

# The effectiveness of diagnostic tests for the assessment of shoulder pain due to soft tissue disorders: a systematic review

J Dinnes\*  
E Loveman  
L McIntyre  
N Waugh



Southampton Health Technology Assessments Centre,  
University of Southampton, UK

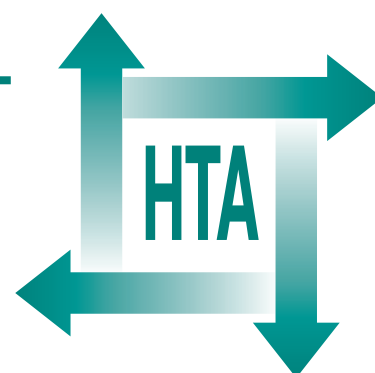
\*Corresponding author

**Declared competing interests of authors:** none

## *Executive summary*

*Health Technology Assessment 2003; Vol. 7: No. 29*

Health Technology Assessment  
NHS R&D HTA Programme





## Executive summary

### Background

Shoulder pain is a significant cause of morbidity; the prevalence of self-reported pain is estimated to be between 16 and 26%, and it is the third most common cause of musculoskeletal consultation in primary care. The cause can be difficult to diagnose owing to the complex anatomy of the shoulder and the spectrum of underlying disorders. Most shoulder problems fall into three major categories: soft tissue disorders, articular injury or instability, and arthritis. The incidence of lesions increases with age as tendon tissue progressively weakens or degenerates, but repeated microtrauma or overuse from professional or athletic activity can also cause soft tissue problems in all age groups.

There are no clear national guidelines for the diagnosis of shoulder pain. Several diagnostic tests are used for the diagnosis of soft tissue disorders, including clinical assessment, ultrasonography, magnetic resonance imaging (MRI), magnetic resonance arthrography (MRA) and arthroscopy, yet their relative accuracy, cost-effectiveness and impact on quality of life are uncertain.

### Objective

To evaluate the evidence for the effectiveness and cost-effectiveness of the newer diagnostic imaging tests as an addition to clinical examination and patient history for the diagnosis of soft tissue shoulder disorders.

### Methods

#### Literature search

Literature was identified from several sources including general medical databases.

#### Inclusion criteria

The primary inclusion criteria for the assessment of test accuracy were studies of clinical examination, ultrasound, MRI or MRA in patients suspected of having soft tissue shoulder disorders. Outcomes assessed were clinical impingement syndrome or rotator cuff tear (RCT) (full, partial or any). Only cohort studies were included.

#### Quality assessment and data extraction

The methodological quality of included test accuracy studies was assessed using a formal

quality assessment tool for diagnostic studies developed by the NHS Centre for Reviews and Dissemination at the University of York. The extraction of study findings was conducted in duplicate using a pre-designed and piloted data extraction form to avoid any errors.

#### Data synthesis

For each test, sensitivity, specificity and positive and negative likelihood ratios (LRs) with 95% confidence intervals were calculated for each study. Where no trade-off between sensitivity and specificity was revealed, and studies were otherwise sufficiently homogeneous, pooled estimates of sensitivity, specificity and LRs were calculated using random effects methods. Potential sources of heterogeneity were investigated by conducting subgroup analyses according to features of the population (spectrum), test, and reference test, and study quality.

### Results

#### Quality of included studies

The prevalence of rotator cuff (RC) disorders was high in most studies (overall mean prevalence over 50% for all tests), although it varied according to the setting and outcome used. The study setting was not always reported, but where it was, only two were conducted in centres other than hospital radiology or orthopaedics departments. Partial verification of patients was common and in many studies patients were selected because they had undergone the reference test. Sample sizes were generally very small, with overall means of less than 100.

The reference tests used in the studies were often inappropriate with many studies (especially ultrasound studies) using arthrography alone, despite problems with its sensitivity. Others used more than one reference test, in some cases clearly stating that the test used was based on the result of the index test.

Few studies reported details of those interpreting the tests other than that they were orthopaedists or radiologists, often specialising in shoulder disorders.

#### Clinical assessment

Ten cohort studies were included: seven examined the accuracy of individual clinical examination tests and six estimated the accuracy of clinical

examination *per se* or the combination of two or more positive test results. Individual tests were either good at ruling out RCTs when negative (high sensitivity) or at ruling in such disorders when positive (high specificity), but small sample sizes mean that there was no conclusive evidence for any single test that can conclusively diagnose RC disorders. Pooled results from four studies that evaluated clinical examination as a whole indicated overall sensitivity and specificity to be 0.90 (95% CI: 0.87 to 0.93) and 0.54 (95% CI: 0.47 to 0.61) for detection of full-thickness RCTs.

