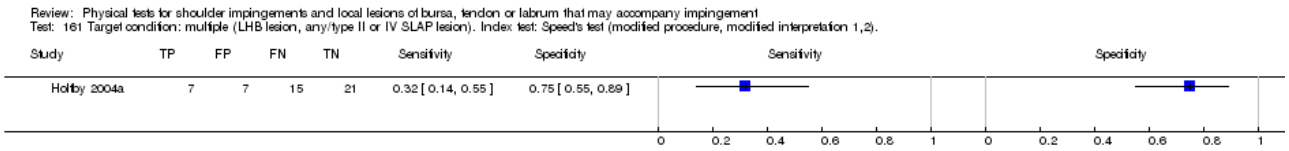
Test 159. Target condition: labrum, type II SLAP lesion of. Index test: Yergason's test (modified interpretation 1,2)..



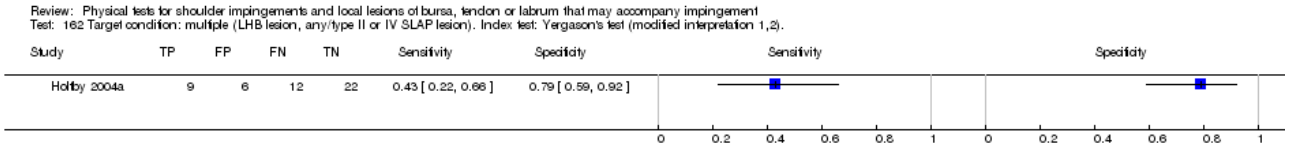
Test 160. Target condition: multiple (LHB tendinitis/LHB avulsion/SLAP lesion, any). Index test: Speed's test (modified procedure, modified interpretation 1)..



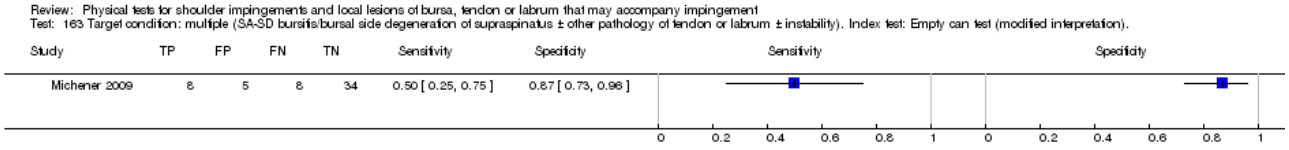
Test 161. Target condition: multiple (LHB lesion, any/type II or IV SLAP lesion). Index test: Speed's test (modified procedure, modified interpretation 1,2)..



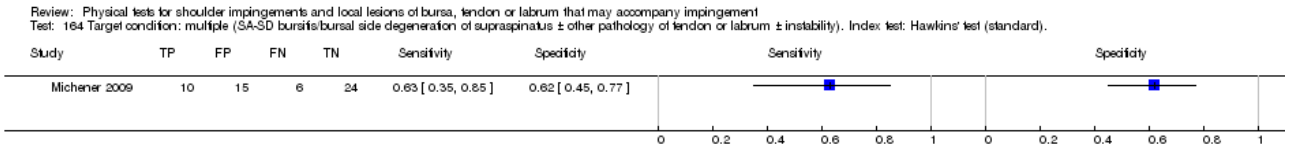
Test 162. Target condition: multiple (LHB lesion, any/type II or IV SLAP lesion). Index test: Yergason's test (modified interpretation 1,2)..



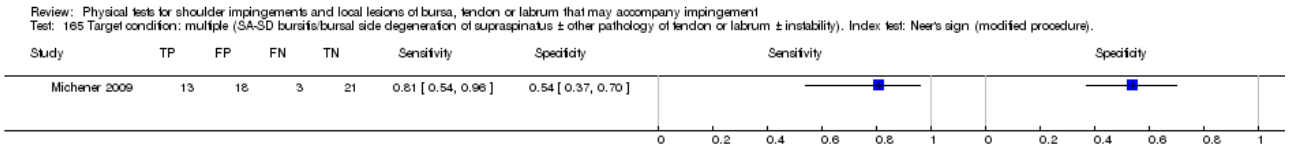
Test 163. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Empty can test (modified interpretation)..



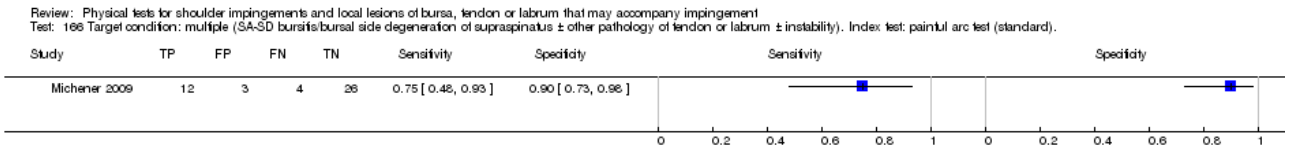
Test 164. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Hawkins' test (standard)..



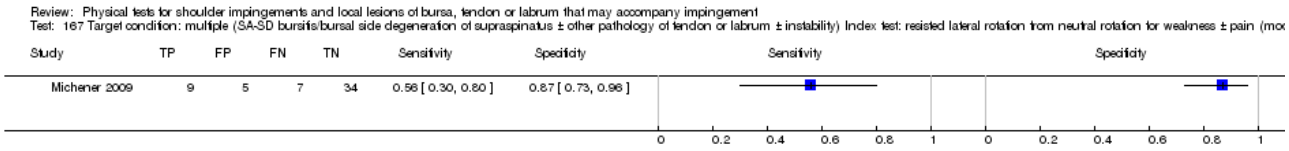
Test 165. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: Neer's sign (modified procedure)..



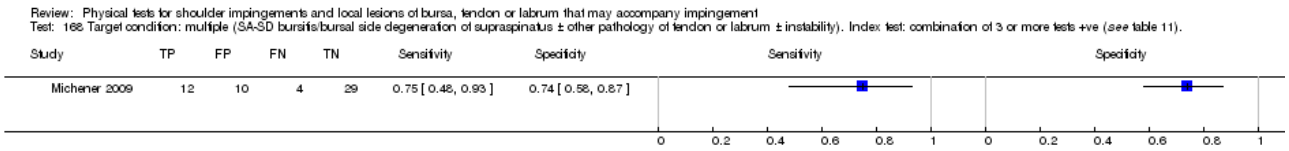
Test 166. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: painful arc test (standard)..



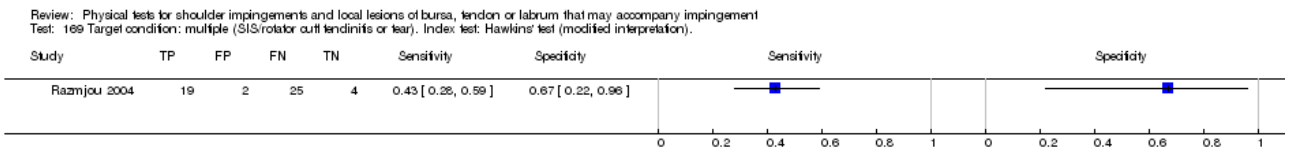
Test 167. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability) Index test: resisted lateral rotation from neutral rotation for weakness ± pain (modified interpretation 1,2)..



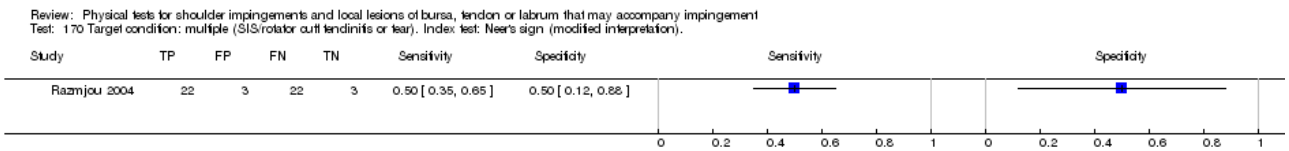
Test 168. Target condition: multiple (SA-SD bursitis/bursal side degeneration of supraspinatus ± other pathology of tendon or labrum ± instability). Index test: combination of 3 or more tests +ve (*see* table 11)..



Test 169. Target condition: multiple (SIS/rotator cuff tendinitis or tear). Index test: Hawkins' test (modified interpretation)..



Test 170. Target condition: multiple (SIS/rotator cuff tendinitis or tear). Index test: Neer's sign (modified interpretation)..



ADDITIONAL TABLES

Table 1. Index tests for impingement and secondary disorders

Tests intended to identify impingement in general					
Test	*Reference*	*Specified prerequisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, according to the author(s)*
Painful arc test	Cyriax 1982	None	The patient actively elevates, then lowers, the shoulder through **abduction**.	Onset and offset of pain during elevation, during lowering, or both.	**Subacromial impingement**; **calcific tendonitis**; pain secondary to shoulder joint instability; or **internal impingement** (involving the deep aspect of the **rotator cuff** or the **LHB** tendon)
Tests intended to identify subacromial impingement					
Test	*Reference*	*Specified prerequisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, according to the author(s)*
Hawkins' test	Hawkins 1980	None	The upright patient's arm is passively positioned in 90° of **flexion** at shoulder and elbow. The tester then forcibly **medially rotates** the patient's shoulder.	Reproduction of the patient's pain	**Subacromial impingement**
Neer's sign	Neer 1977; Neer 1983 (Neer 1972a, sometimes cited, does not give a clear account of this test)	None	The tester forcibly elevates the sitting patient's arm through **scaption**, preventing **scapular** movement by pressing down on the **clavicle** and **acromion** with the other hand.	Pain constitutes a positive Neer's sign.	**Subacromial Impingement** and 'many other shoulder conditions, including stiffness (partial frozen shoulder), instability (e.g. anterior subluxation), arthritis, calcium deposits, and bone lesions'.
Neer's test	Neer 1977; Neer 1983	None	The tester forcibly **flexes** the sitting patient's arm, preventing **scapular** movement by pressing down on the **clavicle** and **acromion** with the other hand (*Neer's sign). The patient is given an injection of 10 ml, 1% xylocaine beneath the **anterior acromion** before the manoeuvre is repeated.	A positive *Neer's sign which is abolished by the injection is termed a positive Neer's test.	**Subacromial impingement**

Table 1. Index tests for impingement and secondary disorders *(Continued)*

'Yocum's (**impingement**) test'	Leroux 1995 and Naredo 2002 cite Yocum 1983: apparently a misconception (*see* under technique).	None	Yocum did not describe a novel impingement test in the article cited (but *see* comment relating to the *empty can test, further in this table). Leroux 1995 and Naredo 2002 may have misinterpreted a photograph depicting Hawkins' test. According to Naredo 2002, the patient places the hand of the affected arm on his or her other shoulder and, keeping the point of the affected shoulder down, raises the elbow of the same limb.	Reproduction of the patient's pain	**Subacromial impingement**
Tests intended to identify internal impingement					
Test	*Reference*	*Specified prerequisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, according to the author(s)*
Anterior apprehension test at 90° for pain	Krishnan 2004	None	The test may be performed with the patient sitting or **supine**. In the latter position the test may be termed the fulcrum test. With the elbow flexed 90, the patient's shoulder is positioned in 90° **abduction** and full **lateral rotation**. (As distinct from the version of this test described by Jobe 1989, no **anterior** pressure is applied to the humeral head (*see* below).	Pain is considered a positive result	**Internal impingement**
Anterior apprehension test at 90° for pain	Jobe 1989	None	The supine patient's shoulder is placed in in 90° **abduction** and full **lateral rotation,** with the elbow **flexed** 90°. Maintaining this position, the tester applies an **anterior** pressure to the **posterior** aspect of the **humeral head**.	Pain but no apprehension. (Note that Rowe 1981 described a test which, apart from the patient being in sitting, was performed comparably to that presented here. However, Rowe's test, which was for **subluxation**, required that both pain *and* apprehension be present for a positive result.)	Pain associated with **anterior subluxation**. Since the original description of this test, this pain has more specifically been ascribed to **posterosuperior glenoid impingement** (Jobe 1995, Jobe 1996).
Anterior release test	Gross 1997	None	The patient lies **supine**, affected shoulder over the edge of the examination couch. The shoulder is passively **abducted** to 90° while the tester applies a **posteriorly** directed force to the **humeral** head. Maintaining this force, the tester brings the arm into full **lateral rotation**. Then	Sudden pain, an increase in pain or reproduction of symptoms [on release]	Primarilrily occult instability; but the authors link this to **posterosuperior glenoid impingement**.

