Concurrent Criterion-Related Validity of Acromioclavicular Joint Physical Examination Tests: A Systematic Review

Janette W. Powell, PT, BEd (Eng), BAppSc (Pty), OCS, STC, EMT-B
Peter A. Huijbregts, PT, MSc, MHSc, DPT, OCS, MTC, FAAOMPT, FCAMT

Abstract: This article systematically reviews the available research on concurrent criterion-related validity of physical examination tests for the diagnosis of acromioclavicular joint (ACJ) dysfunction. A literature search yielded four research studies on the topic of concurrent criterion-related validity of physical examination tests of the ACJ. These studies had various methodological shortcomings. Methodological scores on the STARD (Standards for Reporting of Diagnostic Accuracy) criteria yielded scores from 1/22 to 16/22. All studies examined pain provocation tests only. The currently available best research evidence supports the inclusion of a number of tests with a specific interpretation in a physical examination format for the diagnosis of painful ACJ dysfunction. A negative finding on the cross-body adduction test, tenderness on palpation of the ACJ, and the Paxinos sign may serve to rule out a painful ACJ dysfunction. A positive finding on the active compression test, the cross-body adduction test, and the acromioclavicular resisted extension test may serve to rule in a painful ACJ dysfunction. A positive finding on all three tests for the cross-body adduction, active compression, and resisted acromioclavicular extension may be relevant when the physical therapist is considering a medical-surgical referral and associated higher-risk interventions. This review indicates that future research is required 1) to evaluate the diagnostic utility of the gold standard tests used in the studies retrieved; 2) to examine the reliability and concurrent criterion-related validity (with validated gold standard tests) of these and other physical tests and history items commonly used in the diagnosis of ACJ lesions, both isolated and in the form of multi-test regimens; and 3) to study predictive validity of findings on tests and multi-test regimens for ACJ dysfunction coupled to outcomes with diagnosis-specific (orthopedic manual) physical therapy, medical, and surgical interventions.

Key Words: Concurrent Criterion-Related Validity, Acromioclavicular Joint, Physical Examination, Systematic Review, STARD Criteria

Shoulder pain is a common reason for patients to seek physical therapy (PT) services. Dysfunction of the acromioclavicular joint (ACJ) is a common component of shoulder pain[1-7]. ACJ separations (grades I and II) have been described as accounting for 45% of all athletic shoulder injuries[4]. The incidence of injuries to the clavicle and the associated joints has been reported to be as high as 23/1000 athletic exposures for ice hockey and 17/1000 athletic exposures for lacrosse[5]. The prevalence of atraumatic osteolysis has been reported to be as high as 27% in weightlifters[6]. Kiner[7] noted that over half of the ACJ injuries occur in the under-30 population. The ACJ is one of the most frequently injured joints in certain sports, e.g., football, ice hockey, skiing, and rugby[1,2].

Table 1 contains pathologies that may affect the ACJ[1-3,8-11]. Dislocations of the ACJ account for 12% of all dislocations affecting the shoulder girdle[1]. Rapid degeneration of the intra-articular disc commences in the second decade of life and is significant by the fourth decade[2,3]. A lack of intra-articular disc development may play a significant role in the development of osteoarthritis. The ACJ is also prone to inflammatory, septic, and crystalline arthropathy[2,3]. The deltoid, trapezius, and

Address all correspondence and request for reprints to:
Peter Huijbregts
Shelbourne Physiotherapy Clinic
#100B-3200 Shelbourne Street
Victoria, BC V8P 5G8
Canada
shelbournephysio@telus.net

The Journal of Manual & Manipulative Therapy
Vol. 14 No. 2 (2006), E19 - E29

Concurrent Criterion-Related Validity of Acromioclavicular Joint Physical Examination Tests: A Systematic Review / E19

Table 1: Pathologies/dysfunctions affecting the acromioclavicular joint[1-11]

Traumatic conditions
• Separation/Dislocation (types I –VI)   • Fracture
Infectious conditions
• Septic arthritis
Inflammatory conditions
• Rheumatoid arthritis   • Systemic lupus erythematosus   • Ankylosing spondylitis   • Subacromial bursitis   • Rotator cuff pathology
Degenerative joint disease
• Osteoarthritis   • Osteolysis
Metabolic conditions
• Gout

pectoralis major muscles may contribute to pathologic conditions including osteolysis of the distal clavicle via compressive forces that these muscles place on the ACJ during repeated forceful contraction[2].

