

Diagnostic Utility of Clinical Tests for SLAP Lesions: A Systematic Literature Review

JANETTE W. POWELL, PT, OCS, STC¹, PETER A. HUIJBREGTS, PT, DPT, OCS, FAAOMPT, FCAMT²,
RICHARD JENSEN, PT, PhD³

Superior Labrum Anterior-to-Posterior (SLAP) lesions have been discussed, defined, and investigated since Andrews et al¹ first described this pathology in 1985. These lesions, which involve the superior region of the glenoid labrum, with or without the attachment of the long head of biceps, have been noted to be challenging to diagnose conservatively²⁻¹⁸. Non-surgical assessment tools, i.e., manual clinical tests and imaging studies, do not seem to paint a clear diagnostic picture and many authors

have noted the need for arthroscopic visualization for sufficient diagnostic accuracy^{2,3,5-13,15-19}.

SLAP lesions are often described as complex injuries with a spectrum of locations and varied types of tissue defects in the glenoid labrum and its associated structures¹³. Snyder et al²⁰ classified these pathological variations into four types on the basis of morphology (Figure 1):

- Type I: degenerative fraying with no detachment of the biceps insertion

- Type II: detachment of the biceps insertion
- Type III: a bucket-handle tear of the superior aspect of the labrum with an intact biceps tendon insertion to bone
- Type IV: an intra-substance tear of the biceps tendon with a bucket-handle tear of the superior aspect of the labrum

The Type II SLAP lesions have been further divided into three subtypes depending on whether the detachment of the labrum involves the anterior aspect of the labrum alone, the posterior aspect alone, or both aspects^{8,10,11,13,14,21,22}. The above classification system has been expanded to include an additional three types^{8,10,11,21-23}:

- Type V: a Bankart lesion that extends superiorly to include a Type II SLAP lesion
- Type VI: an unstable flap tear of the labrum in conjunction with a biceps tendon separation
- Type VII: a superior labrum and biceps tendon separation that extends anteriorly, inferiorly to the middle glenohumeral ligament

It is likely that the varied patho-anatomy and/or patho-mechanics of these differ-

ABSTRACT: SLAP lesions are often complex injuries with varied defects and tissue involvement that are challenging to diagnose clinically. The literature notes the need for visualization under arthroscopy for adequate diagnostic accuracy. The goal of this article is to provide a current best-evidence synthesis with regard to physical examination tests used for the diagnosis of SLAP lesions. A literature search yielded 17 studies that investigated the diagnostic utility of clinical tests for SLAP lesions. These studies investigated 19 clinical tests. A narrative review and a systematic review of methodological quality using the QUADAS methodological quality assessment tool yielded 3 high-quality diagnostic utility studies. Current best evidence indicates that a negative finding for the passive compression test provides the therapist with the greatest evidence-based confidence that a SLAP lesion is absent. A positive finding on the anterior apprehension maneuver, the anterior slide test, the Jobe relocation test, the passive compression test, the Speed test, and the Yergason test or a combination of positive findings on the Jobe relocation test and the active compression test or the Jobe relocation test and the anterior apprehension maneuver provides the therapist with the research-based confidence required to rule in a SLAP lesion. For ruling in a SLAP lesion, the greatest diagnostic value should likely be placed on a positive finding on the passive compression test. Suggestions for future research are provided.

KEYWORDS: Best-evidence Synthesis, Diagnostic Utility, QUADAS, SLAP Lesion, Systematic Review

¹Supervisor of Performance Medicine, Cirque du Soleil, "O", Las Vegas, NV.

²Assistant Professor, Online Education Department, University of St. Augustine for Health Sciences, St. Augustine, FL.

³Dean, Division of Advanced Studies, University of St. Augustine for Health Sciences, St. Augustine, FL.

Address all correspondence and requests for reprints to: Dr. Peter Huijbregts, shelbournphysio@telus.net

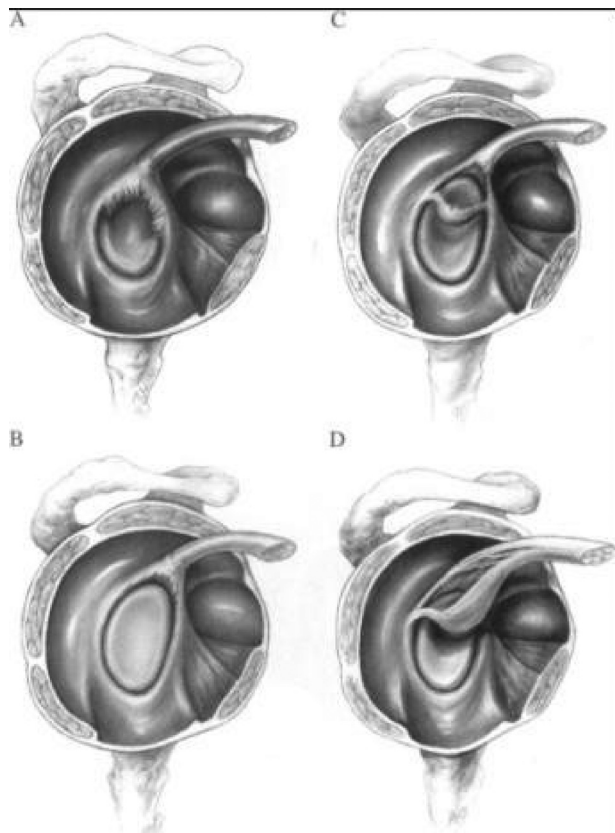


FIGURE 1. Classification of superior glenoid or superior labrum and biceps anchor (SLAP) lesions.

A, Type I: degenerative fraying of the superior labrum with the edge still firmly attached to the glenoid.

B, Type II: detachment of the superior labrum and biceps tendon from the glenoid with resultant destabilization of the biceps anchor.

C, Type III: bucket-handle tear of superior labrum; remaining labrum and biceps anchor are stable.

D, Type IV: bucket-handle tear of superior labrum with extension into the biceps tendon. From D'Allessandro et al¹³, p. 288.

ent types of SLAP lesions significantly alter the clinical presentation^{10,14}.

Further complicating the clinical presentation and correct diagnosis is the noted variability with regard to “normal” anatomy of the superior labrum including the appearance and degree of attachment of the labrum to the glenoid rim with a spectrum of foramina and recesses and the variability in the surrounding glenohumeral ligament structures^{8,10,13,14,21-23}. Adding to these intrinsic complexities these lesions are also often associated with additional concomitant shoulder pathology^{10-12,14,21}. SLAP lesions commonly accompany other extra- and intra-articular shoulder pathology further complicating the clinical diagnostic process^{14,18,23-25}.

The incidence of SLAP lesions has been reported in the literature to range between 6–26% of all shoulder injuries evaluated arthroscopically^{10,13,14,21-23}. A variety of etiologic mechanisms have been implicated; including fall on an outstretched arm, overhead work, internal impingement, traction/tension on the biceps tendon due to overhead athletic activity, and instability resulting in a “peel-back” of the superior aspect of the labrum when the biceps insertion is twisted as the arm is brought into abduction and external rotation^{3,8,10,11,13,14,21-23}. Musgrave and Rodosky²³ suggested that different mechanisms of injury were likely to be operational in the various SLAP lesion types. Huijbregts²¹ reported on the high

sensitivity of non-specific posterior shoulder pain and the inability to perform overhead athletic activities but specificity of such history findings is clearly very limited. Similarly, a patient report of mechanical symptoms was also found lacking in diagnostic utility²¹.

In the currently predominant evidence-based practice paradigm, research-based data on utility of diagnostic tests supplement and replace the pathophysiologic rationale previously used for the interpretation of these tests. Because data on epidemiology, mechanism of injury, and other history findings with regard to SLAP lesions clearly provide insufficient information to allow for confident research-based clinical diagnosis the clinician typically turns to the physical examination to gather more data. Research studies on diagnostic accuracy of individual tests provide statistics such as sensitivity, specificity, predictive values, and likelihood ratios. However, the numerical value for these statistics and hence their clinical interpretation are highly dependent on the methodological quality of the study in which they were derived. Various authors have written narrative reviews discussing the diagnostic accuracy of physical examination tests for SLAP lesions but only few have combined this report of diagnostic accuracy statistics with a methodological quality assessment of the diagnostic accuracy studies in the form of a systematic review^{5,7,26,27}. The systematic reviews by Dinnes et al²⁶ and by Luime et al⁷ were published in 2003 and 2004, respectively. With 4 of the 17 studies retrieved for this review published after 2004 and another two published in the years that these reviews were published, an update was obviously justified. However, 2 more recent systematic reviews both published in 2007 seemed to question the need for this review. While Jones and Galluch⁵ provided descriptions of the tests as we did in this review, Hegedus et al²⁷ did not, thereby limiting the clinical utility of their review as a stand-alone paper. Neither review provided a summary description of the studies retrieved, which we did because it allowed for a discussion of research validity other than solely relying

on the methodological quality assessment tool as both other reviews did. Jones and Galluch used a non-validated assessment tool scoring only three items, whereas Hegedus et al used the same tool as in this study. Aside from the fact that our review showed that this tool does not adequately address some critical methodological errors leading us to discard in this review one of the studies deemed a higher-quality study by Hegedus et al, comparison of our scores on this tool with other sources scoring the same studies forces us to question reliability of the tool and thereby sole reliance on this tool. Finally, literature search strategies used by the other recent reviews and this review were substantially different yielding five extra references as compared to Jones and Galluch and two additional references as compared to Hegedus et al; it should be noted that the latter authors did retrieve one study we did not find.

Having thus justified the need for this review, the goal for this article is to provide the most comprehensive and clinically useful synthesis with regard to the quality of studies addressing physical examination tests used specifically for the diagnosis of SLAP lesions.

Methods and Materials

For our literature review, we searched the PubMed, Proquest, and Cumulative Index to Nursing and Allied Health (CINAHL) online databases for peer-reviewed references. In the Proquest database, we used the Health and Medical Complete Database and the Nursing and Allied Health Source Database.

Our search used the key words: *SLAP, Superior Labrum, Superior Labral, Glenoid Labrum, and Glenoid Labral*. These were noted to be the key terms in the literature during our preliminary reading on SLAP lesions. The search was not limited by adding other terms to increase the chance of identifying relevant studies. After reviewing abstracts of the literature selected, full-text format of those studies that appeared to quantitatively investigate diagnostic utility of physical examination diagnostic test(s) of SLAP lesions as compared to a gold standard test or test regimen were re-

trieved. We then did a hand search of the reference lists of the articles retrieved to locate further relevant references fitting these same inclusion criteria.

We limited the literature search from 1985 to July 2007. The 1985 start date was chosen for the literature search, as this was when SLAP lesions were first described in the literature¹. An increased interest in research-based data on diagnostic utility has led to the development of methodological quality assessment tools such as the STARD (Standards for Reporting of Diagnostic Accuracy)²⁸⁻³⁰ and QUADAS (Quality Assessment of Diagnostic Accuracy Studies)^{28,31-33} criterion lists. The STARD list is a prospective tool designed to outline the features required for an unbiased diagnostic accuracy study. In contrast, the QUADAS list is a retrospective tool used to critique the methodological rigor of a diagnostic accuracy study²⁸. For this review we chose the retrospective QUADAS tool to systematically evaluate the methodological quality of the studies selected in this literature review.

