

Reliability and Diagnostic Accuracy of 5 Physical Examination Tests and Combination of Tests for Subacromial Impingement

Lori A. Michener, PhD, PT, ATC, Matthew K. Walsworth, MD, PT, William C. Doukas, MD, Kevin P. Murphy, MD

ABSTRACT. Michener LA, Walsworth MK, Doukas WC, Murphy KP. Reliability and diagnostic accuracy of 5 physical examination tests and combination of tests for subacromial impingement. *Arch Phys Med Rehabil* 2009;90:1898-903.

Objective: To investigate the reliability and diagnostic accuracy of individual tests and combination of tests for subacromial impingement syndrome (SAIS).

Design: A prospective, blinded study design.

Setting: Orthopedic surgeon shoulder clinic.

Participants: Patients with shoulder pain (n=55, mean age=40.6y).

Interventions: Patients were evaluated with 5 physical examination tests for SAIS: Neer, Hawkins-Kennedy, painful arc, empty can (Jobe), and external rotation resistance tests. Surgical diagnosis was the reference standard.

Main Outcome Measures: Diagnostic accuracy calculated with a receiver operating characteristic (ROC) curve and sensitivity, specificity, positive likelihood ratio (+LR), and negative likelihood ratio (-LR). A forward stepwise binary logistic regression analysis was used to determine the best test combination for SAIS. An ROC curve analysis was also used to determine the cut point of the number of tests discriminating between the presence and absence of SAIS. Kappa coefficients and percent agreement assessed interrater reliability.

Results: The ROC analyses revealed a significant area under the curve (AUC) (AUC=.67-.72, $P<.05$) for all tests, except for the Hawkins-Kennedy. The tests with a +LR greater than or equal to 2.0 were the painful arc (+LR=2.25; 95% CI, 1.33-3.81), empty can (+LR=3.90; 95% CI, 1.5-10.12), and the external rotation resistance tests (+LR=4.39; 95% CI, 1.74-11.07). Tests with -LR less than or equal to 0.50 were the painful arc (-LR=.38; 95% CI, .16-.90), external rotation resistance (-LR=.50; 95% CI, .28-.89), and Neer tests (-LR=.35; 95% CI, .12-.97). The regression analysis had no specific test combinations for confirming or ruling out SAIS. The ROC analysis was significant (AUC=.79, $P=.001$), with a cut point of 3 positive tests out of 5 tests. Reliability was moderate to substantial agreement ($\kappa=.45-.67$) for the painful

arc, empty can, and external rotation resistance tests and fair strength of agreement ($\kappa=.39-.40$) for the Neer and Hawkins-Kennedy tests.

Conclusions: The single tests of painful arc, external rotation resistance, and Neer are useful screening tests to rule out SAIS. The single tests of painful arc, external rotation resistance, and empty can are helpful to confirm SAIS. The reliability of all tests was acceptable for clinical use. Based on reliability and diagnostic accuracy, the single tests of the painful arc, external rotation resistance, and empty can have the best overall clinical utility. The cut point of 3 or more positive of 5 tests can confirm the diagnosis of SAIS, while less than 3 positive of 5 rules out SAIS.

Key Words: Diagnosis; Rehabilitation; Rotator cuff; Sensitivity and specificity; Shoulder; Shoulder impingement syndrome; Tendonitis.

© 2009 by the American Congress of Rehabilitation Medicine

SHOULDER PAIN ACCOUNTS for 33.2% of physician office visits for musculoskeletal pain in the United States.¹ The prevalence of shoulder pain is 7% to 27% for adults and is increasing for women and older individuals.¹⁻³ The most frequent cause of shoulder pain is SAIS³⁻⁵; however, there is no consensus as to the diagnostic criteria for SAIS. The physical examination tests used to diagnose SAIS vary widely in clinical trials.^{6,7} This lack of consensus for diagnosing SAIS is an obstacle to investigating treatment interventions and prognosis for this disorder. There are several reasons for the lack of diagnostic consensus, including the multifactorial nature of this disorder along with the limited and conflicting evidence as to the diagnostic capability of physical tests for SAIS.