There is increasing focus within the medical and allied health community to substantiate current practice with scientific evidence. This is often referred to as evidence-based practice (EBP). EBP stresses the examination and clinical application of scientific research evidence. Within the EBP paradigm, the emergence of new evidence in literature can and should change the way patients are evaluated and treated. Sackett et al[12] described EBP as "…the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients…" Current research commonly delves into the reliability and accuracy of diagnostic tests (including the history and physical examination), the predictive value of prognostic markers, and the efficacy and safety of therapeutic, rehabilitative, and preventive regimens. Such research has the potential to invalidate previously accepted diagnostic tests and therapeutic and preventive interventions and to replace them with new ones that are more accurate, efficient, effective, and safer[12]. EBP is constructed from the best available research evidence, clinician expertise, and patients' values[12]. This composite approach to diagnosis, prognosis, and management holds the potential to optimize and progressively evolve the evaluation and treatment provided in the medical community.

In this article we discuss the evidence base for diagnosis of ACJ dysfunction, specifically research into the concurrent criterion-related validity of physical examination tests for the ACJ. We first briefly review the statistical concept of concurrent criterion-related validity and the associated relevant statistics, followed by a narrative description of the studies that have researched ACJ physical examination test validity. We then discuss research validity of the studies reviewed both in a narrative format and based on the Standards for Reporting of Diagnostic Accuracy (STARD) statement[13,14]. A clinical interpretation with a suggestion for a physical examination format based on best available evidence as discussed in this article and suggestions for future research conclude the article.

Concurrent Criterion-Related Validity

Concurrent criterion-related validity evaluates the extent to which a test or measure can be used as a substitute for an established gold standard test. In studies researching this type of validity, two tests are performed at approximately the same time and the researchers evaluate whether the test studied can be used as a clinical alternative to the gold standard test[15]. This type of validity is particularly relevant for physical therapists as many of the gold standard assessment techniques (e.g., visualization of tissue integrity via surgical procedures or intra-articular anaesthetic injections) fall outside of the PT scope of practice.

Diagnostic tests and measures frequently yield dichotomous results such that the patient either has or does not have the disease or dysfunction. When comparing a dichotomous clinical test or measure to a dichotomous gold standard test, there are four possible outcomes[16]:
- True positive (TP): The test indicates that the patient has the disease or dysfunction and this is confirmed by the gold standard test.
- False positive (FP): The clinical test indicates that the disease or dysfunction is present, but this is not confirmed by the gold standard test.
- False negative (FN): The clinical test indicates absence of the disorder, but the gold standard test shows that the disease or dysfunction is present.
- True negative (TN): The clinical and the gold standard test agree that the disease or dysfunction is absent.

These values are used to calculate the statistical measures of accuracy, sensitivity, specificity, negative and positive predictive values, and negative and positive likelihood ratios as indicated in Table 2[15-20].

The statistical measure of accuracy provides a quantitative measure of the overall value of a diagnostic test, but it has minimal value in the diagnostic decisions, as it does not differentiate between the diagnostic value of positive and negative test results. The usefulness of predictive values seems great but is limited by the fact that for predictive values to apply, the prevalence in the

clinical population being examined has to be identical to the prevalence in the study population from which the predictive values were derived[16,20]. Davidson[20] noted that because of this issue, positive and negative predictive values should be disregarded in the diagnostic process.

Interpretation of sensitivity and specificity values is easiest when their values are high[16,20]. When a test has high sensitivity, negative test results will likely rule out the disease or dysfunction, as there are very few false negatives when sensitivity is high[16,20]. Similarly, when a test has high specificity, a positive test result will likely rule in the disease or dysfunction as there are very few false positives when specificity is high[16,20]. Without providing specific quantitative cut-off points, Davidson[20] used the mnemonics:

- SnOUT: With highly Sensitive tests, a negative result will rule a disorder OUT.
- SpIN: With highly Specific tests, a positive result will rule a disorder IN.

For most diagnostic procedures, the statistical measures of sensitivity and specificity are inversely related: Tests with high sensitivity often have lower specificity, and vice versa[16]. A diagnostic test can only be 100% sensitive and 100% specific if there is no overlap between the population that has the disease or dysfunction and the population that does not[16]. Davidson[20] noted another problem with the measures of sensitivity and specificity in that these measures tell us how often a test will be positive or negative in patients who we already know have or do not have the disease or dysfunction. Obviously, this does not correspond with the clinical situation where it is not known if the disease or dysfunction is present.