### Ultrasound

Thirty-eight cohort studies investigating the accuracy of ultrasound were identified. Ultrasound was most accurate when used for the detection of full-thickness tears, although results were heterogeneous: pooled sensitivity 0.87 (95% CI: 0.84 to 0.89) and specificity 0.96 (0.94 to 0.97). Sensitivity was lower for detection of partial-thickness tears (0.67, 95% CI: 0.61 to 0.73) although specificity remained high, and studies were again very heterogeneous. Statistically, several possible reasons for the differences in sensitivity estimates between studies were identified, including prevalence and mean age. The number of studies available limited the power of the subgroup analyses. It remains to be determined whether or not ultrasound can provide such conclusive evidence for the value of a negative ultrasound finding in ruling out the presence of a tear.

### MRI

Twenty-nine cohort studies were included, most using conventional MRI pulse sequences as opposed to fat-suppressed MRI. For full-thickness tears, overall pooled sensitivities and specificities were high (0.89, 95% CI: 0.86 to 0.92; and 0.93, 95% CI: 0.91 to 0.95, respectively) and the studies were not statistically heterogeneous. For detection of partial-thickness RCTs, pooled sensitivity estimate was much lower (0.44, 95% CI: 0.36 to 0.51) although specificity again remained high (0.90, 95% CI: 0.87 to 0.92). Where tear prevalence is relatively high, a negative magnetic resonance finding may be sufficient to rule out the presence of a full-thickness tear, but between study heterogeneity means that similar conclusions cannot yet be drawn regarding a positive test result.

### MRA

Six studies investigating the accuracy of MRA were included. The type of MRI, views and contrast used varied considerably between studies, making any conclusions difficult. The pooled results suggest that MRA may be very accurate for detection of full-thickness RCTs [overall pooled sensitivity 0.95

(95% CI: 0.82 to 0.98) and specificity 0.93 (95% CI: 0.84 to 0.97), both estimates homogeneous]. Its performance for the detection of partial-thickness tears is less consistent. There is also some suggestion that MRA performs better than ultrasound or MRI, but any such benefit must be set against the invasiveness and potential discomfort to patients of the procedure

### Direct comparisons of tests

Direct evidence for the performance of one test compared with another is very limited. Further research is needed to determine the place of these imaging tests in the diagnosis of RC disorders.

## Conclusions

Our results suggest that clinical examination by specialists can rule out the presence of a RCT, and that either MRI **or** ultrasound could equally be used for detection of full-thickness RCTs. Although still not by any means accurate, ultrasound **may** be better at picking up partial tears. Given the large differential in the cost of the two procedures, the implication from current evidence is that ultrasound is the more cost-effective test to use in a specialist hospital setting for identification of full-thickness tears. Whether or not these results are transferable to settings with lower prevalence, different spectra of disease and less-specialised clinicians, such as in primary care, remains to be determined.

## Implications for further research

There is a need for large, well-designed, prospective studies of the diagnosis of shoulder pain. In particular, a follow-up study of patients with shoulder pain in primary care is needed to inform our understanding of the natural history and epidemiology of shoulder pain and, for those patients referred to secondary care, a prospective cohort study of clinical examination, ultrasound and MRI, alone and/or in combination is also needed. The ability of these tests not only to diagnose the spectrum of soft tissue shoulder disorders (not just RCT) but also to inform treatment decisions remains to be determined.

## Publication

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

# NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

The research reported in this monograph was commissioned by the HTA Programme and was funded as project number 01/39/01. Technology assessment reports are completed in a limited time to inform decisions in key areas by bringing together evidence on the use of the technology concerned.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

## Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA Programme Director:

Professor Kent Woods

Series Editors:

Professor Andrew Stevens, Dr Ken Stein, Professor John Gabbay,  
Dr Ruairidh Milne, Dr Chris Hyde and Dr Rob Riemsma

Managing Editors:

Sally Bailey and Sarah Llewellyn Lloyd

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report.

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2003

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to HMSO, The Copyright Unit, St Clements House, 2-16 Colegate, Norwich, NR3 1BQ.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.

Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.