The spectrum of SAIS involves injury of the subacromial space structures to include rotator cuff tendinosis, rotator cuff partial thickness tear, and bursitis. Injury to the rotator cuff most frequently involves the supraspinatus tendon caused by direct compression to the tendon within the subacromial space and by tension overload applied to the tendon leading to intrinsic tendon degeneration during repetitive activities. The tests used to differentially diagnose SAIS via compression of the tendon in the subacromial space include the Neer,⁸ Hawkins-Kennedy,⁹ and the painful arc¹⁰ tests. Tests used to

From the Department of Physical Therapy, Virginia Commonwealth University, Medical College of Virginia (Michener); Department of Radiology, Virginia Commonwealth University Health System (Walsworth), Richmond, VA; Department of Orthopedic and Rehabilitation, Walter Reed Army Medical Center, Washington, DC (Doukas); Heekin Orthopaedics Specialists, Jacksonville, FL (Murphy); and Uniformed Services University of the Health Sciences, Bethesda, MD (Murphy).

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Correspondence to Lori A. Michener, PhD, PT, ATC, Department of Physical Therapy, Virginia Commonwealth University, Medical College of Virginia Campus, PO Box 980224, Richmond, VA 23298, e-mail: lamichen@vcu.edu. Reprints are not available from the author.

0003-9993/09/9011-0023\$36.00/0
doi:10.1016/j.apmr.2009.05.015

List of Abbreviations

AUC	area under the curve
CI	confidence interval
-LR	negative likelihood ratio
+LR	positive likelihood ratio
ROC	receiver operating characteristic
SAIS	subacromial impingement syndrome

apply tension to the supraspinatus tendon via supraspinatus muscle contraction include resisted shoulder elevation with the empty can (Jobe Test)¹¹ and resisted shoulder external rotation, which is also known as the external rotation resistance test.¹² The diagnostic ability of these tests has been examined,¹²⁻¹⁶ but these studies have provided inconsistent results and have major methodologic limitations that include a retrospective design, nonblinding of the examiner who determined the reference standard, and use of a reference standard that has its own error. Previous studies¹⁶⁻²⁰ investigating reliability of SAIS tests have also shown inconsistent results, with Kappa values ranging from .18 to 1.0, indicating poor to almost perfect agreement.

Clinicians make a diagnosis from a combination of findings from the history and physical examination. Three studies¹²⁻¹⁴ have examined the combination of tests for diagnosing SAIS. The combination of the external rotation resistance test (weakness or pain), Hawkins-Kennedy, and painful arc tests showed moderate to large shifts in probability to confirm the diagnosis and screen for SAIS.¹² Conversely, another study¹⁴ found the combination of positive findings for 5 tests of 7 tests (Hawkins-Kennedy, Neer, Horizontal Adduction, Speeds, Yeagason, Painful Arc, Drop Arm) was only minimally helpful to diagnose SAIS with a small shifts in posttest probability. In a third study¹³ investigating the combination of the Neer and Hawkins-Kennedy tests, the diagnostic capability was not improved when compared with either test individually.

To examine the effectiveness of treatment for SAIS, further evidence is needed to determine those tests with the strongest ability to screen for and diagnose SAIS. Moreover, further examination of the usefulness of the combination of tests for SAIS is needed. The purposes of this study were to determine (1) interrater reliability of SAIS physical examination tests, (2) diagnostic accuracy of the SAIS physical examination tests, and (3) a specific cluster of tests that are optimal to confirm or rule out SAIS. These results will aid clinicians in diagnosing SAIS. Optimizing the diagnosis of SAIS will improve the ability to guide treatment decision making and ultimately lead to improved medical management.

METHODS

This was a prospective blinded cohort study approved by the Medical Center Investigational Review Board. Consecutive patients presenting with shoulder pain to an orthopedic surgeon's office were invited to participate in this study. To be included in this study, patients had to report shoulder pain for at least 1 week, and shoulder pain had to be their primary complaint. A total of 65 patients were invited to participate; 3 refused to participate, and 7 did not undergo the reference standard (surgery), leaving a total of 55 subjects for the sample. The sample (n=55) consisted of 47 men and 8 women with an average age of 40.6 ± 15.1 years (range 18–83y) and an average symptom duration of 33.8 ± 48.9 months (range 2–230mo).