Likelihood ratios (LR) summarize the data of sensitivity and specificity in a more clinically relevant format[16,20]. Jaeschke et al[21] provided guidelines for the clinical interpretation of positive and negative LR data (Table 3). Davidson[20] provided the following "nutshell" clinical summary:

- A positive LR ≥ 10 provides a clinically significant degree of certainty that the patient with a positive test has the disorder for which you are testing.
- A negative LR ≤ 0.1 provides a clinically significant degree of certainty that the patient with a negative test result does not have the disorder for which you are testing.
- A positive or negative LR close to 1.0 provides little change in the probability that the patient has or does not have a disease or dysfunction; i.e., this test is of little diagnostic value.

Pretest probability can be defined as how likely a clinician thinks it is that a person has a specific disease

Table 2: Definition and calculation of statistical measures of concurrent criterion-related validity[16,18,20]

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

TP- true positive; TN- true negative; FP- false positive; FN- false negative

Table 3: Interpretation of likelihood ratios on changes from pretest to posttest probability[21]

	Numerical value	Change from pretest to posttest 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 0.5-1	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

or dysfunction before doing a diagnostic test[20]. To determine pretest probability, clinicians can use personal clinical experience, information on pathophysiology, or data from studies on prevalence in a specific population. We discussed data relevant to ACJ dysfunction above[1-7]. Likelihood ratios are then used to calculate posttest probability or the likelihood that a patient has a specific disease or dysfunction after a diagnostic test is done[20]. We can calculate posttest probability by:

1. Establishing or estimating pretest probability
2. Calculating pretest odds: pretest odds = pretest probability / (1-pretest probability)
3. Calculating posttest odds: posttest odds = pretest odds x LR
4. Calculating posttest probability: posttest probability = posttest odds / (posttest odds+1)

Ideally, a high pretest probability is combined with a diagnostic test that has a high positive LR to rule in a diagnosis. Alternatively, a low pretest probability and a test with a low negative LR combine to confidently rule out a diagnosis.

Methods

We searched the PubMed, Proquest, Cumulative Index to Nursing and Allied Health (CINAHL), Index to Chiropractic Literature (ICL), and Ostmed Osteopathic Literature Database online databases from 1990 to March 2006 for peer-reviewed references using the key word "acromioclavicular." We chose not to limit the search by adding other terms to increase our chance of identifying relevant studies. After reviewing abstracts, we retrieved in full-text format only those studies that quantitatively investigated diagnostic accuracy of physical examination diagnostic tests of the ACJ as compared to a gold standard test or test regimen. We then did a hand search of the reference lists of the retrieved articles to locate further relevant references fitting these same inclusion criteria.

Results

The PubMed search yielded 791 references of which three[22-24] met the inclusion criteria. The Proquest search provided 41 articles: one[22] met inclusion criteria. The CINAHL search produced 153 articles: three met our criteria[22-24]. The ICL and Ostmed search yielded ten and two articles, respectively, but none met our inclusion criteria. Eliminating duplications from these database searches, we retrieved three studies that met our inclusion criteria[22-24]. A hand search of the reference lists of these articles provided one additional reference[25]. Our search thus yielded a total of four relevant articles[22-25].

Table 4 provides data on the diagnostic utility of the individual ACJ physical examination tests studied in the retrieved articles[22-25]. Table 5 provides data on the diagnostic utility of multiple-test regimens. 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 for the tests studied. Where provided by the authors, raw data is added in Tables 4 and 5 between brackets.

Narrative Description of Retrieved Studies

O'Brien et al[22] investigated the active compression test for its diagnostic utility with regard to glenoid labrum tears and ACJ abnormalities. They performed a prospective study using 318 patients including 268 consecutive patients presenting with shoulder pain and 50 control subjects who presented at their clinic with knee pain and who denied shoulder pain. The active compression test involved the client standing with the affected arm straight and forward flexed to 90°. The arm was then horizontally adducted 10-15° and maximally internally rotated. The patient then resisted a downward force applied by the examiner to the distal arm (Figure 1).