Table 1. Index tests for impingement and secondary disorders *(Continued)*

			the **posteriorly** directed force is released.		
Modified relocation test for **postero-superior glenoid impingement**	Hamner 2000	None	The patient lies**supine**. The shoulder is held by the tester in full **lateral** rotation and positioned at each of 90°, 100° and 120° of **abduction**. In each of these positions the tester applies a force to the patient's upper **humerus**, first directed **anteriorly**, then **posteriorly**.	Pain on the **anteriorly** directed force which is relieved by the **posteriorly** directed force	**Internal impingement**
Posterior impingement test	Meister 2004	None	The **supine** patient's shoulder is placed into 90°-110° degrees of **abduction** and 10°-15° **extension**. Full **lateral rotation** is then added.	Pain felt deeply within the **posterior** aspect of the shoulder joint	**Posterior glenoid impingement** and concomitant tear of the internal surface of the **rotator cuff**, of the **posterior glenoid labrum**, or both.
Relocation test for pain	Jobe 1989	Positive apprehension test	This is an extension of the apprehension test for pain at 90°, which it immediately follows. With the patient's arm still **abducted** and **laterally rotated**, **posterior** pressure is applied to the **humeral head**.	The pain of the apprehension test is relieved. While posterior pressure is maintained, reduced pain may allow greater lateral rotation.	Pain associated with **anterior subluxation**. Since the original description of this test, this pain has more specifically been ascribed to **posterosuperior glenoid impingement** (Jobe 1995, Jobe 1996).
Tests intended to differentiate between subacromial and internal impingement					
Test	*Reference*	*Specified prerequisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, according to the author(s)*
Internal rotation resistance strength test	Zaslav 2001	Positive Neer's sign	The patient and tester stand, the tester to the rear. The patient's elbow is **flexed** to about 90°, and the shoulder positioned at 90° **abduction** and 80°**lateral rotation**. In this position, lateral- and **medial rotation** are manually, **isometrically** resisted.	**Lateral rotation** is strong. **Medial rotation** is weak.	**Internal impingement.** The converse is a 'negative' finding, and signifies **subacromial outlet impingement**
Tests intended to diagnose rotator cuff tears or tendinosis					
Test	*Reference*	*Specified prerequisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, ac-*

Table 1. Index tests for impingement and secondary disorders *(Continued)*

					cording to the author(s)
Bear-hug test	Barth 2006a	None	The patient places the palm of the affected limb, fingers **extended**, on the opposite shoulder. The patient is asked to hold this position, while the tester, by applying a force perpendicular to the forearm, attempts to **laterally rotate** the shoulder.	The patient is unable to hold the hand in contact with the shoulder, or is > 20% weaker than on the unaffected side.	Tear of **subscapularis**
Belly-press test	Gerber 1996	Inadequate range of motion to perform the *lift-off test (see below)	The patient, in a sitting position, presses against the abdomen with the palm of the hand while trying to keep the shoulder in full **medial rotation**.	Full **medial rotation** cannot be maintained. The patient feels weak and the shoulder drops back into **extension**. The patient tries to exert pressure by **extending** the elbow and **flexing** the wrist.	Weakness of **subscapularis**, implying a partial or complete tear
Drop arm test	Codman 1934	None	This test was not clearly described in its primary source. By convention, it is applied in the plane of **abduction**, with the patient's arm placed passively above 90° by the tester; the support is removed, and the patient attempts to lower the arm actively.	The patient is unable to actively lower the arm under control beyond the horizontal, and it drops to his or her side.	Tear of **supraspinatus**
Drop sign	Hertel 1996a	Normal passive range of movement at the shoulder is required: capsular contracture (hypomobility) or ruptured subscapularis (hypermobility) might cause false -ve and false +ve results, respectively. The authors suggest proceeding to this test if the external rotation lag sign is positive.	The patient sits. The tester stands behind the patient, supports the arm with the elbow **flexed** to 90° and the shoulder elevated to 90° in the **plane of the scapula**, then **laterally rotates** the shoulder to just short of full range. The tester continues to support the elbow while releasing the wrist and asking the patient to maintain the lifted-off position.	The patient cannot maintain the position and there is a 'drop' or 'lag', which is recorded to the nearest 5°.	Tear of postero-superior rotator cuff, particularly **infraspinatus**, or **neuropathy**. The authors suggest that the value of the test is in assessing involvement of infraspinatus having established the presence of a poster-superior cuff tear using the external rotation lag sign.
Empty can test (Jobe's test, **supraspinatus** test).	Jobe 1983	None	There are two stages. Preliminarily, the tester evaluates the **deltoid**, with the patient's arm at 90° of **abduction** and **neutral rotation**. To evaluate supraspinatus, the arm is then moved	Pain or weakness on testing supraspinatus	**Supraspinatus impingement** (pain) or tear (weakness)

Table 1. Index tests for impingement and secondary disorders *(Continued)*

Note that Yocum 1983 described the same test (minus the preliminary deltoid component) in the same year, apparently derived from the same studies at the Centinela Hospital Medical Centre Biomechanics Laboratory, California. Thus the empty can test has also been termed 'Yocum's test' (REF and *see* separate entry for *'Yocum's **impingement** test' above). Jobe 1982 is often cited as the source of this test, but the manoeuvre described in that report was a strengthening exercise, not a diagnostic test.			into **medial rotation** (thumb pointing down) and 90° of **scaption**, where the patient is asked to **isometrically resist** a downward pressure applied by the tester.		
External rotation lag sign	Hertel 1996a	Normal passive range of movement at the shoulder is required: capsular contracture (hypomobility) or ruptured subscapularis (hypermobility) might cause false -ve and false +ve results, respectively.	The patient sits. The tester stands behind the patient, supports the arm with the elbow **flexed** to 90° and the shoulder in 20° of elevation (in the **plane of the scapula**), then **laterally rotates** the shoulder to 5° short of full range. The tester asks the patient to maintain the **lateral rotation** and, while continuing to support the elbow, releases the wrist.	An angular 'drop' or 'lag', which is recorded to the nearest 5°	Tear of **supraspinatus** ± **infraspinatus**. A 15° lag or greater signifies a complete tear of both or a **neuropathy**.

Table 1. Index tests for impingement and secondary disorders *(Continued)*

Full can test	Kelly 1996	None	The patient sits, arm **laterally rotated** (thumb pointing up) and in 90° of **scaption**. The patient is then asked to **isometrically** resist a downward pressure applied on the arm by the tester.	Weakness (the test was described in the context of strength assessment, not pain-provocative testing). However, by convention, the test is often interpreted as for the *empty can test.	**Supraspinatus** dysfunction
Gum-turn test	Gumina 2008a	None	Starting in the *empty can test position, the patient traces a 20-cm wide spiral drawn on the wall, from centre to periphery and back 10 times, resting for one minute, then repeating the procedure.	The test is positive if weakness or pain prevent completion. (For positive results, the number of turns completed were recorded, but it is unclear how these data were used. Results were compared with the contralateral arm but, again, it is unclear how these data were used.)	**Postero-superior rotator cuff** tear
Internal rotation lag sign. (Also *see* *lift-off test, Gerber 1991a; and *lift-off test, Gerber 1996.)	Hertel 1996a	Adequate range of **medial rotation**. If this is not available, the belly press test (*see* above) should be used.	The patient sits. The tester, standing to the rear, brings the patient's hand behind the back and **flexes** the elbow to 90°, so that the back of the hand rests on the spine at waist level. Gripping the patient's wrist, the tester then lifts the back of the hand clear of the spine until the shoulder is in almost full **medial rotation**. The tester, who continues to support the elbow but releases the wrist, asking the patient to actively maintain this position.	A lag occurs, the magnitude of which is recorded to the nearest 5°.	'An obvious drop of the hand may occur with large tears. A slight lag indicates a partial tear of the **cranial** part of the **subscapularis** tendon.'
Lift-off test. (Also *see* ***internal rotation** lag sign, Hertel 1996a, and *lift-off test, Gerber 1996.)	Gerber 1991a	Adequate passive range of **medial rotation**. Active **medial rotation** not inhibited by pain.	The arm is brought passively behind the patient's body into **medial rotation**, such that the hand rests against the spine at waist level, palm backwards. The patient attempts to lift the hand off his or her back.	Inability to lift the hand off the back	Tear of **subscapularis**
Lift-off test. (Also *see* ***internal rotation** lag sign, Hertel 1996a, and *lift-off	Gerber 1996	Adequate range of internal rotation. If this is not available, the *belly-press test should be used instead.	The arm is brought passively behind the patient's body into full internal rotation. The hand, palm facing backwards, is at waist level but not in contact with the spine. The patient attempts to maintain this position. (This description differs slightly from that above, despite apparently relat-	(a) The patient cannot maintain the position: the hand drops back to the body and cannot be actively lifted off without elbow extension;	Tear of **subscapularis.** No information is given on differential interpretation of (a) and (b).

Table 1. Index tests for impingement and secondary disorders *(Continued)*

test, Gerber 1991a.)			ing to the same patient sample, but tallies with the internal rotation lag sign.)	or (b) the patient is weak, so that the hand drops back more than 5°, but not all the way to the spine.	
Lift-off test with force	Kelly 1996	Adequate **medial rotation**. If this is not available, the *belly press test should be used instead.	As above, except the patient is asked to maintain the lift-off position against manually applied resistance.	Weakness (the test was described in the context of strength assessment, not pain-provocative testing).	**Subscapularis** dysfunction
Napoleon test	Schwamborn 1999 [German] Burkhart 2002	None	This is a modification of the belly-press test. The patient adopts a Napoleonic pose, palm on abdomen and with the elbow positioned **laterally**.	Burkhart 2002 refined the test's interpretation thus. A negative (normal) result is where the patient can press against the abdomen without wrist **flexion**. A positive result is an inability to press against the abdomen without wrist **flexion** to 90°. Intermediate results may occur.	**Subscapularis** tear (positive result) or partial tear (intermediate result)
Passive horizontal adduction (scarf test)	Cyriax 1982	None	The patient's arm is passively **horizontally adducted** across the chest.	Pain	Lesions of the **ACJ**, but also of the lower part of the tendon of **subscapularis**
Patte's test	Patte 1987 [French], Leroux 1995	None	With the arm supported in 90° of **scaption**, the patient is asked to **laterally rotate** maximally against the tester's **isometric** resistance. The starting position in terms of the degree of rotation was not specified.	There are three possible responses: (A) strong and painless; (B) normal ability to resist despite pain; and (C) inability to resist, with gradual lowering of the forearm. (C) is subcategorised as follows: (1) decreased resistance compared to the other side, allowing the tester to lower the forearm; (2) the patient can perform the test against gravity but	(1) Normal; (2) simple **tendinitis** of **infraspinatus**; (3) ruptured **infraspinatus** tendon. The score 1-3 'has been claimed to increase in parallel with the severity of muscle atrophy and the size of the tear'.