The rights of the subjects were protected by an informed consent, which was read and signed by all participants. All patients then underwent a standardized history and physical examination by 2 clinicians: 1 orthopedic surgeon board certified in orthopedic surgery and a physical therapist board certified in orthopedics, with 17 and 8 years experience in musculoskeletal examinations, respectively. Each clinician independently performed a standardized history and physical examination and was blinded to each other's findings and without knowledge of any imaging studies. Before the start of the study, the clinicians underwent training to standardize technique and interpretation of the tests.

The Neer test⁸ was performed with the examiner stabilizing the scapula with a downward force while fully flexing the

humerus overhead maximally while applying overpressure. A positive test was reproduction of pain of the superior shoulder.

The Hawkins-Kennedy test⁹ was performed by the examiner flexing the humerus and elbow to 90° and then maximally internally rotating the shoulder and applying overpressure. A positive test was reproduction of pain of the superior shoulder.

The painful arc¹⁰ was performed by asking the patient to actively abduct his/her shoulder and report any pain during abduction. If pain of the superior shoulder was noted between 60° and 120° of abduction, the test was considered positive.

The empty can test,¹¹ also known as the Jobe test, was performed by the examiner elevating the shoulder to 90° in the scapular plane (30°–40° anterior to the coronal plane) and then placing the shoulder in internal rotation by asking the patient to rotate the shoulder so that his/her thumb was pointing toward the floor. The examiner then applied a downward directed force at the wrist while the patient attempted to resist. A positive test was considered if weakness was detected of the involved shoulder as compared bilaterally.

The external rotation resistance test¹² was performed by placing the arm at the patient's side and flexing their elbow to 90°. A medially directed force was exerted on the distal forearm to resist shoulder external rotation. A positive test was considered if weakness was detected of the involved shoulder as compared bilaterally.

There were 5 tests for SAIS; results were recorded as "positive" or "negative" for each test as described. After completion of the history and physical examination, the patients underwent an arthroscopic examination within an average of 2.6 months (± 2.7 mo, range: 1d–8mo) after the clinical examination. The reference standard was determined via operative findings reported by an operative surgeon blinded to the clinical examination findings. This method was used to ensure blinding of the reference standard to the clinical examination to eliminate the risk for review bias. The intraoperative reference standard criteria for SAIS were the presence of any of the following: visually enlarged bursa, fibrotic appearing bursa, or degeneration of the supraspinatus tendon at the superficial aspect. Patients with additional shoulder pathologies such as partial or full-thickness rotator cuff tears, labral tears, or fraying and instability were not excluded.

Data Analysis

The first aim was to determine the intertester reliability for the 5 impingement tests between 2 clinicians. Reliability was

Table 1: Diagnoses of Subjects by the Gold Standard of Surgical Findings

Diagnosis	Number of Subjects (%)
Negative for SAIS	39/55 (71)
Positive for SAIS	16/55 (29)
SAIS only	1 (1.8)
SAIS and	
PT-RCT	4 (7.3)
FT-RCT	3 (5.5)
Labral tear (any type)	11 (20)
Instability	1 (1.8)
Acromio-clavicular joint	
Degenerative joint disease	3 (5.5)
Biceps tendon tear	3 (5.5)

NOTE. n=55. Some patients had more than 1 additional diagnosis beyond SAIS.

Abbreviations: PT-RCT, partial-thickness rotator cuff tear; FT-RCT, full-thickness rotator cuff tear.

Table 2: Intertester Kappa Reliability Coefficients and Agreement From 2 Clinicians for 5 Impingement Shoulder Tests

Test	Kappa Coefficient (95% CI)	Percentage Agreement
Hawkins-Kennedy	.39 (.12-.65)	69
Neer	.40 (.13-.67)	71
Painful arc	.45 (.18-.72)	73
Empty can (Jobe)	.47 (.22-.72)	76
External rotation resistance	.67 (.40-.94)	87

NOTE. n=55.

calculated for each test by calculating a kappa coefficient and percent agreement.