The QUADAS tool evaluates 14 items. Cook and Hegedus³⁴ noted that a score of less than 10/14 was where they noted changes in meta-analytic values for diagnostic accuracy summary statistics in the poorly designed studies³⁴. We therefore adopted a score minimum of 10/14 as the "cut-off" score for methodological quality in this systematic review. Use of the QUADAS tool facilitates the evaluation of the internal and external validity of each study and thereby allows the clinician to avoid exaggerated reporting of diagnostic utility from poorly designed studies that might otherwise lead to errors in clinical decision-making, including inaccurate diagnosis, inappropriate treatment, and premature adoption of a special test that provides little value²⁸. However, literature on reliability and validity of the QUADAS tool is very limited³². Although Cook and Hegedus have suggested a cut-off value for acceptable methodological quality, this value is not based on a broad consensus. This led us to decide to assess methodological quality both systematically with the QUADAS tool and also narratively looking for possible methodological errors specific to the studies re-

trieved before providing a current best-evidence synthesis.

Results

The PubMed Search retrieved 977 articles. The Proquest Search produced 338 articles. The CINAHL Search yielded 93 articles. Of these studies, 17 studies met the inclusion criteria^{4,15-19,24,25,35-43}. The hand search of the reference lists of these articles yielded no further references. The 17 studies investigated 19 clinical tests for the diagnosis of SLAP lesions. Figure 2 illustrates the literature search and screening process leading to the best-evidence synthesis presented. Table 1 provides data on the diagnostic utility of the various clinical tests studied. Table 2 provides data on the diagnostic utility of multiple-test regimens for the diagnosis of SLAP lesions^{4,40}. Where not provided by the authors but where sufficient data was available, we calculated the values for sensitivity, accuracy, and positive and negative likelihood ratios using the formulas provided in Table 3²⁸. Where provided by the authors, we added between brackets in Tables 1 and 2 the raw data from which these diagnostic utility statistics were calculated. Table 4 provides the analysis and scoring of each study on its methodological quality using the QUADAS tool^{28,31-33}.

Anterior Apprehension Maneuver

Guanche and Jones⁴ prospectively investigated 59 consecutive patients (60 shoulders) undergoing shoulder arthroscopy with a seven-test regimen that included the anterior apprehension maneuver in an orthopaedic surgery setting. They provided a photograph and description of test performance and interpretation identical to those provided in various textbooks⁴⁴⁻⁴⁶. The patient was supine and the examiner abducted the arm to 90° with the elbow in 90° of flexion and then progressively externally rotated the involved shoulder. A positive test was indicated by a look or feeling of apprehension or alarm on the patient's face and the patient's resistance to further motion. The patient might also state that the feeling experienced was what it felt like when the shoulder was previ-

ously dislocated. The rater for the clinical test portion was one orthopaedic surgeon. The gold standard test was arthroscopy, and the surgical procedure was described in moderate detail. The study is unclear on the number of raters involved in the arthroscopic diagnosis. It is also unclear if the arthroscopic examiners were blinded to the clinical test results. Of 60 shoulders, 33 shoulders were arthroscopically diagnosed with SLAP lesions. However, this study did not use the diagnostic classification of SLAP lesions as described by Snyder et al²⁰ and introduced above. The study analyzed diagnostic utility of single tests and/or combinations. However, the number of patients who tested positive on the anterior apprehension maneuver was not provided in addition to the diagnostic utility statistics.

Nakagawa et al¹⁷ prospectively studied 54 throwing athletes undergoing shoulder arthroscopy in an orthopaedic and sports medicine surgical setting. This study investigated 18 clinical tests but only one was operationally defined. Nakagawa et al¹⁷ did not provide operational definitions for test performance or interpretation of the anterior apprehension maneuver. Four clinical fellows in orthopaedic surgery clinically evaluated the shoulders. The gold standard test was arthroscopic findings but the surgical procedure was not described. One surgeon blinded to the clinical test results assessed all shoulders intra-operatively. Of 54 patients, 24 were arthroscopically diagnosed with SLAP lesions. However, the number of patients who tested positive on the anterior apprehension maneuver was not provided.

Active Compression Test

O'Brien et al⁴² introduced the active compression test with a prospective study of 318 patients consisting of 268 consecutive shoulder pain patients and 50 controls (knee pain patients) in an orthopaedic surgical setting. The authors defined this test as one where the patient flexed the involved arm forward 90° with the elbow in full extension. The patient then adducted the arm 10° to 15° medial to the sagittal plane of the body. The arm was internally rotated so the

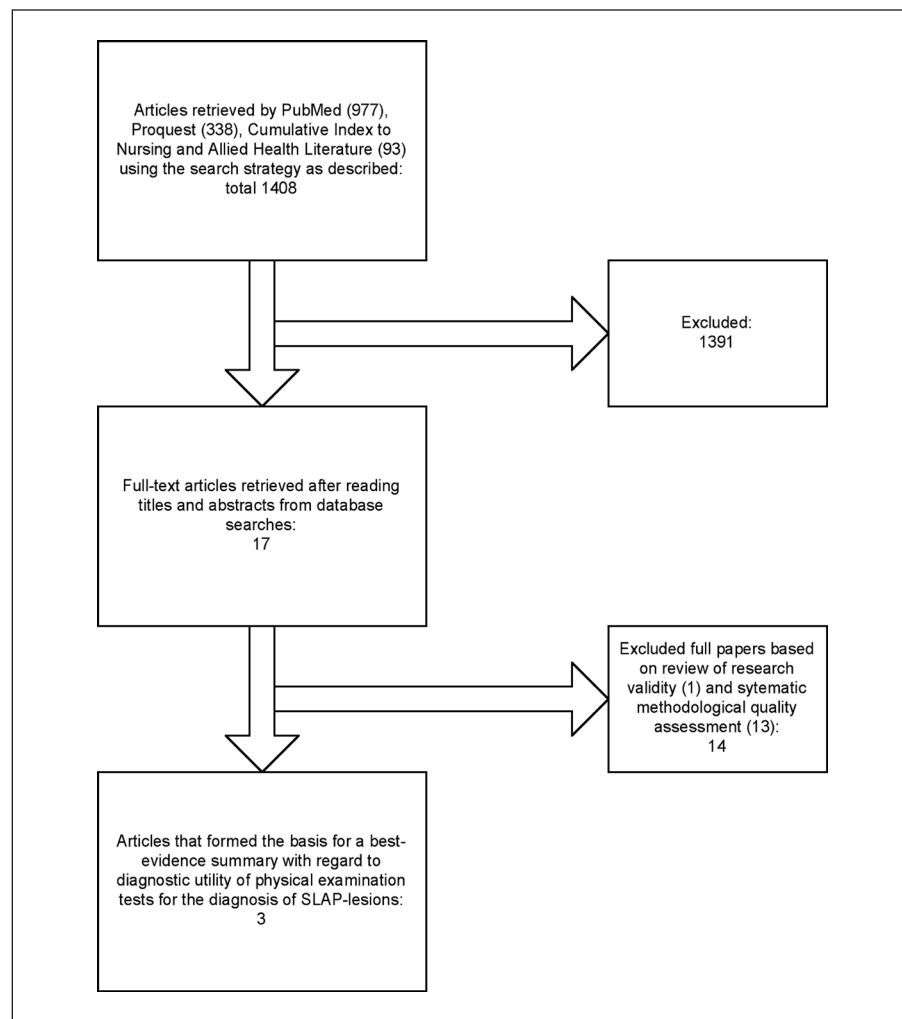


FIGURE 2. Flow diagram of literature search and screening process leading to a best evidence summary.

patient's thumb pointed downward. The examiner then applied a downward force to the arm. With the arm in the same position, the arm was fully externally rotated so the patient's palm faced upward and the same downward force was again applied. The test was considered positive if pain elicited with the first maneuver was reduced or eliminated with the second maneuver. Pain or painful clicking described as inside the glenohumeral joint was indicative of a labral tear. Pain localized to the acromioclavicular joint or on top of the shoulder was considered diagnostic of acromioclavicular joint abnormality. From the study, it remained unclear how many raters were involved in the clinical and surgical assessments. The 50 control patients did not undergo the gold standard arthroscopic examina-

tion. The study also did not clarify whether the arthroscopic examiner was blinded to the clinical test results. Labral tears were not defined in this study. The surgical procedure was not described. Of all patients, 56 had positive active compression tests and 53 had labral tears at surgery.

McFarland et al²⁴ retrospectively investigated 426 consecutive shoulder arthroscopy patients in an orthopaedic surgical setting. The control group consisted of 387 patients and included patients with type I SLAP lesions; 39 patients were included in the SLAP lesion group defined as types II, III, and IV lesions. The physical examination done up to 4 weeks prior to surgery included varied combinations of three SLAP lesion clinical tests: 294 patients had all three

TABLE 1. Diagnostic utility SLAP lesion clinical tests^{4,15-19,24,25,36-39,41,43}.