Table 1. Index tests for impingement and secondary disorders *(Continued)*

				is cannot resist the pressure applied by the tester; and (3) the patient cannot perform the test against gravity.	
Rent test (transdeltoid palpation)	Codman 1934	None	The tester draws the upright patient's shoulder into **extension**, palpating **anterior** to the **acromion**.	There is a tender depression (rent) **anterior** to the **acromion** and, just **distal** to this, an eminence.	A rent represents a full thickness tear of **supraspinatus**; the associated eminence is the **greater tuberosity**, and possibly a stump of **supraspinatus**' attachment distal to the tear. If portions of the adjacent **rotator cuff** tendons are torn, the tenderness, eminence and rent may be a little internal or external to the mid-point of the insertion of **supraspinatus**itself.
Resisted **abduction**	Cyriax 1982	None	The patient stands, arm at side, and is asked to **abduct** the arm maximally against the tester's **isometric resistance**, which is applied at the elbow.	Pain or weakness (either or both)	**Supraspinatus lesion**. (1) pain: minor **lesion**; (2) painful weakness: partial tear; (3) painless weakness: complete tear or **neuropathy**.
Resisted **lateral rotation** from **neutral rotation**	Cyriax 1982	None	The patient stands, elbow at side and flexed to 90°, shoulder in **neutral rotation**. He or she is then asked to **laterally rotate** the shoulder maximally against the tester's **isometric resistance**, which is applied at the wrist.	Pain or weakness (either or both)	**Infraspinatus** or (less likely) **teres minor lesion**. (1) pain: minor **lesion**; (2) painful weakness: partial tear; (3) painless weakness: complete tear or **neuropathy**.
Resisted **medial rotation** from **neutral rotation**	Cyriax 1982	None	The patient stands, elbow at side and flexed to 90°, shoulder in **neutral rotation**. He or she is then asked to **medially rotate** the shoulder maximally against the tester's **isometric resistance**, which is applied at the wrist.	Pain or weakness (either or both)	Lesion of **subscapularis** or another **medial rotator**. (1) Pain: minor **lesion**; (2) painful

Table 1. Index tests for impingement and secondary disorders *(Continued)*

					weakness: partial tear; (3) painless weakness: complete tear or **neuropathy**.
Whipple test	Savoie 2001	None	The patient horizontally **adducts** the straight arm, so that the hand, palm down is in front of the unaffected shoulder. In this position the tester applies a downwards force at the wrist, which the patient isometrically resists.	No details of interpretation were given.	Tear of anterior supraspinatus
Tests intended to diagnose LHB tears or tendinosis					
Gilcreest's test (Gilcreest's palm up test)	Gilcreest 1936	None	The patient elevates the arms in full lateral rotation, holding a weight (e.g. 5 lb dumbbells) in each hand. The tester palpates the LHB while the patient, maintaining full lateral rotation, lowers both arms through abduction. Occasionally the vibrations produced by the snap may be visible in the LHB.	When the arms reach an angle of from 110° to 90° degrees, a definite snap may be audible and/or palpable, and a sharp pain is elicited both in the shoulder and in the region of the bicipital groove.	Recurrent dislocation of LHB tendon. Since used in a modified form for LHB tendinitis (Naredo 2002).
Speed's test	Crenshaw 1966	None	The patient **flexes** his or her shoulder against **isotonic resistance** with the elbow extended and the forearm **supinated**.	Pain localised to the bicipital groove	Degenerative changes of the **LHB**, or **synovitis** of its **tendon sheath.** Recently the test has also been applied to the diagnosis of **SLAP lesions** (*see* below).
Upper cut test	Kibler 2009	None	The patient, elbow at the side and flexed to 90°, palm upwards and with the shoulder in neutral rotation, is asked to make a fist. The tester, with a hand placed over the fist, applies **isotonic resistance** as the patient attempts to rapidly bring the hand up towards the chin, in the manner of a boxing upper cut.	Pain or a painful pop over the **anterior** portion of the involved shoulder during the resisted movement is interpreted as a positive result.	**LHB** or **SLAP lesions** (*see* below)
Yergason' test (supination sign)	Yergason 1931	None	The patient's elbow is **flexed** to 90° and the forearm **pronated**. The patient then actively **supinates** against the tester's resistance.	Pain localised to the **bicipital groove**.	Degenerative changes of the **LHB**, or **synovitis** of its **tendon sheath**. Recently, the test has also been applied to the diagnosis

Table 1. Index tests for impingement and secondary disorders *(Continued)*

of **SLAP lesions** (*see* below).

Tests intended to diagnose tears of the glenoid labrum					
Test	*Reference*	*Specified pre-requisites*	*Technique*	*Definition of positive response*	*Specific implication of a positive response, according to the author(s)*
Active compression test	O'Brien 1998a	None	The patient, who is standing, **flexes** his or her shoulder to 90°, then **adducts** 10-15° and **medially rotates** fully. The elbow remains **extended** throughout. The tester stands behind the patient and applies a uniform downward force to the arm. This is repeated in full **lateral rotation**.	Pain on the 1st manoeuvre, reduced or eliminated on the 2nd	**SLAP lesion**
Anterior slide test	Kibler 1995a	None	The patient sits or stands, hands on hips and thumbs pointing **posteriorly**. One of the tester's hands is placed across the top of the shoulder from behind, with the last segment of the index finger extending over the anterior aspect of the **acromion** at the shoulder joint. The tester's other hand is placed behind the elbow, and a forward and slightly **superiorly** directed force is applied to the elbow and upper arm. The patient is asked to push back against this force.	Pain localised to the front of the shoulder under the tester's hand, and/or a pop or click in the same area, or reproduction of the symptoms felt during overhead activity	Unstable **SLAP lesion**
Biceps load II test	Kim 2001	None	The patient lies **supine**. The tester gently grips his or her wrist and elbow, **elevating** the shoulder to 120° and **laterally rotating** it fully. The patient's forearm is **supinated**, and elbow **flexed** to 90°. The patient is now asked to flex his or her elbow against the tester's**isometric resistance**.	Pain provoked by resisted elbow **flexion**.	**SLAP lesion**
Biceps tension test	Snyder 1990a	None	Probably as for *Speed's test, but whether resistance is **isometric** or **isotonic** was not made clear	Not defined	Unstable **SLAP lesion**
Compression-rotation test	Snyder 1990a	None	The patient lies supine, shoulder **abducted** to 90° and elbow **flexed** to 90°. The tester holds the patient's wrist with one hand, while cradling the elbow with the other. The tester then applies a compression force along the line of the **humerus** while rotating the shoulder, in an attempt to trap the torn **labrum**.	Palpable catching & snapping, analogous to that felt during a positive McMurray's test for a torn meniscus at the knee	Unstable **SLAP lesion**

Table 1. Index tests for impingement and secondary disorders *(Continued)*

Crank test	Liu 1996c	None	The patient lies **supine**. The tester, holding the patient's arm and wrist, forward **flexes** the shoulder fully (c.f. the entry below) and, while axially loading the shoulder through the **humerus**, rotates it **medially** and **laterally**.	Clicking, apprehension or both (c.f. the entry below).	Tear of the **glenoid labrum**
Crank test	Liu 1996b	None	The patient sits or lies (the lying variant is stated to be the more sensitive test: c.f. the entry above) with the elbow **flexed** 90° and the shoulder **elevated** 160° in the **plane of the scapula** (c.f. the entry above). The tester compresses the joint along the line of the humerus with one hand, while fully rotating the shoulder in either direction with the other.	Pain, usually during **lateral rotation**, with or without a click; or reproduction of symptoms (usually pain or a sensation of catching: c.f. the entry above).	Tear of the glenoid labrum. Interpretation is confused by the discrepancies with the entry above, but also by the recommendation, here, to conduct the test in sitting as well as in supine, especially since, 'frequently, a positive crank test in the upright position will also be positive in the supine position'. If the supine test is more accurate, the rationale for additionally testing in sitting is unclear
Modified dynamic **labral** shear	Kibler 2009	None	The patient stands. The elbow is **flexed** and the shoulder **elevated** to above 90° of **scaption**, then **externally rotated** to the point of tightness. The shoulder is then guided into maximal **horizontal abduction**. The tester then applies a shear load by maintaining **external rotation** and **horizontal abduction** while lowering the arm to 60° of **scaption**. Reportedly, this differs from the test described by O'Driscoll (no further citation information given) in that the arm is not placed into maximal **horizontal abduction** until it is **elevated** above 120°. (Reportedly, in pilot testing this modification was found to reduce the high number of false positive tests due to pain through the whole motion.)	Reproduction of the pain and/or a painful click or catch along the **posterior** joint line between 120° and 90° of **scaption** is interpreted as a positive result.	**SLAP lesion**
Pain provocation test	Mimori 1999a	None	The sitting patient's shoulder is passively **abducted** to between 90 & 100° & fully **externally rotated**. With the patient's elbow **flexed** to 90°, his or	Pain, greater in the **pronated** position	Unstable **SLAP lesion**

Table 1. Index tests for impingement and secondary disorders *(Continued)*

			her forearm is fully **pronated**, then **supinated**, by the tester.		
Palpation for **bicipital groove** tenderness	Morgan 1998a	None	Deep pressure applied to the **bicipital groove** on the symptomatic and (for comparison) the asymptomatic arm	Pain elicited by deep pressure on the symptomatic arm, compared to no pain on the asymptomatic arm	**SLAP lesion**
Passive compression test	Kim 2007b	None	The patient is in side-lying, affected arm uppermost. The tester places one hand over the **acromion**, using the other to cradle the elbow, which is **flexed** to 90°. The shoulder is abducted to 30° and laterally rotated. The tester then applies a compressive force through the axis of the humerus while drawing the shoulder into extension.	Pain or a painful click	**SLAP lesion**
Passive distraction test	Rubin 2002	None	The patient lies **supine** with the shoulder off the examining table. The arm is **elevated** "in the plane of the trunk" with the elbow **extended**, and the forearm held in neutral or slight **supination**. The forearm is then gently **pronated** without rotating the **humerus**.	Pain. If asked, the patient will frequently indicate with accuracy the anterior or **posterior** location of the **lesion**.	**SLAP lesion**
SLAPprehension test	Berg 1998a	None	The arm of the seated or standing patient is **horizontally adducted** across the chest with the elbow **extended** and the shoulder **medially rotated**. The test is repeated with the shoulder **laterally rotated**.	'SLAPprehension' (meaning unclear), pain which may be referred to the **bicipital groove**, and sometimes an audible or palpable click. Repeating the manoeuvre in **lateral rotation** must be less painful, or the test is negative or indeterminate.	Unstable **SLAP lesion**
Speed's test	Crenshaw 1966	None	The patient **flexes** his or her shoulder against **isotonic resistance** with the elbow extended and the forearm **supinated**.	Pain	Originally developed to diagnose **LHB lesions** (*see* above), the test has recently also been applied to the diagnosis of **SLAP lesions**.
Upper cut test	Kibler 2009	None	The patient, elbow at the side and flexed to 90°, palm upwards and with the shoulder in neutral rotation, is	Pain or a painful pop over the **anterior** portion of the	**SLAP** or**LHB lesions** (*see* above)

Table 1. Index tests for impingement and secondary disorders *(Continued)*

			asked to make a fist. The tester, with a hand placed over the fist, applies **isotonic resistance** as the patient attempts to rapidly bring the hand up towards the chin, in the manner of a boxing upper cut.	involved shoulder during the resisted movement is interpreted as a positive result.	
Yergason' test, Supination sign	Yergason 1931	None	The patient's elbow is **flexed** to 90° and the forearm **pronated**. The patient then actively **supinates** against the tester's resistance.	Pain localised to the **bicipital groove**.	Originally developed to diagnose **biceps lesions** (*see* above), the test has recently also been applied to the diagnosis of **SLAP lesions**.

Table 2. Glossary. Terms marked * have their own entries

Abduction. Sideways movement of a limb away from the body, as in flapping the arms. The opposite of *adduction. The range of abduction is measured from the arm-at-side position (0°).

Adduction. Movement of a limb towards the midline of the body. The opposite of *abduction.

Accuracy. Formally, the proportion of all cases correctly identified by the test. Estimated as (TP+TN)/(TP+FP+FN+TN).

ACJ. See ACROMIOCLAVICULAR JOINT.

Acromioclavicular joint. The joint between the outer end of the *clavicle and the *acromion.

Acromion. A bony process that projects from the *scapula and forms the point of the shoulder. It lies above the shoulder joint.

Anterior. Towards the front. The opposite of *posterior.

Arthrography. A diagnostic technique in which X-rays are taken after injection of a contrast material into a joint.

Biceps. See LONG HEAD OF BICEPS.

Bicipital groove. A groove on the front of the upper *humerus that accommodates the Tendon of the *long head of biceps.

Bursa. A lubricating sac. Bursae are often found where ligaments, muscles, tendons or bones rub together.

Bursal-side. Pertaining to the outer (superficial) aspect of the *rotator cuff: the aspect adjacent to the *subacromial-subdeltoid bursa.