A second aim was to determine the diagnostic accuracy for the 5 physical examination tests. The operative findings were used as the reference standard, and the patients were classified as positive or negative for SAIS based on the surgical findings. An ROC curve analysis for each physical examination test was used to calculate the AUC, which represents the probability that the test can discriminate between healthy and disease states. The AUC values range from 0 to 1, with an AUC of 1 indicating 100% probability that a given test can discriminate between healthy and SAIS.²¹ The AUC, sensitivity, specificity, +LR and -LR, and the corresponding 95% CI were calculated for each test.^{22,23} Prevalence was calculated by dividing the number of patients positively confirmed via the reference standard with SAIS by the total number of patients.

A third aim was to determine the ability of a combination of tests to diagnose the presence or absence of SAIS. A binary logistic regression analysis was used to determine the parsimonious set of predictor variables (test cluster) to predict the presence or absence of SAIS. This method was used to control for the shared variance among the tests, thus allowing for the independent value of each of the SAIS tests when applied in combination with all of the other SAIS tests. To determine the optimal specific test cluster to confirm SAIS, a forward stepwise logistic regression was performed by entering only those tests with a +LR greater than or equal to 2.0 because those tests were likely to result in an important increase in the likelihood of the presence of SAIS.²⁴ To determine if there was an optimal specific test cluster to rule out SAIS, tests with -LR less than or equal to .50 were entered into a forward stepwise logistic regression. Liberal criterion for entrance of .15 and exit of .20 was used for all binary logistic regression analysis. A second analysis, an ROC curve analysis, was performed to determine if there was a threshold of the number of positive tests that discriminated between patients with and without

SAIS. If the AUC was significant, the point nearest the uppermost left-hand corner of the graph was identified visually, which is the cut point of the number of positive tests for distinguishing between patients with and without SAIS using a combination of tests.

An a priori power analysis was performed to determine sample size needed for the binary logistic regression analysis. We estimated the R^2 among the variables to be .20 from a pilot study on 10 subjects and a prior study using binary logistic regression for diagnosing SAIS.¹² Sample size calculations based on an alpha of .05, 5 predictor variables, effect size (f^2) of .25, and power of .80 indicated a required sample size of 57 subjects.

RESULTS

The reference standard surgical findings are summarized in table 1. The prevalence of SAIS was 29%; 16 of 55 patients were confirmed via the gold standard of surgical findings with SAIS either in isolation as the primary diagnosis or in combination with another glenohumeral joint diagnosis. The 39 subjects who did not have a confirmed diagnosis of SAIS were diagnosed via surgical findings with (in order of frequency) glenohumeral instability, glenoid labral tear, rotator cuff tear, acromion-clavicular joint arthritis, and adhesive capsulitis. There were no missing values for the sample of 55 subjects for either of the 2 clinicians for the 5 impingement tests or for the surgical reference standard. The intertester reliability kappa coefficients for the 2 clinicians ranged from .39 to .67, and the percent agreement ranged from 69% to 87% (table 2). The strength of agreement ranged from moderate to substantial except for the Hawkins-Kennedy and Neer Tests with a kappa of .39 and .40, respectively, according to the scale described by Landis and Koch.²⁵

The AUC result from the ROC curve analysis for each of the 5 tests was significant ($P \leq .05$), except for the Hawkins-Kennedy test. The values for the AUC, sensitivity, specificity, and +LR and -LR values are reported in table 3 for each of the 5 tests.

To determine the optimal test combination for ruling in SAIS, a binary logistic regression analysis was performed by using a forward stepwise method. The tests with +LR greater than or equal to 2.0 entered into the regression analysis were the painful arc, external rotation resistance test, and empty can tests. The regression analysis yielded no optimal combination of tests for diagnosing SAIS. Similarly, to determine an optimal test combination for screening SAIS, tests with -LR less than or equal to .50 were entered into a forward stepwise binary logistic regression analysis. The painful arc, external rotation resistance test, and Neer were entered, resulting in no optimal test combination for ruling out SAIS.