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Anterior apprehension maneuver							
• Guanche & Jones ⁴	-----	0.40	0.87	0.90	0.33	3.077	0.722
• Nakagawa et al ¹⁷	0.59	0.83	0.40	0.53	0.75	1.38	0.425
Active compression test							
• O'Brien et al ⁴²	0.988 (253/256)	1.00 (53/53)	0.985 0.63	0.946 (200/203)	1.00 (53/56)	66.66 (200/200)	0.00
• Guanche & Jones ⁴	-----	0.47	0.73	0.87	0.40	2.33	0.507
• McFarland et al ²⁴	0.54 (221/409)	(18/38) 0.778	0.55 (203/371)	0.10 (18/186)	0.91 (203/223)	1.044	0.964
• Myers et al ⁴³		0.54	0.111	0.70	0.143	0.875	2.00
• Nakagawa et al ¹⁷	0.595	0.63	0.60	0.52	0.62	1.35	0.767
• Parentis et al ¹⁹	0.57	0.54	0.50	0.355	0.75	1.25	0.75
• Stetson & Templin ¹⁸	-----		0.31	0.34	0.50	0.783	1.48
Anterior slide test							
• Kibler ³⁷	0.858 (194/226)	0.784 (69/88)	0.915 (125/138)	0.64 (69/107)	0.868 (125/144)	9.22	0.236
• McFarland et al ²⁴	0.77 (322/419)	0.08 (3/38)	0.84 (319/381)	0.05 (3/65)	0.90 (319/354)	0.50	1.095
• Nakagawa et al ¹⁷	0.54	0.05	0.93	0.33	0.56	0.714	1.021
• Parentis et al ¹⁹	-----	0.10	0.815	0.19	0.676	0.55	1.104
Biceps groove tenderness							
• Guanche & Jones ⁴	-----	0.44	0.40	0.69	0.19	0.733	1.4
• Nakagawa et al ¹⁷	0.56	0.25	0.80	0.50	0.57	1.25	0.938
Biceps load test							
• Kim et al ³⁸	0.96 (72/75)	0.909 (10/11)	0.969 (62/64)	0.83 (10/12)	0.98 (62/63)	29.32	0.094
Biceps load test II							
• Kim et al ³⁹	0.944 (120/127)	0.897 (35/39)	0.966 (85/88)	0.921 (35/38)	0.955 (85/89)	26.38	0.107
Clunk test							
• Nakagawa et al ¹⁷	0.57	0.44	0.68	0.53	0.59	1.375	0.824
Compression rotation test							
• McFarland et al ²⁴	0.71 (214/303)	0.24 (7/29)	0.76 (207/274)	0.09 (7/74)	0.90 (207/229)	1.0	1.0
• Nakagawa et al ¹⁷	0.63	0.25	1.0	1.0	0.58	0.00	0.75
Crank test							
• Liu et al ⁴¹	0.919 (57/62)	0.906 (29/32)	0.933 (28/30)	0.935 (29/31)	0.903 (28/31)	13.52	0.101
• Guanche & Jones ⁴	-----	0.40	0.73	0.82	0.29	1.48	0.822
• Myers et al ⁴³	0.444	0.346	0.70	0.75	0.292	1.153	0.939
• Nakagawa et al ¹⁷	0.66	0.58	0.72	0.63	0.68	2.072	0.583
• Parentis et al ¹⁹	-----	0.125	0.826	0.238	0.684	1.436	1.059
• Stetson & Templin ¹⁸	-----	0.46	0.56	0.41	0.61	1.364	0.964

TABLE 1. Diagnostic utility SLAP lesion clinical tests^{4,15-19,24,25,36-39,41,43}. (Continued)

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Forced abduction test							
• Nakagawa et al ¹⁷	0.67	0.67	0.67	0.62	0.71	2.03	0.492
Impingement sign (Hawkins)							
• Nakagawa et al ¹⁷	0.59	0.50	0.67	0.55	0.63	1.515	0.746
• Parentis et al ¹⁹	-----	0.675	0.304	0.297	0.683	0.969	1.083
Impingement sign (Neer)							
• Nakagawa et al ¹⁷	0.48	0.33	0.60	0.40	0.53	0.825	1.12
• Parentis et al ¹⁹	-----	0.50	0.522	0.313	0.706	1.046	0.958
Jobe relocation test							
• Guanche & Jones ⁴	-----	0.44	0.87	0.91	0.34	3.385	0.643
• Hammer et al ⁸	-----	1.00 (13/13)	-----	-----	-----	-----	-----
• Nakagawa et al ¹⁷	0.56	0.75	0.40	0.50	0.67	1.25	0.625
• Parentis et al ¹⁹	-----	0.50	0.533	0.317	0.71	1.07	0.938
Pain provocation test							
• Mimori et al ¹⁶							
—MRA	0.97	1.00	0.90	-----	-----	10	0.0
—Arthroscopy	1.00 (11/11)	1.00 (11/11)	----- (0/0)	1.00 (11/11)	----- (0/0)	-----	-----
• Parentis et al ¹⁹	-----	0.15	0.902	0.40	0.709	1.53	0.94
Passive compression test							
• Kim et al ²⁵	0.836 (51/61)	0.818 (27/33)	0.857 (24/28)	0.871 (27/31)	0.80 (24/30)	5.72	0.212
Resisted supination external rotation test							
• Myers et al ⁴³	0.825	0.828	0.818	0.923	0.643	4.549	0.21
SLAPprehension test							
• Berg & Ciullo ³⁶	-----	0.819 (54/66)	-----	-----	-----	-----	-----
Speed test							
• Guanche & Jones ⁴	-----	0.18	0.87	0.80	0.26	0.138	0.943
• Holtby & Razmjou ¹⁵	0.56 (28/50)	0.32 (7/22)	0.75 (21/28)	0.50 (7/14)	0.58 (21/36)	1.28	0.91
• Nakagawa et al ¹⁷	0.57	0.04	1.00	1.00	0.57	0.00	0.6
• Parentis et al ¹⁹	-----	0.478	0.674	0.348	0.721	1.466	0.774
Yergason test							
• Guanche & Jones ⁴	-----	0.09	0.93	0.80	0.25	0.978	1.2857
• Holtby & Razmjou ¹⁵	0.63 (31/49)	0.43 (9/21)	0.79 (22/28)	0.60 (9/15)	0.65 (22/34)	2.05 0.00	0.72 0.87
• Nakagawa et al ¹⁷	0.61	0.13	1.00	1.00	0.59	1.92	0.935
• Parentis et al ¹⁹	-----	0.125	0.935	0.455	0.711	1.92	0.936

TABLE 2. Diagnostic utility multi-test regimens (modified from Guanche & Jones⁴ and Liu et al⁴⁰).

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
2 positive tests							
• Jobe & O'Brien ⁴	-----	0.41	0.91	0.93	0.34	4.56	0.648
• Jobe & apprehension ⁴	-----	0.38	0.93	0.94	0.33	5.43	0.667
• O'Brien & apprehension ⁴	-----	0.38	0.82	0.86	0.31	2.11	0.756
3 positive tests							
• Jobe, O'Brien & apprehension ⁴	-----	0.34	0.91	0.92	0.32	3.78	0.725
≥5 positive tests							
• Apprehension, relocation, load & shift, sulcus, & crank ⁴⁰	0.89 (48/54)	0.90 (37/41)	0.85 (11/13)	0.95 (37/39)	0.73 (11/15)	6.0	0.12

TABLE 3. Definition and calculation of statistical measures of concurrent criterion-related validity²⁸.

Statistical measure	Definition	Calculation
Accuracy	The proportion of people who were correctly identified as either having or not having the disease or dysfunction	$(TP + TN) / (TP + FP + FN + TN)$
Sensitivity	The proportion of people who have the disease or dysfunction who test positive	$TP / (TP + FN)$
Specificity	The proportion of people who do not have the disease or dysfunction who test negative	$TN / (FP + TN)$
Positive predictive value	The proportion of people who test positive and who have the disease or dysfunction	$TP / (TP + FP)$
Negative predictive value	The proportion of people who test negative and who do not have the disease or dysfunction	$TN / (FN + TN)$
Positive likelihood ratio	How likely a positive test result is in people who have the disease or dysfunction as compared to how likely it is in those who do not have the disease or dysfunction	$Sensitivity / (1 - specificity)$
Negative likelihood ratio	How likely a negative test result is in people who have the disease or dysfunction as compared to how likely it is in those who do not have the disease or dysfunction	$(1 - sensitivity) / specificity$

tests done, whereas 132 underwent one or two tests. The active compression test was done in 409 patients (371 control and 38 SLAP lesion patients). The active compression test was defined and interpreted as outlined by O'Brien et al⁴². It is unclear on how many testers were involved in the clinical assessments. The arthroscopic reference test was described in moderate detail and performed by only one surgeon. It was also unclear whether the arthroscopic examiner was blinded to the clinical test results. The active compression test was

positive in 168 of the control patients and 18 of the SLAP lesion patients; 39 patients were arthroscopically diagnosed with SLAP lesions (types II, III, and IV) and 387 patients had type I SLAP lesions or no lesion.

Stetson and Templin¹⁸ prospectively studied 65 patients with shoulder pain patients in an orthopaedic surgical setting. One orthopaedic surgeon performed the physical examination which included two SLAP lesion clinical tests. The authors defined and interpreted the active compression test as outlined by

O'Brien et al⁴². The surgical assessments were described in moderate detail but it was unclear how many assessors were involved in the surgical assessments and if they were blinded to the clinical test results. Of the patients, 41 tested positive with the clinical test, 14 of which had arthroscopically confirmed tears of the glenoid labrum.

Myers et al⁴³ prospectively studied 40 overhead throwing athletes in a sports medicine and orthopaedic surgical setting. The physical examination included 3 SLAP lesion clinical tests and was per-

TABLE 4. QUADAS values^{28,31-33} of the SLAP lesion diagnostic accuracy studies^{4,15-19,24,25,35-43}.

Item	Berg & Ciullo ³⁶	Guanche & Jone ³⁴	Hamner et al ³⁵	Holtby & Razmjou ¹⁵	Kibler ³⁷	Kim et al ²⁵	Kim et al ³⁸	Kim et al ³⁹	Liu et al ⁴⁰	Liu et al ⁴¹	McFarland et al ²⁴	Mimori et al ¹⁶	Myers et al ⁴³	Nakagawa et al ¹⁷	O'Brien et al ⁴²	Parentis et al ¹⁹	Stetson & Templin ¹⁸
1. Was the spectrum of patients representative of the patients who will receive the test in practice?	U	Y	U	U	Y	Y	N	Y	N	Y	Y	U	Y	Y	Y	Y	U
2. Were selection criteria clearly described?	N	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	N	Y	N	Y	Y	N
3. Is the reference standard likely to classify the target condition correctly?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4. Is the period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	U	Y	U	N	N	U	U	U	U	U	Y	U	Y	U	U	U	U
5. Did the whole sample or random selection of the sample receive verification using a reference standard of diagnosis?	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	U	Y	Y
6. Did patients receive the same reference standard regardless of the index test result?	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	U	Y	Y
7. Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8. Was the execution of the index test described in sufficient detail to permit its replication?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
9. Was the execution of the reference standard described in sufficient detail to permit its replication?	N	Y	N	Y	N	Y	N	N	U	Y	Y	N	N	N	N	N	N
10. Were the index test results interpreted without knowledge of the results of the reference test?	U	Y	Y	Y	U	Y	U	Y	U	Y	Y	U	Y	U	Y	Y	U
11. Were the reference standard results interpreted without knowledge of the results of the index test?	U	U	U	Y	U	U	U	U	U	N	U	U	N	U	N	N	U
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	N	Y	U	Y	Y	Y	U	U	U	U	U	U	U	U	N	Y	Y

TABLE 4. QUADAS values^{28,31-33} of the SLAP lesion diagnostic accuracy studies^{4,15-19,24,25,35-43}. (Continued)

13. Were uninterpretable/intermediate test results reported?	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
14. Were withdrawals from the study explained?	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U
TOTAL # of “Yes” items	3	11	7	9	8	10	6	8	5	9	10	3	9	5	6	9	6

Y = Yes, N = No, U = Unclear

formed by 12 orthopaedic sports medicine fellows. The authors defined and interpreted the active compression test as outlined by O’Brien et al⁴². One surgeon blinded to the clinical test findings completed and interpreted all of the arthroscopic gold standard assessments. The surgical procedure was not described. Of 40 subjects, 29 athletes had type II to IV SLAP lesions on arthroscopy. The number of subjects, who tested positive on the clinical test, was not provided.