Bursography. A diagnostic technique in which X-rays are taken after injection of a contrast material into a *bursa.

Calcific tendonitis. An inflammation of tendon characterised by deposition of calcium within the tendon's substance. The tendon of *supraspinatus is commonly affected in this way.

Clavicle. The collarbone.

Cranial. Towards the head.

Caudal. Away from the head.

Deltoid. The muscle which gives rise to the rounded contour of the shoulder. Its major function, in concert with *supraspinatus, is to *abduct the shoulder.

Distal. The direction away from the body.

Table 2. Glossary. Terms marked * have their own entries

Elevate. To move upwards. At the shoulder, elevation may be through *flexion, *abduction or in the *plane of the scapula. In each case the range of the movement is measured from the arm-at-side position (0°).

Extend. See EXTENSION.

Extension. In general terms, straightening a joint to lengthen a limb. The opposite of *flexion. At the shoulder, it denotes movement backwards. The range of shoulder extension is measured from the arm-at-side position (0°).

External rotation. See LATERAL ROTATION.

False Negative (FN). The cases which a test incorrectly classifies as not having a disease.

False Positive (FP). The cases which a test incorrectly classifies as having a disease.

Flex. See FLEXION.

Flexion. In general terms, bending a joint to shorten a limb (as in bending the arm up at the elbow).The opposite of *extension. At the shoulder it denotes movement forwards. The range of shoulder flexion is measured from the arm-at-side position (0°).

FN. See FALSE NEGATIVE.

FP. See FALSE POSITIVE.

Glenoid. The socket of the shoulder joint.

Glenoid labrum. A fibrocartilage (gristly) extension of the *glenoid rim that deepens the socket of the shoulder joint.

Gold standard. A reputedly optimal *reference standard.

Greater tuberosity. A protuberance on the upper *humerus to which *supraspinatus attaches.

Horizontal abduction. The movement in which the arm is positioned parallel to the ground and brought backwards. The opposite of *horizontal adduction.

Horizontal adduction. The movement in which the arm is positioned parallel to the ground and brought forwards. The opposite of *horizontal abduction.

Humerus. The upper arm bone.

Humeral head. The rounded upper part of the *humerus, which forms the ball of the shoulder joint.

Impingement. Pinching. This causes ‘catching’ or aching pain without appreciable joint stiffness, and may lead to local inflammation and tissue damage. Subcategories include *internal impingement, *subacromial outlet impingement.

Index test. The test undergoing evaluation against a *reference standard.

Inferior. Relating to the lower portion of a structure. Opposite of *superior.

Inferiorly. Downwards. Opposite of *superiorly.

Infraspinatus. See ROTATOR CUFF.

Internal rotation. See MEDIAL ROTATION.

Internal impingement. Pinching of structures inside the shoulder joint at the extremes of movement. The *glenoid rim, the *glenoid labrum and the deep surface of the *rotator cuff are vulnerable to this type of *impingement, and may be affected singly or in combination.

Isometric resistance. Tester-applied resistance that prevents an attempted movement.

Isotonic resistance. Tester-applied resistance that allows an attempted movement

Joint-side. Pertaining to the inner (deep) aspect of the *rotator cuff: the aspect adjacent to the shoulder joint.

Labrum. See GLENOID LABRUM.

Lateral. Away from the midline of the body. The opposite of *medial.

Table 2. Glossary. Terms marked * have their own entries

***Lateral rotation.** At the shoulder this denotes a twisting movement as in unfolding the arms. The opposite of *medial rotation.

Lesion. An area of tissue damage.

LHB. See LONG HEAD OF BICEPS.

Long head of biceps (LHB). The portion of the biceps that arises inside the shoulder joint. The tendon arches over the *humerus to pass into the arm.

LR-. See NEGATIVE LIKELIHOOD RATIO.

LR+. See POSITIVE LIKELIHOOD RATIO.

Magnetic resonance arthrography (MRA). *MRI following injection of a contrast material into a joint.

Magnetic resonance Imaging (MRI). A non-invasive diagnostic technique. Tissues' differing responses in a strong electromagnetic field are analysed by computer and translated into an accurate anatomical image.

Medial. Towards the midline of the body. The opposite of *lateral.

Medial rotation. At the shoulder, a twisting movement as in folding the arms or bringing the hand behind the back. The opposite of *medial rotation.

MRA. See MAGNETIC RESONANCE ARTHROGRAPHY.

MRI. See MAGNETIC RESONANCE IMAGING.

Negative likelihood ratio (LR-). The ratio between the probability of a negative test result when the disease is present, and the probability of a negative test result when the disease is absent; estimated as (1-Sn)/Sp.

Negative predictive value. The probability that the disease is absent when the test is negative; estimated as TN/(FN+TN).

Neuropathy. A disorder of a nerve that may result in muscle weakness.

Neutral rotation. A position of neither *lateral nor *medial rotation.

Plane of the scapula. A plane of shoulder movement between *flexion/*extension and *abduction/*adduction.

Posterior. Towards the back. The opposite of *anterior.

Positive likelihood ratio (LR+). The ratio between the probability of a positive test result when the disease is present, and the probability of a positive test result when the disease is absent; estimated as Sn/(1-Sp).

Positive predictive value (PPV). The probability that the disease is present when the test is positive; estimated as TP/(TP+FP).

PPV. See POSITIVE PREDICTIVE VALUE.

Pronation. The movement of the forearm that, in relaxed standing, would bring the palm to face backwards.

Prone. Lying face downwards.

Proximal. The direction towards the body.

Reference standard. A highly accurate method of diagnosis. It provides a benchmark against which other methods are judged.

Rheumatoid disease. A systemic disease, one manifestation of which is inflammation of joints.

Rotator cuff. A musculotendinous cuff that surrounds and blends with the shoulder joint, contributing to stability as well as producing movements. It comprises four overlapping units: supraspinatus, which lies on top of the joint and produces *abduction is the most commonly damaged; infraspinatus lies behind the joint, produces *lateral rotation and is the second most commonly damaged; subscapularis lies in front of the joint, produces *medial rotation and is damaged comparatively rarely. The fourth unit, teres minor, lies below *infraspinatus. It is relatively unimportant.

SA-SD *bursa. See SUBACROMIAL-SUBDELTOID BURSA.

**Scaption. ***Elevation of the arm in the *plane of the scapula.

Table 2. Glossary. Terms marked * have their own entries

Scapula. Shoulder blade.

Scapular. Relating to the *scapula.

Sensitivity (Sn). The proportion of cases with the disease that are correctly identified by the *index test i.e. the true positive rate; estimated as TP/(TP+FN).

SIS. See SUBACROMIAL IMPINGEMENT SYNDROME.

SLAP lesion (Superior Labrum Anterior to Posterior *lesion). A tear in the upper part of the *glenoid labrum that extends forwards and backwards (Snyder 1990a; see Footnotes). It may result from *internal impingement.

Sn. See SENSITIVITY.

Sp. See SPECIFICITY.

Specificity (Sp). The proportion of cases without the disease that are correctly identified by the *index test i.e. the true negative rate; estimated as TN/(FP+TN).

Subacromial impingement. Pinching of the *subacromial-subdeltoid bursa, the *rotator cuff, the *long head of biceps, or a combination of these, between the *humerus and the *acromion.

Subacromial impingement syndrome. A collection of signs and symptoms considered characteristic of *subacromial impingement.

Subacromial-subdeltoid *bursa. A palm-sized *bursa centred deep to the anterolateral tip of the *acromion. Extending *distally - under the *deltoid - as well as *proximally, and being superficial to the tendons of the *rotator cuff, it facilitates movement at the shoulder.

Subacromial outlet impingement. See SUBACROMIAL IMPINGEMENT.

Subluxation. A loss of joint congruity lesser in degree than in dislocation.

Subscapularis. See ROTATOR CUFF.

Superior. Relating to the upper portion of a structure. Opposite of *inferior.

Superiorly. Upwards. Opposite of *inferiorly.

Supination. The movement of the forearm that, in relaxed standing, brings the palm to face forwards.

Supine. Lying flat with face upwards.

Supraspinatus. See ROTATOR CUFF.

Synovitis. Inflammation of *synovium.

Synovium. Slippery tissue that lines joints, bursae and the sheaths that surround some tendons, such as the *long head of biceps.

Systemic. Body-wide, as opposed to local.

Tendon Sheath. See SYNOVIUM.

Teres minor. See ROTATOR CUFF.

Tendinitis. Inflammation affecting a tendon.

Tendinosis. Degenerative changes affecting a tendon.

TN. See TRUE NEGATIVE.
TP. See TRUE POSITIVE.

True Negative (TN). The cases which a test correctly identifies as not having a disease.

True Positive (TP). The cases which a test correctly identifies as having a disease.

Ultrasonography. A non-invasive diagnostic technique in which high- frequency sound waves are bounced from the tissues in order to form images of the body's internal structures.

Table 2. Glossary. Terms marked * have their own entries

Xylocaine. A local anaesthetic.

Snyder 1990a
Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ. SLAP lesions of the shoulder. Arthroscopy 1990;6(4):274-9.

Table 3. Reference tests for impingement and secondary disorders

Test	Definition	Adequate reference standard for:	Qualifications
Open surgery	**A diagnostic 'gold' standard.** An invasive procedure during the course of which the interior of the shoulder joint and subacromial-subdeltoid bursa may be directly visualised through an open incision.	(1) Subacromial impingement. (2) Subacromial-subdeltoid bursitis. (3) Bursal side rotator cuff tears. (4) Full thickness rotator cuff tears.	(1) Tears of the rotator cuff's internal substance and joint side may be missed, as may SLAP lesions and disorders of the LHB. (2) Rotator cuff tears may be missed if obscured e.g. by inflammation. (3) Not applicable to primary care.
Arthroscopy	**A diagnostic 'gold' standard.** A 'keyhole' surgical procedure, in which the interior of the shoulder joint and subacromial-subdeltoid bursa may be visualised through a flexible fibre-optic tube.	(1) Subacromial-subdeltoid bursitis. (2) Subacromial impingement. (3) Anterosuperior glenoid impingement. (4) Posterosuperior glenoid impingement. (5) Bursal side rotator cuff tears. (6) Full thickness rotator cuff tears. (7) Joint side rotator cuff tears. (8) Disorders of LHB. (9) SLAP lesions.	(1) There is a technical and interpretive learning curve. (2) Tears of the rotator cuff's internal substance may be missed. (3) Rotator cuff tears may be missed if obscured, e.g. by inflammation. (4) Not applicable to primary care.
Ultrasonography	A non-invasive diagnostic technique in which high-frequency sound waves are bounced (reflected) from the tissues in order to form images of the body's internal structures.	(1) Full thickness rotator cuff tears.	(1) Technique and interpretation are highly operator-dependent. The presence/absence of data/material confirming accuracy in individual diagnostic studies should be taken into account. (2) SLAP lesions cannot be visualised using ultrasound.

Table 3. Reference tests for impingement and secondary disorders *(Continued)*

Magnetic Resonance Imaging (MRI)	A non-invasive diagnostic technique. Tissues' differing responses in a strong electromagnetic field are analysed by computer and translated into an accurate anatomical image.	(1) Full thickness rotator cuff tears.	This applies in settings (such as general primary care) where there is likely to be a low incidence of this disorder.
Arthrography	A diagnostic technique in which X-rays are taken after injection of a fluid contrast material into a joint.	(1) Joint side rotator cuff tears. (2) Full thickness rotator cuff tears.	
Magnetic Resonance Arthrography (MRA)	A combination of Magnetic Resonance Imaging (MRI) and arthrography. An MRI scan is done after injection of contrast material into a joint.	(1) Joint side rotator cuff tears. (2) Full thickness rotator cuff tears. (3) SLAP lesions.	
Bursography	A diagnostic technique in which X-rays are taken after injection of a contrast material into a bursa.	(1) Bursal side rotator cuff tears.	
Local anaesthesia	A minimally invasive procedure in which a local anaesthetic is injected, usually into the subacromial space (this is the second part of Neer's impingement test) and the effect on signs and/or symptoms noted.	(1) Subacromial outlet impingement.	(1) Correct interpretation is dependent on the injection's accuracy. 'Guided' injection, using fluoroscopy or ultrasound, is therefore preferable to 'blind' injection technique.