Table 3: Diagnostic Accuracy for Impingement Shoulder Tests

Test	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	-LR (95% CI)	AUC (95% CI)
Hawkins-Kennedy	.63 (.39-.86)	.62 (.46-.77)	1.63 (.94-2.81)	.61 (.31-1.20)	.62 (.46-.79) <i>P</i> =.17
Neer	.81 (.62-1.0)	.54 (.38-.69)	1.76 (1.17-2.66)	.35 (.12-.97)	.68 (.53-.83) <i>P</i> =.04
Painful arc	.75 (.54-.96)	.67 (.52-.81)	2.25 (1.33-3.81)	.38 (.16-.90)	.71 (.56-.86) <i>P</i> =.02
Empty can (Jobe)	.50 (.26-.75)	.87 (.77-.98)	3.90 (1.50-10.12)	.57 (.35-.95)	.69 (.52-.85) <i>P</i> =.03
External rotation resistance	.56 (.32-.81)	.87 (.77-.98)	4.39 (1.74-11.07)	.50 (.28-.89)	.72 (.55-.88) <i>P</i> =.01

NOTE. n=55 subjects.

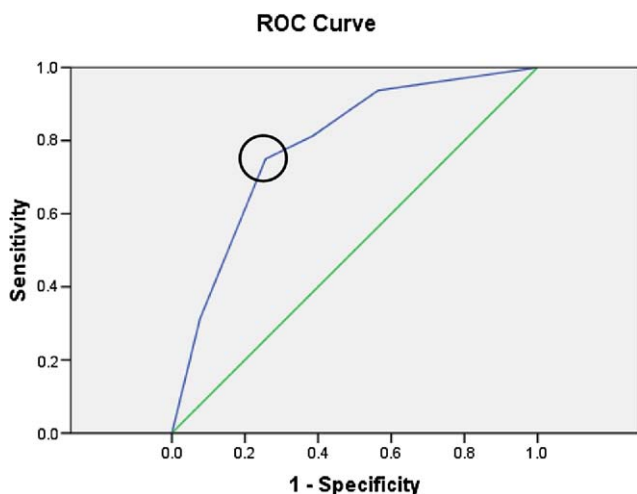


Fig 1. The ROC curve for determining the cut point for the number of the 5 impingement tests to confirm and screen for SAIS. The circle indicates the point nearest the uppermost left-hand corner of the graph, which represents the cut point. The cut point is 3+ of 5 tests for discrimination of the presence or absence of SAIS using 5 impingement tests. The 5 tests were the Hawkins-Kennedy, Neer, painful arc, empty can (Jobe), and external rotation resistance tests.

The ROC curve analysis to determine the threshold for the number of positive tests to discriminate between patients with and without SAIS was significant (AUC=.79, 95% CI, .66–.92; $P=.001$) (fig 1). The cut point for discrimination was 3 positive tests out of 5; the associated diagnostic accuracy values for this cut point is depicted in table 4.

To represent the meaningfulness of each of the tests individually and the combination of tests from the ROC curve analysis, the posttest probabilities were calculated by using the respective +LR and –LR. The pretest probability for this study was 29%, and this was used for the calculations. Table 5 depicts the posttest probabilities for diagnosing SAIS, and the posttest probabilities for screening SAIS.

DISCUSSION

To effectively treat patients with shoulder pain, they must first be accurately diagnosed. The physical examination tests used to diagnose shoulder patients should be reliable and accurate. This study has provided clinically useful information to diagnose SAIS. The interrater reliability was fair to substantial for the 5 tests. The painful arc, empty can, and external rotation resistance test individually and the test combination of any 3 positive tests out of 5 have the best ability to confirm SAIS, with small to moderate shifts in pretest to posttest probability. Tests useful for screening to rule out SAIS, the Neer, painful arc, external rotation resistance test, and the test combination of less than 3 positive of the 5 tests have the best ability to rule out SAIS, with small shifts in pretest to posttest probability. The Hawkins-Kennedy test as an

individual test was not able to discriminate between those with and without SAIS (AUC=.62, $P=.17$).