Parentis et al¹⁹ studied 132 consecutive patients in an orthopaedic surgical setting. The study provided insufficient detail to determine whether it was prospective or retrospective. Three surgeons performed a standard pre-operative evaluation including nine clinical tests for SLAP lesions. During the active compression test the raters purposely did not ask the patients to differentiate pain “on top” of the shoulder from pain “inside” the shoulder in an attempt to diminish the amount of subjectivity from this test. It was unclear how many surgeons were involved in the assessment/interpretation of the arthroscopic diagnosis or whether they were blinded to the clinical test results. The surgical procedure was not described. Arthroscopically, 40 patients were diagnosed with SLAP lesions but the specific number of patients who tested positive on the active compression test was not provided.

Guanche and Jones⁴ both illustrated and described the active compression test; pain in the shoulder was considered positive. Again, the number of patients who tested positive on the active compression test was not provided. Also, the question as to the location of pain was omitted in 27 patients causing the authors to exclude these cases from statistical analysis.

Nakagawa et al¹⁷ also did not operationally define performance or interpretation for the active compression test. Also, the number of patients who tested positive on the active compression test was not provided.

Anterior Slide Test

Kibler³⁷ introduced the anterior slide test with a study of 226 athletes in a sports medicine/orthopaedic surgical setting. The study provided insufficient detail to determine whether the study was prospective or retrospective. The study operationally defined the anterior slide test whereby the patient was examined in either standing or sitting with hands on hips and the thumbs pointing posteriorly. One of the examiner’s hands was positioned at the top of the shoulder from the posterior direction, with the last segment of the index finger extending over the anterior aspect of the acromion at the glenohumeral joint. The examiner’s other hand was placed behind the elbow and a forward and a slightly superiorly directed force was applied to the elbow and upper arm (Figure 3). The patient was asked to push back against this force. 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 a positive result. This test was also considered positive if patients reported this testing maneuver reproduced their presenting symptoms as associated with overhead activity. The study was unclear as to how many testers were involved in the clinical and surgical assessments or if arthroscopic examiner was blinded to the clinical test results. The gold standard arthroscopic examination was not described and was only used on 98 of the athletes. The clinical

test was positive in 82 patients, and 69 of these patients had superior glenoid labral tears upon surgery.

McFarland et al²⁴ defined the anterior slide test as positive if there was a click or pain deep in the shoulder thereby differing from the Kibler³⁷ definition. The anterior slide test was performed in 419 patients (381 control and 38 SLAP lesion patients). The clinical test was positive in 62 of the control patients and 3 of the SLAP lesion patients. Data on arthroscopic findings were provided above.

Nakagawa et al¹⁷ also provided no operational definition for the anterior slide test. Again, the specific number of patients who tested positive on the anterior slide test was not provided.

Parentis et al¹⁹ also included the anterior slide test and defined and interpreted the anterior slide test as outlined by Kibler. Data on arthroscopic findings are noted above but again the specific number of patients who tested positive on the clinical test was not provided.

Biceps Groove Tenderness

Guanche and Jones⁴ also studied the tenderness upon palpation of the bicipital groove. Although we can assume that this test entails the examiner palpating the region of the bicipital groove on the involved shoulder and the test being considered positive if the patient reports pain or tenderness in this region, a specific description of test performance and interpretation was not provided. Also, the number of patients who tested positive for biceps groove tenderness was not provided.

The Nakagawa et al¹⁷ study did not include an operational definition of test performance and interpretation. Again,



FIGURE 3. Anterior slide test.

no information was available on the number of patients who tested positive for biceps groove tenderness.

Biceps Load Test

Kim et al³⁸ prospectively studied 75 consecutive patients with unilateral recurrent anterior shoulder dislocations and a Bankart lesion, in an orthopaedic surgical setting. They defined this test where the patient was examined supine with the examiner seated on the involved side grasping the patient's wrist and elbow gently. The involved arm was abducted to 90° with the forearm in the supinated position. The patient was allowed to relax and an anterior apprehension test was performed. When the patient became apprehensive during the external rotation of the shoulder, external rotation was stopped. The patient was then asked to flex the elbow while the examiner resisted flexion with one hand and asked how the apprehension had changed, if at all. The test was repeated and the patient was instructed not to pull the entire upper extremity, but to bend the elbow against the examiner's resistance. The direction of the examin-

er's resistance should be on the same plane as the patient's arm so as not to change the degree of abduction and rotation of the shoulder. The forearm was kept supinated during the test. If the apprehension lessened with the resisted flexion, or if the patient was more comfortable than before the test, the test was considered negative for a SLAP lesion. If the apprehension had not changed, or if the test became more painful, the test was considered positive. Two testers were involved in the clinical testing but it was unclear from the study how many were involved in the arthroscopic assessment. It was also unclear whether the arthroscopic examiner was blinded to the clinical test results. The gold standard arthroscopic examination was not described. Of 75 patients, 22 patients had a positive biceps load test and 8 of these patients had surgically confirmed SLAP lesions. In total, 11 patients had SLAP lesions on surgery yielding three false negative clinical test results.

Biceps Load Test II

Kim et al³⁹ prospectively studied 127 patients in an orthopaedic surgical setting.

With this test the patient was supine with the examiner seated on the involved side grasping the patient's wrist and elbow gently. The involved arm was elevated to 120° and maximally externally rotated with the elbow in 90° of flexion and the forearm supinated (Figure 4). The patient was asked to flex the elbow while resisting elbow extension by the examiner. The test was considered positive if the patient reported (more) pain during the resisted elbow flexion. The test was negative if pain was not elicited by the resisted elbow flexion or if the preexisting pain during the elevation and external rotation of the arm was unchanged or diminished by the resisted elbow flexion. Two testers were involved in the clinical testing but it was unclear how many were involved in the arthroscopic assessment or whether the arthroscopic examiner was blinded to the clinical test results. The gold standard arthroscopic examination was not described. Of 127 subjects, 38 had a positive biceps load test II and 35 of these patients had type II SLAP lesions upon surgical evaluation; in total, 39 subjects had surgically confirmed SLAP lesions yielding four false negative findings.

Clunk Test

Nakagawa et al¹⁷ did not define the clunk test in their investigation. However, other sources^{1,10,46} have defined this test where the patient is supine with the examiner placing one hand on the posterior aspect of the glenohumeral joint while the other hand grasps the bicondylar aspect of the humerus at the elbow and fully abducts the arm. The examiner's hand at the posterior glenohumeral joint provides an anterior translation of the humeral head while simultaneously rotating the humerus externally with the hand holding the elbow (Figure 5). The proposed mechanism of this test is similar to the McMurray test of the knee where the examiner attempts to trap the torn labrum between the glenoid and the humeral head. A positive test is produced by the presence of a clunk or grinding sound and is indicative of a labral tear. Again, Nakagawa et al¹⁷ did not specify the number of patients who tested positive on the clunk test.



FIGURE 4. Biceps load test II.

Compression Rotation Test

McFarland et al²⁴ described the compression rotation test where the patient's shoulder is abducted 90°. A compression force is axially applied to the humerus, which is then rotated internally and externally to trap the torn labrum. Labral tears may be felt to catch and snap during the test, as meniscal tears do with the McMurray test. The compression-rotation test was performed on 303 patients (274 control and 29 SLAP lesion patients). The test was positive in 67 of the control patients and 7 of the SLAP lesion patients. Data on arthroscopic findings are provided above.

Nakagawa et al¹⁷ also did not define the compression rotation test. Also, the number of patients who tested positive on the compression rotation test was not provided.

Crank Test

Liu et al⁴¹ retrospectively studied 63 patients in an orthopaedic surgical setting. The test was performed with the patient's

involved arm elevated to 160° in the scapular plane. The arm was then loaded axially along the humerus with one of the examiner's hands while the examiner's other hand performed humeral rotation (internal rotation and/or external rotation). The test was performed with the patient in standing and supine positions. It was considered positive for a SLAP lesion if the maneuver produced pain, with or without a click, or reproduced mechanical symptoms similar to the patient's symptoms. It was unclear how many assessors were involved in the clinical or surgical tests. The surgical gold standard procedure was described in detail. It was unclear if the arthroscopic examiner was blinded to the clinical test results. In this study 31 patients had positive crank tests, whereas 32 patients were noted to have glenoid labral tears on arthroscopy.

Stetson and Templin¹⁸ interpreted the crank test as outlined by Liu et al⁴¹, but rather than testing in the supine and sitting position these authors tested in the supine or sitting position. Of their patients, 29 tested positive with the

crank test and 12 of had arthroscopically confirmed tears of the glenoid labrum.

Guanche and Jones⁴ used a picture and short description of performance and interpretation of the crank test. The test was performed with the shoulder in full abduction and with the patient only in a supine position. The crank test was considered positive with pain and/or a click. The authors did not provide information on the specific number of patients who tested positive on this clinical test.

Nakagawa et al¹⁷ again did not define the crank test or its interpretation and also did not provide data on the number of patients who tested positive.

Myers et al⁴³ interpreted the crank test as outlined by Liu et al⁴¹, but rather than testing in the supine and sitting position they tested in the supine or sitting position. Arthroscopic findings are provided above. The number of subjects positive on the crank test was not provided.

Parentis et al¹⁹ defined and interpreted the crank test as outlined by Liu et al⁴¹, with the exception of an undefined start position for the patient. Arthroscopic findings are provided above. The number of subjects positive on the crank test was not provided.

Forced Abduction Test

Nakagawa et al¹⁷ introduced this clinical SLAP lesion test whereby the patient's arm was forced into maximal abduction in the vertical position. The patient was asked about pain. The elbow was then flexed. A positive test was indicated by pain at the postero-superior aspect of the shoulder with maximal abduction that was relieved or diminished with elbow flexion. The number of patients who tested positive for the forced abduction test was not provided.

Hawkins Impingement Sign

Parentis et al¹⁹ described the Hawkins test where the patient's involved shoulder is forward flexed, adducted, and internally rotated with the elbow at 90° flexion. This test was considered positive when the patient's pain was recreated with this maneuver. Arthroscopic find-



FIGURE 5. Clunk test.

ings are provided above. The number of subjects positive on the clinical test was not provided.

Nakagawa et al¹⁷ did also not define performance and interpretation of the Hawkins impingement test nor did they provide data on how many patients tested positive on this clinical test.

Neer Impingement Sign

Parentis et al¹⁹ described the Neer test where the patient's arm is passively elevated through forward flexion. This test is considered positive when the patient's pain is reproduced with this maneuver. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Nakagawa et al¹⁷ also did not define the Neer impingement test or its interpretation. Also, the number of patients who tested positive on this clinical test was not provided.