Table 4. Summary of target conditions, studies, and patients/shoulders

Target condition	Studies	Shoulders/patients
Subacromial or internal impingement	5	471/466
Rotator cuff tendinopathy or tears	18	2477/2337
LHB tendinopathy or tears	3	660/557
Glenoid labral lesions	11	1245/1236
Multiple undifferentiated target conditions*	4	201/200

*LHB/labral pathology; LHB/SLAP lesions; SA-SD bursitis/bursal-side degeneration of supraspinatus; and SIS/rotator cuff tendinitis or tear.

Table 5. Summary: studies of tests for subacromial and internal impingement

Study ID	Shoulders (patients, if different)	Specific target condition	Index test name, provenance (where clarification is required) and manner of use compared to original description (standard/ modified procedure/modified interpretation)	Discrepancies between reported and back-calculated summary statistics (Sn, Sp, PPV, NPV or accuracy)	
				Yes	No

Table 5. Summary: studies of tests for subacromial and internal impingement *(Continued)*

Subacromial impingement					
Calis 2000	125 (120)	SIS	• Combination: ALL 7 +ve • Drop arm test (modified interpretation 2) • Hawkins' test (standard) • Neer's sign (standard) • Painful arc test (standard) • Passive horizontal adduction (modified interpretation) • Speed's test (modified interpretation 2) • Yergason's test (modified interpretation 2)	D	
Gumina 2008	120	SIS	• Gum-Turn test (novel)	E	
MacDonald 2000	85	SA-SD bursitis	• Hawkins' test (standard) • Neer's sign (modified procedure) • Hawkins' test OR Neer's sign (modified procedure) • Hawkins' test AND Neer's sign (modified procedure)		No
Naredo 2002	31	SA-SD bursitis Subacromial impingement in real time (dynamic ultrasonography)	• Combination: Hawkins' test, Neer's sign, 'Yocum's (impingement) test' (overall criterion for +ve result not stated)		No
Differentiating subacromial from internal impingement					
Zaslav 2001	110	Subacromial *versus* internal impingement	• Internal rotation resistance strength test (novel)		No
Internal impingement					
None	None				

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source
A: Isolated absolute discrepancy of 1% to <5% - a suspected or confirmed typographical error
B: Isolated absolute discrepancy of 5% to <10% - a suspected or confirmed typographical error
C: Isolated discrepancy of 10% or more - a suspected or confirmed typographical error
D: Multiple absolute discrepancies of which the greatest is 1% to <5%
E: Multiple absolute discrepancies of which the greatest is 5% to <10%
F: Multiple absolute discrepancies of which the greatest is 10% or more
?: 2 X 2 table not reported and cannot be deduced with certainty.

NR: Summary statistics not reported

Table 6. Summary: studies of tests for rotator cuff tears or tendinopathy

Study ID	Shoulders (patients, if different)	Specific target condition	Index test name, provenance (where clarification is required) and manner of use compared to original description (standard/ modified procedure/modified interpretation)	Discrepancies between reported and back-calculated summary statistics (Sn, Sp, PPV, NPV or accuracy)	
				Yes	No
Barth 2006	68	Subscapularis, any tear of Subscapularis, complete tear of Subscapularis, partial tear of	• Bear-hug test (novel) • Belly-press test (modified procedure) • Lift-off test (Gerber 1991a: modified interpretation 1) • Napoleon test (Burkhart 2002: standard)		No
Castoldi 2009	395 (390)	Supraspinatus, FTT of, full-width Supraspinatus, PTT of, isolated	• External rotation lag sign (standard)		No
Calis 2000	125 (120)	Supraspinatus, FTT of	• Drop arm test (standard) • Hawkins' test (modified interpretation 2) • Neer's sign (modified interpretation 2) • Painful arc test (modified interpretation 2) • Passive horizontal adduction (modified interpretation 2) • Speed's test (modified interpretation 2) • Yergason's test (modified interpretation 2)	F	
Frost 1999	73	Supraspinatus, FTT, degeneration or tendinitis of Supraspinatus, FTT or degeneration of Supraspinatus, FTT of	• Hawkins' test (modified procedure, modified interpretation 2)	NR	NR
Gumina 2008	120	Rotator cuff, postero-superior, FTT of	• Gum-Turn test (novel)	E	

Table 6. Summary: studies of tests for rotator cuff tears or tendinopathy *(Continued)*

		Rotator cuff, postero-superior, supraspinatus AND infraspinatus, FTT of			
		Supraspinatus, FTT of			
Hertel 1996	100	Rotator cuff, postero-superior FTT or PTT of	• Drop sign (novel) • Empty can test for weakness ± pain (modified interpretation 2) • External rotation lag sign (novel)	C	
		Subscapularis, any tear of	• Internal rotation lag sign (novel) • Lift-off test (Gerber 1991a: probably standard)		
Holtby 2004b	50	Supraspinatus, PTT or tendinitis of	• Empty can test for pain WITHOUT weakness (standard)		No
		Supraspinatus, FTT of	• Empty can test for weakness ± pain (standard)		
		Rotator cuff, large or massive FTT of	• Empty can test for weakness ± pain (modified interpretation 2)		
Iagnocco 2003	528 (425)	Supraspinatus, any disease of, including calcification	• Empty can test (no reference or details given)	NR	NR
		Infraspinatus, any disease of, including calcification	• Resisted lateral rotation from neutral rotation (no reference or details given)		
		Subscapularis, any disease of, including calcification	• Resisted medial rotation from neutral rotation (no reference or details given)		
Itoi 1999	143 (136)	Supraspinatus, FTT of	• Empty can test for pain ± weakness (modified interpretation 1) • Empty can test for pain AND/OR weakness (modified interpretation 1) • Empy can test for weakness ± pain (standard) • Full can test for pain ± weakness (modified interpretation 1)		No

Table 6. Summary: studies of tests for rotator cuff tears or tendinopathy *(Continued)*

			• Full can test for pain AND/OR weakness (modified interpretation 1) • Full can test for weakness ± pain (standard)	
Itoi 2006	160 (149)	Supraspinatus FTT or PTT of	• Empty can test for pain ± weakness (modified interpretation 1) • Empty can test for weakness ± pain (standard) • Empty can test for weakness < grade 3 ± pain (modified interpretation 1) • Full can test for pain ± weakness (modified interpretation 1) • Full can test for weakness ± pain (standard) • Full can test for weakness < grade 3 ± pain (modified interpretation 1)	B
		Infraspinatus, FTT or PTT of	• Resisted external rotation from neutral rotation for weakness < grade 3 (modified interpretation 1)	
		Subscapularis, any tear of	• Lift-off test with force for weakness < grade 2 ± pain (Gerber 1991a: modified procedure; modified interpretation 1)	
Kim 2006	200	Rotator cuff, FTT or PTT of	• Empty can test for pain ± weakness (modified interpretation 1,2)	B
		Rotator cuff, FTT of	• Empty can test for pain OR weakness (ONE ONLY) (modified interpretation 1,2) • Empty can test for pain AND weakness (BOTH) (modified interpretation 1,2) • Empty can test for weakness ± pain (modified interpretation 2) • Full can test for pain ± weakness (modified interpretation 1,2) • Full can test for pain OR weakness (ONE ONLY) (modified interpretation 1,2) • Full can test for pain AND weakness (BOTH) (modified interpretation 1,2) • Full can test for weakness ± pain (modified interpretation 2)	
MacDonald 2000	85	Rotator cuff, FTT or PTT of	• Combination: Hawkins' test (modified interpretation 2) OR Neer's sign (modified procedure, modified interpretation 2) +ve • Combination: Hawkins' test (modified interpretation 2) AND Neer's sign (modi-	C

Table 6. Summary: studies of tests for rotator cuff tears or tendinopathy *(Continued)*

			fied procedure, modified interpretation 2) +ve • Hawkins' test (modified interpretation 2) • Neer's sign (modified procedure, modified interpretation 2)		
Miller 2008b	46 (37)	Rotator cuff, postero-superior, FTT of	• Drop sign (modified interpretation 2) • External rotation lag sign (modified interpretation 2)	A	
		Subscapularis, FTT of	• Internal rotation lag sign (modified interpretation 2)		
Naredo 2002	31	Supraspinatus, FTT, PTT or tendinitis of	• Empty can test for pain AND/OR weakness (standard)	E	
		Infraspinatus, FTT, PTT or tendinitis of	• Patte's test for pain AND/OR weakness (Leroux 1995: standard)		
		Subscapularis, any tear or tendinitis of	• Combination: lift-off test (Gerber 1991a cited, but Gerber 1996/Hertel 1996a described); resisted medial rotation from neutral rotation. Overall criterion for +ve result not given.		
		Supraspinatus, FTT or PTT of	• Empty can test for weakness ± pain (standard)		
		Infraspinatus, FTT or PTT of	• Patte's test for weakness ± pain (Leroux 1995: standard)		
		Subscapularis, any tear of	• Combination: lift-off test (Gerber 1991a cited, but Gerber 1996/Hertel 1996a described), resisted medial rotation from neutral rotation. Overall criteria for +ve result not given.		
		Supraspinatus, tendinitis of	• Empty can test for pain WITHOUT weakness (standard)		
		Infraspinatus, tendinitis of	• Patte's test for pain WITHOUT weakness (Leroux 1995: standard)		
		Subscapularis, tendinitis of	• Combination: lift-off test (Gerber 1991a cited, but Gerber 1996/Hertel 1996a described, modified interpretation 2), resisted medial rotation from neutral rotation. Overall criterion for +ve result not given.		
Norwood 1989	103	Rotator cuff, FTT of, multiple- *versus* single-tendon	• Active abduction to < 90° (novel)	NR	NR

Table 6. Summary: studies of tests for rotator cuff tears or tendinopathy *(Continued)*

Speer 1994	100	Rotator cuff, any disease of	• Relocation test for pain (Jobe 1989: modified procedure) • Relocation test for pain (Jobe 1989: standard)		No
Suder 1994	31	Rotator cuff, FTT or PTT of	• 'Impingement sign' (no reference or details given) • 'Impingement test' (no reference or details given)	NR	NR
		Rotator cuff, FTT of	• 'Impingement sign' (no reference or details given) • 'Impingement test' (no reference or details given)		
		Rotator cuff, PTT of	• 'Impingement sign' (no reference or details given) • 'Impingement test' (no reference or details given)		
Wolf 2001	119	Rotator cuff, FTT of	• Rent test (standard)		No

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source
A: Isolated absolute discrepancy of 1% to <5% - a suspected or confirmed typographical error
B: Isolated absolute discrepancy of 5% to <10% - a suspected or confirmed typographical error
C: Isolated discrepancy of 10% or more - a suspected or confirmed typographical error
D: Multiple absolute discrepancies of which the greatest is 1% to <5%
E: Multiple absolute discrepancies of which the greatest is 5% to <10%
F: Multiple absolute discrepancies of which the greatest is 10% or more
?: 2 X 2 table not reported and cannot be deduced with certainty.
NR: Summary statistics not reported

Table 7. Summary: studies of tests for LHB tears or tendinopathy

Study ID	Shoulders (patients, if different)	Specific target condition	Index test name, provenance (where clarification is required) and manner of use compared to original description (standard/ modified procedure/modified interpretation)	Discrepancies between reported and back-calculated summary statistics (Sn, Sp, PPV, NPV or accuracy)	
				Yes	No
Iagnocco 2003	528 (425)	LHB, any lesion of	• Speed's test (standard)	NR	NR
Kibler 2009	101	LHB, any lesion of	• Active compression test (modified interpretation 2) • Anterior slide test (modified procedure, modified interpretation 1,2) • Bear-hug test (modified interpretation 1,2)	E	

Table 7. Summary: studies of tests for LHB tears or tendinopathy *(Continued)*

			• Belly-press test (modified interpretation 2) • Modified dynamic labral shear (novel) • Speed's test (modified procedure) • Upper cut test (novel) • Yergason's test (modified procedure)	
Naredo 2002	31	LHB, any lesion of	• Combination: Yergason's test (standard), Gilcreest's palm up test (modified procedure, modified interpretation 1,2). Criteria for +ve result not given.	No