The interrater reliability kappa values indicate that 3 tests, the painful arc, empty can, and external rotation resistance test, have moderate to substantial strength of agreement. Prior studies of reliability agree with our findings for the painful arc of moderate to substantial agreement^{17,18} and for the empty can of moderate^{16,17} to almost perfect²⁰ strength of agreement. For the external rotation resistance test, the level of agreement was higher in this study than in prior studies.¹⁷ The Hawkins-Kennedy and Neer tests had only a fair strength of agreement. Previous studies investigating the reliability of the Neer and Hawkins-Kennedy tests vary widely, with the Neer test showing poor to almost perfect agreement¹⁸⁻²⁰ and the Hawkins-Kennedy showing fair to almost perfect agreement.^{17,19,20}

Sensitivity and specificity values provide an estimate of diagnostic ability. Likelihood ratios combine the sensitivity and specificity values as a ratio and quantify the posttest probability of the disease with use of a given test. Using likelihood ratios allows the clinician to interpret the test findings for individual patients by calculating the shift in pretest to posttest probability of SAIS. Generally, the +LR and –LR can be interpreted according to the criteria set forth by Jaeschke et al,²⁴ whereby using a test to confirm a diagnosis (eg, ruling in the disease) with a +LR greater than or equal to 2.0 will result in small but sometimes important increases in the likelihood of the disease. Conversely, using a test to rule out the disease that has a –LR less than or equal to .50 will result in small but sometimes important decreases in the likelihood of the disease.

Three tests were found to be useful to confirm the diagnosis of SAIS. Specifically, the painful arc, external rotation resistance, and empty can tests had significant AUC with +LRs ranging from 2.25 to 4.39, which produce small to moderate shifts in pretest to posttest probability of the presence of SAIS. Although these shifts in probability are considered small to moderate, they nonetheless may be helpful in confirming the diagnosis of SAIS. Previous studies and a systematic review^{12,16,26} are in agreement with our findings for the external rotation resistance and empty can tests, indicating diagnostic accuracy values useful for confirming SAIS. However, the painful arc has shown only limited usefulness for confirming SAIS.^{12,14,26}

Tests useful for screening to rule out SAIS, the Neer, painful arc, and external rotation resistance tests are the best of the 5 tests with significant AUC and –LRs ranging from .35 to .50. These tests yield small but likely clinically useful shifts in pretest to posttest probability. Our study indicates a somewhat stronger ability of the external rotation resistance test and painful arc to screen for SAIS than prior studies.^{12,14} Prior evidence for the Neer test consistently indicates usefulness for screening for SAIS,¹²⁻¹⁵ but pooled diagnostic values from a systematic review indicate limited utility of this test.²⁶ The Neer test also had the lowest level of reliability. Methodologic differences may account for the contradictory findings with prior studies. These include a lack of clinician blinding to the test results when determining the reference standard, differ-

Table 4: Diagnostic Accuracy for Any Test Combination From the ROC Curve Analysis

	Sensitivity (95% CI)	Specificity (95% CI)	+LR (95% CI)	–LR (95% CI)	AUC (95% CI)
Any test combination	.75 (.54–.96)	.74 (.61–.88)	2.93 (1.60–5.36)	.34 (.14–.80)	.79 (.66–.92)
Cut point: 3+ of 5 tests					$P=.001$

NOTE. 3+ of 5 tests is the cut point for the discrimination of the presence or absence of SAIS using 5 impingement tests. The 5 tests were the Hawkins-Kennedy, Neer, painful arc, empty can (Jobe), and external rotation resistance. Total subjects n=55.