Jobe Relocation Test

Hamner et al³⁵ prospectively investigated the role of the Jobe relocation test in 13

overhand throwing athletes in a biomechanical lab and orthopaedic surgical setting. They described the test with the patient supine and the test performed at 90°, 110°, and 120° of shoulder abduction with the involved shoulder in maximal external rotation. An anterior force and then a posterior force were applied to the proximal humerus. Pain that occurred with the anterior force and that was relieved or diminished with the posterior force was considered a positive test finding. It was unclear how many raters were involved in the clinical or surgical assessment. The gold standard arthroscopic procedure was described in moderate detail but it was unclear if the arthroscopic examiner was blinded to the clinical test results. Of 13 subjects, 6 had positive relocation tests in all positions; the other 7 had a negative test at 90° but positive tests at 110° and 120° abduction. During arthroscopy, 2 athletes had SLAP lesions confirmed and 9 subjects were described as having postero-superior labral fraying, which was not included in the diagnosis of SLAP lesions.

Guanche and Jones⁴ used a picture

and a short description of performance and interpretation of the Jobe relocation test, as described by Hammer et al¹⁹. Guanche and Jones positioned the shoulder in 90° abduction only and did not investigate the additional positions described by Hammer et al¹⁹. However in this study, specific data on the number of patients positive on the clinical test was not provided.

Nakagawa et al¹⁷ did not define the Jobe relocation test or its interpretation nor was data provided on the number of patients who tested positive on this clinical test.

Parentis et al¹⁹ defined and interpreted the Jobe relocation test as outlined by Hamner et al³⁵. Arthroscopic findings are provided above. The number of subjects who tested positive on the clinical test was not provided.

Pain Provocation Test

Mimori et al¹⁶ introduced the pain provocation test in a prospective study of 32 throwing athletes in an orthopaedic surgical setting. They defined the pain provocation wherein with the patient sitting the examiner maintained the patient's upper arm at 90–100° of abduction and then externally rotated the shoulder. The pain provocation test was performed with the forearm in two different positions: maximal pronation (Figure 6A) and supination (Figure 6B). The examiner evaluated the patient report of the severity of provoked pain during shoulder external rotation with the forearm in the two positions. Patients were asked in which position of the forearm they felt more pain. The test was considered positive when the pain was provoked only when the forearm was pronated or when the pain was more severe in this position than with the forearm supinated. It was unclear how many raters were involved for the clinical or reference tests or whether they were blinded. The reference test used was magnetic resonance arthrography (MRA) for all subjects: 15 subjects also had an arthroscopic examination. Of the subjects, 23 tested positive on the pain provocation test. Of the 15 arthroscopically evaluated, 11 were diagnosed with SLAP lesions and all of these

11 subjects had a positive pain provocation test. The authors provided diagnostic utility statistics based on MRA findings only; we also calculated statistics based on the arthroscopic findings.

Parentis et al¹⁹ defined and interpreted the pain provocation test as outlined by Hamner et al³⁵. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Passive Compression Test

Kim et al²⁵ investigated the passive compression test on 61 consecutive patients in an orthopaedic surgical setting. They defined the test with the patient side lying with the involved shoulder up and the examiner standing behind the patient. The examiner stabilized the involved shoulder by grasping the acromioclavicular joint region with a cupped hand. The examiner grasped the patient's bent elbow to position the shoulder in external rotation with 30° abduction (Figure 7A). The examiner then provided a proximal compression force while extending the humerus (Figure 7B). The test was considered positive if pain or a painful click was elicited in the glenohumeral joint. It was unclear if this study was prospective or retrospective. Two physicians completed the clinical assessments that included the passive compression test. The gold standard arthroscopic procedure was described in detail, but it was unclear how many raters were involved in the arthroscopic assessment of the patients or whether there was any blinding in this study. Of 61 subjects, 33 were arthroscopically diagnosed with SLAP lesions, whereas 31 patients had a positive passive compression test.

Resisted Supination External Rotation Test

Myers et al⁴³ studied the resisted supination external rotation test, wherein the patient was supine on the examination table with the scapula near the edge of the table. The examiner supported the patient's involved arm at the elbow and hand while standing at the patient's side. The involved limb was then placed in the

starting position with the shoulder abducted to 90°, the elbow flexed to 65-70°, and the forearm in neutral or slight pronation. The patient was asked to attempt to supinate the hand with maximal effort against examiner resistance, while the shoulder was gently externally rotated to maximum. The patient was asked to describe the symptoms at maximum external rotation. The test was considered positive if the patient had anterior or deep shoulder pain, clicking or catching in the shoulder, or reproduction of the symptoms that occurred during throw-

ing. The test was deemed negative if the patient had posterior shoulder pain, apprehension, or no pain. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

SLAPprehension Test

Berg and Ciullo³⁶ retrospectively investigated 66 consecutive SLAP-lesion patients in an orthopaedic/sports medicine setting. For the SLAPprehension test, the patient was seated or standing



FIGURE 6A. Pain provocation test (start with forearm pronated).



FIGURE 6B. Pain provocation test (start with forearm supinated).

with the involved arm horizontally abducted across the chest with the shoulder internally rotated, the elbow extended, and the forearm pronated (thumb down). The test was then repeated with the shoulder externally rotated and the forearm supinated (thumb up). If pain was felt in the bicipital groove, with or without an audible or palpable click, in the first test position and was diminished or removed in the second test position, the test was considered positive. It is unclear how many raters were involved in the clinical or arthroscopic assessment. The surgical assessment procedure was not described. It was also unclear whether raters were blinded in this study. All patients were arthroscopically diagnosed with SLAP lesions; 54 patients had a positive clinical test.

Speed Test

Guanche and Jones⁴ used a picture and a short description of test performance and interpretation. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Holtby and Razmjou¹⁵ investigated 152 consecutive patients in a prospective study in an orthopaedic surgical setting. The Speed test was referenced, but not described. The arthroscopic procedure was described in detail. It was unclear on how many assessors were involved in the clinical test or surgical assessment. It was also unclear if the arthroscopic examiner was blinded to the clinical test results. Of 22 patients arthroscopically diagnosed with SLAP lesions, 14 had a positive Speed test¹⁵.

Nakagawa et al¹⁷ again did not define the performance or interpretation of this clinical test. The number of patients positive on the Speed test was also not provided.

Parentis et al¹⁹ defined the Speed test as the patient fully supinating the forearm and extending the elbow while resisting forward flexion of the involved shoulder. The test was considered positive if the patient's pain was reproduced. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Yergason Test

Parentis et al¹⁹ also included the Yergason test. The test was defined as the patient resisting forearm supination with the involved elbow flexed at 90°. The test was considered positive if the patient's pain was reproduced particularly in the region of the bicipital groove. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Holtby and Razmjou¹⁵ referenced but did not describe the Yergason test. Arthroscopic findings are provided

above; 15 patients had a positive Yergason test.

Guanche and Jones⁴ used a picture to illustrate the test and a short description of test performance and interpretation. The test was considered positive if unusual pain in the biceps tendon was produced during the maneuver. Arthroscopic findings are provided above. The number of subjects positive on the clinical test was not provided.

Nakagawa et al¹⁷ also did not define the Yergason test or its interpretation, nor was the number of patients positive on the Yergason test provided.



FIGURE 7A. Passive compression test start position.

FIGURE 7B. Passive compression test end position.



Multi-Test Regimens

Liu et al⁴⁰ retrospectively studied 54 patients in an orthopaedic/sports medicine setting. They investigated five clinical tests: the anterior apprehension maneuver, Jobe relocation test, load and shift test, sulcus sign, and crank test. Liu et al defined the anterior apprehension maneuver, the Jobe relocation test, and the crank test. However, for the Jobe relocation test, Liu et al only tested in the 90° abduction position. The authors described the load and shift test with the patient seated or supine and the arm positioned in 20° of abduction. One of the examiner's hands stabilized the patient's scapula while the other hand grasped the proximal humerus and loaded the humeral head medially into the glenoid while performing anterior-to-posterior translation. The sulcus sign was undefined in the study. However, it has been described elsewhere as the examiner grasping the arm below the elbow and pulling the arm distally^{19,44,46}. The test is considered positive if a sulcus is noted to be present below the acromion^{19,44,46}. It was unclear as to how many testers were involved in the physical examination and/or the surgical assessments or whether the arthroscopic examiner was blinded to the clinical test results. SLAP lesions were clearly defined in this study, as was the surgical procedure. In this study, 41 subjects were arthroscopically diagnosed with SLAP lesions, whereas 37 had positive clinical tests.

Guanche and Jones⁴ also statistically analyzed the clinical SLAP lesion tests included in their study (i.e., the active compression test, apprehension, and Jobe relocation test) in pairs and/or combinations, using AND and OR Boolean operators to establish test combinations. They noted that no combination of these three tests was of value for the clinical diagnosis of SLAP lesions. Gold standard arthroscopic findings are provided above.

Discussion

As noted in the introduction, the goal for this article is to provide a current best-evidence synthesis with regard to

physical examination tests used for the diagnosis of SLAP lesions. Central to achieving this stated goal is an evaluation of the research validity of the articles retrieved. Research validity is the extent to which conclusions of a study are believable and useful⁴⁷. There are three areas specific to these diagnostic utility studies on physical examination tests for SLAP lesions where research validity can be threatened: construct validity, external validity, and statistical conclusion validity. In addition to this narrative review, we will use the implications of the systematic review using the QUADAS criterion list to provide a current best-evidence summary.

Construct Validity

A construct is an artificial framework that is not directly observable. The main threat to construct validity in diagnostic utility studies is the discrepancy between the construct as labeled and the construct as implemented⁴⁷. In the studies retrieved, this discrepancy takes various forms.

SLAP lesion as a pathology requiring surgery

Various authors^{8,10,11,13,19,21} have noted that different types of SLAP lesions seem to respond to different types of interventions. Lesions that disrupt the biceps attachment more commonly require surgical intervention for optimal restoration of function^{8,10,11,13,19}. SLAP lesions that are complex, such as those involving rotator cuff tears, those associated with glenohumeral joint instability, or those with concomitant biceps tendon pathology, may respond to different management strategies than isolated SLAP lesions^{8,10,13,14}. The main, although not verbalized, construct in nearly all of the studies retrieved (except for the studies by McFarland et al²⁴, Myers et al⁴³, and Stetson and Templin¹⁸) is that all SLAP lesions require surgical intervention.

Basic science evidence with regard to the extent and localization of vascularization of the glenoid labrum is contradictory with some authors describing only vascularization of the peripheral

labrum that in addition is more extensive posteriorly and inferiorly than (anterior) superiorly⁴⁸, whereas other authors have described more global vascularization⁴⁹. Vascularization and thereby healing potential is an obvious prerequisite for successful conservative management. Some authors have suggested that type I SLAP lesions do not necessarily require surgical intervention^{10,21}. Expert opinion and equivocal basic science evidence indicate that some SLAP lesions may benefit from non-surgical management.