Table 4: Diagnostic utility acromioclavicular joint physical examination tests[22-25]

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Active compression test							
• O'Brien et al[22]	0.97 (255/262)	1.00 (55/55)	0.925 (200/207)	0.915 (55/62)	1.00 (200/200)	13.3	0.0
• Chronopoulos et al[23]	0.92 (298/325)	0.41 (7/17)	0.95 (291/308)	0.29 (7/24)	0.97 (291/301)	8.2	0.6
• Walton et al[24]	0.53	0.16	0.90	0.62	0.52	1.60	0.9
• Maritz and Oosthuizen[25]	---	0.68 (15/22)	---	---	---	---	---
Cross-body adduction test							
• Chronopoulos et al[23]	0.79 (437/553)	0.77 (27/35)	0.79 (410/518)	0.20 (27/135)	0.98 (410/418)	3.7	0.3
• Maritz and Oosthuizen[25]	---	1.00 (22/22)	---	---	---	---	---
Acromioclavicular resisted extension							
• Chronopoulos et al[23]	0.84 (292/348)	0.72 (13/18)	0.85 (279/330)	0.20 (13/64)	0.98 (279/284)	4.8	0.3
Acromioclavicular joint tenderness							
• Walton et al[24]	0.53	0.96	0.10	0.52	0.71	1.07	0.4
• Maritz and Oosthuizen[25]	---	0.95 (21/22)	---	---	---	---	---
Paxinos sign							
• Walton et al[24]	0.65	0.79	0.50	0.61	0.70	1.58	0.4

Table 5: Diagnostic utility multi-test regimens consisting of cross-body adduction stress, active compression, and acromioclavicular resisted extension tests (modified from Chronopoulos et al[23])

	Accuracy	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
≥1 positive test	0.75 (237/315)	0.00 (16/16)	0.74 (221/299)	0.17 (16/94)	1.00 (221/221)	0.00	1.4
≥2 positive tests	0.89 (279/315)	0.81 (13/16)	0.89 (266/299)	0.28 (13/46)	0.99 (266/269)	7.4	0.2
3 positive tests	93 (294/315)	0.25 (4/16)	0.97 (290/299)	0.31 (4/13)	0.96 (290/302)	8.3	0.8

Fig. 1: Active compression test, maximal internal rotation

Fig. 2: Active compression test, maximal external rotation

The test was then repeated in the same position with the arm maximally externally rotated (Figure 2). The authors provided no data on the amount of force used. This test was considered positive for ACJ dysfunction if the pain was localized to the ACJ on the first position and relieved or eliminated on the second position. Pain "deep inside the shoulder," with or without a click, in the first position and eliminated or reduced in the second position was considered indicative of a glenoid labrum tear. The gold standard test used consisted of various combinations of radiography, MRI, intra-operative confirmation, and a positive outcome after diagnosis-specific surgical intervention. No data was provided to clarify these findings constituting a positive gold standard test. The authors concluded that the active compression test

was a clinically valuable tool being both highly sensitive and specific for diagnosing ACJ pathology.

Chronopoulos et al[23] evaluated the cross-body adduction stress test, the active compression test, and the acromioclavicular resisted extension test for their isolated and combined diagnostic utility with regard to chronic isolated ACJ lesions. The study was a retrospective case-control study that used 35 patients diagnosed with chronic isolated ACJ lesions and 580 control subjects who had undergone surgical procedures for other shoulder conditions. Patients with non-isolated chronic ACJ lesions and patients who did not have a diagnostic arthroscopy were excluded from the study. The cross-body adduction stress test was described as a test where the client's arm is forward flexed to 90° and

Fig. 3: Cross-body adduction stress test

Fig. 4: Acromioclavicular resisted extension test

then horizontally adducted across the body (Figure 3). The authors did not specify whether this test was active or passive. This test was considered positive if it caused pain localized to the ACJ. The acromioclavicular resisted extension test was performed with the client's shoulder flexed to 90° combined with maximal internal rotation and 90° of elbow flexion. The client was then asked to horizontally abduct the arm against resistance (Figure 4). This test was considered positive if it caused pain at the ACJ. The active compression test has already been described above. The authors provided no data on the amount of force applied with the tests. The gold standard test used in this study for the diagnosis of ACJ lesions was pain localized to the top of the shoulder or the ACJ region, local tenderness on ACJ palpation, at least one diagnostic injection into the ACJ with complete or nearly complete pain relief, and arthroscopic confirmation of the diagnosis of an ACJ lesion. Therefore, the final diagnosis for all patients was based on history, examination, and arthroscopic findings. No further data was provided to clarify this gold standard test regimen. The authors concluded that the three tests studied had isolated clinical utility. They also analyzed the diagnostic utility of multi-test regimens based on these three tests (Table 5) and suggested that a clinician should use a criterion of one positive test when high sensitivity is required whereas a criterion of three positive tests is appropriate when high specificity is necessary.