Table 5: Posttest Probabilities for Confirming and Ruling Out the Diagnosis of SAIS for Each Test and Combination of 3+ Tests of 5 Tests

Test	Pretest Probability (%)	+LR (95% CI)	+LR Posttest Probability (%)	-LR (95% CI)	-LR Posttest Probability (%)
Hawkins-Kennedy	29	1.63 (.94–2.81)	39.96	.61 (.31–1.20)	19.95
Neer	29	1.76 (1.17–2.66)	41.82	.35 (.12–.97)	12.50
Painful arc	29	2.25 (1.33–3.81)	47.89	.38 (.16–.90)	13.44
Empty can (Jobe)	29	3.90 (1.5–10.12)	61.43	.57 (.35–.95)	18.88
External rotation resistance	29	4.39 (1.74–11.07)	64.20	.50 (.28–.89)	16.96
Test combination: $\geq 3+$ tests/5 tests	29	2.93 (1.60–5.36)	54.40	—	—
Test combination: $< 3+$ tests/5 tests	29	—	—	.34 (.14–.80)	12.10

NOTE. n=55.

ences in positive criteria for the empty can and ERRT, and differences in the reference gold standard. In particular, Calis et al¹⁴ confirmed the diagnosis of SAIS with a 10-mL lidocaine injection into the subacromial space (impingement test), which is considered the clinical confirmation of SAIS. Additionally, the presence of additional noncuff shoulder pathologies varied across studies, and the spectrum of rotator cuff disease varied to include SAIS only or the entire spectrum of rotator cuff disease of SAIS to full-thickness rotator cuff tears.

A diagnostic process involves the integration of multiple findings from the history and physical examined to arrive at a diagnosis. Studies are needed to provide evidence as to the likelihood of a disorder based on the combination of history and clinical examination items. In this study, test combinations were examined in combination by 2 methods. A binary logistic regression analysis was used to determine if there was a specific set of tests that would be useful to rule in or rule out SAIS. Interestingly, no combination of tests was revealed to predict the absence or presence of SAIS. Only a single prior retrospective study¹² identified a combination of tests, the external rotation resistance test, Hawkins test, and a painful arc test, as useful to confirm and screen for SAIS with a +LR of 10.56 and -LR of .17. Our analysis did not confirm these results of this specific combination.

A second method was used to determine if any combination of tests would be useful in determining the presence or absence of SAIS. For any combination of positive test results, the ROC curve analysis indicated a cut point of 3 tests, which means that 3 or more of 5 tests positive discriminated between those with and without SAIS. A +LR of 2.93 for this test combination increases the likelihood that a patient has SAIS. Conversely, less than 3 positive tests yielded a -LR of .34, decreasing the likelihood of SAIS. The diagnostic utility of this combination of tests is approximately equal to that of the prior indicated single tests for confirming and screening for SAIS. This is the first study to determine a cut point to establish the meaningfulness of a test combination. A prior study¹⁴ found that the combination of 5 positive of 7 tests had the best +LR (+LR=2.75), with only a corresponding small shift in posttest probability; while screening for SAIS, less than 3 positive tests of 7 had the lowest -LR (-LR=.36), indicating a small shift in probability. An ROC curve analysis was not performed, so it unknown if these results indicate a significant ability to discriminate between the presence and absence of SAIS.¹⁴

Study Limitations

This study has several limitations. Other shoulder pathologies may influence the ability of the tests to confirm or rule out SAIS. However, the presence of other shoulder pathologies is a clinical reality, thereby allowing for generalizability of our

results to similar patients presenting with shoulder pain. Future studies should examine the ability of tests to diagnose SAIS in the presence and absence of other shoulder pathologies. The confidence intervals of the +LR and -LR point estimates are wide, indicating less precision in estimating the shift in pretest to posttest probability of the disease. Lastly, a combination of history and physical examination tests is likely more powerful than any single test for diagnosing SAIS. We did not include all potential history questions, impairments, and tests used to diagnose SAIS. Further studies investigating all tests used to diagnoses SAIS are needed to determine a parsimonious set of clinical examination items to confirm and screen for SAIS.

CONCLUSIONS

The results of this study provide new and expanded evidence as to the ability of 5 commonly used clinical tests. The single tests of painful arc, external rotation resistance test, and empty can provide the best diagnostic utility and reliability. The Neer test has clinical utility to screen for SAIS but has only fair reliability. Also of diagnostic utility is the use of the cut point of 3+/5 tests, with 3 or more tests positive of 5 useful in confirming SAIS, whereas less than 3 positive of the 5 tests is helpful in decreasing the likelihood of SAIS.