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source
A: Isolated absolute discrepancy of 1% to <5% - a suspected or confirmed typographical error
B: Isolated absolute discrepancy of 5% to <10% - a suspected or confirmed typographical error
C: Isolated discrepancy of 10% or more - a suspected or confirmed typographical error
D: Multiple absolute discrepancies of which the greatest is 1% to <5%
E: Multiple absolute discrepancies of which the greatest is 5% to <10%
F: Multiple absolute discrepancies of which the greatest is 10% or more
?: 2 X 2 table not reported and cannot be deduced with certainty.
NR: Summary statistics not reported

Table 8. Summary: studies of tests for labral lesions

Study ID	Shoulders (patients, if different)	Specific target condition	Index test name, provenance (where clarification is required) and manner of use compared to original description (standard/ modified procedure/modified interpretation)	Discrepancies between reported and back-calculated summary statistics (Sn, Sp, PPV, NPV or accuracy)	
				Yes	No
Guanche 2003	60 (59)	Labrum, any SLAP lesion of	• Active compression test (modified procedure, modified interpretation 2: 2 x 2 table not calculable for this test) • Anterior apprehension test at 90° for pain (Krishnan 2004: modified interpretation 2) • Anterior release test described as in Gross 1997 with modified interpretation 2, but erroneously labelled as Jobe's relocation test. • Crank test (Liu 1996b: modified procedure, modified interpretation 2) • Palpation for bicipital groove tenderness (standard) • Speed's test (modified interpretation 1,2) • Yergason's test (modified interpretation 1,2)		No
Kibler 2009	101	Labrum, any SLAP lesion of	• Active compression test (modified interpretation 2)	E	

Table 8. Summary: studies of tests for labral lesions *(Continued)*

			• Anterior slide test (modified procedure, modified interpretation 1, 2) • Bear-hug test (modified interpretation 1,2) • Belly-press test (modified interpretation 2) • Modified dynamic labral shear (novel) • Speed's test (modified procedure, modified interpretation 2) • Upper cut test (novel) • Yergason's test (modified procedure, modified interpretation 2)		
Kim 2001	127	Labrum, type II SLAP lesion of	• Biceps load II test (novel)		No
Kim 2007b	61	Labrum, any SLAP lesion of	• Passive compression test (novel)		No
		Labrum, type II-IV SLAP lesion of	• Passive compression test (novel, modified interpretation 2)		
Liu 1996b	62	Labrum, any tear of	• Crank test (novel)		No
O'Brien 1998	206	Labrum any tear of	• Active compression test (novel)		No
Oh 2008	146	Labrum, type II SLAP lesion of	• Active compression test (modified interpretation 2) • Anterior apprehension test at 90° for pain OR apprehension (Rowe 1981: modified interpretation 1,2) • Anterior slide test (modified interpretation 2) • Biceps load II test (standard) • Compression-rotation test (modified interpretation 2) • Palpation for bicipital groove tenderness (modified interpretation 2) • Relocation test for pain OR apprehension (modified interpretation 2) • Speed's test (modified procedure, modified interpretation 1,2) • Whipple test (modified interpretation 2) • Yergason's test (modified interpretation 2)	E	

Table 8. Summary: studies of tests for labral lesions *(Continued)*

Parentis 2006	132	Labrum, type II SLAP lesion of	• Active compression test (modified interpretation 1,2) • Anterior slide test (modified interpretation 2) • Crank test (Liu 1996b: modified procedure, modified interpretation 2) • Hawkins' test (modified procedure, modified interpretation 2) • Neer's sign (modified procedure, modified interpretation 2) • Pain provocation test (modified interpretation 2) • Modified relocation test for posterosuperior glenoid impingement (modified interpretation 2) mislabelled as Jobe's relocation test • Speed's test (modified interpretation 1,2) • Yergason's test (modified interpretation 1,2)	A	
Schlechter 2009	254 (246)	Labrum, type II-IV SLAP lesion of	• Active compression test (modified interpretation 2) • Anterior slide test (modified procedure) • Combination: active compression test (modified interpretation 2) OR passive distraction test (standard) • Passive distraction test (standard)	E	
Stetson 2002	65	Labrum, any tear of	• Active compression test (modified interpretation 1) • Crank test (standard)		No
Suder 1994	31	Labrum, any tear of	• 'Impingement sign' (no reference or details given) • 'Impingement test' (no reference or details given)	NR	NR

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source
A: Isolated absolute discrepancy of 1% to <5% - a suspected or confirmed typographical error
B: Isolated absolute discrepancy of 5% to <10% - a suspected or confirmed typographical error
C: Isolated discrepancy of 10% or more - a suspected or confirmed typographical error
D: Multiple absolute discrepancies of which the greatest is 1% to <5%
E: Multiple absolute discrepancies of which the greatest is 5% to <10%
F: Multiple absolute discrepancies of which the greatest is 10% or more
?: 2 X 2 table not reported and cannot be deduced with certainty.
NR: Summary statistics not reported

Table 9. Summary: studies of tests for multiple, undifferentiated target conditions

Study ID	Shoulders (patients, if different)	Specific target condition	Index test name, provenance (where clarification is required) and manner of use compared to original description (standard/ modified procedure/modified interpretation)	Discrepancies between reported and back-calculated summary statistics (Sn, Sp, PPV, NPV or accuracy)	
				Yes	No
Bennett 1998	46 (45)	LHB tendinitis/LHB avulsion/SLAP lesion, any	• Speed's test (modified procedure, modified interpretation 1)	C	
Holtby 2004a	50	LHB lesion, any/ SLAP lesion, any	• Speed's test (modified procedure, modified interpretation 1, 2)		No
			• Yergason's test (modified interpretation 1, 2)		
Michener 2009	55	SA-SD bursitis/ bursal-side degeneration of supraspinatus (but patients with PTT or FTT *inter alia* were not excluded)	• Empty can test (modified interpretation 2)		No
			• Hawkins' test (standard)		
			• Neer's sign (modified procedure)		
			• Painful arc test (standard)		
			• Resisted lateral rotation from neutral rotation for weakness ± pain (modified interpretation 1,2)		
			• 3 or more tests +ve		
Razmjou 2004	50	SIS/rotator cuff tendinitis or tear	• Hawkins' test (modified interpretation 2)		No
			• Neer's sign (modified interpretation 2)		

Modified interpretation 1: criteria for a positive test result not as described in the primary source
Modified interpretation 2: target condition of test not as described in the primary source
A: Isolated absolute discrepancy of 1% to <5% - a suspected or confirmed typographical error
B: Isolated absolute discrepancy of 5% to <10% - a suspected or confirmed typographical error
C: Isolated discrepancy of 10% or more - a suspected or confirmed typographical error
D: Multiple absolute discrepancies of which the greatest is 1% to <5%
E: Multiple absolute discrepancies of which the greatest is 5% to <10%
F: Multiple absolute discrepancies of which the greatest is 10% or more
?: 2 X 2 table not reported and cannot be deduced with certainty.
NR: Summary statistics not reported

Table 10. Reasons for excluded trials

Main reason	N	Details
Not DTA study	93	Not shown

Table 10. Reasons for excluded trials *(Continued)*

Not DTA study. Systematic review only	11	Beaudreuil 2009; Calvert 2009; Dessaur 2008; Hegedus 2008; Hughes 2008; Jones 2007; Luime 2004; Meserve 2009; Munro 2009; Pugh 2009; Walton 2008
Abstract only / no published full report	3	Ansara 2006; Morrissey 2005; Sileo 2006
Not DTA study of physical tests	5	Birtane 2001; Cullen 2007; El Dalati 2005; Jee 2001; O'Connor 2005
Not DTA study of included condition	8	Chronopoulos 2004; Kim 2004b; Kim 2005 (all not impingement); Kim 2007a (rheumatoid disease); Lafosse 2007 (timing of surgery); Lewis 2007 (not impingement); Odom 2001; Walton 2004 (not impingement)
Highly selected population - sports / lesion	9	Berbig 1999 (post traumatic dislocation); Brasseur 2004 (veteran tennis players); Hamner 2000 (overhead throwing athletes); Kibler 2006a (athletes); Kim 1999 (all post anterior dislocation); Meister 2004 (all overhead athletes); Mimori 1999(throwing injuries); Myers 2005 (all athletes); Walsworth 2008 (military)
Highly selected pop: 100% prevalence or by exclusion	12	Berg 1998 (slap); Burkhart 2000 (slap), Burkhart 2002 (rotator cuff); Fukuda 1996 (rotator cuff); Gschwend 1988 (no disease negative; no specificity); Liu 1996a (tears removed); Lyons 1992 (all RCT - tear size study); Morgan 1998 (slap); Pandya 2008 (slap); Read 1998 (100% prevalence by exclusion); Rhee 2005a (all slap); Watson 1989 (all subacromial impingement - no specificity)
Special equipment used	2	McCabe 2005; Osbahr 2006
Unsatisfactory / unacceptable reference test / control: 4	4	Gerber 1991; Lo 2004; Scheibel 2005 (control); Silva 2008 (MRI for impingement)
Unclear reporting of tests, testing and/or population	7	Adolfsson 1991; Ardic 2006; Malhi 2005; Miller 2008a; Murrell 2001; Norregaard 2002; Wnorowski 1997
Lack of, incomplete or grossly inconsistent data:	8	Ebinger 2008; Fodor 2009; Leroux 1995 (very large discrepancies); Litaker 2000 (large discrepancies); Polimeni 2003 (no data); Rowan 2007 (no test-specific data); Sandenbergh 2006 (no 2 x 2); Sorensen 2007 (data presentation)

DTA = Diagnostic test accuracy, N = number of studies

Table 11. Summary of systematic reviews

#	Target condition	Study ID	Search period	Included studies	N	Main conclusions	Notes and comments
1	Any	Hegedus 2008a	1966 to October 2006	45	999	**Conclusion 1** 'There is a lack of clarity with regard to whether common orthopaedic special tests are useful in differentially diagnosing pathologies of the shoulder'	**General** • Included studies with highly selected populations. • Pooled some clinically heterogeneous data. **Re conclusion 1** • Our conclusions broadly agree.
2	Impingement, Any tear, FTT, PTT	Dinnes 2003	1985 to October 2001	10	1235	**Conclusion 1** 'Few tests provided convincing evidence of the presence or absence of disease ... [Although] individual tests did perform well in the study by [Hertel 1996] ... the sample size was small and CIs were very wide'. The internal rotation lag sign also had a very low negative LR ... Other tests demonstrating high positive and negative LRs were the rent test and internal rotation resistance strength test.' **Conclusion 2** 'In four studies [Litaker 2000; Lyons 1992; MacDonald 2000; Read 1998], negative LRs were sufficiently low to confirm that disease is absent in those with a negative diagnosis'. **Conclusion 3**	**Re conclusion 1** • Our conclusions broadly agree. • We excluded three (Litaker 2000; Lyons 1992; Read 1998) of the four studies underpinning this conclusion on clinical or methodological grounds. **Re conclusion 2** • We excluded those studies evaluating generic examination.

Table 11. Summary of systematic reviews *(Continued)*

						'The results suggest that [generic] clinical examination by specialists can rule out the presence of a rotator cuff tear'.	
3	SIS	Alqunaee 2012	to January 2011	16	2390	**Conclusion 1** 'The Hawkins-Kennedy test, Neer's sign and empty can test are ... more useful for ruling out rather than ruling in SIS.' **Conclusion 2** 'The drop arm test and lift-off test ... are more useful for ruling in SIS if the test is positive.'	**General** • Pooled some clinically heterogeneous data. • Included several studies which we excluded on clinical or methodological grounds (Leroux 1995; Lyons 1992; Malhi 2005; Murrell 2001; Scheibel 2005; Walch 1998) or categorised as 'awaiting classification' pending clarification (Nanda 2008; Park 2005). • Included one study which postdated our search (Fowler 2010), but which concerned a selected population. **Re conclusion 1** • Our conclusions partially agree. We suggest cautious interpretation, as the point estimates are small, the 95% CIs wide, and the pooled data clinically heterogeneous. **Re conclusion 2** • The pooled point estimate for the drop arm test is small. That for the lift-off test is large, but the 95% CIs are wide. Again, the pooled data are clinically heterogeneous.
4	Rotator cuff disease	Beaudreuil 2009a	to June 2006	9	2116	*See* notes and comments	**General** • A descriptive review with transliteration of data from the primary studies and no quality assessment.

