

## CONCISE REPORT

# Diagnosing patients with longstanding shoulder joint pain

J Nørregaard, M R Krogsgaard, T Lorenzen, E M Jensen

*Ann Rheum Dis* 2002;61:646–649

**Objective:** To examine the interobserver agreement of commonly used clinical tests and diagnoses in patients with shoulder pain, and the accuracy of these tests and ultrasonographic findings in comparison with arthroscopic findings.

**Methods:** Eighty six patients with longstanding shoulder joint pain were “blindly” examined by two trained doctors using several clinical tests. In all patients an ultrasonographic examination was performed, and in 42 (49%) an arthroscopy.

**Results:** Tests for impingement showed poor to moderate agreement. Tenderness of muscles, muscle weakness, and tests for labral lesion also showed poor agreement. Pain during muscle contraction showed moderate agreement. The agreement of clinical diagnoses was poor and the accuracy was low in comparison with arthroscopy. Ultrasonography was accurate in full thickness supraspinatus tendon tears, but inaccurate for partial tears and labral lesions.

**Conclusions:** Most clinical tests showed poor agreement. Clinical and ultrasonographic diagnoses had low accuracy in comparison with arthroscopy.

More research is needed to establish uniform methods of defining shoulder disorders.<sup>1</sup> Good agreement should be obtainable when different clinicians use clinical classification criteria of shoulder problems, and if possible these criteria should be validated—for example, by arthroscopy.

In a study in which diagnosis of shoulder pain was intended, only three clinical diagnoses emerged from cluster analysis<sup>2</sup>: one group with movement restriction (synovial group), a small group with restriction in all directions (frozen shoulder), and a group without restriction of shoulder joint movements (shoulder girdle group).

Even though most shoulder problems resolve in a few weeks to months, longstanding shoulder problems are common.<sup>3</sup> Many clinical tests are used to establish the different shoulder diagnoses (impingement/tendinitis, different tendon lesions, acromioclavicular disease, labral lesions). These are often supplemented by magnetic resonance imaging or an ultrasonographic examination, both of which have been found to be accurate for diagnosing full thickness rotator cuff tears.<sup>4,5</sup> We wanted to find answers to the following questions about the process of diagnosing chronic shoulder pain:

- What is the interobserver agreement of historical information, clinical tests, and the final diagnosis?
- To what extent are clinical and ultrasonographic diagnoses accurate in comparison with arthroscopic diagnoses as the “gold standard”?

## METHODS

The study group comprised consecutive patients with shoulder pain who fulfilled the following inclusion criteria:

- Clinical suspicion of labral or rotator cuff lesion or
- Shoulder pain originating from the shoulder joint/tendons, with a duration of at least two months (severe pain), three months (moderate pain), or six months (slighter pain). Steroid injections should have been tried without lasting effect or found not to be indicated. This primary evaluation was based upon history, symptoms, and tests as described below and was performed by trained specialists.

Exclusion criteria included diffuse pain syndromes, disc, joint or nerve disease, severe medical diseases, psychiatric disease, pregnancy or lactation, previous shoulder operation, or fracture.

## Study plan

Consecutive patients from the participating departments were included. An ultrasonographic examination of the shoulder was first performed by one of two experienced ultrasonographers who was unaware of the clinical condition of the patient. Thereafter, patients were examined clinically by an orthopaedic surgeon and a rheumatologist randomly, both having long experience with shoulder disorders. Before starting the study each clinical test was carefully discussed to optimise interobserver agreement.

The clinical examinations were performed on the same day and were initially performed without prior knowledge of the patient's history. History and symptoms were then registered by a standardised interview. After this, the examiners were given the result of the ultrasonography. After each step the examiners registered suggestions for the diagnosis. Then the two clinicians discussed their findings and decided on a diagnosis and a treatment plan. Arthroscopy was found to be indicated in 42 patients.