References

- Wofford JL, Mansfield RJ, Watkins RS. Patient characteristics and clinical management of patients with shoulder pain in U.S. primary care settings: secondary data analysis of the National Ambulatory Medical Care Survey. *BMC Musculoskelet Disord* 2005;6:4.
- Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population: a systematic review. *Scand J Rheumatol* 2004;33:73-81.
- Ostor AJ, Richards CA, Prevost AT, Speed CA, Hazleman BL. Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology (Oxford)* 2005;44:800-5.
- van der Windt DA, Koes BW, de Jong BA, Bouter LM. Shoulder disorders in general practice: incidence, patient characteristics, and management. *Ann Rheum Dis* 1995;54:959-64.
- Vecchio P, Kavanagh R, Hazleman BL, King RH. Shoulder pain in a community-based rheumatology clinic. *Br J Rheumatol* 1995;34:440-2.
- Ejnisman B, Andreoli CV, Soares BG, et al. Interventions for tears of the rotator cuff in adults. *Cochrane Database Syst Rev* 2004;(1):CD002758.
- Michener LA, Walsworth MK, Burnet EN. Effectiveness of rehabilitation for patients with subacromial impingement syndrome: a systematic review. *J Hand Ther* 2004;17:152-64.

8. Neer CS. Impingement lesions. *Clin Orthop Relat Res* 1983; March(173):70-7.
9. Hawkins RJ, Kennedy JC. Impingement syndrome in athletes. *Am J Sports Med* 1980;8:151-8.
10. Kessel L, Watson M. The painful arc syndrome. Clinical classification as a guide to management. *J Bone Joint Surg Br* 1977; 59:166-72.
11. Jobe FW, Moynes DR. Delineation of diagnostic criteria and a rehabilitation program for rotator cuff injuries. *Am J Sports Med* 1982;10:336-9.
12. Park HB, Yokota A, Gill HS, El RG, McFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. *J Bone Joint Surg Am* 2005;87:1446-55.
13. MacDonald PB, Clark P, Sutherland K. An analysis of the diagnostic accuracy of the Hawkins and Neer subacromial impingement signs. *J Shoulder Elbow Surg* 2000;9:299-301.
14. Calis M, Akgun K, Birtane M, Karacan I, Calis H, Tuzun F. Diagnostic values of clinical diagnostic tests in subacromial impingement syndrome. *Ann Rheum Dis* 2000;59:44-7.
15. Leroux JL, Thomas E, Bonnel F, Blotman F. Diagnostic value of clinical tests for shoulder impingement syndrome. *Rev Rhum Engl Ed* 1995;62:423-8.
16. Holtby R, Razmjou H. Validity of the supraspinatus test as a single clinical test in diagnosing patients with rotator cuff pathology. *J Orthop Sports Phys Ther* 2004;34:194-200.
17. Ostor AJ, Richards CA, Prevost AT, Hazleman BL, Speed CA. Interrater reproducibility of clinical tests for rotator cuff lesions. *Ann Rheum Dis* 2004;63:1288-92.
18. Nomden JG, Slagers AJ, Bergman GJ, Winters JC, Kropmans TJ, Dijkstra PU. Interobserver reliability of physical examination of shoulder girdle. *Man Ther* 2009;14:152-9.
19. Nørregaard J, Krogsgaard MR, Lorenzen T, Jensen EM. Diagnosing patients with longstanding shoulder joint pain. *Ann Rheum Dis* 2002;61:646-9.
20. Johansson K, Ivarson S. Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. *Man Ther* 2009;14:231-9.
21. Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. New York: Wiley-Intersciences; 2002.
22. Roux C, Priol G, Fechtenbaum J, Cortet B, Liu-Léage S, Audran M. A clinical tool to determine the necessity of spine radiography in postmenopausal women with osteoporosis presenting with back pain. *Ann Rheum Dis* 2007;66:81-5.
23. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean J Radiol* 2004;5:11-8.
24. Jaeschke R, Guyatt GH, Sackett DL. *Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group.* *JAMA* 1994;271:703-7.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
26. Hegedus EJ, Goode A, Campbell S, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med* 2008;42:80-92.