Table 11. Summary of systematic reviews *(Continued)*

							• The conclusions are not contentious.
5	Rotator cuff disease	Hughes 2008a	January 1966 to April 2007	13	2010	**Conclusion 1** One test, the rent test in Wolf 2001, is identified with LR+ >10 and LR - < 0.1; but a contradictory result in Lyons 1992 is noted. **Conclusion 2** Other tests with LR+ > 10 or LR- < 0.1 are listed. **Conclusion 3** Hertel 1996 was excluded on the grounds of arithmetical discrepancies.	**Re conclusion 1** • We agree, but excluded Lyons 1992 on clinical grounds. **Re conclusion 2** • We excluded four studies underpinning these conclusions on clinical or methodological grounds (Ardic 2006; Leroux 1995; Lyons 1992; Murrell 2001) and categorised one (Park 2005) as 'awaiting classification' pending clarification. **Re conclusion 3** • Our back-calculations identified only one discrepancy in Hertel 1996, which we attributed to a typographical error.
6	Rotator cuff disease	Longo 2011	Not reported	Not reported	Not reported	*See* notes and comments.	**General** • Included for completeness, but not a systematic review
7	Labral disease	Munro 2009a	1995 to June 2007	15	Numbers reported by test, not by study.	**Conclusion 1** The biceps load II and internal rotation resistance strength tests were identified as having large LR+ and moderate LR-, based on single studies of good quality.	**General** • Pooled some clinically heterogeneous data. **Re conclusion 1** • Our conclusions are broadly agree.
8	SLAP	Calvert 2009a	January 1970 to June 2004	15	Unclear	**Conclusion 1** 'The current literature being used as a resource for teaching in medical schools and continuing education	**General** • Included studies with highly selected populations. **Re conclusion 1**

Table 11. Summary of systematic reviews *(Continued)*

						lacks the validity necessary to be useful.' **Conclusion 2** 'There are no good physical examination tests that exist for effectively diagnosing a SLAP lesion.'	• While sharing concerns as to the validity of much of the diagnostic test accuracy literature, we consider this an over generalisation. **Re conclusion 2** • We distinguish 'limited or contradictory evidence for accuracy' from 'evidence of inaccuracy', and place a number of tests former category.
9	SLAP	Dessaur 2008a	1996 to 2006	17	2148	**Conclusion 1** 'It appears that no single test is sensitive or specific enough to to determine the presence of a SLAP lesion accurately'.	**General** • Included studies with highly selected populations. **Re conclusion 1** • We distinguish between 'limited or contradictory evidence for accuracy' from 'evidence of inaccuracy', and place a number of tests in the former category.
10	SLAP	Jones 2007a	January 1 1966 to July 1 2006	12	2260	**Conclusion 1** 'SLAP-specific physical examination results cannot be used as the sole basis of a diagnosis of a SLAP lesion.'	**General** • A descriptive review with transliteration of data from the primary studies and limited quality assessment. • Included studies with highly selected populations. **Re conclusion 1** • Given the current state of knowledge, we agree with this conclusion.
11	SLAP	Luime 2004b	1966 to 2003	17	1901	**Conclusion 1** 'Most promising [tests] for establishing labral tears are currently the biceps load I [not relevant to this review]	**General** • Included studies with highly selected populations. **Re conclusion 1**

Table 11. Summary of systematic reviews *(Continued)*

						and II, pain provocation of Mimori, and the internal rotation resistance strength tests.'	• We agree regarding the biceps load II and internal rotation resistance strength tests. However, in the population relevant to the present review, the pain provocation test did not perform well.
12	SLAP	Meserve 2009a	1966 to June 2007	6	777	**Conclusion 1** The anterior slide test is a poor test for predicting the presence of a labral lesion in the shoulder. **Conclusion 2** Active compression, crank, and Speed tests are more optimal choices. **Conclusion 3** Clinicians should choose the active compression test first, crank second and Speed test third when a labral lesion is suspected.	**General** • Included studies with highly selected populations. • Pooled some clinically heterogeneous data. **Re conclusion 1** • Based on the results of Schlechter 2009 (notwithstanding that this study was prone to arithmetical error), which post-dated the search of Meserve 2009a, we cannot unconditionally agree. **Re conclusions 2 & 3** • Based on current knowledge, we agree concerning the active compression test. • In relatively unselected populations, we found the crank test inferior to biceps load II for ruling in labral tears; and LRs for Speed's test did not suggest that it would be clinically useful.
13	SLAP	Walton 2008a	To May 2006	7	Numbers reported by test, not by study.	**Conclusion 1** 'Yergason's test is the only [test] that shows a significant ability to influence clinical decision making, based on the results of the current analysis. **Conclusion 2** 'Methodologic inadequacies in the report-	**General** • Included studies with highly selected populations. • There was limited quality assessment. • Pooled some clinically heterogeneous data.

Table 11. Summary of systematic reviews *(Continued)*

		ing of the publications are common, and caution must be exercised when drawing inferences from the results of these studies.'	**Re conclusion 1** • In relatively unselected populations, LRs for Yergason's test did not suggest that it would be clinically useful. **Re point 2** • We agree.

APPENDICES

Appendix 1. Search strategy for MEDLINE, EMBASE, CINAHL and AMED (OVID Web format): to November 2005

Index tests	
Diagnostic tests of interest	
	General terms associated with diagnostic tests
1	Diagnos$.mp.
2	Examin$.mp.
3	Man?euv$.mp.
4	Sign$.mp.
5	Test$.mp.
6	or / 1-5
	Named diagnostic tests
7	Active compression.mp.
8	(Anterior adj (release or slide or apprehension)).mp.
9	(Biceps adj (load or tension)).mp.
10	Bicipital groove.mp.
11	Compression rotation.mp.
12	Crank.mp.
13	Drop arm.mp.
14	Empty can.mp.
15	External rotation.mp.
16	External rotation lag.mp.
17	Full can.mp.
18	Gerber$.mp.
19	Hawkins.mp.
20	Hawkins Kennedy$.mp.
21	Impingement$.mp.
22	Infraspinatus.mp.

(Continued)

23	Internal rotation.mp.
24	Internal rotation resistance strength.mp.
25	IRRST.mp.
26	Jobe$.mp.
27	Lag.mp.
28	Lift off.mp.
29	Mimori$.mp.
30	Modified relocation.mp.
31	Neer$.mp.
32	O'Brien$.mp.
33	Pain provocation.mp.
34	Painful arc.mp.
35	Patte$.mp.
36	Physical.mp.
37	Posterior impingement.mp.
38	Relocation.mp.
39	Rent.mp.
40	Rotator cuff.mp.
41	SLAP.mp.
42	Snyder$.mp.
43	SLAPPrehension.mp.
44	Speed$.mp.
45	Subscapularis.mp.
46	Yergason$.mp.
47	Yocum$.mp.
48	Zaslav$.mp.
49	Or / 7-48
	General terms associated with diagnostic tests combined with named diagnostic tests (and /)

(Continued)

50	and / 6,49
Conditions of interest	
51	Arc.mp.
52	Bursitis.mp.
53	Impingement$.mp.
54	Lesion$.mp.
55	Patholog$.mp.
56	Tear$.mp.
57	Tend#nitis.mp.
58	Or / 51-57
	Structures at risk
59	Biceps.mp.
60	Bicipital.mp.
61	Glenoid.mp.
62	Infraspinatus.mp.
63	Intraarticular.mp.
64	Labr$.mp.
65	Rotator cuff.mp.
66	SA SD.mp.
67	Shoulder.mp.
68	Subacromial.mp.
69	Subdeltoid.mp.
70	Subscapular$.mp.
71	Subcoracoid.mp.
72	SLAP.mp.
73	Teres minor.mp.
74	Or / 59-73
	Classes of disorder combined with structures at risk (and /)

(Continued)

75	and / 58,74
Authors of eponymous tests	
76	Crenshaw A$.au.
77	Gerber C.au.
78	Hawkins R.au.
79	Jobe C$.au.
80	Mimori K.au.
81	Neer C$.au.
82	O'Brien S$.au.
83	Patte D.au.
84	Snyder S$.au.
85	Yergason R$.au.
86	Yocum L$.au.
87	Zaslav K$.au.
88	Or / 76-87
	Authors combined with classes of disorder (and /)
89	and / 88,58
	Authors combined with classes of disorder or structures at risk (and /)
90	and / 88,74
	Authors combined with classes of disorder or structures at risk (or /)
91	Or / 89,90
Diagnostic tests of interest combined with conditions of interest (and/)	
92	and / 50,75
Synthesis	
93	or / 91-92
94	limit 93 to English language
95	remove duplicates from 94

Appendix 2. Search strategy for MEDLINE, EMBASE and AMED (OVID Web format): 2005 to February 2010

Search Strategy run 15/02/2012:

---

1 Diagnos$.mp. (3120783)
2 Examin$.mp. (2921664)
3 Man?euv$.mp. (35699)
4 Sign$.mp. (6361696)
5 Test$.mp. (3682306)
6 or/1-5 (11970722)
7 Active compression.mp. (312)
8 (Anterior adj (release or slide or apprehension)).mp. (368)
9 (Biceps adj (load or tension)).mp. (25)
10 Bicipital groove.mp. (187)
11 Compression rotation.mp. (41)
12 Crank.mp. (1153)
13 Drop arm.mp. (15)
14 Empty can.mp. (128)
15 External rotation.mp. (5940)
16 External rotation lag.mp. (16)
17 Full can.mp. (2648)
18 Gerber$.mp. (432)
19 Hawkins.mp. (566)
20 Hawkins Kennedy$.mp. (18)
21 Impingement$.mp. (7596)
22 Infraspinatus.mp. (1793)
23 Internal rotation.mp. (3969)
24 Internal rotation resistance strength.mp. (3)
25 IRRST.mp. (1)
26 Jobe$.mp. (150)
27 Lag.mp. (39297)
28 Lift off.mp. (662)
29 Mimori$.mp. (17)
30 Modified relocation.mp. (2)
31 Neer$.mp. (1309)
32 O'Brien$.mp. (1052)
33 Pain provocation.mp. (318)
34 Painful arc.mp. (81)
35 Patte$.mp. (1269379)
36 Physical.mp. (798454)
37 Posterior impingement.mp. (86)
38 Relocation.mp. (5150)
39 Rent.mp. (853)
40 Rotator cuff.mp. (10385)
41 SLAP.mp. (887)
42 Snyder$.mp. (1231)
43 SLAPPrehension.mp. (3)
44 Speed$.mp. (133045)
45 Subscapularis.mp. (1571)
46 Yergason$.mp. (24)
47 Yocum$.mp. (35)
48 Zaslav$.mp. (23)
49 or/7-48 (2202477)
50 and/6,49 (1286501)
51 Arc.mp. (20009)
52 Bursitis.mp. (4827)
53 Impingement$.mp. (7596)
54 Lesion$.mp. (915638)
55 Patholog$.mp. (790962)
56 Tear$.mp. (47604)
57 Tend#nitis.mp. (6231)
58 or/51-57 (1673209)

59 Biceps.mp. (12979)
60 Bicipital.mp. (613)
61 Glenoid.mp. (4590)
62 Infraspinatus.mp. (1793)
63 Intraarticular.mp. (9261)
64 Labr$.mp. (8170)
65 Rotator cuff.mp. (10385)
66 SA SD.mp. (25)
67 Shoulder.mp. (69932)
68 Subacromial.mp. (2381)
69 Subdeltoid.mp. (254)
70 Subscapular$.mp. (4647)
71 Subcoracoid.mp. (165)
72 SLAP.mp. (887)
73 Teres minor.mp. (366)
74 or/59-73 (100979)
75 and/58,74 (21410)
76 Crenshaw A$.au. (110)
77 Gerber C.au. (467)
78 Hawkins R.au. (336)
79 Jobe C$.au. (51)
80 Mimori K.au. (344)
81 Neer C$.au. (67)
82 O'Brien S$.au. (3422)
83 Patte D.au. (136)
84 Snyder S$.au. (2886)
85 Yergason R$.au. (0)
86 Yocum L$.au. (43)
87 Zaslav K$.au. (9)
88 or/76-87 (7870)
89 and/58,88 (657)
90 and/74,88 (628)
91 or/89-90 (968)
92 and/50,75 (8459)
93 or/91-92 (9221)
94 limit 93 to english language (7748)
95 limit 94 to yr="2005 -Current" (3000)
97 remove duplicates from 96 (1888)

Appendix 3. Quality assessment tool* and coding manual

*Adapted from Whiting (2003), Cochrane Diagnostic Reviewers Handbook version 0.3 (2005)

1. Was the spectrum of patients representative of the patients who will receive the test in practice? [To define spectrum bias]

Though clinical examination can be applied at all stages, our target population is the relatively unselected one in a primary care setting. This level of care may involve self-referral to a physiotherapist or, more usually, consultation with a general medical practitioner and possible cross-referral to a physiotherapist (often located in the community) or for imaging tests.