Relevant to a possible subgroup of patients with SLAP lesions that might be amenable to conservative management is the consideration that diagnostic utility statistics for type-specific clinical tests would allow the clinician to differentiate within SLAP lesions those best managed conservatively and surgically. Mirkovic et al⁶ categorized clinical tests into two groups. One proposed group consisted of tests that compress the bicipital-labral complex with or without humeral head translation, e.g., the crank, clunk, compression-rotation, anterior slide, and load and shift tests. The second proposed group comprised those tests that may or may not compress but add tension to the biceps complex: active compression, SLAPprehesion, Speed, biceps load I and II, and pain provocation tests. Mirkovic et al⁶ suggested that the SLAP compression tests might be more accurate for type III and IV SLAP lesions and that those that involve the biceps tendon might be more accurate for type II and IV SLAP lesions. However, type-specific diagnostic utility statistics have yet to be reported. Research-based type-specific clinical diagnosis would enable future studies to investigate more homogenous subgroups allowing for the development of clinical prediction rules and randomized controlled trials of conservative and surgical management using the same more homogenous groups.

SLAP lesion as a diagnosis guiding conservative management

The intent of the best-evidence summary given below is to provide the clinical

cian with a research-based method of confidently ruling in or ruling out the diagnosis of SLAP lesion. In case a clinician is able to rule in this diagnosis, ideally this knowledge would somehow guide further conservative management. However, within a mechanism-based classification system a clinician diagnoses joint dysfunction. Paris and Loubert⁵⁰ defined joint dysfunction as the presence of hypomobility, hypermobility, or aberrant motion. The SLAP lesion could consist of any combination of these motion abnormalities: torn/loose labral tissue may block and limit motion, and/or allow excessive motion, and/or alter normal arthrokinematic motion at the glenohumeral interface. The construct as implemented in these studies is that of a patho-anatomical abnormality. With its specific effect on presenting dysfunction unknown as would be relevant in a mechanism-based classification system or in the absence of clinical prediction rules as used in a treatment-based classification system, the relevance for the clinician of a confident clinical diagnosis of SLAP lesion remains unknown.

Reference test: Definition SLAP lesion

All studies retrieved used arthroscopic findings as the gold standard although Mimori et al¹⁶ provided diagnostic utility statistics based on MRA findings. Either reference test depends on the operational definition for what raters consider a SLAP lesion or a specific type of this lesion. However, the studies retrieved show that only a broad consensus exists on these operational definitions. Guanche and Jones⁴ and Kim et al²⁵ used the classical definition for a SLAP lesion: labral disruptions involving the superior labrum from the 10 to 2 o'clock position on the glenoid face²⁰. In contrast, Kibler³⁷ defined a SLAP lesion when occurring from the 9.30 to 2.30 o'clock position on the glenoid face. Most of the diagnostic studies did not define their SLAP lesion descriptions or criteria. McFarland et al²⁴, Myers et al⁴³, and Stetson and Templin¹⁸ defined type I SLAP lesions as normal and catego-

rized patients with type I SLAP lesions with those who had no SLAP lesions. Devaluation of the reference test by an unclear operational definition of the reference test scoring criteria calls diagnostic utility statistics into question.

Reference test: Imaging or arthroscopy

Arthroscopic findings are considered the gold standard test for the diagnosis of SLAP lesions^{2,3,5-13,15-19}. The orthopaedic literature^{3,8,13} has questioned the diagnostic accuracy of imaging tests (including MRI, CT, and ultrasonography). Burkhart and Morgan³ reported that MRI scans, including gadolinium-enhanced MRI, usually do not show SLAP lesions. Luime et al⁷ noted that approximately 10-20% of patients with a normal reading on shoulder MRI or ultrasonography may still have a labral tear. False positive imaging results also occur, typically due to normal congenital labral variants as described above^{8,11}. Our calculation of diagnostic utility statistics for the Mimori et al¹⁶ shown in Table 1 suggest that MRA may also be a less than optimal reference test for the diagnosis of SLAP lesions. This calls into question the findings of this particular study.

Reference test: Arthroscopy

Although arthroscopy is the acknowledged gold standard test, approximately half of the studies did not describe the procedure in sufficient detail to permit replication^{16,17,19,37-39,42,43}. By not having such descriptions (patient position, location of portals, order of visualization, etc.), it is not clear if cross-comparisons between the studies are valid. Only two studies^{25,41} described their arthroscopic diagnostic procedure in enough detail to allow replication.

Lack of rater blinding

Cook et al²⁸ explained how an unblinded rater may have a predisposition to select a positive or negative test interpretation based on knowledge of the results of other test findings, additional diagnostic information, past experience, or per-

sonal preference. Only two of the studies^{17,43} blinded the surgeons to the results of the clinical diagnostic tests. In no study was it clear whether raters of the clinical test were blinded to other findings. This means that the construct as labeled might be the diagnostic utility of a single physical examination test but access to other data makes the construct as implemented diagnostic utility of the whole examination process including said clinical test.

Rater preference might have played a role in studies introducing a new clinical test for SLAP lesions. Unwittingly the researchers might be biased to producing convincing statistics as to why their new test should be added to clinical examination for SLAP lesions. This might also explain to some extent why on the publication describing diagnostic utility of a new clinical test the accuracy statistics always are more favorable than in follow-up validation studies. Many of the SLAP clinical tests, when initially introduced to the scientific literature, demonstrate clinically relevant positive likelihood ratios (Table 1: the active compression test (66.66), the anterior slide test (9.2), the biceps load test (29.32), the biceps load test II (26.38), the crank test (13.52), the pain provocation test (10.0), the passive compression test (5.72)) yet when subsequently investigated by other groups these tests have not produced such clinically relevant likelihood ratios (0.5-2.3). It is important to note that a clinical test that shows sufficient diagnostic utility in more than one study is more relevant for clinical diagnosis than a test that has only shown sufficient diagnostic utility in one study. None of the research proposing new clinical SLAP lesion tests blinded the surgeons to the clinical preoperative findings.

Time delay between clinical and reference test

Marked variability existed among the studies with regard to the delay between performance of the clinical test(s) and the diagnostic arthroscopic surgery. Guanche and Jones⁴ and Kim et al³⁸ minimized this time delay to 24 hours or

less. In the Myers et al⁴³ study this delay was 1 week or less, McFarland et al²⁴ had 4 weeks or less between clinical and reference test. In the Liu et al⁴⁰ study the delay was up to 48 weeks. Much of the research^{15-19,25,37,39,41,42} provided insufficient detail on the length of this delay. A protracted delay makes it possible, and even probable, that the SLAP lesion may change (for worse with additional wear and tear or for the better due to natural healing and possible other conservative management) leading to an invalid categorization and correlation with the clinical test at the time of surgery.

External validity

External validity covers the extent to which the research results can be generalized beyond the internal specifications of the study sample to other subjects, settings and times⁵¹.

Spectrum bias

Diagnostic utility studies frequently suffer from spectrum bias. This bias occurs when the subjects studied are not representative of the subjects to which the results of the diagnostic utility study are going to be applied²⁸. There are two reasons for spectrum bias in the diagnostic utility studies discussed above: retrospective study design and study setting.

Of the 17 studies retrieved, 10 described prospective research, 4 were retrospective, and 3 provided insufficient detail to determine the nature of the study. Whereas the prospective studies include subjects that may have a positive clinical test but not a positive gold standard test, retrospective tests only have subjects who were shown to have SLAP lesions on arthroscopic evaluation. Needless to say, prospective studies, will have a population that is more representative of the populations seen in physical therapy clinical practice. Study setting for all studies retrieved were sports medicine/orthopaedic surgical settings. Using patients seen in a secondary-care medical specialist environment limits external validity with regard to the primary-care level settings, such as physical therapy environments.

High prevalence of a pathology or dysfunction among study subjects reduces the opportunity to detect false positive and true negative results. This inflates test sensitivity and at the same time undervalues test specificity, when the test is applied to patient populations with a lower prevalence⁷. Jones and Galuch⁵ noted how prevalence rates have an even greater impact when small populations are studied, as is the case in most of the studies retrieved. Prevalence rates also have a substantial impact on positive and negative predictive values and thus strongly influence clinical decisions from clinical tests⁵. We noted earlier how prevalence numbers for SLAP lesions in the literature cover quite a range: 6–26%^{10,13,14,21-23}. Prevalence rates for the literature retrieved in this review demonstrate an even greater range with SLAP lesions present in 9–73% of the subjects studied⁵. Extrapolating results from these diagnostic utility studies to a different population, such as a typical physical therapy setting, thereby becomes problematic⁵. We should consider that clinical findings in a more general setting have lower sensitivity but higher specificity than suggested in the available literature⁷.

Type of SLAP lesion

Another noted variation among the studies was the specific type of SLAP lesion investigated. Kim et al²⁵ and Guanche and Jones⁴ investigated the diagnostic accuracy for types I-IV SLAP lesions. Liu et al⁴⁰ provide unclear operational definitions for interpretation of arthroscopic findings and referred to grades rather than the classification in types. Holtby and Razmjou¹⁵ investigated the diagnostic accuracy of clinical tests for type II and IV lesions. Parentis et al¹⁹ investigated the diagnostic accuracy for types I and II lesions only. Kim et al³⁸ investigated the diagnostic accuracy for type II lesions in recurrent anteriorly dislocated shoulders. Kim et al³⁹ investigated the diagnostic accuracy for isolated type II lesions. The diagnostic utility data for each study is only applicable to the type of SLAP lesion investigated in that particular study and cannot

be generalized to all types of SLAP lesions.

Operational definition of test performance and interpretation

An additional region of variability in the studies was with regard to operational definition of performance and/or interpretation of the clinical tests studied. Nakagawa et al¹⁷ included minimal to no definitions of the clinical tests under investigation and their study was, therefore, discarded from further consideration.

At times there was great variability in the operational definitions of test performance among the studies. Liu et al⁴¹ introduced the crank test and defined it as a test to be performed in the standing and supine positions. Stetson and Templin¹⁸ and Myers et al⁴³ tested in the supine or sitting position. Guanche and Jones⁴ only tested in supine. Parentis et al¹⁹ used an undefined start position for the patient. Nakagawa et al¹⁷ did not clarify test performance at all. In contrast, at times there was good consistency for operational definitions between studies: 6 of the 8 studies that researched the active compression test had identical definitions on the test and its interpretation. For diagnostic utility data to apply to the clinical situation, the clinician needs to replicate both test performance and interpretation.

Indeterminate findings

None of the 17 studies mention unclear or indeterminate test results. It is difficult to imagine that such a volume of clinical tests and arthroscopic procedures produced such clear definitive findings.

Statistical Conclusion Validity

Statistical conclusion validity concerns the use of statistical procedures for analyzing the data, leading to invalid conclusions about relationship between the index and reference tests. All studies reviewed used appropriate statistical measures to determine concurrent criterion-related validity. A number of the studies

did not provide sufficient data for the reader to calculate the clinical test diagnostic utility score (5/17 studies). Many studies provided the diagnostic utility statistics but did not provide any raw data allowing the reader to check the data provided (10/17 studies).

With regard to interpretation and clinical use of diagnostic utility statistics, we need to consider that retrospective research works backwards from subjects who were arthroscopically diagnosed with SLAP-lesions. In addition to issues related to spectrum bias, this devaluates the study because of statistical conclusion validity concerns due to the inherent decreased level of control as compared to prospective studies over the variables of operational definitions and the reliability with which the data were collected⁵¹.