## Clinical examination

Range of motion was clinically assessed in the standing position without the use of a goniometer for abduction in the plane of scapula (scaption), and flexion (active and passive) and external rotation. Tenderness was assessed by manual pressure at the muscles and tendons of the shoulder—pressure was not standardised. Pain and weakness during isometric contraction was assessed for the abduction elevation, internal and external rotation, and elbow supination. Impingement tests were carried out by active and passive flexion and scaption.<sup>6</sup> A 90 degrees flexion or scaption movement was combined with internal rotation (Hawkins test).<sup>7</sup> These impingement tests were regarded as positive if they elicited pain. The impingement release test was performed by pressing the humerus head downward during abduction.<sup>8</sup>

A number of instability/labrum tests were carried out. An apprehension test was performed with the patient in a supine position.<sup>9</sup> The relocation test was performed by applying a posteriorly directed force during a supine apprehension test.<sup>9</sup>

**Table 1** Historical information obtained by the two clinicians by use of the standardised questionnaire (number indicates percentage positive) and the agreement between observers ( $\kappa$  and asymptotic standard error)

	Rheumatologist	Orthopaedic surgeon	Agreement
<b>Mechanism of symptom start</b>			
Trauma with acute pain	40	40	0.75(0.07)
Dislocation	2	2	0.49(0.35)
Fall on stretched arm	6	7	0.90(0.21)
Fall on the shoulder	15	10	0.58(0.14)
Heavy load/pull/throw	18	8	0.39(0.14)
Repeated loads	36	49	0.51(0.08)
Activity over the shoulder	20	23	0.62(0.11)
<b>Where do you feel the pain?</b>			
Deep in the shoulder	62	72	0.15(0.08)
Shoulder upper side	21	60	0.09(0.07)
In front of the shoulder	30	42	0.15(0.08)
In back of the shoulder	28	24	0.49(0.10)
At the clavicle	3	10	0.30(0.20)
Neck	13	8	0.38(0.16)
Upper arm anterior	15	17	0.40(0.12)
Upper arm lateral	35	47	0.69(0.08)
Upper arm posterior	8	7	0.42(0.19)
Elbow/forearm/hand	16	17	0.71(0.12)
Axillary region	2	1	-0.02(0.40)
<b>Symptom description</b>			
Pain by lifting/throwing	87	92	0.26(0.16)
Dead-arm syndrome throwing	14	13	0.55(0.14)
Feeling of locking of the shoulder	20	16	0.57(0.12)
Resting pain	64	66	0.54(0.08)
Nocturnal pain	74	76	0.72(0.10)

The anterior release test was performed by pressing the humerus backward in the supine position followed by a release of the force.<sup>10</sup> Crank and anterior slide tests were performed by applying a force in the axis of the humerus.<sup>11</sup> In these tests laxity, click, and pain were registered. The "load-shift" test was used to examine anterior and posterior laxity by pulling and pushing the humerus head forwards and backwards. The sulcus test was performed by applying a firm downward pressure on the arm during slight abduction.

Abnormal subluxation and tenderness of the acromioclavicular and sternoclavicular joints was registered. Laxity was registered by applying a downward pressure on the shaft of the clavicle irrespective of whether arm adduction or scapula elevation provoked pain from the joints.

For ultrasonography a Siemens Versa-pro equipment was used with a 7.5 MHz linear array and a curvilinear 5 MHz transducer. Rotator cuff tendons, the posterior labrum, and the subacromial bursa were examined with the patients sitting.<sup>3</sup> The anterior labrum was assessed with the patients supine.<sup>12</sup>

Anterior and posterior portals were used for arthroscopy, and the humeroscapular joint was examined systematically in all patients.

### Statistics

Interobserver agreement for assessment of the range of motion (degrees) was assessed by calculation of the standard deviation of the difference.<sup>13</sup> All other clinical tests were scored as either positive or negative. Interobserver agreement (reliability) of the different tests was determined by  $\kappa$  statistics.<sup>14</sup>  $\kappa$  Values <0.4 indicate low agreement, values 0.4–0.6 moderate agreement, and values >0.6 good agreement. The clinical diagnoses and ultrasonography were related to the arthroscopic findings, allowing calculation of the sensitivity, specificity, and accuracy.

### Ethics

The project was approved by the local ethics committee, and informed consent was obtained from the patients.

### RESULTS

Eighty six patients (52% women), with a mean symptom duration of 25 months, were included in the study.