Walton et al[24] evaluated the diagnostic utility of clinical and imaging tests for ACJ pain. The clinical tests included local ACJ tenderness, the active compression test as described above, and the Paxinos sign. Subjects were 38 patients selected from a group of 1037 consecutive patients with shoulder pain. The inclusion criterion for these 38 patients was pain indicated on a pain drawing that was located between the mid-portion of the clavicle and the deltoid insertion. Exclusion criteria were previous distal clavicle or ACJ surgery, clavicular fracture, pregnancy, allergy to lidocaine or contrast medium, contra-indication to MRI or bone scan, refusal to participate in the study, and markings in the pain drawing beyond the area indicated above. For the Paxinos sign, the patient sat with the arm relaxed by his or her side. The examiner's thumb was placed under the postero-lateral aspect of the acromion; the index and long fingers (same or contralateral hand) were then placed superior to the mid-portion of the ipsilateral clavicle (Figure 5). The thumb then applied an antero-superior force concurrently while the fingers applied an inferior force. This test was considered positive if it caused or increased pain localized to the ACJ. The gold standard test was ≥ 50% pain reduction after imaging-guided intra-articular anaesthetic infiltration of the ACJ. Of 38 patients, 28 scored positive on this gold standard test; i.e., prevalence in this population was 74%. The authors noted that most of the clinical and imaging tests studied did not significantly add to the ability to predict who had ACJ pain. Only combinations of the results of the bone scan and the Paxinos sign provided clinically relevant diagnostic utility with a positive likelihood ratio of 55 for both tests to be positive and a negative likelihood ratio of 0.44 when both tests were negative. The authors also noted that all individual clinical tests were either highly sensitive or highly specific but not both and that each individual test had only limited diagnostic utility for diagnosing ACJ pain.

In a case series design without an asymptomatic control group, Maritz and Oosthuizen[25] studied the active compression test, the cross-body adduction stress test, and local ACJ tenderness in 22 patients. The gold standard test was an unspecified percentage of pain relief after intra-articular infiltration of the ACJ. No test description or definition of what was considered a

Fig. 5: Paxinos sign

positive response was provided. The authors concluded that as no test was 100% accurate, the total clinical presentation should be taken into account for diagnosis with the greatest diagnostic utility attached to intra-articular infiltration.

Research Validity

When interpreting research, we need to first examine its statistical conclusion validity, external validity, and construct validity[19]. With regard to statistical conclusion validity, all studies reviewed here used appropriate statistical measures to determine concurrent criterion-related validity. However, the lack of a control group in the Maritz and Oosthuizen[25] study prevents calculation of

data other than sensitivity, which by itself provides only limited information for the clinical diagnostic process.

As to external validity, the subjects in the O'Brien et al[22] and Chronopoulos et al[23] studies were all surgical patients, which limits the external validity with regard to the conservatively managed, at least initially, patient population seen in PT clinical practice. In addition, the definition of the patients diagnosed with ACJ lesions involved either an unspecified[22] or insufficiently specified[23] multi-test regimen or tests that are outside of the PT scope of practice[22,23,25]. The Walton et al[24] study provides the most clinically useful and operationally well-defined inclusion and exclusion criteria for physical therapists. In particular, the criterion of pain on a pain drawing localized between the mid-portion of the clavicle and the deltoid insertion would allow the physical therapist to identify the patient population to which the study results might apply.

When assessing construct validity, we have to compare the construct as labeled to the construct as implemented[19]. A physical therapist diagnoses joint dysfunction. Paris and Loubert[26] defined joint dysfunction as the presence of hypomobility, hypermobility, or aberrant motion. The construct as labeled in the studies reviewed was ACJ abnormality or lesion. The construct as implemented in these studies --as evidenced by the fact that all physical examination tests studied were pain provocation tests-- was that of a painful ACJ abnormality or lesion. However, the construct as implemented in orthopedic manual physical therapy (OMPT) practice is that of a patho-kinesiological entity that may or may not be painful, making the results of these studies irrelevant to the OMPT diagnosis of a non-painful ACJ dysfunction.