Considering the varied time delay from clinical testing to arthroscopic surgery it is also interesting to note that out of 9 prospective studies^{4,15,18,19,25,38,39,42,43} (5 of which investigated consecutive patients^{4,15,19,25,38}) there were no subject withdrawals reported.

Systematic Review: QUADAS Criterion List

Table 4 provides the QUADAS scores for the included studies. The range (3–10 out of a possible score of 14) of these scores indicates the varied methodological quality of the studies. Six of the eight studies that scored $\geq 9/14$ were published during or after 2002 indicating promising improvement in the methodological quality of diagnostic utility studies for clinical tests for the diagnosis of SLAP lesions over the last five years.

Of note: The QUADAS scores in this review vary from the QUADAS scores allocated by Cook and Hegedus³⁴. Cook and Hegedus scored 14 of the studies included here: their QUADAS scoring totals are included in Table 5. These authors did not provide data on QUADAS item scoring. Of these 14 studies only the Hamner et al³⁵ study received a similar score. In the recent systematic review discussed in the introduction, Hegedus et al²⁷ also used the QUADAS tool for the methodological

quality assessment of studies on diagnostic accuracy of physical examination tests for SLAP lesions. The fact that there were again discrepancies in the QUADAS scores assigned in this review as compared to the QUADAS scores assigned in the Cook and Hegedus study³⁴ (Table 5) further puts the reliability—and thereby also the validity—of this tool in question.

Whiting et al³¹ provided a users' guide with procedures for scoring on each item in their tool. In an interrater reliability study, Whiting et al³² reported a median of 90% agreement for the QUADAS individual items scoring; variability in the scoring ranged between 50–100%. Score differences between Cook and Hegedus³⁴, Hegedus et al²⁷, and the current study seem to indicate the need for further study into the reliability of this tool and, for now, careful interpretation of the scores provided by the QUADAS tool.

Best-Evidence Summary

Although the narrative review above discussed many issues with the research validity of the studies retrieved, there was only one study that we felt we needed to discard due to what we considered a fatal flaw: the lack of operational definitions with regard to test performance and interpretation in the study by Nakagawa et al¹⁷ did not allow for test replication and clinical interpretation. We also discarded the diagnostic utility findings based on the MRA reference test provided by Mimori et al¹⁶.

The QUADAS tool and the cut-off score of 10/14 provided by Cook and Hegedus³⁴ allowed us to discard in total 14 studies^{4,11,15-18,24,25,36-39,41-43} due to their lack of methodological quality including the two noted above. However, we should note that the narrative review uncovered methodological weaknesses in all three remaining studies.

Guanche and Jones⁴, Kim et al²⁵, and McFarland et al²⁴ did not blind the surgeon as the reference test administrator from the clinical test findings. Nor did any of the studies provide data on the number of surgeons involved in the arthroscopic procedures. Two of the re-

maining studies^{4,25} had moderate sample sizes (59 and 61 subjects, respectively) and, therefore, their rather high prevalence numbers (56% and 54%, respectively) may have inflated test sensitivity and undervalued test specificity. Of the two studies that were follow-up studies of multiple clinical SLAP lesion tests, Guanche and Jones did not provide any of the raw data collected and the McFarland et al² study categorized type I SLAP lesions most likely amenable to conservative management in the control group. Kim et al²⁵ provided insufficient data on the duration of the time delay between the clinical tests and the surgical diagnostic procedure. This same study²⁵ also was an original study for a newly introduced clinical SLAP lesion test and so further investigations are warranted on this test prior to integrating it wholeheartedly into clinical practice. Tables 6 and 7 collate the diagnostic utility data from the methodologically sound studies as determined by the QUADAS criteria ($\geq 10/14$)³⁴.

Although provided in the various tables, the statistics of accuracy and predictive values are less relevant to our best-evidence synthesis²⁸. The accuracy of a diagnostic test provides a quantitative measure of its overall value, but because it does not differentiate between the diagnostic value of positive and negative test results, its value with regard to the diagnostic decision-making process is minimal. The prevalence in the clinical population being examined with a specific test has to be identical to the prevalence in the study population from which the predictive values were derived before we can justifiably use predictive values as a basis for diagnostic decisions. Considering the issue of spectrum bias discussed above, the usefulness is limited to those situations where we can justifiably make assumptions on similarity of prevalence, allowing us to virtually disregard these statistical data in our diagnostic decision-making process²⁸.

Likelihood ratios can be either positive or negative. A positive likelihood ratio indicates a shift in probability favoring the existence of a disorder if the test is found to be positive. Conversely, a negative likelihood ratio indicates a shift

TABLE 5. Total number of “yes” items scored with the QUADAS criteria for the literature retrieved.

Item	Berg & Ciullo ³⁶	Guanche & Jones ⁴	Hammer et al ³⁵	Holtby & Razmjou ¹⁵	Kibler ³⁷	Kim et al ²⁵	Kim et al ³⁸	Kim et al ³⁹	Liu et al ⁴⁰	Liu et al ⁴¹	McFarland et al ²⁴	Mimori et al ¹⁶	Myers et al ⁴³	Nakagawa et al ¹⁷	O’Brien et al ⁴²	Parentis et al ¹⁹	Stetson & Templin ¹⁸
TOTAL number of “yes” items	3	11	7	9	8	10	6	8	5	9	10	3	9	5	6	9	6
TOTAL number of “yes” items according to Cook & Hegedus ³⁴	---	12	7	12	6	---	9	10	---	11	11	7	8	10	3	5	10
TOTAL number of “yes” items according to Hegedus et al ²⁷	8	12	7	11	7	---	9	11	---	11	11	7	8	9	5	10	11

TABLE 6. Diagnostic utility SLAP lesion clinical tests from methodologically sound studies^{4,24,25}.

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Anterior apprehension maneuver							
• Guanche & Jones ⁴	----	0.40	0.87	0.90	0.33	3.077	0.722
Active compression test							
• Guanche & Jones ⁴	----	0.63	0.73	0.87	0.40	2.33	0.507
• McFarland et al ²⁴	0.54 (221/409)	0.47 (18/38)	0.55 (203/371)	0.10 (18/186)	0.91 (203/223)	1.044	0.964
Anterior slide test							
• McFarland et al ²⁴	0.77 (322/419)	0.08 (3/38)	0.84 (319/381)	0.05 (3/65)	0.90 (319/354)	0.50	1.095
Biceps groove tenderness							
• Guanche & Jones ⁴	----	0.44	0.40	0.69	0.19	0.733	1.4
Compression rotation test							
• McFarland et al ²⁴	0.71 (214/303)	0.24 (7/29)	0.76 (207/274)	0.09 (7/74)	0.90 (207/229)	1.0	1.0
Crank test							
• Guanche & Jones ⁴	----	0.40	0.73	0.82	0.29	1.48	0.822
Jobe relocation test							
• Guanche & Jones ⁴	----	0.44	0.87	0.91	0.34	3.385	0.643
Passive compression test							
• Kim et al ²⁵	0.836 (51/61)	0.818 (27/33)	0.857 (24/28)	0.871 (27/31)	0.80 (24/30)	5.72	0.212
Speed test							
• Guanche & Jones ⁴	----	0.18	0.87	0.80	0.26	0.138	0.943
Yergason test							
• Guanche & Jones ⁴	----	0.09	0.93	0.80	0.25	0.978	1.2857

TABLE 7. Diagnostic utility multi-test regimens from methodologically sound studies (modified from Guanche & Jones⁴).

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
2 positive tests							
• Jobe & O'Brien		0.41	0.91	0.93	0.34	4.56	0.648
• Jobe & apprehension		0.38	0.93	0.94	0.33	5.43	0.667
• O'Brien & apprehension		0.38	0.82	0.86	0.31	2.11	0.756
3 positive tests							
• Jobe, O'Brien & apprehension		0.34	0.91	0.92	0.32	3.78	0.725

TABLE 8. Interpretation of likelihood ratios on changes from pre-test to post-test probability⁵².

	Numerical Value	Change from pre-test probability to post-test probability
Positive likelihood ratio	> 10	Large and often conclusive
	Between 5–10	Moderate change
	Between 2–5	Small but sometimes important
	Between 1–2	Small and rarely important
Negative likelihood ratio	Between .5–1	Small and rarely important
	Between 0.2–0.5	Small but sometimes important
	Between 0.1–0.2	Moderate change
	< 0.1	Large and often conclusive

in probability favoring the absence of a disorder if the test is found to be negative. Table 8 provides the shifts in probability that a patient does or does not have a particular disorder given a positive or negative test associated with specific range of positive and negative likelihood ratios⁵². Likelihood ratios summarize data on sensitivity and specificity and are considered the most useful statistical measures to express criterion-related validity. However, this review only yielded the passive compression test as a test demonstrating as much as a moderate negative likelihood ratio (0.21) indicating the possible significance of a negative finding on this test for ruling out a SLAP lesion. The passive compression test also demonstrated at best a moderate positive likelihood ratio (5.72), indicating the possible significance of a positive test for diagnosing a SLAP lesion. The absence of clinically clearly relevant likelihood ratio values forced greater emphasis on sensitivity and specificity data in constructing a current best-evidence summary.

Relevant to our current best-evidence synthesis based on these remaining three studies is the fact that physical examination tests that demonstrate high sensitivity are clinically useful screening tools in that they can be used for ruling out a selected diagnosis. With a highly sensitive test, there are few false negatives. On the other hand, highly specific tests are appropriate for “ruling in” a finding because the likelihood of a false positive finding is low. This ability of highly sensitive and highly specific tests to rule out or rule in a condition, respectively, is captured in a mnemonic:

- **SnNOUT:** With highly **S**ensitive tests, a **N**egative result will rule a disorder **OUT**.
- **SpPIN:** With highly **S**pecific tests, a **P**ositive result will rule a disorder **IN**.

With regard to a best-evidence synthesis, we therefore seek to identify highly sensitive and specific tests to rule out or rule in a diagnosis, respectively²⁸.

Based on the results of the three remaining higher methodological quality studies, only the passive compression test has a sufficiently high sensitivity of 0.82. This means that a negative result on the passive compression test is clinically significant in that it allows the therapist to confidently rule out a SLAP lesion. Six tests have demonstrated high specificity: the anterior apprehension maneuver (0.87), the anterior slide test (0.84), the Jobe relocation test (0.87), the passive compression test (0.86), the Speed test (0.87), and the Yergason test (0.93). Two multi-test regimens also demonstrate high specificity (when both tests are positive): the Jobe relocation test and the active compression test (0.91), and the Jobe relocation test and the anterior apprehension maneuver (0.93). A positive result on these tests, or multi-test regimens, is significant in that it provides the therapist with research-based confidence in diagnosing a SLAP lesion. It is worth noting that the data providing support for the passive compression test come from a study²⁵ that introduces the test to

the medical community and that this test has not yet been investigated for its diagnostic value in additional studies.