Agreement of the responses to different interview questions was variable (table 1). There was generally low agreement between the two observers for clinical tests (table 2). The results of some tests are not shown in table 2: (a) The orthopaedic surgeon generally scored labrum tests positive 1.3–3 times more often than the rheumatologist did. All showed poor agreement ( $\kappa < 0.4$ ); (b) tests for muscle weakness and abnormality of the acromioclavicular and sternoclavicular joints all showed poor agreement (all  $\kappa < 0.4$ ).

The interobserver agreement of diagnoses based on clinical examination alone was poor (all  $\kappa < 0.4$ ) and did not improve when the patient's history was known.

### Ultrasonography and arthroscopy

The results of ultrasonography were full thickness supraspinatus rupture in four patients, partial rupture in seven, anterior labrum lesion in five, infraspinatus rupture in one. In 57 (66%) no lesions were registered. Arthroscopy showed supraspinatus full thickness rupture in three patients, partial rupture in four, infraspinatus tendon rupture in two, anterior labral tears in eight, and superior labral tears in three.

### Accuracy of clinical tests and diagnoses compared with arthroscopy

Only the total supraspinatus ruptures had reasonable sensitivity (67%) and a specificity of 90% when examined by the surgeon (accuracy 79%). The accuracy of the diagnoses of partial supraspinatus ruptures and labral lesions and all the rheumatologist's diagnoses was <65%. The accuracy was not improved by adding information of the patient's history, and only marginally higher when the result of the ultrasonography was available, and only for supraspinatus ruptures.

### Accuracy of clinical ultrasonography

Ultrasonography was reasonably accurate for full thickness rotator cuff tears (two out of three, sensitivity 67%)—

**Table 2** Performance of the clinical test. For “mobility”, data are presented as degrees of movement and SD of the difference as a measure of the agreement. For other variables the percentage of positive clinical tests is presented and the agreement between observers determined by  $\kappa$  with asymptotic standard error (ASE)

	Rheumatologist	Orthopaedic surgeon	Measure of agreement
Mobility (degrees)			SD of difference
Abduction actively	151	158	21
Abduction passively	163	166	12
Flexion active	154	162	20
Flexion passive	163	167	12
External rotation	64	60	20
			$\kappa$ (ASE)
Painful arch	29	66	0.14 (0.07)
Tenderness			
Supraspinatus muscle	34	33	0.29(0.09)
Supraspinatus tendon	28	35	0.25(0.09)
Infraspinatus muscle	28	17	0.51(0.10)
Infraspinatus tendon	19	16	0.35(0.12)
Trapezius biceps	21	14	0.28(0.12)
Biceps tendon	8	23	-0.05(0.12)
Pain by			
Abduction	53	51	0.43(0.08)
Internal rotation	14	17	0.43(0.13)
External rotation	43	43	0.67(0.08)
Supination	6	26	0.14(0.12)
Fast passive flexion	47	49	0.28(0.08)
Flexion, internal rotation	41	42	0.07(0.08)
Scaption, internal rotation	45	30	0.40(0.08)
Impingement relief	22	25	0.34(0.10)
Laxity			
Anterior drawer	12	8	0.67(0.16)
Sulcus sign	12	23	0.13(0.12)

specificity 100%. The sensitivity of ultrasonography for detecting partial thickness tears and labral tears was <35% for both.

## DISCUSSION

The historical information, obtained by standardised interview by the two examiners, had low  $\kappa$  values. The  $\kappa$  value was >0.4 for only a few of the diagnostic tests; this is usually regarded as the lowest acceptable level. Only “Hawkins test”, “pain by abduction/internal rotation/external rotation”, “tenderness in the infraspinatus muscle”, and “anterior drawer” had acceptable agreement, while the agreement between the two observers for other tests was poor. Of the range of motion tests passive abduction and flexion had the best agreement. The two examiners had gone through each test carefully before starting the study to make sure that they performed them similarly. Therefore, the poor agreement for the result of the tests is probably not caused by completely different ways of performing the tests, but may be partly explained by differences in interpreting the patients’ responses to the tests. The different clinical backgrounds (rheumatology, orthopaedic surgery) may also have influenced the results.