Another issue relevant to construct validity is the gold standard test used. Three of the studies reviewed used single diagnostic intra-articular anaesthetic infiltrations[23-25]. Walton et al[24] used image-guided infiltrations; it is unclear whether the blocks used in the Chronopoulos et al[23] and Maritz and Oosthuizen[25] studies were image-guided. Parlington and Broome[27] noted that non-image-guided intra-articular infiltrations were placed successfully in the ACJ in only 16/24 (67%) of cadaveric shoulders. Schwarzer et al[28] reported a false-positive rate of single, uncontrolled intra-articular blocks of 38% in patients with chronic low-back pain. We found no data specific to the false-positive rate of single intra-articular ACJ blocks. The gold standard test used in the O'Brien et al[22] study consisted of various combinations of radiography, MRI, intra-operative confirmation, and a positive outcome after diagnosis-specific surgical intervention. Walton et al[24] noted that all isolated tests, including radiographic evaluation, had only limited diagnostic utility. They also noted that the high sensitivity but low specificity established in their study for the diagnosis of ACJ lesions by means of MRI meant that a positive MRI finding couldn't be used to establish the presence

of ACJ-related pain[24]. Stein et al[29] also reported on the high age-related prevalence of arthritic changes found during MRI-evaluation in the ACJ of possibly asymptomatic subjects. Using surgical outcome as a gold standard may be invalidated by the placebo effect of surgery. Moseley et al[30] reported on the placebo effect of arthroscopic surgery for osteoarthritis of the knee. We found no data on a possible placebo effect for ACJ surgery. We also found no data in the literature on the diagnostic utility of intra-operative evaluation of the ACJ. However, it is obvious that the construct as labeled in all four studies was that of a painful lesion of the ACJ. The discussion of the diagnostic utility of the gold standard tests used for this study implies that a number of the patients may have in fact been diagnosed inappropriately invalidating, to some extent, the study results.

Another way in which to assess research validity is by applying established criteria for quality assessment of different research formats. Criteria for the systematic assessment of the methodology of studies into diagnostic accuracy have been described in the STARD (Standards for Reporting of Diagnostic Accuracy) statement[13,14]. Table 6 provides the proposed STARD criteria and the scores attained by the studies retrieved for this article. This score provides additional data on the research validity that supplements but –in the authors' opinion-- cannot replace the information discussed above. Although no cut-off values have been established for the STARD score, it would seem obvious we can place no value on the findings reported by Maritz and Oosthuizen[25] and only limited value on the findings reported by O'Brien et al[22].

Clinical Interpretation

Despite shortcomings with regard to research validity, the four studies reviewed do represent the best available research evidence. As noted above, EBP is a composite of best available research evidence, clinician expertise, and patient values[12]. Obviously, clinician expertise also applies to critical interpretation and application of potentially methodologically flawed studies based on the clinician's knowledge of methodology and statistics.

Based on the results of the studies reviewed, three tests have consistently high sensitivity: the cross-body adduction test (0.77-1.00), tenderness on palpation of the ACJ (0.95-0.96), and the Paxinos sign (0.79). Even if we disregard the sensitivity values established in the Maritz and Oosthuizen[25] study due to its low methodological score, this still means that a negative result on these three tests may be clinically significant in that it **decreases** the likelihood of painful ACJ dysfunction.

Three other tests have demonstrated high specificity: the active compression test (0.90-0.95), the cross-body adduction test (0.79), and the acromioclavicular resisted extension test (0.85). A positive result on these tests may be clinically significant in that it **increases** the likelihood

Table 6: STARD score of the studies retrieved[13,14]