In summary, research evidence at this time supports the inclusion of the following tests with the following interpretation in a physical examination format for the diagnosis of SLAP lesion:

- A negative finding on the passive compression test to rule out a SLAP lesion.
- A positive finding on the anterior apprehension maneuver, the anterior slide test, the Jobe relocation test, the passive compression test, the Speed test, and the Yergason test or on a combination of positive findings on the Jobe relocation test and the active compression test or the Jobe relocation test and the anterior apprehension maneuver to rule in a SLAP lesion; the greatest value should be placed on a positive finding on the passive compression test due to the strength of its positive likelihood ratio in the diagnostic utility data.

Limitations

Limitations of this systematic literature review and the subsequent best-evidence summary involve the literature search strategy presented. Articles may have been missed based on the omission of certain search phrases and key words. Limiting the search to English-language articles only may have omitted relevant articles written in another language. The final limitations concern the QUADAS methodological quality assessment tool. Some of our findings question this tool's reliability and, therefore, validity. In addition, there is no consensus-based cut-off value for the QUADAS tool even though we implemented the cut-off value used by Cook and Hegedus³⁴.

Conclusion

Current best evidence indicates that a negative finding for the passive compression test provides the therapist with the greatest evidence-based confidence that a SLAP lesion is absent. A positive

finding on the anterior apprehension maneuver, the anterior slide test, the Jobe relocation test, the passive compression test, the Speed test, and the Yergason test or on a combination of positive findings on the Jobe relocation test and the active compression test or the Jobe relocation test and the anterior apprehension maneuver provides the therapist with the research-based confidence required to rule in a SLAP lesion. For ruling in a SLAP lesion, the greatest diagnostic value should likely be placed on a positive finding on the passive compression test.

This review of the literature and critical analysis of research validity has also provided directions for future research. Most importantly, future research needs to provide clear operational definitions of test performance and interpretation of test findings. Also, we strongly suggest that description and interpretation of arthroscopic findings as the appropriate gold standard test needs to be standardized to a greater degree. A third category of indeterminate findings on the clinical tests studied would provide for more realistic diagnostic utility data. Also important is that future studies be done prospectively in primary-care settings, thereby decreasing spectrum bias but also allowing the research to produce information more relevant to such settings. Solely studying physical examination tests or test clusters also would provide for data on diagnostic utility that were not influenced by other findings in history and physical examination. Most important is likely the need to produce SLAP lesion type-specific diagnostic utility statistics. Research-based type-specific clinical diagnosis would enable future studies to investigate more homogenous subgroups allowing for the development of clinical prediction rules and randomized controlled trials of conservative and surgical management using the same more homogenous groups. Finally, this review has shown that further study is required with regard to the reliability and validity of the QUADAS methodological quality assessment tool prior to its more widespread use in systematic review of diagnostic utility studies.

REFERENCES

1. Andrews JR, Carson WG, McLeod WD. Glenoid labrum tears related to the long head of the biceps. *Am J Sports Med* 1985;13:337-341.
2. 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;19:531-539.
3. Burkhart SS, Morgan C. SLAP lesions in the overhead athlete. *Orthop Clin North Am* 2001;32:431-441.
4. Guanche CA, Jones DC. Clinical testing for tears of the glenoid labrum. *Arthroscopy* 2003;19:517-523.
5. Jones GL, Galluch DB. Clinical assessment of superior glenoid labral lesions. *Clin Orthop* 2007;455:45-51.
6. Mirkovic M, Green R, Taylor N, Perrot M. Accuracy of clinical tests to diagnose superior labral anterior and posterior (SLAP) lesions. *Phys Ther Rev* 2005;10:5-14.
7. Luime JJ, Verhagen AP, Miedema HS, et al. Does this patient have an instability of the shoulder or a labrum lesion? *JAMA* 2004;292:1989-1999.
8. Nam EK, Snyder SJ. The diagnosis and treatment of superior labrum, anterior and posterior (SLAP) lesions. *Am J Sports Med* 2003;31:798-810.
9. Swearingen JC, Mell AG, Langenderfer J, LaScala S, Hughes RE, Kuhn JE. Electromyographic analysis of physical examination tests for type II superior labrum anterior-posterior lesions. *J Shoulder Elbow Surg* 2006;15:576-579.
10. Wilk KE, Reinold MM, Dugas JR, Arrigo CA, Moser MW, Andrews JR. Current concepts in the recognition and treatment of superior labral (SLAP) lesions. *J Orthop Sports Phys Ther* 2005;35:273-291.
11. Parentis MA, Mohr KJ, El Attrache NS. Disorders of the superior labrum: Review and treatment guidelines. *Clin Orthop* 2002;400:77-87.
12. Mileski RA, Snyder SJ. Superior labral lesions in the shoulder: Pathoanatomy and surgical management. *J Am Acad Orthop Surg* 1998;6:121-131.
13. D'Alessandro DF, Fleischli JE, Connor PM. Superior labral lesions: Diagnosis and management. *J Athl Train* 2000;35:286-292.
14. Kim TK, Queale WE, Cosgarea AJ, McFarland EG. Clinical features of the different types of SLAP lesions: An analysis of one

- hundred and thirty-nine cases. *J Bone Joint Surg Am* 2003;85:66–71.
15. 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).
 16. Mimori K, Muneta T, Nakagawa T, Shinomiya K. A new pain provocation test for superior labral tears of the shoulder. *Am J Sports Med* 1999;27:137–142.
 17. Nakagawa S, Yoneda M, Hayashida K, Obata M, Fukushima S, Miyazaki Y. Forced shoulder abduction and elbow flexion test: A new simple clinical test to detect superior labral injury in the throwing shoulder. *Arthroscopy* 2005;21:1290–1295.
 18. Stetson WB, Templin K. The crank test, the O'Brien test, and routine magnetic resonance imaging scans in the diagnosis of labral tears. *Am J Sports Med* 2002;30:806–809.
 19. Parentis MA, Mohr KJ, Yocum LA. An evaluation of the provocative tests for superior labral anterior posterior lesions. *Am J Sports Med* 2006;34:265–268.
 20. Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ. SLAP lesions of the shoulder. *Arthroscopy* 1990;6:274–279.
 21. Huijbregts PA. SLAP lesions: Structure, function, and physical therapy diagnosis and treatment. *J Manual Manipulative Ther* 2001;9:71–83.
 22. Higgins LD, Warner JP. Superior labral lesions: Anatomy, pathology, and treatment. *Clin Orthop* 2001;390:73–82.
 23. Musgrave DS, Rodosky MW. SLAP lesions: Current concepts. *Am J Orthop* 2001;30:29–38.
 24. McFarland EG, Kim TK, Savino RM. Clinical assessment of three common tests for superior labral anterior-posterior lesions. *Am J Sports Med* 2002;30:810–815.
 25. Kim YS, Kim JM, Ha KY, Choy S, Joo MW, Chung YG. The passive compression test: A new clinical test for superior labral tears of the shoulder. *Am J Sports Med* 2007;35:1489–1494.
 26. 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 Technol Assess* 2003;7(29), iii, 1–166.
 27. Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT, Cook C. Physical examination tests of the shoulder: A systematic review with meta-analysis of individual tests. *Br J Sports Med* 2007; doi: 10.1136/bjism.2007.038406.
 28. Cook C, Cleland J, Huijbregts P. Creation and critique of studies of diagnostic accuracy: Use of the STARD and QUADAS Methodological Quality Assessment Tools. *J Manual Manipulative Ther* 2007;15:93–102.
 29. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003;326:41–44.
 30. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clin Chem* 2003;49:7–18.
 31. Whiting P, Rutjes AWS, 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 Med Res Methodol* 2003;3:25.
 32. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006;6:9.
 33. Tatsioni A, Zarin DA, Aronson N, et al. Challenges in systematic reviews of diagnostic technologies. *Ann Intern Med* 2005;142:1048–1055.
 34. Cook C, Hegedus E. *Orthopaedic Physical Examination Tests: An Evidence-Based Approach*. Upper Saddle River, NJ: Prentice Hall, 2007.
 35. Hamner DL, Pink MM, Jobe FW. A modification of the relocation test: Arthroscopic findings associated with a positive test. *J Shoulder Elbow Surg* 2000;9:263–267.
 36. Berg EE, Ciullo JV. A clinical test for superior glenoid labral or SLAP lesions. *Clin J Sport Med* 1998;8:121–123.
 37. Kibler WB. Specificity and sensitivity of the anterior slide test in throwing athletes with superior glenoid labral tears. *Arthroscopy* 1995;11:296–300.
 38. Kim SH, Ha K, Han KY. Biceps load test: A clinical test for superior labrum anterior and posterior lesions in shoulders with recurrent anterior dislocations. *Am J Sports Med* 1999;27(3):300–303.
 39. Kim SH, Ha KL, Ahn JH, Kim SH, Choi HJ. Biceps load test II: A clinical test for SLAP lesions of the shoulder. *Arthroscopy* 2001;17:160–164.
 40. Liu SH, Henry MH, Nuccion S, Shapiro MS, Dorey F. Diagnosis of glenoid labral tears: A comparison between magnetic resonance imaging and clinical examinations. *Am J Sports Med* 1996;24:149–154.
 41. Liu SH, Henry MH, Nuccion SL. A prospective evaluation of a new physical examination in predicting glenoid labral tears. *Am J Sports Med* 1996;24:721–725.
 42. 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. *Am J Sports Med* 1998;26:610–613.
 43. Myers TH, Zemanovic JR, Andrews JR. The resisted supination external rotation test: A new test for the diagnosis of superior labral anterior posterior lesions. *Am J Sports Med* 2005;33:1315–1320.
 44. Magee DJ. *Orthopaedic Physical Assessment*. 4th ed. Philadelphia, PA: W.B. Saunders Company, 2002.
 45. Cleland J. *Orthopaedic Clinical Examination: An Evidence-Based Approach for Physical Therapists*. Philadelphia, PA: Saunders, 2007.
 46. Malanga GA NS. *Musculoskeletal Physical Examination: An Evidence-Based Approach*. Philadelphia, PA: Mosby, 2006.
 47. Domholdt E. *Physical Therapy Research: Principles and Applications*. Philadelphia, PA: W.B. Saunders Company, 1993.
 48. Cooper DE, Arnoczky SP, O'Brien SJ, Warren RF, DiCarlo E, Allen AA. Anatomy, histology, and vascularity of the glenoid labrum. *J Bone Joint Surg Am* 1992;74:46–52.
 49. Prodromos CC, Ferry JA, Schiller AL, Zarins B. Histological studies of the glenoid labrum from fetal life to old age. *J Bone Joint Surg Am* 1990;72:1344–1348.
 50. Paris SV, Loubert P. *Foundations of Clinical Orthopaedics*. 3rd ed. St. Augustine, FL: Institute Press, 1999.
 51. Portney LG, Watkin MP. *Foundations of Clinical Research: Application to Practice*. Upper Saddle River, NJ: Prentice-Hall, 2000.
 52. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271:703–707.