One other study of interobserver agreement had results comparable with ours.<sup>15</sup> In that study a long duration of symptoms was associated with disagreement on diagnoses. It should be emphasised that patients with primarily shoulder girdle pain or frozen shoulder were not included in our study. Higher interobserver agreement would probably be obtainable if these patient groups were included in the study and if fewer diagnostic groups were used.<sup>2</sup>

The accuracy of clinical tests and diagnoses in comparison with the arthroscopic findings was low and only slightly better when the results of ultrasonographic evaluation became available in addition to the clinical examination. Better results for both agreement and accuracy of clinical and ultrasonographic diagnosis are probably obtainable in patients with a shorter history and with more clearly defined conditions—for

example, in patients with traumatic shoulder instability or full thickness tears of the rotator cuff.

## ACKNOWLEDGEMENTS

We thank the Danish Rheumatism Association for financial support.

## Authors’ affiliations

**J Nørregaard, T Lorenzen, E M Jensen**, Department of Rheumatology, Copenhagen University Hospital, Bispebjerg, Denmark  
**M R Krogsgaard**, Department of Orthopaedic Surgery, Copenhagen University Hospital, Bispebjerg, Denmark

Correspondence to: Dr J Nørregaard, Bomlærkevej 14, DK-2970 Hørsholm, Denmark; j.norregaard@dadlnet.dk

Accepted 28 January 2002

## REFERENCES

- Green S, Buchbinder R, Glazier R, Forbes A. Systematic review of randomised controlled trials of interventions for painful shoulder: selection criteria, outcome assessment, and efficacy. *BMJ* 1998;316:354–60.
- Winters JC, Groenier KH, Sobel JS, Arendzen HH, Meyboom-de Jongh B. Classification of shoulder complaints in general practice by means of cluster analysis. *Arch Phys Med Rehabil* 1997;78:1369–74.
- MacFarlane GJ, Hunt IM, Silman AJ. Predictors of chronic shoulder pain: a population based prospective study. *J Rheumatol* 1998;25:1612–15.
- Robertson PL, Schweitzer ME, Mitchell DG, Schlesinger F, Epstein RE, Frieman BG, et al. Rotator cuff disorders: interobserver and intraobserver variation in diagnosis with MR imaging. *Radiology* 1995;194:831–5.
- Teefey SA, Hasan SA, Middleton WD, Patel M, Wright RW, Yamaguchi K. Ultrasonography of the rotator cuff. A comparison of ultrasonographic and arthroscopic findings in one hundred consecutive cases. *J Bone Joint Surg Am* 2000;82:498–504.
- Neer CS. Impingement lesions. *Clin Orthop* 1973;70–7.
- Hawkins RJ, Kennedy JC. Impingement syndrome in athletes. *Am J Sports Med* 1980;8:151–8.
- Corso G. Impingement relief test: an adjunctive procedure to traditional assessment of shoulder impingement syndrome. *J Orthop Sports Phys Ther* 1995;22:183–92.
- Speer KP, Hannafin JA, Altchek DW, Warren RF. An evaluation of the shoulder relocation test. *Am J Sports Med* 1994;22:177–83.

- 10 **Gross ML**, Distefano MC. Anterior release test. A new test for occult shoulder instability. *Clin Orthop* 1997;339:105-8.
- 11 **Liu SH**, Henry MH, Nuccion S. A prospective evaluation of a new physical examination in predicting glenoid labral tears. *Am J Sports Med* 1996;24:721-5.
- 12 **Schydrowsky P**, Strandberg C, Galbo H, Krogsgaard M, Jørgensen U. The value of ultrasonography in the diagnosis of labral lesions in patients with anterior shoulder dislocation. *Eur J Ultrasound* 1998;8:107-13.
- 13 **Bland JM**, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* 1986;i:307-10.
- 14 **Silman A**. *Epidemiological studies - a practical guide*. Cambridge: Cambridge university Press; 1995.
- 15 **de Winter AF**, Jans MP, Scholten RJ, Deville W, van Schaardenburg D, Bouter LM. Diagnostic classification of shoulder disorders: interobserver agreement and determinants of disagreement. *Ann Rheum Dis* 1999;58:272-7.

Want full text but don't have  
a subscription?

Pay per view

For just \$8 you can purchase the full text of individual articles using our secure online ordering service. You will have access to the full text of the relevant article for 48 hours during which time you may download and print the pdf file for personal use.

[www.annrheumdis.com](http://www.annrheumdis.com)