STARD Items	O'Brien et al[22]	Chronopoulos et al[23]	Walton et al[24]	Maritz and Oosthuizen[25]
Identifies article as a study of diagnostic accuracy	0	1	1	0
States research questions or aims	1	1	1	0
Describes study population (inclusion, exclusion criteria, settings, locations)	1	1	1	0
Describes participant recruitment	1	1	1	0
Describes participant sampling	0	0	1	0
Describes data collection (prospective or retrospective)	1	1	1	1
Describes reference standard and rationale	0	1	1	0
Describes technical specifications of material and methods involved	0	0	1	0
Describes definition and rationale for units, cut-off points, or categories of results of tests	0	0	0	0
Describes number, training, and expertise of raters	0	1	1	0
Were the raters blinded to the results of the other test? Describes clinical information available to raters	1	1	1	0
Describes statistical methods for comparing diagnostic accuracy and expressing uncertainty	1	1	1	0
Describes methods for calculating test reproducibility, if done	NA	NA	NA	NA
Reports when study was done with start and end dates for recruitment	0	1	1	0
Reports clinical and demographic characteristics subjects	0	1	0	0
Reports how many subjects satisfying inclusion criteria did not undergo the tests; describes why these subjects were not tested	1	1	1	0
Reports time interval between researched and reference test and any treatment provided in between tests	0	0	0	0
Reports disease severity in subjects with target condition and other diagnoses in subjects without target condition	0	1	1	0
Reports cross-tabulation of researched and reference test	0	0	0	0
Reports adverse effects from researched and reference tests	0	0	0	0
Reports estimates of diagnostic accuracy and measures of statistical uncertainty	0	0	0	0
Reports how indeterminate test results, missing responses, and outliers of researched test were handled	1	1	1	0
Reports estimates of variability between raters, centers, or subject subgroups, if done	NA	NA	NA	NA
Reports estimates of test reproducibility, if done	NA	NA	NA	NA
Discusses clinical applicability of study findings	0	1	1	0
Total Score	**8/22**	**15/22**	**16/22**	**1/22**

NA-not applicable

of painful ACJ dysfunction.

None of the tests evaluated has demonstrated a consistently relevant negative likelihood ratio. However, the active compression test was demonstrated in two studies to be a clinically relevant positive LR (8.2-13.3), again indicating the possible significance of a positive test for ruling **in** the diagnosis of a painful ACJ lesion.

As noted above, Chronopoulos et al[23] also analyzed the diagnostic utility of multi-test regimens consisting of the cross-body adduction stress, active compression, and resisted acromioclavicular extension tests (Table 5). They suggested that a clinician should use a criterion of one positive test when high sensitivity is required whereas a criterion of three positive tests is appropriate when high specificity is needed. High sensitivity would be relevant if the physical therapist did not want to run the risk of missing a possible contribution of the ACJ to the patient's complaint of shoulder pain. High specificity is required if the clinician is considering an intervention with higher inherent risks. This will likely not be relevant to any PT intervention but may be relevant when the PT is considering referring a patient out for medical (e.g., intra-articular infiltration) or surgical interventions.

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 painful ACJ dysfunction:

- A negative finding on the cross-body adduction test, tenderness on palpation of the ACJ, and the Paxinos sign to rule **out** a painful ACJ dysfunction
- A positive finding on the active compression test, the cross-body adduction test, and the acromioclavicular resisted extension test to rule **in** painful ACJ dysfunction
- A positive finding on all three tests (the cross-body adduction stress, active compression, and resisted acromioclavicular extension tests) may be relevant when considering a medical-surgical referral and associated higher-risk interventions

Conclusion

Research into the diagnostic utility of physical examination tests to detect ACJ dysfunction is limited to pain provocation tests. No research has been done into ACJ examination techniques needed to validate or guide either an OMPT patho-kinesiological diagnosis or intervention of a painful or non-painful ACJ dysfunction. Application of the existing research evidence to PT clinical practice is also limited by research validity issues with regard to limited external validity to the patient population seen in PT for this problem and insufficiently validated gold standard tests. Methodological quality scores based on the STARD criteria varied greatly from 1/22 to 16/22, invalidating at least one of the identified studies[25]. Furthermore, we could find no research that evaluated the predictive validity of tests for ACJ dysfunction with regard to the outcome with PT, medical, or surgical interventions.

Directions for future research should include:

- Studies to evaluate the diagnostic utility of current gold standard tests
- Studies to evaluate reliability and concurrent criterion-related validity (with validated gold standard tests) of these and other commonly used physical tests and history items, both isolated and in the form of multi-test regimens
- Studies of predictive validity of findings on tests and multi-test regimens for ACJ dysfunction coupled to outcomes with diagnosis-specific OMPT and other PT, medical, and surgical interventions