Clearly defined patient populations are unlikely in retrospective studies (Bossuyt 2003; van der Schouw 1995: see Footnotes for citations).

Yes (Y)	(a) The setting was primary care AND (b) the population was unselected but defined by age and gender AND (c) the reference test was non- or minimally invasive (physical tests plus local anaesthesia, ultrasound, MRI) AND (d) there was diagnostic uncertainty AND (e) the study was prospective and (f) recruitment was consecutive
No (N)	*General factors*

(Continued)

	(a) There was no diagnostic uncertainty i.e. the study compared diseased- with healthy subjects (case-control study) OR (b) the study was not prospective OR (c) recruitment was not consecutive
	Review-specific factors (a) The setting was secondary or tertiary care OR (b) the population was clearly selected OR (c) the reference test was more than minimally invasive (surgery, arthroscopy, arthrography, MRA, CT)
Unclear (?)	Insufficient information

2. Were selection criteria clearly described?

This criterion is omitted from the Cochrane Diagnostic Reviewers' Handbook (Cochrane Diagnostic Review Group 2005) but considered important in the present context, in which pain may arise from a number of conditions other than the target condition.

Y	(a) The selection criteria were clearly described (e.g. pain in the shoulder/ deltoid region, painful arc of motion, pain on overhead activities contributing to a clinical suspicion of impingement) AND (b) the exclusion criteria were clearly described (e.g. referred pain, gross restriction of movement, inflammatory disease, fracture)
N	(a) The selection criteria were undescribed/ very unclearly described (e.g. "shoulder pain") OR (b) the exclusion criteria were undescribed/ very unclearly described
?	(a) The selection criteria were described AND (b) the exclusion criteria were described BUT (c) the description of the selection criteria was not completely clear (e.g. an unqualified statement such as, "patients with suspected impingement") OR (d) the description of the exclusion criteria was not completely clear

3. Is the reference standard likely to correctly classify the target condition?

The generally recognised 'gold' standards are inapplicable to primary care. In general, the diagnostic tests that are applicable to primary care are less likely to correctly classify the target conditions. There two exceptions:

a. Since structural abnormalities of the rotator cuff are common in asymptomatic shoulders (MacDonald 2000; Milgrom 1995; Sher 1995), subacromial local anaesthesia may be more relevant to the symptoms of subacromial outlet impingement than diagnostic imaging, arthroscopy or open surgery (Dinnes 2003). However, since the site of anaesthesia would be critical, only subacromial bursal injections performed under guidance (fluoroscopic or ultrasonographic) will be accepted as a satisfactory reference test.
b. Based on data from eight primary studies (N = 687) that used arthroscopy and/or open surgery as reference standards for full thickness rotator cuff tears in low-prevalence samples (range 3 to 37%; mean 25% (16.32 to 33.68%)), MRI had a pooled sensitivity of 0.90 (0.84 to 0.94) and specificity of 0.95 (0.92 to 0.96) (Dinnes 2003). On these grounds, MRI appears sufficiently accurate for use as a reference test for full thickness rotator cuff tears in settings (such as general primary care) where there is likely to be a low prevalence of this disorder

Y	The reference standard was (a) arthroscopy OR (b) surgery OR (c) a combination of these OR (d) local anaesthesia of the subacromial bursa by guided injection OR (e) the target condition was full thickness rotator cuff tears in a sample with a likely low prevalence of this condition and the reference standard was MRI
N	Not applicable
?	The reference standard was (a) arthrography OR (b) subacromial local anaesthesia by 'blind' injection OR (c) MRA OR (d) MRI, except as defined above OR (e) ultrasonography

4. Is the time period between the reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? [To identify disease progression bias]

The acceptable interval would vary according to the average duration of symptoms.

(Continued)

Y	The average interval was ≤ (a) the average duration of symptoms OR (b) 1 month (whichever was the shorter)
N	The conditions for 'Y' were expressly not met
?	Insufficient information

5. Did the whole sample, or a random selection of the sample, receive verification using a reference standard? [To identify partial verification bias]

Y	(a) All patients were accounted for as having undergone a reference test OR (b) a randomly selected sample of patients underwent a reference test. (Score 'Y' even if different reference tests were used)
N	(a) Not all patients were accounted for as having undergone a reference test OR (b) a non-random selection of patients underwent a reference test
?	Insufficient information

6. Did patients receive the same reference standard regardless of the index test result? [To identify differential verification bias]

Y	(a) All patients underwent the same reference test OR (b) patients underwent different reference tests, but these were probably equivalent (e.g. arthroscopy and open surgery)
N	Patients underwent different reference tests, which were probably not equivalent (e.g. arthrography and surgery)
?	Insufficient information

7. Was the reference standard independent of the index test? [To identify incorporation bias]

Y	Self explanatory
N	Self explanatory
?	Self explanatory

8. Was the execution of the index test described in sufficient detail to permit replication of the test?

This criterion is omitted from the Cochrane Diagnostic Reviewers' Handbook (Cochrane Diagnostic Review Group 2005) but included here because minor technical variations may affect physical tests' outcomes, and interpretation may not be straightforward.

Y	(a) A clear, detailed description was given enabling replication and interpretation OR (b) a reference was given to an adequate source of this information
N	(a) The description lacked sufficient clarity to enable replication or interpretation AND (b) no reference was given to an adequate source of this information
?	Not applicable

9. Was the execution of the reference standard described in sufficient detail to permit its replication?

This criterion is omitted from the Cochrane Diagnostic Reviewers' Handbook (Cochrane Diagnostic Review Group 2005) but included here because the reference tests' interpretation is ultimately subjective.

Y	(a) A clear, detailed description was given enabling replication and interpretation OR (b) a reference was given to an adequate source of this information

(Continued)

N	(a) The description lacked sufficient clarity to enable replication or interpretation AND (b) no reference was given to an adequate source of this information
?	Not applicable

10. Were the index test results interpreted without knowledge of the results of the reference standard? [To identify test review bias?]

Clinical examination is highly subjective, and retrospective interpretation is a potential concern.

Y	There was a clear statement of blinding
N	There does not appear to have been blinding
?	The study was prospective and it is unclear whether there was blinding, but the index test preceded the reference standard. This does not apply to retrospective studies, in which both tests are likely to have been re-interpreted at the same time (Whiting 2003). In the absence of a clear statement of blinding, retrospective studies should be scored 'N'

11. Were the reference standard results interpreted without knowledge of the results of the index test? [To identify diagnostic review bias]

Since the clinical relevance of some arthroscopic and surgical findings (e.g. glenoid labral lesions, rotator cuff fraying and even rotator cuff tears) is uncertain, and interpretation of the other reference tests is subjective, foreknowledge of the index test result has potential to influence interpretation.

Y	There was a clear statement of blinding
N	There does not appear to have been blinding
?	The reference test was stated to have been conducted "independently"

12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?

Patients' demographic (age/ sex) and historical data would normally be available when physical test results are interpreted.

Y	Demographic and historical data were available when index test/s was/were interpreted
N	Demographic or historical data were not available when index test/s was/were interpreted
?	Insufficient information

13. Were uninterpretable/ intermediate test results reported?

Y	(a) The study was prospective AND (b) recruitment was consecutive AND (c) test results were reported for all initially included patients
N	(a) Recruitment was not consecutive OR (b) test results were not reported for all initially included patients
?	(a) Insufficient information OR (b) the study was not prospective (due to inconsistent reporting in clinical records, uninterpretable/ intermediate test results are sometimes not identified in retrospective studies (van der Schouw 1995))

14. Were withdrawals from the study explained?

(Continued)

Y	(a) The study was prospective AND (b) recruitment was consecutive AND (c) withdrawals were reported AND (d) withdrawals were explained (ideally by a flow chart)
N	(a) The study was not prospective OR (b) recruitment was not consecutive (unexplained non-recruitment equating to unreported/explained withdrawal) OR (c) withdrawals did not appear to have been reported OR (d) withdrawals were unexplained
?	Insufficient information

Footnotes

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Annals of Internal Medicine 2003;138(1):W1-12.

van der Schouw 1995

van der Schouw YT, Van Dijk R, Verbeek AL. Problems in selecting the adequate patient population from existing data files for assessment studies of new diagnostic tests. Journal of Clinical Epidemiology 1995;48(3):417-22.

WHAT'S NEW

Date	Event	Description
3 September 2014	Amended	Republished (September 2014) to include a Plain Language Summary.

CONTRIBUTIONS OF AUTHORS

Nigel Hanchard co-ordinated the production of the protocol and produced the first drafts at all stages of its development. All drafts were critically reviewed by Helen Handoll.

Nigel Hanchard co-ordinated the production of the review, screened search results, retrieved papers, performed study selection, designed the data extraction form, appraised quality and extracted data of all the included studies, wrote to authors for additional information, collated the review data, entered data into RevMan, analysed the data, interpreted the results, provided a clinical perspective, and drafted most of the writing of the review.

Mario Lenza screened search results, retrieved papers, performed study selection, appraised quality and extracted data of half of the included studies, gave key input into discussions on the structuring of the review, interpreted the results, provided a clinical perspective, and provided feedback on draft versions of the review.

Helen Handoll screened search results, retrieved papers, performed study selection, appraised quality and extracted data of half of the included studies, gave key input into discussions on the structuring of the review, interpreted the results, gave general advice, drafted select sections of the review and provided feedback on draft versions of the review.

Yemisi Takwoingi provided key input and feedback on review structure, data analysis and statistical methods, provided a methodological perspective, and gave feedback on draft versions of the review.

Nigel Hanchard is the guarantor of the review.

DECLARATIONS OF INTEREST

None known. Nigel Hanchard has recently completed a study on the diagnostic accuracy of physical tests for subacromial and internal shoulder impingement. The ‘reference’ test was ultrasound guided injection of local anaesthetic into the subacromial bursa and shoulder joint.

SOURCES OF SUPPORT

Internal sources

- University of Teesside, Middlesbrough, UK.

External sources

- Department of Health, UK.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Subsequent to the protocol we excluded studies with highly selected populations, such as overhead throwing athletes.

We did not contact all authors as planned in respect of missing key information such as 2 x 2 table data, study participant numbers and application of blinding. This is because we judged it unlikely that a request would be successful in obtaining reliable data. We left eight studies apparently drawing on the same database in Studies awaiting classification as it would require a major effort by ourselves and the study authors to clarify statistical and procedural inconsistencies in these studies.

Discrepancies in 2 x 2 tables due to rounding errors were a common finding. A rule was devised for the inclusion of data from studies with minor discrepancies in their 2 x 2 tables (Data extraction and management).

INDEX TERMS

Medical Subject Headings (MeSH)

Arthroscopy; Bursa, Synovial [injuries]; Bursitis [*diagnosis]; Glenoid Cavity; Joint Instability [diagnosis]; Physical Examination [*methods]; Prospective Studies; Randomized Controlled Trials as Topic; Rotator Cuff Injuries; Rupture [diagnosis]; Shoulder Impingement Syndrome [*diagnosis]; Tendinopathy [*diagnosis]

MeSH check words

Humans