Acknowledgments

REFERENCES

1. Magee DJ, Reid DC. Shoulder injuries. In: Zachazewski JE, Magee DJ, Quillen WS, eds. *Athletic Injuries and Rehabilitation*. Philadelphia, PA: W.B. Saunders Company, 1996.
2. Renfree KJ, Wright TW. Anatomy and biomechanics of the acromioclavicular and sternoclavicular joints. *Clin Sports Med* 2003;22:219-237.
3. Garretson RB, Williams GR. Clinical evaluation of injuries to the acromioclavicular and sternoclavicular joints. *Clin Sports Med* 2003;22:239-254.
4. Debski RE, Parsons IM, Woo SLY, Fu FH. Effect of capsular injury on acromioclavicular joint mechanics. *J Bone Joint Surg* 2001;83A:1344-1351.
5. Hutchinson MR, Ahuja GS. Diagnosing and treating clavicle injuries. *Phys Sportsmed* 1996;24(3):26-36.
6. Auge WK, Fischer RA. Arthroscopic distal clavicle resection for isolated atraumatic osteolysis in weight lifters. *Am J Sports Med* 1998;26:189-192.
7. Kiner A. Diagnosis and management of grade II acromioclavicular joint separation. *Clin Chiro* 2004;7:24-30.
8. Lehtinen JT, Lehto MUK, Kaarela K, Kautiainen HJ, Belt EA, Kauppi MJ. Radiographic joint space in rheumatoid acromioclavicular joints: A 15-year prospective follow-up study in 74 patients. *Rheumatology* 1999;38:1104-1107.
9. Santis DD, Palazzi C, D'Amico E, Di Mascio DE, Pace-Palitti V,

Petricca A. Acromioclavicular cyst and "porcupine shoulder" in gout. *Rheumatology* 2001;40,11:1320-1321.

10. Shaffer BS. Painful conditions of the acromioclavicular joint. *J Am Acad Orthop Surg* 1999;7:176-188.

11. Lemos MJ. The evaluation and treatment of the injured acromio-clavicular joint in athletes. *Am J Sports Med* 1998;26:137-144.

12. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence-based medicine: What it is and what it isn't. *BMJ* 1996;312:71-72.

13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clin Chem* 2003;49:7-18.

14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *BMJ* 2003;326:41-44.

15. Huijbregts PA. Sacroiliac joint dysfunction: Evidence-based diagnosis. *Orthopaedic Division Review* 2004;May/June:18-32, 41-44.

16. Huijbregts PA, Hobby M, Salmas P. Scaphoid fracture: A case report illustrating evidence-based diagnosis and discussing measures of reliability and concurrent criterion-related validity. *Interdivisional Rev* 2005:Jan/Feb:14-18.

17. Greenfield MLVH, Kuhn JE, Wojtys EM. A statistics primer: Validity and reliability. *Am J Sports Med* 1998;26,3:483-485.

18. Fritz JM, Wainner RS. Examining diagnostic tests: An evidence-based perspective. *Phys Ther* 2001;81:1546-1564.

19. Huijbregts PA. Spinal motion palpation: A review of reliability studies. *J Manual Manipulative Ther* 2002;10:24-39.

20. Davidson M. The interpretation of diagnostic tests: A primer for physiotherapists. *Aust J Physiother* 2002;48:227-233.

21. 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.

22. 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.

23. Chronopoulos E, Kim TK, Park HB, Ashenbrenner D, McFarland EG. Diagnostic value of physical tests for isolated chronic acromioclavicular lesions. *Am J Sports Med* 2004;32:655-661.

24. Walton J, Mahajan S, Paxinos A, Marshall J, Bryant C, Shnier R, Quinn R, Murrell GA. Diagnostic values of tests for acromioclavicular joint pain. *J Bone Joint Surg* 2004;86A:807-812.

25. Maritz NGJ, Oosthuizen PJ. Diagnostic criteria for acromioclavicular joint pathology. *J Bone Joint Surg* 2002;84B(Suppl.1):78.

26. Paris SV, Loubert PV. *Foundations of Clinical Orthopaedics*. 3rd ed. St. Augustine, FL: Institute Press; 1999.

27. Parlington PF, Broome GH. Diagnostic injection around the shoulder: Hit and miss. A cadaveric study of injection accuracy. *J Shoulder Elbow Surg* 1998;7:147-150.

28. Schwarzer AC, April CN, Derby R, Fortin J, Kine G, Bogduk N. The false-positive rate of uncontrolled diagnostic blocks of the lumbar zygapophysial joints. *Pain* 1994;58:195-200.

29. Stein BE, Wiater JM, Pfaff HC, Bigliani LU, Levine WN. Detection of acromioclavicular joint pathology in asymptomatic shoulders with magnetic resonance imaging. *J Shoulder Elbow Surg* 2001;10:204-208.

30. Moseley JB, Wray NP, Kuykendall D, Willis K, Landon G. Arthroscopic treatment of osteoarthritis of the knee: A prospective, randomized, placebo-controlled trial. Results of a pilot study. *Am J Sports Med* 1996;24:28-34.